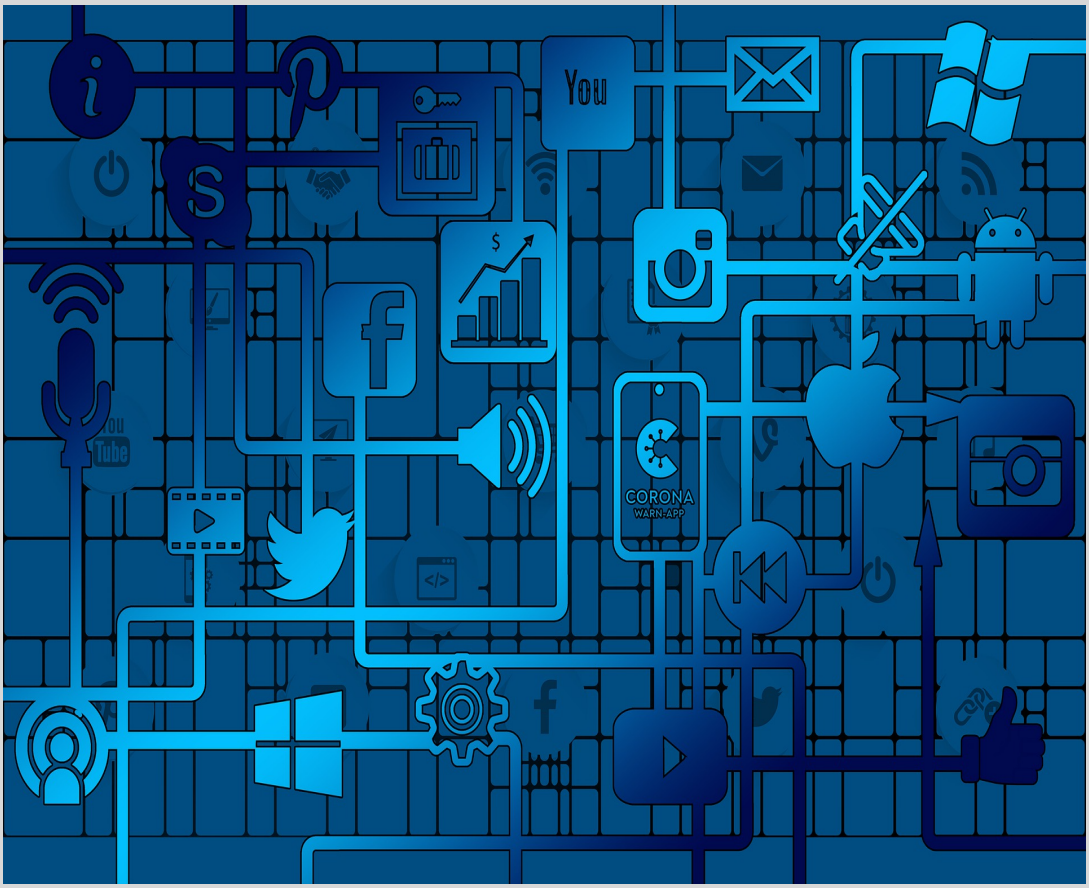
ELECTRONIC WORD-OF-MOUTH (eWOM) AND CUSTOMER OPINIONS IN HOSPITALITY AND TOURISM

Babajide Abubakr Muritala

UNIVERSITAT ROVIRA i VIRGILI

ELECTRONIC WORD-OF-MOUTH (eWOM) AND CUSTOMER OPINIONS IN HOSPITALITY AND TOURISM

Babajide Abubakr Muritala



DOCTORAL THESIS
2022

UNIVERSITAT ROVIRA I VIRGILI
ELECTRONIC WORD-OF-MOUTH (eWOM) AND CUSTOMER OPINIONS IN HOSPITALITY AND TOURISM
Babajide Abubakr Muritala

Babajide Abubakr Muritala

ELECTRONIC WORD-OF-MOUTH (eWOM) AND CUSTOMER OPINIONS IN HOSPITALITY AND TOURISM

DOCTORAL THESIS

PhD Supervisors

Dra. Maria Victòria Sánchez Rebull

Dra. Ana Beatriz Hernández Lara

Department of Business Management

UNIVERSITAT ROVIRA i VIRGILI

REUS, 2022

DEPARTAMENT DE GESTIÓ D'EMPRESES
Universitat Rovira i Virgili
Facultat d'Economia i Empresa

FAIG CONSTAR que aquest treball, titulat **"Electronic word-of-mouth (eWOM) and customer opinions in hospitality and tourism"**, que presenta **Babajide Abubakr Muritala** per a l'obtenció del títol de Doctor, ha estat realitzat sota la meva direcció al **Departament de Gestió d'Empreses** d'aquesta universitat.

HAGO CONSTAR que el presente trabajo, titulado **"Electronic word-of-mouth (eWOM) and customer opinions in hospitality and tourism"**, que presenta **Babajide Abubakr Muritala** para la obtención del título de Doctor, ha sido realizado bajo mi dirección en el **Departamento de Gestión de Empresas** de esta universidad.

I STATE that the present study, entitled **"Electronic word-of-mouth (eWOM) and customer opinions in hospitality and tourism"**, presented by **Babajide Abubakr Muritala** for the award of the degree of Doctor, has been carried out under my supervision at the **Department of Business Management** of this university.

Reus, 14 de maig de 2022/ Reus, 14 de mayo de 2022/ Reus, 14th of May, 2022

María Victoria Sánchez Rebull - DNI 39875081G (TCAT)
Signat digitalment per María Victoria Sánchez Rebull - DNI 39875081G (TCAT)
Data: 2022.05.19 13:33:01 +02'00'

Ana Beatriz Hernandez Lara - DNI 31863206H (TCAT)
Firmado digitalmente por Ana Beatriz Hernandez Lara - DNI 31863206H (TCAT)
Fecha: 2022.05.19 13:53:30 +02'00'

Les directores de la tesis doctoral

Las directoras de la tesis doctoral

Doctoral Thesis Supervisors

La lucha termina cuando comienza la gratitud

~ Neale Donald Walsch

Acknowledgements

I will like to express my sincere gratitude to my supervisors Professor Maria Victòria Sánchez Rebull and Professor Ana Beatriz Hernández Lara for their unwavering support throughout this journey and for going above and beyond what was required for doctoral supervision. I thank you both for giving me the opportunity and selecting me for the fellowship during the COFUND candidate competition, the assistance before and after arrival in Spain, and for the kindness throughout the course of the PhD. Thank you for providing the academic guidance, support, and necessary resources to carry out my research and for giving me the freedom to explore in different directions throughout the PhD. I will forever be grateful to you both for your kindness. Gracias por todo, estoy muy agradecido.

I am grateful to Professor Alexandre Perera Lluna for hosting me at the Universitat Politècnica de Catalunya (UPC) in Barcelona and for giving me access to the B2S laboratory's data infrastructure, which was essential for my research.

I thank Professor Raffaele Scuderi for hosting me at the Kore University of Enna in Italy for my international research secondment and for the ongoing research collaboration.

I thank Ms. Laura Castillejo, the UGAD coordinator in charge of financial arrangements for her continuous support with all required payments during the PhD.

I am grateful to the Faculty of Economy and Business and the Department of Business Management of the URV, from the Dean Antonio Terceño, to the heads of department during my study Ana Beatriz Hernández Lara and Xavier Càmara Turull, the heads of the PhD program in

the faculty during my study Mireia Valverde and Agustí Segarra, the teachers of the PhD bridging courses; Maria Victòria Sánchez Rebull, Mireia Valverde, Gerard Ryan, María Teresa Sorrosal, and all the administrative staff in the Department of Business Management; Cristina Porres Folch, Montserrat Bruno Saltó, and Ramon Puig Gimenez.

I deeply appreciate all my fellow colleagues during the course of the PhD, starting from Dra. Valeria Ferreira and Dr. Ignacio Sánchez, thanks for the guidance and information from the time I arrived till the submission of this thesis through different means, including helping to receive those phone calls in Spanish ☺. I am grateful to have shared the journey with everyone in seminari AS11 and elsewhere; Dra. Maria Boada, Maria Biendicho, Tahereh Maghsoudi, Yaghoub Abdi, Othmane Aride, Marzhan Biesenbina, Génesis Guarimata, Nadima Hassan, Joanna Andraszak, Tasneem Binte Morshed, Samson Fashola, Lucy Victoria Ojeda, Claudia Leyva, and Bilal Afzal. It was a great pleasure knowing and working with you all.

I thank fellow COFUND colleagues like Michael Bamiloshin, Alfonsina Abat, and others for their advices and support during the PhD.

I will like to thank the Martí i Franquès COFUND coordinators Dra. Oana Moldovan, Dra. Paloma Ponton Merino, and the former coordinator Dr. Bogdan Mihai Nae. Thank you all for the constant support.

I thank Ana Benages and Sandra Galves in the human resources department of the URV for all their support along the way.

I am grateful to the European Union and the URV, whose joint generous funding of the Martí i Franquès COFUND scholarship made it possible for me to study with this grant.

Finally, I will like to express my deep gratitude and love to my family for being there right from the beginning up to this point.

Electronic word-of-mouth (eWOM) and customer opinions in hospitality and tourism

Abstract

In the contemporary digital age, electronic word-of-mouth (eWOM) by customers is important in virtually every customer-facing industry. It is even more important in the hospitality and tourism sector, where excellent service is praised and poor service criticized on the Internet via social media or online customer reviews. These comments, reviews, and ratings have an edge over traditional word-of-mouth because they can potentially reach more people, persist for longer, and possibly go viral on the Internet. Hence, they can influence the perception and performance of hospitality and tourism businesses, with customers making purchase decisions based on what their online peers say. The aim of this doctoral thesis was to investigate electronic word-of-mouth (eWOM) and customer opinions in the hospitality and tourism industry. To achieve this aim, a literature review is conducted and electronic word-of-mouth (eWOM) data from social media is collected and analyzed in the context of the COVID-19 pandemic, which broke out during the course of this study and immediately became the foremost issue in the global economy as well as in this industry. This thesis revealed the current state of this research area, the perception of cruise tourism as a result of COVID-19 outbreaks on cruise ships, and the practice of staycations during the first two years of the pandemic. The main conclusions include a reaffirmation of the importance of electronic word-of-mouth (eWOM) in the hospitality and tourism with, in addition to key insights in the respective study areas.

Resumen

En la era digital contemporánea, el boca-oreja electrónico de los clientes es importante en prácticamente todas las industrias de cara al cliente. Es aún más importante en el sector de la hostelería y del turismo, donde la opinión sobre el servicio recibido se muestra en Internet a través de las redes sociales o las reseñas de los clientes en línea. Estos comentarios, reseñas y calificaciones tienen una ventaja sobre el boca-oreja tradicional porque potencialmente pueden llegar a más personas, perdurar en el tiempo e incluso pueden volverse virales en Internet, con las consecuencias asociadas a ello. Por lo tanto, pueden influir en la percepción y el desempeño de los negocios de hotelería y turismo, ya que los clientes pueden tomar decisiones de compra en función de lo que dicen sus pares en línea. El objetivo de esta tesis doctoral es investigar el boca-oreja electrónico y las opiniones de los clientes en la industria hotelera y turística. Para lograr este objetivo, se realiza una revisión de la literatura y se recopilan y analizan datos relacionados con la industria hotelera y turística extraídos de las redes sociales en el contexto de la pandemia de COVID-19, que estalló durante el curso de este estudio e inmediatamente se convirtió en el problema más importante de esta industria, así como de la economía global. Esta tesis revela el estado actual de esta área de investigación, la percepción del turismo de cruceros como resultado de los brotes de COVID-19 en los mismos y la práctica de estancias vacacionales en los dos primeros años de la pandemia, cuando las limitaciones para viajar y realizar turismo eran muy significativas. Las principales conclusiones incluyen una reafirmación de la importancia del boca-oreja electrónico en la hotelería y el turismo, además de información clave en las respectivas áreas de estudio.

Resum

En l'era digital contemporània, el boca-orella electrònic dels clients és important a pràcticament totes les indústries orientades al client. És encara més important en el sector de l'hostaleria i el turisme, on es lloa el servei excel·lent i es critica el servei dolent a Internet mitjançant les xarxes socials o les opinions dels clients en línia. Aquests comentaris, ressenyes i puntuacions tenen un avantatge respecte al boca-orella tradicional perquè poden arribar a més persones, persistir durant més temps i possiblement tornar-se virals a Internet. Per tant, poden influir en la percepció i el rendiment de les empreses d'hostaleria i del turisme, i els clients prenen decisions de compra en funció d'allò que diuen els seus companys en línia. L'objectiu d'aquesta tesi doctoral és investigar el boca-orella electrònica i les opinions dels clients en el sector de l'hostaleria i el turisme. Per aconseguir aquest objectiu, es realitza una revisió de la literatura i es recullen i s'analitzen dades electròniques de boca a boca de les xarxes socials en el context de la pandèmia de la COVID-19, que va esclatar durant el transcurs d'aquest estudi i es va convertir de seguida en el principal problema d'aquesta indústria i de l'economia global. Aquesta tesi revela l'estat actual d'aquesta àrea de recerca, la percepció del turisme de creuers com a conseqüència dels brots de COVID-19 als creuers i la pràctica de les estades vacacionals en els dos primers anys de la pandèmia. Les principals conclusions inclouen una reafirmació de la importància del boca a boca electrònic a l'hostaleria i el turisme, a més d'informació clau en les àrees d'estudi respectives.

Electronic word-of-mouth (eWOM) and customer opinions in hospitality and tourism

Table of Contents

List of Tables

List of Figures

CHAPTER 1. INTRODUCTION

Justification and Scope of Study

In the first iteration of the Internet, websites were mostly static and website administrators created most of the website content by themselves. The second generation of the Internet usually referred to as Web 2.0 was more interactive and encouraged user participation, giving end users the ability to add and share their own content to websites in the form of text, pictures, and videos (Toledano, 2013). Web 2.0 was coined to describe the growing concept in the early to mid-2000s (Madadipouya, 2013) where the website acted as a platform and the website content was co-created by users. Examples of Web 2.0 websites founded during this period include; TripAdvisor (2000), Wikipedia (2001), Yelp (2004), YouTube (2005), and social networks like LinkedIn (2002), Myspace (2003), Facebook (2004), Reddit (2005), and Twitter (2006). In the present day, reviews and ratings are usually presented in association with e-commerce. For example, in the hospitality and tourism industry, popular platforms such as TripAdvisor, Booking.com, and Airbnb display reviews and ratings from past customers on the same webpage or with links to where online reservations can be made.

In the same vein, social media networks created during this early period provided a means to connect, form communities, and have conversations with other people across distances via the Internet (Xie et al., 2014). It permitted users to share a wide variety of content from their social media accounts, including about their service experiences (Xiang & Gretzel, 2010). Early social media sites like Twitter, Reddit, and Facebook remain popular till the present day, in addition to more recent social media platforms like Instagram, Snapchat, and TikTok. In recent times, social

media content has gone more visual with the explosion of smart devices for people to take pictures and make videos. Faster Internet bandwidth speeds have enabled the consumption of this type of content through the development of 3G, 4G, and most recently, 5G Internet technologies. The combination of all these factors has made social media sites into an important part of the Internet, with Facebook, Instagram, Twitter, Reddit all among the most visited websites as of November 2021 (Statista, 2022). As people interact and engage in conversations on social media, they give their opinions, persuade, and influence others (Fang et al., 2016). Researchers also refer to social media comments and online reviews and ratings published by users as electronic word-of-mouth (eWOM) or user-generated content (UGC). Both words are similar concepts, which encompass more than online reviews and include other forms of electronic communication between producers and consumers, such as virtual communities, blogs, emails, and instant messaging (Muritala et al., 2020).

In this doctoral thesis, we connect eWOM and the tourism and hospitality industry. This industry is an important component of the global economy estimated to contribute up to 10% of GDP and generating 10% of employment worldwide (Kandampully & Solnet, 2019). Spain is not an exception, where it is an important sector of the economy and contributes 12.3% of GDP and accounts for 12.7% of employment (Caixabank Research, 2020). The hospitality and tourism industry is a collective industry where customer service is critical and covers three main areas of activity; accommodation, tourism, and restaurants. Businesses in the hospitality and tourism industry must provide customers with an excellent customer service experience, otherwise such customers would not return in the future (Serra Cantallops & Salvi, 2014). Worse still, they may

spread negative word of mouth to their family, friends, and acquaintances. Before the digital age, this word of mouth was confined to only those within the close circle of the customer, hence the potential risk to businesses from negative word of mouth was relatively contained (Litvin et al., 2008). Indeed, it would be unprecedented for a disaffected customer to take out an advertisement in traditional mass media sources like television, radio, or newspapers to complain about a bad customer service experience. However, with Web 2.0, customers got the ability to share their service experiences, good or bad, through electronic word-of-mouth (eWOM) with a wide audience at very little cost through social media or online review sites.

Social media comments and online reviews provide a real-time source of consumer insight, which is faster and inexpensive compared to most other forms of customer market research (Litvin et al., 2008). Consumers use online reviews to engage in information search, to share experiences, and to give feedback. Prospective consumers can consult the opinions of their online peers before paying for a service with their decision-making and travel experience shaped by the reviews they read and they put more weight on the claims made by their online peers who have used a product or service than the claims made by the business (Papathanassis & Knolle, 2011). Hence, costumers' buying decisions could be influenced much more by comments, reviews, and ratings than by the giant marketing budgets of companies. Businesses use online comments and reviews for customer engagement, to influence booking intentions, build online presence, get customers' opinions and feedback, and earn revenue (Sparks et al., 2013). They provide a means for hospitality and tourism businesses to identify and respond to dissatisfied customers and are a source of insights on how to make service improvements. Consequently, social media

comments and online reviews are useful to both service producers and consumers, hence, they are influential in general, and specifically in the hospitality and tourism industry to both academic researchers and business managers (Muritala et al., 2020).

Another relevant phenomenon that took place in the connection between eWOM and the tourism and hospitality industry is the context determined by the recent health crisis provoked by the COVID-19. The World Health Organization (WHO) declared COVID-19 as a Public Health Emergency of International Concern (PHEIC) on the 30th of January 2020 and as a global pandemic on the 11th of March 2020 (WHO, 2020b, 2020a). Since the early days of the pandemic, the free movement of people locally and especially across national borders was curtailed by national governments to control the spread of the SARS-CoV-2 virus through non-pharmaceutical intervention (NPI) measures like curfews, stay-at-home orders, lockdowns, travel restrictions, and travel bans (Gössling et al., 2020). International travel ground to an almost complete standstill with international arrivals plunging 97% in April 2020, resulting in the loss of international tourism revenues by more than ten times those of the financial crisis (UNWTO, 2020). The hospitality and tourism industry has been one of the industries hit hardest by the COVID-19 pandemic, with global tourism losing trillions of dollars in each of the first two years of the pandemic according to the United Nations World Tourism Organization (UNWTO) (UN, 2021). However, the severe impact of the pandemic on the hospitality and tourism industry also presented a research opportunity to measure its impact on various parts of the industry and to provide insights on how to survive as well as how to recover from the pandemic.

Research Questions

This main aim of this doctoral thesis by a compendium of articles is to contribute to the knowledge about electronic word-of-mouth (eWOM) and customer opinions in hospitality and tourism. The first publication is a bibliometric literature review of online reviews research in hospitality and tourism. Thereafter, this thesis investigated two aspects of the COVID-19 pandemic's impact on the hospitality and tourism industry using social media data. First, there were several outbreaks of COVID-19 on cruise ships during the early part of the pandemic, with the cruise voyage of the *Diamond Princess* holding the second position behind China as the place with the highest number of confirmed cases of the novel coronavirus for almost three weeks. The second publication in this compendium investigated the impact and implications of the COVID-19 outbreaks on cruise ships on the perception of cruise tourism. Second, the pandemic inspired a resurgence of staycations because of the travel restrictions it instigated. Therefore, the third publication in this compendium explored the impact of the growth in staycations on the hospitality and tourism industry during and after the pandemic.

The overall objective of the thesis was divided into research questions based on the study areas of the publications as follows:

For the bibliometric literature review of online reviews in hospitality and tourism:

- What are the general trends in this research area?
- What is the foundational literature?
- What are the major research themes?

- How have the research interests evolved?

For the research on perceptions of cruise tourism after COVID-19 outbreaks:

- What insight can be derived from the public conversation on Twitter about cruising during the COVID-19 outbreaks on cruises?

For the research on COVID-19 era resurgence of staycations:

- How have staycations changed since 2008, i.e. regarding growth and new activities?
- What are the notable aspects of staycations in the first two years of the COVID-19 pandemic?

Structure of the thesis

As mentioned, this thesis is a compendium of three articles. It is based on a "Sandwich Model A" thesis by publication structure with an introductory first chapter, followed by an article in each of the subsequent chapters, and a concluding chapter (Mason & Merga, 2018). Two of these articles are published, one in a journal in the first quartile (Q1) in the Social Sciences Citation Index (SSCI) of the Journal Citation Reports (JCR) and one published in a journal in the second quartile (Q2) of the Social Sciences Citation Index (SSCI). The third article is currently under review in a journal in the second quartile (Q2) of the Social Sciences Citation Index (SSCI). Table 1.1 summarizes the articles in the compendium, their chapter in this thesis, and the journals in which they were published or under review.

Table 1.1 Summary of articles in compendium

	Article 1 (Chapter 2)	**Article 2** (Chapter 3)	**Article 3** (Chapter 4)
Article Title	A Bibliometric Analysis of Online Reviews Research in Tourism and Hospitality	#CoronavirusCruise: Impact and implications of the COVID-19 outbreaks on the perception of cruise tourism	COVID-19 staycations and the implications for hospitality and tourism
Journal	Sustainability	Tourism Management Perspectives	Journal of Leisure Research
Impact Factor 2020	3.251	6.586	2.561
Quartile	Q2	Q1	Q2
Category (Highest rank)	Environmental studies	Management	Sociology
Position within Category	59/125	51/226	52/149

Percentile	53.20	77.65	65.44
Index	SSCI JCR	SSCI JCR	SSCI JCR
Publication Status	Published	Published	Submitted
Keywords	Online reviews; Literature review; VOSviewer; UCINET; Citation analysis; Thematic evolution; Bibliometric analysis; Co-word analysis; Conceptual structure; Science maps	Coronavirus; Crisis communication; Risk perception; Big Data; eWOM; SARF; Information integration theory; Sentiment analysis	Twitter; Topic modeling; eWOM; User-generated content (UGC); Construal level theory

The first chapter of the thesis (Chapter 1) is an introduction. It presents a justification and scope of the study, offering a conceptual background of Web 2.0, online reviews, social media, electronic word-of-mouth, and the impact of the COVID-19 pandemic on hospitality and tourism. This is followed by the research questions of the thesis and an explanation of the structure of the thesis, theoretical frameworks, and analytical methodologies. The collection of articles are each in Chapters 2, 3, and 4, respectively. These chapters are composed of their own introduction,

literature review, methodology, results, discussion, conclusion, and references. The final chapter (Chapter 5) concludes the doctoral thesis by highlighting the most important findings of the study, limitations, and future lines of research resulting from this thesis.

Theoretical frameworks

Theoretical backing is crucial to create useful knowledge in studies utilizing data analytics methodologies from big data sources (Rivera, 2020). Hence, different theoretical frameworks were used as a theoretical support for the studies in Chapters 3 and 4. In the study analyzing the perception of cruise tourism in Chapter 3, Kietzmann's seven functional blocks of social media provided the conceptual framework for understanding the sentiment analysis results. The relevant block for this study is the conversation block of the framework, because Twitter is a conversation-based platform. The conversation block postulates that the conversation velocity and the change in how favorable or unfavorable the sentiment they contain need to be analyzed in order to understand the conversation (Kietzmann et al., 2011). In addition, the social amplification of risk framework (SARF) with the information integration theory (IIT) were combined to explore how risk perceptions during COVID-19 influenced potential tourist attitudes and behavior about cruising. The SARF framework suggests that the social amplification of risk by the media increases its memorability, which leads to increased risk perception (Kasperson et al., 1988). While the IIT states that new information is added into the preliminary beliefs of people, affecting how attitudes and behavior are formed (Anderson, 1981).

In the study investigating the practice of staycations in Chapter 4, construal level theory (CLT) is employed. CLT is the theory backing the notion of psychological distance. This theory from psychology proposes that human beings experience objects beyond their immediate situation by forming abstract mental construals. It postulates that we focus on the abstract at a higher level of construal and on the concrete at lower levels (Trope et al., 2007). Hence, CLT offers a theoretical foundation for interpreting the study findings of how people tried to create a simulacrum of their regular vacation while physically at or close to home.

Analytical Methodologies

VOSviewer software is the main software used for the bibliometric literature analysis in Chapter 2. It is a free to use software, which can construct and display large bibliometric maps in a way that is easy to interpret by paying special attention to the graphical representation of the bibliometric maps. It has been used in several studies across various fields with over 6,000 citations in Google Scholar.

Data analytics of social media data from Twitter is the main methodology for the studies in Chapters 3 and 4. Natural Language Processing (NLP) is a data analytical method which uses algorithms to make sense of unstructured text data. NLP is an automated processing of natural language or human language by algorithms. It is a multidisciplinary field that spans computer science, mathematics, linguistics, electrical engineering, and psychology. The NLP techniques and algorithms used in this work were all implemented using the Python programming language.

For the study analyzing the public perception of cruise tourism in Chapter 3, sentiment analysis is the main NLP technique used to automatically classify the selected tweets to gauge the perception. Sentiment analysis, also referred to as opinion mining, is an automated process of opinion detection which uses semantic relationships to classify the valence or overall polarity of text as positive, neutral, or negative. Sentiment analysis is one of the most common NLP techniques for analyzing huge amounts of text data, which is fast and inexpensive compared to manual processing and has been used in many studies.

For the study evaluating the observance of staycations in Chapter 4, topic modeling is employed as the main NLP technique. Topic modeling of tweets containing staycation was conducted to get the major topics or aspects of the online discourse about staycations using Latent Dirichlet Allocation (LDA). LDA is an efficient topic modeling technique for extracting the hidden topics from large unlabeled textual data with sophisticated machine learning algorithms to enable faster processing and get a better quality of topic classification. Topic modeling has also been applied extensively in the hospitality and tourism literature.

References

Anderson, N. H. (1981). *Foundations of Information Integration Theory*. Academic Press.

Caixabank Research. (2020). *The loss of tourism business is a major blow for the Spanish economy*. https://www.caixabankresearch.com/en/sector-analysis/tourism/loss-tourism-business-major-blow-spanish-economy

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management*, *52*, 498–506. https://doi.org/10.1016/j.tourman.2015.07.018

Gössling, S., Scott, D., & Hall, C. M. (2020). Pandemics, tourism and global change: a rapid assessment of COVID-19. *Journal of Sustainable Tourism*, 1–20. https://doi.org/10.1080/09669582.2020.1758708

Kandampully, J., & Solnet, D. (2019). *Service Management Principles for Hospitality & Tourism*. Goodfellow Publishers Limited. https://books.google.es/books?id=isRLEAAAQBAJ

Kasperson, Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., Kasperson, J. X., & Ratick, S. (1988). The Social Amplification of Risk: A Conceptual Framework. *Risk Analysis*, *8*(2), 177–187. https://doi.org/https://doi.org/10.1111/j.1539-6924.1988.tb01168.x

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241–251. https://doi.org/10.1016/j.bushor.2011.01.005

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, *29*(3), 458–468.

https://doi.org/10.1016/j.tourman.2007.05.011

Madadipouya, K. (2013). Survey on how Web 2.0 can facilitate knowledge management. *IJCER*, *2*(6), 693–695.

Mason, S., & Merga, M. (2018). Integrating publications in the social science doctoral thesis by publication. *Higher Education Research & Development*, *37*(7), 1454–1471. https://doi.org/10.1080/07294360.2018.1498461

Muritala, B. A., Sánchez-Rebull, M.-V., & Hernández-Lara, A.-B. (2020). A Bibliometric Analysis of Online Reviews Research in Tourism and Hospitality. *Sustainability 2020, Vol. 12, Page 9977*, *12*(23), 9977. https://doi.org/10.3390/SU12239977

Papathanassis, A., & Knolle, F. (2011). Exploring the adoption and processing of online holiday reviews: A grounded theory approach. *Tourism Management*, *32*(2), 215–224. https://doi.org/10.1016/j.tourman.2009.12.005

Rivera, M. A. (2020). Big data research in hospitality: From streetlight empiricism research to theory laden research. *International Journal of Hospitality Management*, *86*. https://doi.org/10.1016/j.ijhm.2019.102447

Serra Cantallops, A., & Salvi, F. (2014). New consumer behavior: A review of research on eWOM and hotels. *International Journal of Hospitality Management*, *36*, 41–51. https://doi.org/10.1016/j.ijhm.2013.08.007

Sparks, B. A., Perkins, H. E., & Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tourism Management*, *39*, 1–9. https://doi.org/10.1016/j.tourman.2013.03.007

Statista. (2022). *Most popular websites worldwide as of November 2021, by total visits*. https://www.statista.com/statistics/1201880/most-visited-websites-worldwide/

Toledano, C. A. (2013). Web 2.0: the origin of the word that has changed the way we understand public relations. *Barcelona International PR Conference, Barcelona, Spain*.

Trope, Y., Liberman, N., & Wakslak, C. (2007). Construal levels and psychological distance: Effects on representation, prediction, evaluation, and behavior. In *Journal of Consumer Psychology* (Vol. 17, Issue 2, pp. 83–95). https://doi.org/10.1016/S1057-7408(07)70013-X

UN. (2021). *Coronavirus pandemic could cost global tourism $2 trillion this year | | UN News*. https://news.un.org/en/story/2021/11/1106712

UNWTO. (2020). *Impact assessment of the COVID-19 outbreak on international tourism*. https://www.unwto.org/impact-assessment-of-the-covid-19-outbreak-on-international-tourism

WHO. (2020a). *COVID-19 Public Health Emergency of International Concern (PHEIC) Global research and innovation forum*. https://www.who.int/publications/m/item/covid-19-public-health-emergency-of-international-concern-(pheic)-global-research-and-innovation-forum

WHO. (2020b). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, *31*(2), 179–188. https://doi.org/10.1016/j.tourman.2009.02.016

Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management*, *43*, 1–12. https://doi.org/10.1016/j.ijhm.2014.07.007

CHAPTER 2. BIBLIOMETRIC LITERATURE REVIEW

Preface

This chapter of the thesis reviews the literature on online reviews in tourism and hospitality, and presents the current state of research in the area. A bibliometric approach was used to analyze 632 journal articles on online reviews in tourism and hospitality from 2005 to 2019 from the Scopus Database. This study identifies the most prolific journals, foundational works, and major research themes in the research area. In addition, we analyzed some dimensions of their network structure and the thematic evolution of the research area. The bibliometric method is quantitative and objective, and we carry out an analysis of the area based on citations and keywords. Researchers and business managers can gain useful insights on the current state of the art in this area. There have been only a few literature reviews tracking the growth in this field, and even fewer using bibliometric methods or science maps. Therefore, this work provides an updated review of this fast-growing area with a bibliometric approach to highlight the recent developments with the aid of science maps, and shows the thematic network structure and evolution with an innovative visualization.

The article based on this chapter is titled: “A Bibliometric Analysis of Online Reviews Research in Tourism and Hospitality” was published in 2020 in *Sustainability*, volume 12, issue 23, pages 1 - 18, article number 9977. The journal is indexed in Scopus and both the Social Sciences Citation Index (SSCI) and Science Citation Index Expanded (SCIE) of the Journal Citation Reports. The paper was also presented at the third edition of the International Workshop on Tourism and Hospitality Management conference (IWTHM2021) held in Porto and online on the 7th of May, 2021.

Introduction

Online reviews, also referred to as electronic word-of-mouth (eWOM) or user-generated content (UGC) (Lu and Stepchenkova 2015; Serra Cantallops and Salvi 2014), are all similar concepts with minor differences. eWOM is all electronic communications between producers and consumers and between consumers themselves: emails, websites, consumer review sites, blogs, virtual communities, chat rooms, newsgroups, and instant messaging (Litvin et al. 2008). UGC is data generated online by consumers e.g. text (consumer reviews and blogs) and picture data (Li et al. 2018).

Tourism and hospitality rely very much on word-of-mouth among consumers (Fotis et al. 2012). Before the Internet, this word-of-mouth was usually obtained from friends and acquaintances, but this has now moved to the web, with consumers overwhelmingly consulting the opinions of fellow consumers online (Wang et al. 2002; Xie et al. 2014). Their decision-making and travel experience is shaped by the reviews they read, and this has been a significant development in the industry (Sparks et al. 2013).

Consumers use online reviews at every stage of the travel process; pre-trip, during trip, and post-trip, to engage in information search, share experiences, and give feedback (Fotis et al. 2012). While companies use it for customer engagement, build online presence, get consumers' opinions, influence booking intentions, and earn revenues (He et al. 2017; Leung et al. 2013; Sparks and Browning 2011). With its usefulness to both producers and consumers, it is no

surprise that online reviews are a highly popular topic in tourism and hospitality for both researchers and managers (Kwok et al. 2017).

As a result, there have been many studies on online reviews (Schuckert et al. 2015), but only a few keeping track of the literature. These works presented the state-of-the research at their time of publication, providing a now obsolete image of the field because of the astonishing speed at which the literature is growing. Therefore, recent studies on the application of online reviews in tourism and hospitality have not been integrated with previous studies to understand the present state of research and current research trends. Many of these previous reviews have been systematic (Serra Cantallops and Salvi 2014; Leung et al. 2013) and none have used bibliometric methods or science maps (Centobelli and Ndou 2019) to visually show the development, conceptual structure, and thematic evolution of online reviews in tourism and hospitality.

Hence, the aim of this paper is to integrate and provide an organized summary of the existing research by identifying the foundational and seminal works and the old and new areas of the prior scholarship in order to build knowledge, gain an understanding, and show the future direction in this research area. We achieve this by means of bibliometric analytical techniques that enable us to deconstruct the main anchors and evolution of the research area. We also contribute an innovative way to show the thematic evolution of the key research themes. We draw four research questions based on this research objective to achieve this goal:

RQ1. What are the general trends in this research area?

RQ2. What is the foundational literature?

RQ3. What are the major research themes?

RQ4. How have the research interests evolved?

Literature Review

Even though the research on online reviews in tourism and hospitality is fast-growing (Centobelli and Ndou 2019), it has a relatively brief history (Zeng & Gerritsen 2014). This is because notable online opinion platforms for customers rose to prominence in the early 2000s (Kwok et al. 2017). A few attempts have been made to review the literature and Table 1 presents sixteen previous studies that we identified, which reviewed the literature on online reviews in tourism and hospitality directly or indirectly.

Table 2.1 Previous literature reviews

Authors	**Scope**	**Journal**	**Databases**	**Method**	**Sample Size**	**Years covered**
Leung et al. (2013)	Social media in Tourism & Hospitality	JTTM	Science Direct, EBSCOHost, and Google Scholar	Systematic	44	2007 - 2011
Serra Cantallops	eWOM in Hotel Industry	IJHM	6 selected journals	Systematic	28	2007 - 2011

and Salvi (2014)						
Zeng and Gerritsen (2014)	Social media in Tourism & Hospitality	TMP	Web of Science, EBSCOHost, and Google Scholar	Bibliometric	279	2007 - 2013
Law et al. (2014)	ICT in Tourism & Hospitality	IJCHM	Science Direct, EBSCOHost, Emerald, and Sage	Content analysis	107	2009 - 2013
Schuckert et al. (2015)	Online reviews in Tourism and Hospitality	JTTM	Science Direct, EBSCOHost, and Google Scholar	Content analysis	50	2004 - 2013
Lu and Stepchenkova (2015)	UGC in Tourism and Hospitality	JHMM	Google Scholar, Web of Science, 5 top tourism, and 5 top hospitality journals	Systematic	122	2001 - 2013
Chen and Law (2016)	eWOM in Tourism and Hospitality	IJHTA	Science Direct, EBSCOHost	Systematic	43	2008 - 2013

Kwok et al., (2017)	Online reviews in Tourism and Hospitality	IJCHM	7 top hospitality and tourism journals	Systematic	67	2000 - 2015
Sotiriadis (2017)	Social media in Tourism & Hospitality	IJCHM	Science Direct, Google Scholar	Systematic	146	2009 - 2016
Bore et al. (2017)	eWOM in Hotel Industry	HS	Science Direct, EBSCOHost, Emerald and SpringerLink	Systematic	45	2000 - 2015
Leung et al. (2017)	Social media in Tourism & Hospitality	IJHM	8 top business and 8 top hospitality and tourism journals from Web of Science	Bibliometric	406	2007 - 2016
Li et al. (2018)	Big Data in Tourism	TM	Web of Science, Science Direct, Sage, Emerald, SpringerLink, Wiley Online Library, Google scholar	Systematic	165	2007 - 2017

Mariani et al. (2018)	Big Data in Hospitality and Tourism	IJCHM	Scopus and Web of Science	Systematic	173	2000 - 2016
Lu et al. (2018)	Social media in Tourism & Hospitality	JTTM	Science Direct, EBSCOHost, Google Scholar, and 7 top hospitality and tourism journals	Systematic	105	2004 - 2014
Centobelli and Ndou (2019)	Big Data in Tourism	CIT	Scopus	Bibliometric	109	1990 - 2017
Nusair et al., (2019)	Social media in Tourism & Hospitality	IJCHM	Science Direct, EBSCOHost	Bibliometric	439	2002 - 2016

Table 1 shows that these studies were published in reputable journals, with the International Journal of Contemporary Hospitality Management, Journal of Travel and Tourism Marketing, and International Journal of Hospitality Management having five, three and two publications, respectively. The remaining journals; Tourism Management, Current Issues in Tourism, Journal of Hospitality Marketing and Management, Tourism Management Perspectives, International

Journal of Hospitality and Tourism Administration, and Hospitality and Society, had one publication each.

An analysis of these previous reviews reveals that there are differences in their depth of focus. While some focused directly on online reviews in tourism and hospitality, many considered a larger domain such as social media or big data in tourism and hospitality, with online reviews treated as a subset of this larger domain.

Law et al. (2014) had the biggest scope and explored information and communication technology (ICT) in tourism and hospitality, with online reviews included under ICT. This is followed in scope by Mariani et al. (2018), who investigated business intelligence and big data in tourism and hospitality. Some studies reviewed research focused on types of big data in tourism and hospitality; including UGC e.g. online reviews, device data e.g. GPS and Wi-Fi data, and transaction data e.g. web search data (Centobelli and Ndou 2019; Li et al. 2018). Other studies examined the role of social media in hospitality and tourism, by reviewing the influence of online consumer review platforms, social networking sites, Internet forums, and other online communities (Leung et al. 2013; Leung et al. 2017; Lu et al. 2018; Sotiriadis 2017; Zeng and Gerritsen 2014). Two studies examined eWOM relating only to hotels (Bore et al. 2017; Serra Cantallops and Salvi 2014). While the remaining studies reviewed online reviews/UGC/eWOM in tourism and hospitality (Chen and Law 2016; Kwok et al. 2017; Lu and Stepchenkova 2015; Schuckert et al. 2015).

The previous works can be classified into three groups regarding the academic databases from which they selected their samples. The first group which comprises most of the studies (11), used a keyword search of well-known databases such as Scopus, Google Scholar (Bore et al. 2017; Centobelli and Ndou 2019; Chen and Law 2016; Law et al. 2014; Leung et al. 2013; Li et al. 2018; Mariani et al. 2018; Nusair et al. 2019; Schuckert et al. 2015; Sotiriadis 2017; Zeng and Gerritsen 2014). The second group concentrated on selected top journals like Tourism Management, Journal of Travel Research, etc. (Kwok et al. 2017; Leung et al. 2017; Serra Cantallops and Salvi 2014). While the third group used both selected top journals and databases to get their samples (Lu and Stepchenkova 2015; Lu et al. 2018).

Most of the studies considered only journal articles except those of Li et al. (2018) which used articles and conference papers, Centobelli and Ndou (2019) which used articles, conference papers and book chapters, and Zeng and Gerritsen (2014) which used a lot of grey literature like research degree theses, electronic articles, books, and reports besides articles and conference papers.

The literature reviews in this field started in 2013, and the year of publication explained the number of citations received by the documents, with the older publications having more citations than the newer ones, apart from some exceptions. Five of the sixteen studies reviewed literature from 2007 up till their time of publication (Leung et al. 2013; Serra Cantallops and Salvi 2014; Zeng and Gerritsen 2014; Leung et al. 2017; Li et al. 2018). Three from 2000 (Bore et al. 2017; Kwok et al. 2017; Mariani et al. 2018), two each started from 2004 (Schuckert et al., 2015; Lu et

al., 2018) and 2009 (Law et al., 2014; Sotiriadis, 2017), and one each from 2001 (Lu and Stepchenkova 2015), 2002 (Nusair et al. 2019), and 2008 (Chen and Law 2016). Even though Centobelli and Ndou (2019) searched for the literature from 1990 to 2017, they reviewed documents published from 2011 onwards. In summary, the earliest time of publication of the reviewed literature was from year 2000, with most of the studies considering literature from 2007 onwards. The last year considered in the most recent reviews was 2017 (Centobelli and Ndou 2019; Li et al. 2018).

Regarding method, the previous reviews showed a preference for systematic reviews, with ten of them adopting this approach (Bore et al. 2017; Chen and Law 2016; Kwok et al. 2017; Leung et al. 2013; Li et al. 2018; Lu and Stepchenkova 2015; Lu et al. 2018; Mariani et al. 2018; Serra Cantallops and Salvi 2014; Sotiriadis 2017). Two applied content analysis (Law et al. 2014; Schuckert et al. 2015), and four used a bibliometric approach (Centobelli and Ndou 2019; Leung et al. 2017; Nusair et al. 2019; Zeng and Gerritsen 2014). The bibliometric reviews examined more documents, with three bibliometric reviews having the biggest sample sizes (Zeng and Gerritsen 2014; Leung et al. 2017; Nusair et al. 2019).

We employ the bibliometric analytical method for this study since it provides the tools to answer our research questions on the foundations and themes of this research area. The bibliometric method provides an advantage of objectivity and quantifiability, and helps to avoid subjective biases. It also helps to provide validation for findings which scholars had intuitively inferred in earlier studies and is more informative (Mingers and Leydesdorff 2015; Nerur et al. 2008), which

justifies the suitability of this methodological approach. Bibliometrics is the use of mathematical and statistical methods to quantify and analyze the bibliographic information of publications (Pritchard 1969). This bibliographic information enables researchers to make linkages between authors or papers (Garfield et al. 1983). Thomson Reuters Web of Science (WoS) and Elsevier's Scopus are the traditional databases for carrying out bibliometric studies since they are reliable sources of citation data (Mingers and Leydesdorff 2015).

Many bibliometric studies often utilize science maps (Cobo et al. 2011). The ability to visualize bibliometric networks with science maps was an important technical development in bibliometrics (Mingers and Leydesdorff 2015). Science maps, also known as bibliometric maps or knowledge maps, help to reveal the conceptual, intellectual or social structure of a field (Cobo et al. 2011). It is a spatial representation of the interrelationship between research elements such as authors or citations that facilitates the understanding of the structure and developments in a field (Small 1999). There are different software tools for constructing science maps and Cobo et al. (2011) made a comparative study of these tools. There are also different approaches to extracting a bibliometric network depending on the selected unit of analysis; e.g. authors, documents, journals, cited references or keywords (Cobo et al. 2011).

Science maps were used in just two of the four previous bibliometric reviews. Centobelli and Ndou (2019) used a citation network, while Leung et al. (2017) used a combination of co-citation and co-word analyses to make up for the weaknesses in a single method, and to reveal the

theoretical foundation, and thematic evolution of their area of study. This study is going to follow in the steps of Leung et al. (2017) with this combination for the same reason.

Co-citation analysis is a study of cited documents to get the frequency of citation of two earlier documents together. Hence, a co-citation link is a link between two documents cited together by a later document. When a set of documents is frequently co-cited together in a certain area, this may show that the co-cited documents contain important concepts that peers have recognized. Thus, a co-citation network analysis can identify the core literature in a particular area of study (Small 1973).

Co-word or co-occurrence analysis measures the frequency of co-occurrence of two keywords in the same literature (Whittaker 1989). It relies on the assumption that the co-occurrence of keywords shows a non-trivial relationship (Callon et al. 1991). When sets of keywords are frequently used together by different authors in a certain area, this may show that these keywords have a significant relationship within the research area (Whittaker 1989). This bibliometric method directly extracts the major themes of a research area and the linkages between these themes based on the co-occurrence of word pairs without relying on a priori definitions (He 1999). Hence, help to reveal the conceptual structure and major research themes of the area (Callon et al. 1991).

The strength of a co-citation link is derived from the number of citing authors who cite two earlier works together. Therefore, co-citation is not a permanent relationship and the co-citation network pattern changes with time. The same applies to co-word analysis, which changes with a

change in vocabulary co-occurrences in a field. Therefore, co-citation and co-word patterns in a field will change as the field evolves with a change in the interests and intellectual patterns (Callon et al. 1991; Small 1973).

From the review of the literature, it is obvious that the previous literature review studies have mainly been systematic reviews, and the few bibliometric studies covered a bigger scope, namely social media (Leung et al. 2017; Nusair et al. 2019; Zeng and Gerritsen 2014), and big data (Centobelli and Ndou 2019). Hence, their emphasis on online reviews was only peripheral. Similarly, the two previous bibliometric studies that used science focused on social media (Leung et al. 2017) and big data (Centobelli and Ndou 2019). In summary, there have been no studies to the best of our knowledge providing the foundational literature and major research themes in the maturing research area of online reviews in tourism and hospitality. In addition, our study also covers the recent publications in a fast-growing field, as most of the documents reviewed in this work were published in the last two years. Finally, the thematic evolution chart and the analysis of some dimensions of its network structure is a novel contribution of this work that allows a better understanding of the development and state of the art in the field.

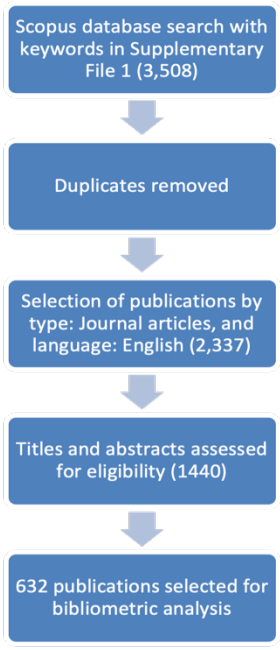
Method

Data collection

We obtained the publications considered for this study from the Scopus database. Scopus was chosen because it is a recognized index with a wide coverage of peer-reviewed publications and provides reliable bibliographic data (Martín-Martín et al. 2018; Mingers & Leydesdorff 2015).

Several search queries with relevant keywords were used (see Supplementary File 1) to search in the title, abstract and keywords of the publications in the database from 2005 to 2019. 2005 was chosen as the start date because most of the previous reviews reported the first publications in this area to be from 2007 or later, two years were added to this to ensure that all relevant documents were found. We filtered the results to include only publications in English and journal articles, resulting in 1440 documents. Their titles and abstracts were reviewed, and we selected 632 relevant publications. We selected theoretical and empirical articles on the application of online reviews in tourism and hospitality, with an emphasis on studies that used online reviews data and excluded technical papers focused on highlighting the performance and accuracy of algorithmic models for prediction, classification, recommendation engines or detection of fake reviews.

Figure 2.1: Systematic process to select literature



Data analysis

To answer the first research question posed by this study, the general literature trends are presented. Co-citation analysis was done to determine the foundational literatures of the study area, which are the most co-cited studies by the selected publications, and answers the second research question. A co-word analysis was also carried out because it enables the determination of the conceptual structure and research themes of the study area and helps to answer the third and fourth research questions. VOSviewer software (version 1.6.13) (van Eck and Waltman 2010) and Ucinet 6.0 (Borgatti et al. 2002) were used to conduct the bibliometric analysis. Data cleaning, an important preprocessing step (Cobo et al. 2011), was performed before data analysis by using a thesaurus file. The thesaurus file was used to combine variants of the same term or related concepts, such as “eWOM,” “e-wom,” “electronic word of mouth,” etc., which were all merged into “eWOM.” The thesaurus file was also used to ignore irrelevant words that did not provide any useful information, e.g., “article.” The first step in the analysis was to produce a chart showing the growth in the number of published articles per year to show the growth in the literature. Next was a study of the most important journals, followed by the co-citation analysis to identify the core documents. Finally, the co-word analysis was carried out to identify the major research themes, with a study of the crucial dimensions of the keyword network structure and the evolution of the research interests.

Results

General literature trends

Figure 2.2 shows the growth in the number of publications within the study period. There was steady growth from 2006 to 2013, a slight drop from 2013 to 2014, followed by rapid growth from 2014 onwards.

Figure 2.2: Growth of literature over time

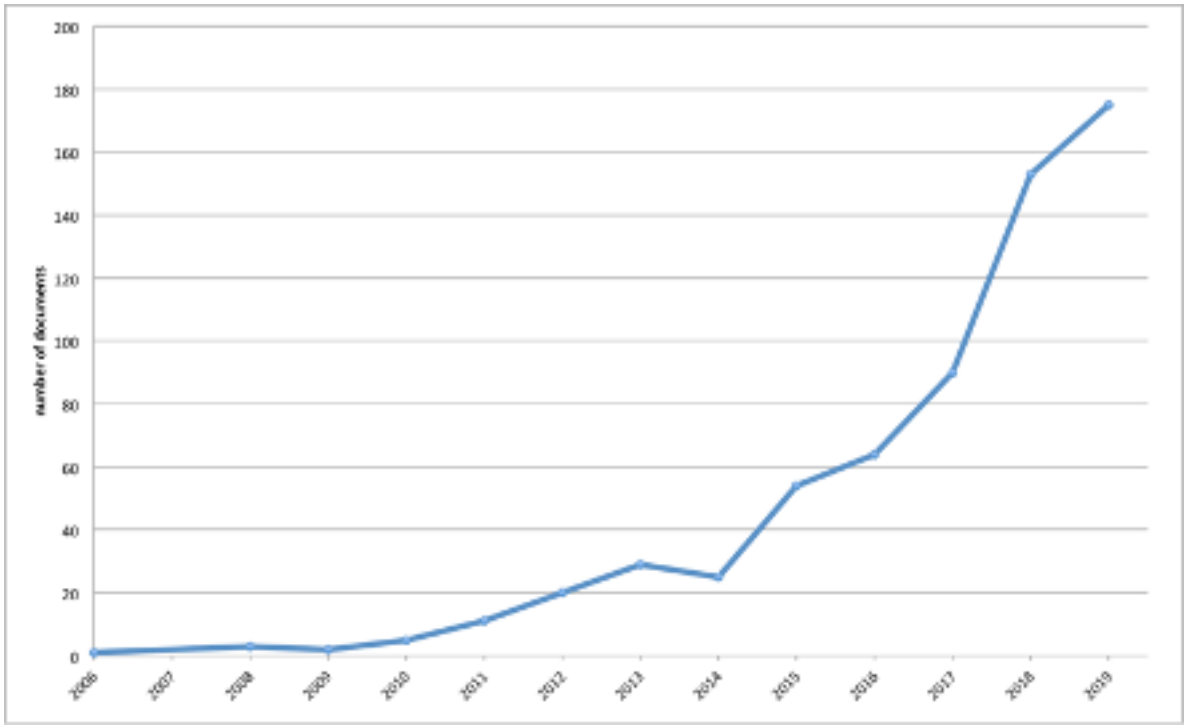


Table 2.2 shows the most prolific journals, using a threshold of six or more publications. The list includes many reputable tourism and hospitality journals, with most of them in the first quartile in the 2018 Scopus CiteScore under the Tourism, Leisure and Hospitality Management category. The 632 articles were from 202 journals.

Table 2.2: List of most prolific journals

	Journal	Total articles	Total citations	Quartile
1	Tourism Management	56	3855	Q1
2	International Journal Of Hospitality Management	53	2644	Q1
3	International Journal Of Contemporary Hospitality Management	42	721	Q1
4	Journal Of Hospitality Marketing And Management	25	675	Q1
5	Current Issues In Tourism	17	98	Q1
6	Journal Of Travel And Tourism Marketing	17	885	Q1
7	Sustainability (Switzerland)	14	40	Q1*
8	Journal Of Travel Research	13	557	Q1
9	Tourism Management Perspectives	13	154	Q1
10	Information Technology And Tourism	12	70	Q1*
11	Journal Of Hospitality And Tourism Technology	12	165	Q1
12	Cornell Hospitality Quarterly	11	641	Q1

13	Journal Of Vacation Marketing	9	150	Q1
14	Journal Of Hospitality And Tourism Research	8	166	Q1
15	Tourism Analysis	8	15	Q3
16	Asia Pacific Journal Of Tourism Research	7	115	Q1
17	Journal Of Business Research	7	153	Q1*
18	Journal Of China Tourism Research	7	15	Q2
19	Anatolia	6	38	Q2*
20	Annals Of Tourism Research	6	131	Q1
21	E Review Of Tourism Research	6	8	Q4
22	International Journal Of Culture Tourism And Hospitality Research	6	26	Q2
23	International Journal Of Hospitality And Tourism Administration	6	37	Q3
24	Journal Of Destination Marketing And Management	6	165	Q1

* *not listed under Tourism, Leisure and Hospitality Management*

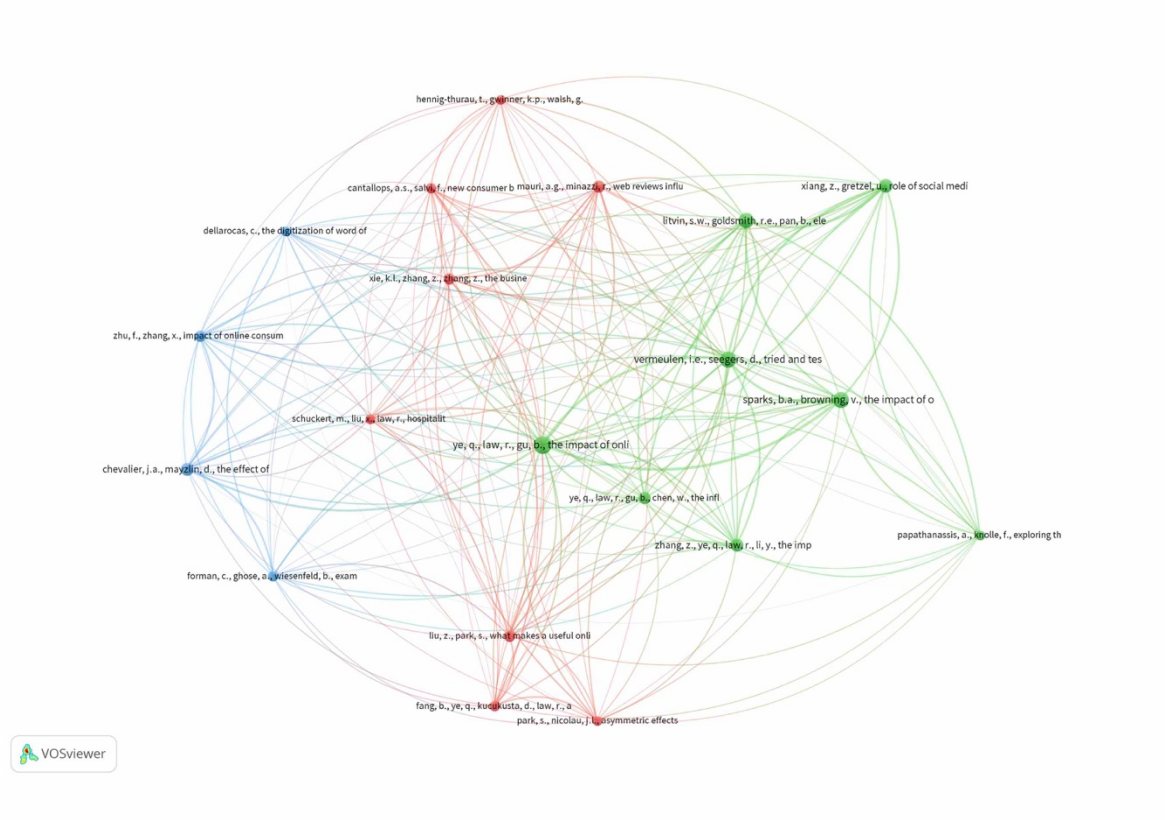
Tourism Management and *International Journal of Hospitality Management* had the most publications; with both of them accounting for 17% of all publications. *Tourism Management* published articles with more impact, which have accumulated more citations than the others with about 26% of the total citations of all selected documents.

Foundational Literature

The co-citation network map was made to identify the foundational literature. To obtain this map, co-citation was selected as the type of analysis, cited references as the unit of analysis, and fractional counting as the counting method, as recommended (van Eck and Waltman 2014). Given the high number of cited references of over 30,000, an adequate minimum threshold of citations was mandatory for clarity. To obtain the top 20 most co-cited articles, as done in previous bibliometric studies (Liu et al. 2015, we set the threshold at 29 citations. The normalization method was association strength and the visualization weight was citations.

Figure 2.3 shows the co-citation network map. The proximity of two papers, the thickness of the lines connecting them, and their size shows the strength of their co-citation links (van Eck and Waltman 2013). These documents in the nodes are the 20 most influential or seminal papers in this area at this time. It is important to point out that even though these papers are not necessarily the most frequently cited papers in terms of bulk citations; they have been highly co-cited by the selected documents and hence represent the foundational literature.

Figure 2.3: Co-citation network map



Tourism Management and *International Journal of Hospitality Management* are the dominant sources of these influential papers with seven and five publications respectively, while other journals have one each of the remaining eight publications. These papers are in three clusters and there is a uniform pattern with the papers in each cluster. The blue cluster has four papers and are the earliest publications of the 20, published between 2003 and 2010. They deal with the effect of eWOM on sales and were written from outside the tourism and hospitality domain, specifically marketing, information systems, and management science (Chevalier and Mayzlin 2006; Dellarocas 2003; Forman et al. 2008; Zhu and Zhang 2010).

The red cluster consists of eight papers. These are the relatively newer documents, and seven of them are the most recent documents of the 20, published between 2013 and 2016. They can be divided into three subgroups; literature reviews (Schuckert et al. 2015; Serra Cantallops and Salvi 2014); a paper on the motivations to write a review (Hennig-Thurau et al. 2004); while the remaining five papers are on the consequences of reviews, e.g. value and usefulness of online reviews (Fang et al. 2016; Liu and Park 2015; Park and Nicolau 2015), the impact of hotel reviews on consumer purchase intention and service expectations (Mauri and Minazzi 2013), and on hotel performance (Xie et al. 2014).

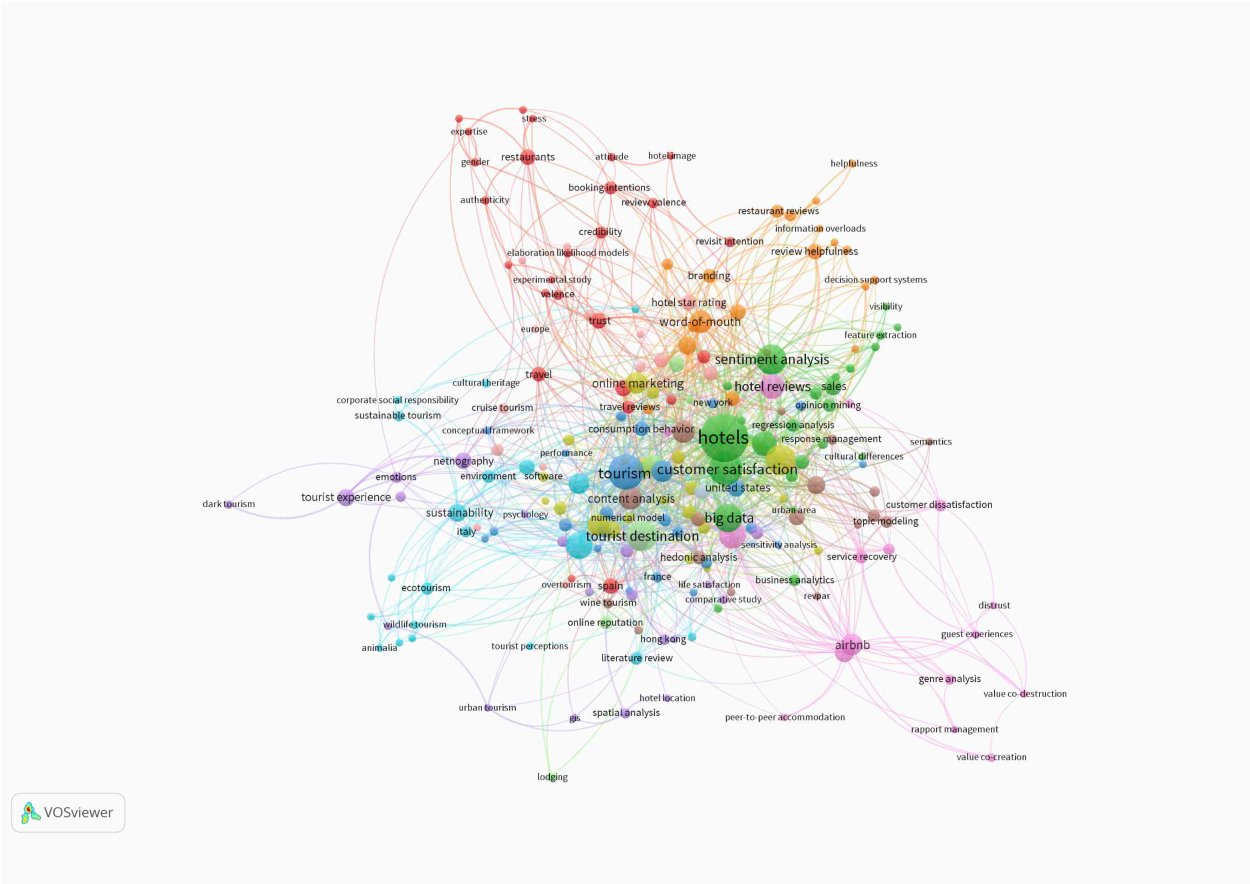
The green cluster is also made up of eight papers and has larger nodes, which implies that they are the most co-cited papers on the map. They were published from 2008 to 2011, and can be situated in time between the early blue cluster and the recent red cluster. The papers in this mid-period were all written from a tourism and hospitality perspective, in contrast to the blue cluster. We can divide this cluster into two subgroups, the first is made up of influential conceptual papers; on the role of social media in online travel information search (Xiang and Gretzel 2010), eWOM in hospitality and tourism (Litvin et al. 2008), and an exploration of the adoption and processing of online reviews (Papathanassis and Knolle 2011). While similar to some papers in the red cluster (Mauri and Minazzi 2013; Xie et al. 2014), the second subgroup comprises papers on the consequences of online reviews on different aspects of the sales process such as consumer consideration (Vermeulen and Seegers 2009), hotel booking and sales (Sparks and Browning 2011; Ye et al. 2011), and popularity of restaurants (Zhang et al. 2010). We find these similar papers in the red and green clusters relatively close on the visualization map.

Main Research Themes

The co-word network map was made to identify the major research themes. To obtain this map, co-occurrence was selected as the type of analysis, all keywords as the unit of analysis, and fractional counting as the counting method, and the thesaurus file was uploaded. The total number of keywords was 2028, which were too many to fit on a chart. Therefore, a threshold of three occurrences was set, which 219 keywords met. The normalization method was fractionalization and the visualization weight was occurrences.

Figure 2.4 shows the main keywords that have been used in online reviews literature (full list in Supplementary File 1). The proximity of two terms and the thickness of the lines connecting them show how frequently they co-occurred as keywords, and the size of a node is determined by the frequency of its occurrence as a keyword.

Figure 2.4: Co-word network map



The most frequent keywords like online reviews, social media, eWOM, UGC, and online ratings had the biggest sizes. To avoid labels from overlapping, VOSviewer does not display labels close to a bigger label in the static image of the map, which prevents some keyword labels from showing. Hence, these oft-recurring keywords that cut across the entire field adding no extra information were removed from the visualisation to simplify it and reveal the structure of the field spatially (He 1999).

We calculated centrality measures to determine the precise structure of this network and identify the keywords that occupy critical positions within it. There are different centrality measures, all of which identify the most central elements in a network. Using Ucinet 6.0, degree, betweenness, and eigenvector centrality were calculated. These measures determine the prominence of a keyword in the network and existing conceptual overlap between them, however some differences may also emerge from their use depending on the network configuration. Thus, considering all of them provides a more complete picture not just about the keywords that occupy highly central positions, but also on the influence of these keywords on others (Valente et al. 2008) The top ten central keywords for each measure are presented in Table 2.3 as done in previous studies (Hu et al. 2013).

Table 2.3: Top 10 keywords with high centrality

Keyword	**Degree centrality**	**Keyword**	**Betweenness centrality**	**Keyword**	**Eigenvector centrality**
Hotels	96	Hotels	3100.50	Hotels	1.00
Tourism	53	Tourism	1453.66	Customer satisfaction	0.65
Customer satisfaction	45	Tourist destination	1240.57	Text analytics	0.61

Text analytics	43	Tourism management	1166.91	Tourism	0.60
Tourist destination	41	Customer satisfaction	1023.74	Sentiment analysis	0.51
Sentiment analysis	40	Hospitality	865.59	Big data	0.47
Tourism management	34	Text analytics	835.86	Hospitality	0.44
Big data	33	Sentiment analysis	767.95	Online ratings	0.42
Hospitality	33	Perception	644.48	Data mining	0.36
Data mining	28	Tourist behavior	584.82	Tourist destination	0.36

Hotels/hospitality (hotels, hotel reviews, hotel star rating, etc.) and tourism (tourism, tourist destination, tourism management, travel, travel reviews, wine tourism, cruise tourism, etc.) are the prominent industry domains and occupy a central position on the map, as expected according to all three centrality measures. Restaurants (restaurants, authenticity, booking intentions, etc.) was also an expected theme because of its relevance in the hospitality industry.

Customer satisfaction is also central, and highly connected on the map as a concept that has been investigated extensively using online reviews. Big data and text analytics techniques (sentiment analysis, topic modeling, etc.), and traditional research and statistical methodologies (regression analysis, sensitivity analysis, content analysis, etc.) also occupy central positions as the processing tools for online reviews. It is worth mentioning that the high values of these keywords in terms of betweenness centrality underline not just their central position in the network, but also their influence on other keywords close to them.

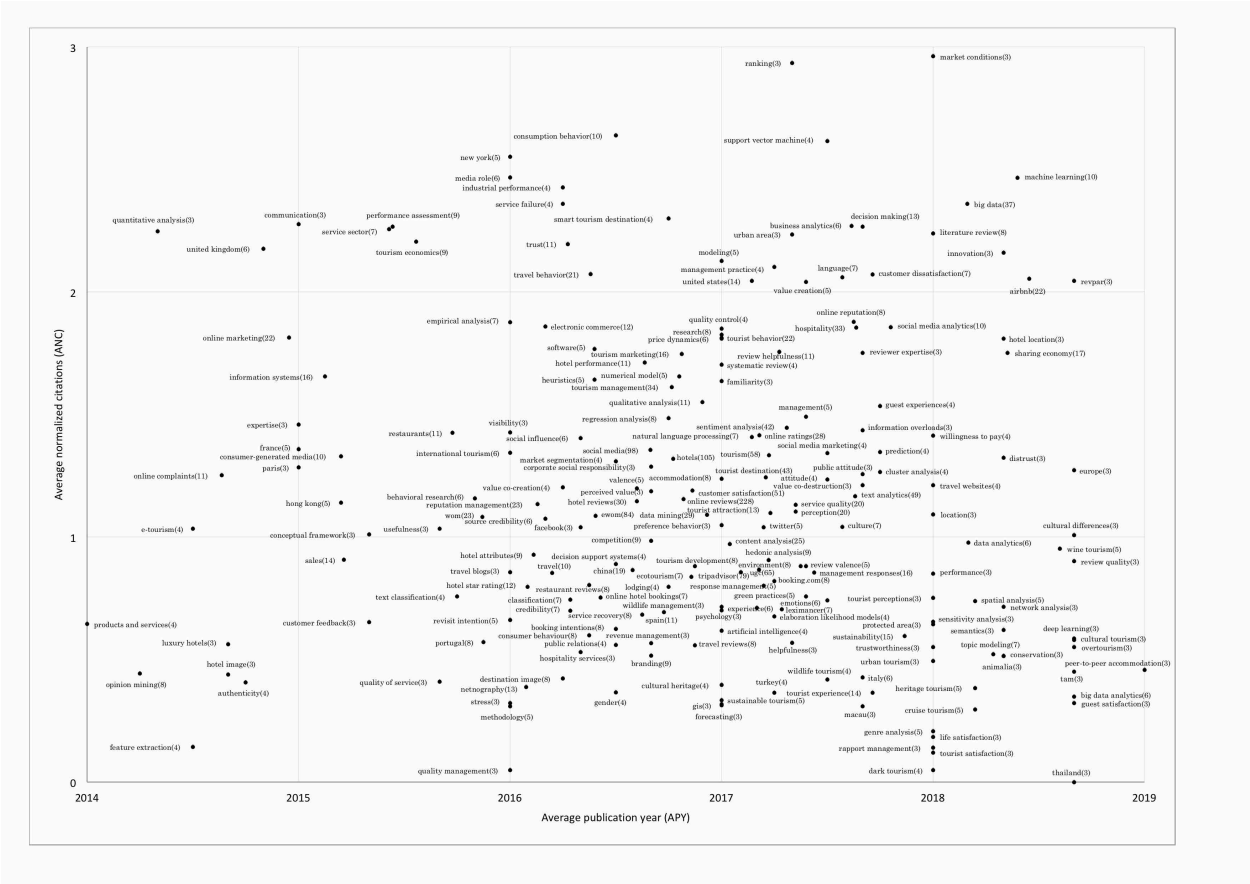
Airbnb (Airbnb, sharing economy, peer-to-peer accommodation, guest experiences, value co-creation and co-destruction, etc.), sustainability (sustainability, sustainable tourism, ecotourism, overtourism, environment, wildlife tourism, corporate social responsibility, etc.), and tourist experience (emotions, dark tourism, heritage tourism, etc.) are also themes that can be identified on the map, showing the underlying structure of the research area.

Evolution of Research Interests

VOSviewer provides an overlay visualization of a network diagram, which shows the nodes on a network map with a color gradation based on the scores calculated by the software for either the average publication year (APY), average number of citations, or average normalized citations (ANC). To show the impact of the keywords over time in the selected literature in an easy-to-understand chart, the calculated figures from VOSviewer were taken and the keyword impact, represented by the ANC, was plotted against publication year (APY).

Figure 2.5 shows the wide range of keywords that have been used in the selected documents and the evolution of research interests in online reviews in tourism and hospitality in a timeline. ANC is the average number of normalized citations received by the documents that the keyword represents. The normalized citation count is the number of citations of the document divided by the average number of all citations of all documents published in the same year it was published, included in VOSviewer data (van Eck and Waltman 2013). ANC was used to compensate for the fact that citation counts favor older publications. APY is the average year of publication of the documents that the keyword represents. We placed the number of occurrences of the keywords in brackets in front of the keywords.

Figure 2.5: Evolution timeline of keywords



This chart originally contained the 219 keywords from the co-word network map after choosing a threshold of three occurrences, with this threshold eliminating many keywords from earlier years. However, placing all 219 keywords on the chart, including those with very high ANC values and very low APY values, made it difficult to visualize most of the keywords. Therefore, five keywords with outlying values were excluded to get a clearer plot, with axes 0–3 for ANC and 2014–2019 for APY. We present these excluded keywords in Table 2.4. In addition, we present the full data of all 219 keywords in Supplementary File 2, i.e., ANC, APY, total number of occurrences, and the breakdown of occurrences by year.

Table 2.4: Outlier values removed from Figure 2.5

Keyword	Occurrences	APY	ANC
Artificial neural network	4	2018.00	4.50
Tourism research	3	2016.67	3.65
Comparative study	4	2018.25	3.57
Experimental study	3	2013.67	1.80
Web 2.0	9	2013.78	1.81

Many of the keywords from the co-word analysis are visible in Figure 2.5, including those that could not be seen on the co-word network map to avoid overlapping. They include keywords about the industry sectors, consumer behavior, management processes to achieve management objectives, online review platforms, research methodologies, and country/city names. The closer to the right side of the plot the keyword appears, the more recent its occurrences in the literature are, and the closer to the top of the plot, the better cited (normalized citations) the articles are that used the keyword. Keywords found to the bottom right of the plot are emergent in the literature, while those at the top right are recent and well cited or hot topics (Ávila-Robinson and Wakabayashi 2018).

The applications of artificial intelligence (AI) or machine learning techniques to the big data of online reviews are prominent themes in the hot topic area, e.g., “big data,” “machine learning,” “artificial neural network,” “support vector machine,” and “business analytics.” Other keywords related to this theme, such as “deep learning,” “text analytics,” “natural language processing,” “data analytics,” “big data analytics,” and “AI,” are found at lower values of ANC, but still in the recent areas. This highlights their current relevance in the literature. The Airbnb and sharing economy theme is also in the hot topic area.

In the emergent area, online reviews are increasingly applied to investigate different niche areas of tourism, such as wine tourism, heritage tourism, cultural tourism, urban tourism, cruise tourism, wildlife tourism, and dark tourism. Sustainability themes such as sustainability, ecotourism, overtourism, environment, green practices, and sustainable tourism can also be found in this emergent area.

Figure 2.5 also shows a subtle evolution in terminology for some terms. For example, the average year of usage of “opinion mining” was 2014.3 (six of its eight appearances were between 2008 and 2016), and it seems to have now been replaced by terms such as “data mining” (2016.9), “text analytics” (2017.6), and “data analytics” (2018.2) in the tourism and hospitality literature. In the same vein, we can also observe an evolution in techniques. For example, feature extraction, a machine learning process used in text analysis to extract the unique features of a document (Liu, 2012), had an average year of 2014.5, and it seems to have now been eclipsed in the tourism and hospitality literature by topic modeling (2018.3), a method for carrying out

feature extraction (Liu, 2012), which suggests that researchers now frequently use this technique for feature extraction. Another way the use of keywords has changed is the gradual change with time from the use of broader keywords to more specific sub-techniques in the literature as the research in the subfield deepens, for example, from a broad keyword like “AI” (2017) to “machine learning” (2018.4), a subset of AI, and then to “artificial neural networks” (2018) and “deep learning” (2018.67), which are both subsets of machine learning.

Discussion and Conclusions

This study used a bibliometric approach to review the research on online reviews in tourism and hospitality by analyzing 632 relevant documents selected from the Scopus database published between 2005 and 2019. The study applied co-citation and co-word analyses, including an innovative visualization of the co-word network to explore the theoretical foundations, network structure, and thematic evolution in this research area. It has successfully carried out an update, reviewed more studies, and revealed the theoretical foundations of the field with the use of science maps and shown the evolution of the major themes in the knowledge area. It has also provided an important quantitative complement with bibliometric methods to the previous literature reviews to enhance the understanding of this fast-growing knowledge area and has validated previous findings, such as the central research themes in online reviews.

Online review research is well established in the tourism and hospitality literature, with the top journals publishing literature in this research area. The earliest foundations of this research area was anchored to the adjacent body of knowledge on online reviews from other disciplines such

as marketing and management, conducted on electronic commerce websites such as Amazon and eBay, before extensive research by tourism and hospitality scholars based on hotel and tourism reviews grew. Besides influential conceptual papers and literature reviews that integrate the knowledge base of the area, most of the foundational papers focused on the consequences, value, usefulness, and effect of online reviews on sales and financial performance, customer booking intention, and service expectations. These publications were important because they answered the fundamental question of whether online reviews were relevant to tourism and hospitality, because it is only when they are relevant that attention needs to be paid to them. Therefore, these papers played a crucial role as foundational works since they tackled the important theoretical concepts as well as the relevance, consequences, and challenges of eWOM. Hence, these documents were often co-cited by many publications in this research area to establish a theoretical base upon which to build their own research.

The co-word analysis confirms the research that has been carried out in the past in the traditional cornerstones of the tourism and hospitality industry; hotels, travel and tourism, and restaurants as established in previous works (Schuckert et al. 2015), with customer satisfaction as a central concept and sentiment analysis as a predominant mode of analysis (Li et al. 2018). These themes not only occupy central positions but also maintain close relationships between themselves.

Big data and the application of machine learning are impacting virtually every industry (Lecun et al. 2015). The tourism and hospitality industry is not an exception, with the application of machine learning techniques such as artificial neural networks, support vector machines, and

deep learning to natural language processing tasks, especially sentiment analysis. Many of these techniques were developed in the 1970s and 1980s, but big data and increased computer-processing capacity has resulted in their revival and wide application in different fields, including natural language understanding (Burges 1998; Lecun et al. 2015). We expect the application of these techniques to grow in the text processing of online reviews, especially in the deployment of novel sub-techniques as they are developed as it has happened in the past i.e. in novel means of implementing neural networks and multilayer neural network (deep learning) models.

There is also an expansion of the application of online reviews beyond the traditional cornerstones of tourism and hospitality research into other themes like Airbnb, sustainability, and tourist experience. Airbnb's emergence with a unique accommodation business model has disrupted the tourism accommodation sector. The platform's celebration in March 2019 of reaching a milestone of 500 million guests, from about 10 million in June 2012, is evidence of its explosive growth (Airbnb 2019a; Guttentag 2015). Guest reviews are a feature of each Airbnb listing, and researchers have used these reviews to investigate guest experiences (Cheng and Jin 2019), host experiences (Cheng and Zhang 2019), and other research perspectives, and is an active area of research in the application of online reviews to tourism and hospitality.

The theme of tourist experience, connected with emotions, heritage and dark tourism, confirms the increasing demand for new forms of tourism and unique experiences, characterized by a departure from mass tourism (Stamboulis and Skayannis 2003), and an increase in demand for experiential tourism (Smith 2006), which these niche areas of tourism provide. The provision of

memorable tourist experiences is directly related to a businesses' ability to generate revenues (Kim et al. 2012), and destination managers and even countries can leverage on this. For example, Airbnb reported a five-fold increase in heritage travel globally since 2014 on its platform (Airbnb 2019b). The application of online reviews to investigate these niche areas of tourism is likely to continue to grow, as there were papers using online reviews to investigate other niche areas like polar tourism (Heimtun 2016), cycling tourism (Chiu and Leng 2017), plantation tours (Carter 2016) etc. among the reviewed papers.

Tourism accounts for about 8% of global greenhouse gas emissions (Lenzen et al. 2018), and sustainability has become an important policy issue in tourism and hospitality (Saarinen 2006). Hence, sustainability, sustainable tourism, ecotourism and other related keywords are an important emergent area and reviews have been used to investigate guest reactions to hotel sustainability measures, and the impact on guest experiences (Gerdt et al. 2019).

In conclusion, the research on online reviews and eWOM is reaching maturity with a rapidly increasing volume of new publications built on a solid base of seminal research and applied to highly varied areas, while increasingly sophisticated analytical methods are employed for analysis. This literature review has shown the foundational literature of this research area, recent growing areas of application in Airbnb and tourist experience, and in niche tourism areas such as wine tourism, dark tourism, etc. It has also revealed the latest analytical techniques being used, such as artificial neural networks and deep learning, and the hot and emergent topic areas, such as machine learning, big data analytics, and sustainability themes.

These findings are useful for scholars in this field to understand the current state of knowledge amidst the overwhelming growth in the literature and for general readers interested in learning about online reviews, and also provide practitioners and business owners a window to understanding a ubiquitous concept that they confront every day in the tourism and hospitality industry. The findings also provide a glimpse of future research growth areas toward the hot and emergent topics and the application of new machine learning sub-techniques using deep learning and neural networks for text processing. The thematic evolution chart, which highlights the hot and emergent topics, can be used in future bibliometric studies in any research area.

Limitations, Theoretical Implications, and Future Research

A limitation of this work is that authors sometime add arbitrary keywords to their papers that do not truly reflect the content of the paper, which could have affected the veracity of the co-word network map and thematic evolution timeline chart. The use of only the Scopus database is another limitation of this study that could have possibly resulted in leaving out a few publications indexed by WoS, but not in Scopus. It was not possible to combine both databases for this study because VOSviewer only allows the use of bibliographic data from either database, but not a combination of both. However, there is a significant overlap between both databases (Martín-Martín et al. 2018) and we recommend future work that can conduct a bibliometric analysis on a combination of both databases.

Our results have important implications for theory and future research. They show the conceptual structure and thematic evolution of this research area, and an individual or

researcher new to this research area can quickly obtain a bird's-eye view of the seminal publications and how the research themes have evolved to date. However, research is a continuous process and as mentioned, the co-citation network map, co-word network map, and, by extension, the thematic evolution chart are not static but change with time. Future studies are thereby needed to track the changes that take place in the conceptual structure and major research themes of this research area with time. In addition, in light of the COVID-19 pandemic, studies employing online reviews and eWOM from various sources such as social media are suggested to understand the various effects of the pandemic.

References

Airbnb. (2019a). *Airbnb Celebrates Half A Billion Guest Arrivals*. https://news.airbnb.com/airbnb-celebrates-half-a-billion-guest-arrivals/

Airbnb. (2019b). *Heritage Travel on the Rise: Airbnb and 23andMe Team Up to Make it Even Easier*. https://news.airbnb.com/heritage-travel-on-the-rise/

Ávila-Robinson, A., & Wakabayashi, N. (2018). Changes in the structures and directions of destination management and marketing research: A bibliometric mapping study, 2005–2016. *Journal of Destination Marketing and Management*, *10*, 101–111.

Bore, I., Rutherford, C., Glasgow, S., Aheri, B., & Antony, J. (2017). A systematic literature review on eWOM in the hotel industry: Current trends and suggestions for future research. *Hospitality and Society*, *7*(1), 63–85.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *UCINET 6 for Windows: Software for Social Network Analysis. Harvard, MA, Analytic Technologies*.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, *2*(2), 121–167.

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics*, *22*(1), 155–205.

Carter, P. L. (2016). Where are the enslaved?: Tripadvisor and the narrative landscapes of southern plantation museums. *Journal of Heritage Tourism*, *11*(3), 235–249.

Centobelli, P., & Ndou, V. (2019). Managing customer knowledge through the use of big data analytics in tourism research. *Current Issues in Tourism*, *22*(15), 1862–1882.

Chen, Y.-F., & Law, R. (2016). A Review of Research on Electronic Word-of-Mouth in Hospitality and Tourism Management. *International Journal of Hospitality and Tourism Administration*, *17*(4), 347–372.

Cheng, M., & Jin, X. (2019). What do Airbnb users care about? An analysis of online review comments. *International Journal of Hospitality Management*, *76*, 58–70.

Cheng, M., & Zhang, G. (2019). When Western hosts meet Eastern guests: Airbnb hosts' experience with Chinese outbound tourists. *Annals of Tourism Research*, *75*, 288–303.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(3), 345–354.

Chiu, W., & Leng, H. K. (2017). Let's go cycling: an analysis of tourists' experience on online user-generated content. *International Journal of Tourism Cities*, *3*(1), 30–42.

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, *62*(7), 1382–1402.

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, *49*(10), 1407–1424.

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management*, *52*, 498–506.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*(3), 291–313.

Fotis, J., Buhalis, D., & Rossides, N. (2012). *Social Media Use and Impact during the Holiday Travel Planning Process BT - Information and Communication Technologies in Tourism 2012* (M. Fuchs, F. Ricci, & L. Cantoni (eds.); pp. 13–24). Springer Vienna.

Garfield, E., Malin, V. M., & Small, H. (1983). Citation Data as Science Indicators. *Toward a Metric of a Science*, *6*, 29.

Gerdt, S.-O., Wagner, E., & Schewe, G. (2019). The relationship between sustainability and customer satisfaction in hospitality: An explorative investigation using eWOM as a data source. *Tourism Management*, *74*, 155–172.

Guttentag, D. (2015). Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism*, *18*(12), 1192–1217.

He, Q. (1999). Knowledge Discovery Through Co-Word Analysis. *Library Trends*, *48*(1), 133–159.

He, W., Tian, X., Tao, R., Zhang, W., Yan, G., & Akula, V. (2017). Application of social media analytics: A case of analyzing online hotel reviews. *Online Information Review*, *41*(7), 921–935.

Heimtun, B. (2016). Emotions and affects at work on Northern Lights tours. *Hospitality and Society*, *6*(3), 223–241.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, *18*(1), 38–52.

Hu, C.-P., Hu, J.-M., Deng, S.-L., & Liu, Y. (2013). A co-word analysis of library and information science in China. *Scientometrics*, *97*(2), 369–382.

Kim, J.-H., Ritchie, J. R. B., & McCormick, B. (2012). Development of a scale to measure memorable tourism experiences. *Journal of Travel Research*, *51*(1), 12–25.

Kwok, L., Xie, K. L., & Richards, T. (2017). Thematic framework of online review research: A systematic analysis of contemporary literature on seven major hospitality and tourism journals. *International Journal of Contemporary Hospitality Management*, *29*(1), 307–354.

Law, R., Buhalis, D., & Cobanoglu, C. (2014). Progress on information and communication technologies in hospitality and tourism. *International Journal of Contemporary Hospitality*

Management, *26*(5), 727–750.

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Lenzen, M., Sun, Y.-Y., Faturay, F., Ting, Y.-P., Geschke, A., & Malik, A. (2018). The carbon footprint of global tourism. *Nature Climate Change*, *8*(6), 522–528.

Leung, D., Law, R., van Hoof, H., & Buhalis, D. (2013). Social Media in Tourism and Hospitality: A Literature Review. *Journal of Travel and Tourism Marketing*, *30*(1–2), 3–22.

Leung, X. Y., Sun, J., & Bai, B. (2017). Bibliometrics of social media research: A co-citation and co-word analysis. *International Journal of Hospitality Management*, *66*, 35–45.

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*, 301–323.

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, *29*(3), 458–468.

Liu, B. (2012). Sentiment analysis and opinion mining. In *Synthesis Lectures on Human Language Technologies* (Vol. 5, Issue 1).

Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, *47*, 140–151.

Liu, Z., Yin, Y., Liu, W., & Dunford, M. (2015). Visualizing the intellectual structure and evolution

of innovation systems research: a bibliometric analysis. *Scientometrics*, *103*(1), 135–158.

Lu, W., & Stepchenkova, S. (2015). User-Generated Content as a Research Mode in Tourism and Hospitality Applications: Topics, Methods, and Software. *Journal of Hospitality Marketing and Management*, *24*(2), 119–154.

Lu, Y., Chen, Z., & Law, R. (2018). Mapping the progress of social media research in hospitality and tourism management from 2004 to 2014. *Journal of Travel and Tourism Marketing*, *35*(2), 102–118.

Mariani, M., Baggio, R., Fuchs, M., & Höepken, W. (2018). Business intelligence and big data in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, *30*(12), 3514–3554.

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, *12*(4), 1160–1177.

Mauri, A. G., & Minazzi, R. (2013). Web reviews influence on expectations and purchasing intentions of hotel potential customers. *International Journal of Hospitality Management*, *34*(1), 99–107.

Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, *246*(1), 1–19.

Nerur, S. P., Rasheed, A. A., & Natarajan, V. (2008). The intellectual structure of the strategic management field: An author co-citation analysis. *Strategic Management Journal*, *29*(3), 319–336.

Nusair, K., Butt, I., & Nikhashemi, S. R. (2019). A bibliometric analysis of social media in hospitality and tourism research. *International Journal of Contemporary Hospitality Management*, *31*(7), 2691–2719.

Papathanassis, A., & Knolle, F. (2011). Exploring the adoption and processing of online holiday reviews: A grounded theory approach. *Tourism Management*, *32*(2), 215–224.

Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research*, *50*, 67–83.

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, *25*, 348–349.

Saarinen, J. (2006). Traditions of sustainability in tourism studies. *Annals of Tourism Research*, *33*(4), 1121–1140.

Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and Tourism Online Reviews: Recent Trends and Future Directions. *Journal of Travel and Tourism Marketing*, *32*(5), 608–621.

Serra Cantallops, A., & Salvi, F. (2014). New consumer behavior: A review of research on eWOM and hotels. *International Journal of Hospitality Management*, *36*, 41–51.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, *24*(4), 265–269.

Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for Information Science*, *50*(9), 799–813.

Smith, W. L. (2006). Experiential tourism around the world and at home: Definitions and standards. *International Journal of Services and Standards*, *2*(1), 1–14.

Sotiriadis, M. D. (2017). Sharing tourism experiences in social media: A literature review and a set of suggested business strategies. *International Journal of Contemporary Hospitality Management*, *29*(1), 179–225.

Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, *32*(6), 1310–1323.

Sparks, B. A., Perkins, H. E., & Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tourism Management*, *39*, 1–9.

Stamboulis, Y., & Skayannis, P. (2003). Innovation strategies and technology for experience-based tourism. *Tourism Management*, *24*(1), 35–43.

Valente, T., Coronges, K., Lakon, C., & Costenbader, E. (2008). How Correlated Are Network

Centrality Measures? *Connections (Toronto, Ont.)*, *28*, 16–26.

van Eck, N J, & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538.

van Eck, Nees Jan, & Waltman, L. (2013). VOSviewer manual. *Leiden: Univeristeit Leiden*, *September*.

van Eck, Nees Jan, & Waltman, L. (2014). Visualizing Bibliometric Networks. In *Measuring Scholarly Impact*.

Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management*, *30*(1), 123–127.

Wang, Y., Yu, Q., & Fesenmaier, D. R. (2002). Defining the virtual tourist community: Implications for tourism marketing. *Tourism Management*, *23*(4), 407–417.

Whittaker, J. (1989). Creativity and Conformity in Science: Titles, Keywords and Co-word Analysis. *Social Studies of Science*, *19*(3), 473–496.

Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, *31*(2), 179–188.

Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management*, *43*, 1–12.

Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior*, *27*(2), 634–639.

Yearofreturn.com. (2019). *Year Of Return – Celebrating 400 years of African Resilience…*. https://www.yearofreturn.com/

Zeng, B., & Gerritsen, R. (2014). What do we know about social media in tourism? A review. *Tourism Management Perspectives*, *10*, 27–36.

Zhang, Z., Ye, Q., Law, R., & Li, Y. (2010). The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management*, *29*(4), 694–700.

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on Sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, *74*(2), 133–148.

Supplementary file 1

Supplementary Information 1

Search Queries

1. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (hotel OR hotels OR hotel AND industry OR hotel AND management))
2. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (tourism OR tourism AND industry OR tourism AND management))
3. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (hospitality OR hospitality AND industry OR hospitality AND management))
4. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (tourist AND destination))
5. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (travel))
6. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (cruises OR camping OR restaurants))
7. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (hotel AND performance))
8. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (hotel AND revenues OR hotel AND revenue AND management OR revenue AND management))
9. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY ("Hotel Revenues" OR "Hotel Revenue Management" OR "Revenue Management"))
10. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (hotel OR hotels OR hotel AND industry OR hotel AND management) AND TITLE-ABS-KEY (profitability OR financial AND performance))
11. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (hotel OR hotels OR "Hotel Industry" OR "Hotel Management") AND TITLE-ABS-KEY (profitability OR "Financial Performance"))
12. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (hotel OR hotels OR hotel AND industry OR hotel AND management) AND TITLE-ABS-KEY (customer AND satisfaction))
13. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (tourism OR "Tourism Industry" OR "Tourism Management") AND TITLE-ABS-KEY (profitability OR "Financial Performance"))
14. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (tourism OR "Tourism Industry" OR "Tourism Management") AND TITLE-ABS-KEY ("customer satisfaction"))
15. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (hospitality OR "Hospitality Industry" OR "Hospitality Management") AND TITLE-ABS-KEY (profitability OR "Financial Performance"))
16. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (hospitality OR "Hospitality Industry" OR "Hospitality Management") AND TITLE-ABS-KEY ("customer satisfaction"))
17. (TITLE-ABS-KEY (tripadvisor OR booking.com OR airbnb OR twitter OR facebook) AND TITLE-ABS-KEY (hotel OR hotels OR hotel AND industry OR "Hotel Industry" OR hotel AND management OR "Hotel Management"))
18. (TITLE-ABS-KEY (tripadvisor OR booking.com OR airbnb OR twitter OR facebook) AND TITLE-ABS-KEY (tourism OR "Tourism Industry" OR "Tourism Management"))

19. (TITLE-ABS-KEY (tripadvisor OR booking.com OR airbnb OR twitter OR facebook) AND TITLE-ABS-KEY (hospitality OR "Hospitality Industry" OR "Hospitality Management"))
20. (TITLE-ABS-KEY (tripadvisor OR booking.com OR airbnb OR twitter OR facebook) AND TITLE-ABS-KEY (profitability OR "Financial Performance"))
21. (TITLE-ABS-KEY (tripadvisor OR booking.com OR airbnb OR twitter OR facebook) AND TITLE-ABS-KEY ("customer satisfaction"))
22. (TITLE-ABS-KEY (tripadvisor OR booking.com OR airbnb OR twitter OR facebook) AND TITLE-ABS-KEY ("Hotel Performance"))
23. (TITLE-ABS-KEY (tripadvisor OR booking.com OR airbnb OR twitter OR facebook) AND TITLE-ABS-KEY (hotel AND revenues OR "Hotel Revenues" OR hotel AND revenue AND management OR "Hotel Revenue Management" OR revenue AND management OR "Revenue Management"))
24. (TITLE-ABS-KEY (tripadvisor OR booking.com OR airbnb OR twitter OR facebook) AND TITLE-ABS-KEY (cruises OR camping OR restaurants))
25. (TITLE-ABS-KEY ("Data Mining" OR "Opinion Mining" OR "Text Mining" OR "Text Analytics" OR "Web Mining") AND TITLE-ABS-KEY (hotel OR hotels OR hotel AND industry OR "Hotel Industry" OR hotel AND management OR "Hotel Management"))
26. (TITLE-ABS-KEY ("Data Mining" OR "Opinion Mining" OR "Text Mining" OR "Text Analytics" OR "Web Mining") AND TITLE-ABS-KEY (tourism OR tourism AND industry OR "Tourism Industry" OR tourism AND management OR "Tourism Management" OR tourist AND destination))
27. (TITLE-ABS-KEY ("Data Mining" OR "Opinion Mining" OR "Text Mining" OR "Text Analytics" OR "Web Mining") AND TITLE-ABS-KEY (hospitality OR hospitality AND industry OR "Hospitality Industry" OR hospitality AND management OR "Hospitality Management"))
28. (TITLE-ABS-KEY ("Data Mining" OR "Opinion Mining" OR "Text Mining" OR "Text Analytics" OR "Web Mining") AND TITLE-ABS-KEY ("customer satisfaction") AND TITLE-ABS-KEY (hotel OR tourism OR hospitality))
29. (TITLE-ABS-KEY ("Data Mining" OR "Opinion Mining" OR "Text Mining" OR "Text Analytics" OR "Web Mining") AND TITLE-ABS-KEY (profitability OR "Financial Performance") AND TITLE-ABS-KEY (hotel OR tourism OR hospitality))
30. (TITLE-ABS-KEY ("Sentiment Analysis") AND TITLE-ABS-KEY (hotel OR hotels OR hotel AND industry OR "Hotel Industry" OR hotel AND management OR "Hotel Management"))
31. (TITLE-ABS-KEY ("Sentiment Analysis") AND TITLE-ABS-KEY (tourism OR tourism AND industry OR "Tourism Industry" OR tourism AND management OR "Tourism Management" OR tourist AND destination))
32. (TITLE-ABS-KEY ("Sentiment Analysis") AND TITLE-ABS-KEY (hospitality OR hospitality AND industry OR "Hospitality Industry" OR hospitality AND management OR "Hospitality Management"))
33. (TITLE-ABS-KEY ("Sentiment Analysis") AND TITLE-ABS-KEY (profitability OR "Financial Performance") AND TITLE-ABS-KEY (hotel OR tourism OR hospitality))
34. (TITLE-ABS-KEY ("Sentiment Analysis") AND TITLE-ABS-KEY ("customer satisfaction") AND TITLE-ABS-KEY (hotel OR tourism OR hospitality))
35. (TITLE-ABS-KEY (ewom OR "Electronic Word of Mouth" OR e-wom OR electronic AND word-of-mouth) AND TITLE-ABS-KEY (hotel OR hotels OR hotel AND industry OR "Hotel Industry" OR hotel AND management OR "Hotel Management"))
36. (TITLE-ABS-KEY (ewom OR "Electronic Word of Mouth" OR e-wom OR electronic AND word-of-mouth) AND TITLE-ABS-KEY (tourism OR tourism AND industry OR "Tourism Industry" OR tourism AND management OR "Tourism Management" OR tourist AND destination))
37. (TITLE-ABS-KEY (ewom OR "Electronic Word of Mouth" OR e-wom OR electronic AND word-of-mouth) AND TITLE-ABS-KEY (hospitality OR hospitality AND industry OR "Hospitality Industry" OR hospitality AND management OR "Hospitality Management"))

38. (TITLE-ABS-KEY (ewom OR "Electronic Word of Mouth" OR e-wom OR electronic AND word-of-mouth) AND TITLE-ABS-KEY (profitability OR "Financial Performance"))
39. (TITLE-ABS-KEY (ewom OR "Electronic Word of Mouth" OR e-wom OR electronic AND word-of-mouth) AND TITLE-ABS-KEY ("customer satisfaction") AND TITLE-ABS-KEY (hotel OR tourism OR hospitality OR lodging))
40. (TITLE-ABS-KEY ("Online Booking" OR "Online Hotel Booking" OR "Booking Intention") AND TITLE-ABS-KEY (hotel OR tourism OR hospitality OR lodging) AND TITLE-ABS-KEY (online AND reviews))
41. (TITLE-ABS-KEY ("Big Data" OR "Data Analytics" OR "Big Data Analytics") AND TITLE-ABS-KEY (hotel OR tourism OR hospitality OR lodging))
42. (TITLE-ABS-KEY ("Responsible Tourism" OR "Sustainable tourism" OR "Tourism Sustainability" OR ecotourism OR "Eco tourism" OR eco-tourism OR "Green Tourism" OR "Green Economy" OR "Green Hotel" OR "Circular Tourism" OR "Circular Economy" OR bioeconomy OR "Bio economy") AND TITLE-ABS-KEY ("Online Reviews" OR "Online Ratings" OR tripadvisor OR booking.com OR airbnb OR twitter))
43. (TITLE-ABS-KEY (blue AND tourism OR "Blue Tourism" OR blue AND economy OR "Blue Economy") AND TITLE-ABS-KEY ("Online Reviews" OR "Online Ratings" OR tripadvisor OR booking.com OR airbnb OR twitter))
44. (TITLE-ABS-KEY ("Online reviews" OR "Online Hotel Reviews" OR "Online Ratings") AND TITLE-ABS-KEY (experience OR "Experience economy" OR "Experience tourism"))
45. (TITLE-ABS-KEY (hotel) AND TITLE-ABS-KEY (tripadvisor))

Supplementary Information 2

Full list of Co-word/Co-occurrence network

Label	Occurrences	Cluster
attitude	4	1
behavioral research	6	1
booking intentions	8	1
booking.com	8	1
branding	9	1
credibility	7	1
data analytics	6	1
empirical analysis	7	1
ewom	84	1
experimental study	3	1
heuristics	5	1
hotel attributes	9	1
hotel image	3	1
hotel performance	11	1
hotel star rating	12	1

industrial performance	4	1
management responses	16	1
numerical model	5	1
online ratings	28	1
performance	3	1
reputation management	23	1
review valence	5	1
reviewer expertise	3	1
revisit intention	5	1
revpar	3	1
tourism marketing	16	1
travel reviews	8	1
trustworthiness	3	1
usefulness	3	1
valence	5	1
artificial intelligence	4	2
classification	7	2
data mining	29	2
decision support systems	4	2
deep learning	3	2
e-tourism	4	2
electronic commerce	12	2
feature extraction	4	2
forecasting	3	2
helpfulness	3	2
hospitality services	3	2
information overloads	3	2
information systems	16	2
natural language processing	7	2
online reviews	228	2
opinion mining	8	2
prediction	4	2
products and services	4	2

quantitative analysis	3	2
response management	5	2
restaurant reviews	8	2
review helpfulness	11	2
review quality	3	2
sentiment analysis	42	2
social influence	6	2
social media analytics	10	2
source credibility	6	2
web 2.0	9	2
animalia	3	3
conservation	3	3
corporate social responsibility	3	3
cultural heritage	4	3
culture	7	3
decision making	13	3
ecotourism	7	3
environment	8	3
italy	6	3
perception	20	3
protected area	3	3
quality control	4	3
quality of service	3	3
sustainability	15	3
sustainable tourism	5	3
tourism development	8	3
tourism management	34	3
tourist attraction	13	3
tourist behavior	22	3
tourist perceptions	3	3
tourist satisfaction	3	3
wildlife management	3	3
wildlife tourism	4	3

accommodation	8	4
conceptual framework	3	4
consumption behavior	10	4
cruise tourism	5	4
europe	3	4
familiarity	3	4
innovation	3	4
leximancer	7	4
lodging	4	4
management	5	4
management practice	4	4
market conditions	3	4
online reputation	8	4
overtourism	3	4
perceived value	3	4
public attitude	3	4
qualitative analysis	11	4
ranking	3	4
software	5	4
tourism	58	4
tourism economics	9	4
twitter	5	4
authenticity	4	5
communication	3	5
dark tourism	4	5
elaboration likelihood models	4	5
emotions	6	5
expertise	3	5
facebook	3	5
gender	4	5
heritage tourism	5	5
media role	6	5
methodology	5	5

netnography	13	5
online marketing	22	5
restaurants	11	5
social media	98	5
stress	3	5
technology acceptance model	3	5
text classification	4	5
tourist experience	14	5
travel	10	5
trust	11	5
china	19	6
comparative study	4	6
cultural tourism	3	6
gis	3	6
hong kong	5	6
hotel location	3	6
life satisfaction	3	6
location	3	6
machine learning	10	6
market segmentation	4	6
preference behavior	3	6
psychology	3	6
smart tourism destination	4	6
spain	11	6
spatial analysis	5	6
tourist destination	43	6
travel behavior	21	6
travel blogs	3	6
urban tourism	3	6
willingness to pay	4	6
airbnb	22	7
consumer behaviour	8	7
customer dissatisfaction	7	7

distrust	3	7
genre analysis	5	7
guest experiences	4	7
hospitality	33	7
hotel reviews	30	7
luxury hotels	3	7
online complaints	11	7
rapport management	3	7
service failure	4	7
service recovery	8	7
sharing economy	17	7
thailand	3	7
travel websites	4	7
tripadvisor	79	7
value co-creation	4	7
value co-destruction	3	7
word-of-mouth	23	7
cluster analysis	4	8
content analysis	25	8
france	5	8
green practices	5	8
hedonic analysis	9	8
online hotel bookings	7	8
paris	3	8
price dynamics	6	8
revenue management	3	8
semantics	3	8
service sector	7	8
social media marketing	4	8
text analytics	49	8
topic modeling	7	8
united kingdom	6	8
wine tourism	5	8

consumer-generated media	10	9
destination image	8	9
experience	6	9
guest satisfaction	3	9
international tourism	6	9
language	7	9
literature review	8	9
macau	3	9
peer-to-peer accommodation	3	9
research	8	9
service quality	20	9
systematic review	4	9
tourism research	3	9
turkey	4	9
user-generated content	65	9
big data	37	10
big data analytics	6	10
business analytics	6	10
competition	9	10
cultural differences	3	10
customer satisfaction	51	10
hotels	105	10
network analysis	3	10
public relations	4	10
regression analysis	8	10
sales	14	10
urban area	3	10
value creation	5	10
visibility	3	10
artificial neural network	4	11
customer feedback	3	11
modeling	5	11
new york	5	11

performance assessment	9	11
portugal	8	11
quality management	3	11
sensitivity analysis	3	11
support vector machine	4	11
united states	14	11

Supplementary Information 3

Foundational literature

	Author(s)	Paper Title	Journal	Year
1	Serra Cantallops, A., Salvi, F.	New consumer behavior: A review of research on eWom and hotels	IJHM	2014
2	Fang, B., Ye, Q., Kucukusta, D., Law, R.	Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics	TM	2016
3	Hennig-Thurau, T., Gwinner, K.P., Walsh, G., Gremler, D.D.	Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet?	JIM	2004
4	Liu, Z., Park, S.	What makes a useful online review? Implication for travel product websites	TM	2015
5	Mauri, A.G., Minazzi, R	Web reviews influence on expectations and purchasing intentions of hotel potential customers	IJHM	2013
6	Park, S., Nicolau, J.L	Asymmetric effects of online consumer reviews	ATR	2015
7	Schuckert, M., Liu, X., Law, R	Hospitality and Tourism Online Reviews: Recent Trends and Future Directions	JTTM	2015
8	Xie, K.L., Zhang, Z., Zhang, Z.	The business value of online consumer reviews and management response to hotel performance	IJHM	2014
9	Litvin, S.W., Goldsmith, R.E., Pan, B.	Electronic word-of-mouth in hospitality and tourism management	TM	2008
10	Papathanassis, A., Knolle, F.	Exploring the adoption and processing of online holiday reviews: A grounded theory approach	TM	2011
11	Sparks, B.A., Browning, V.	The impact of online reviews on hotel booking intentions and perception of trust	TM	2011
12	Vermeulen, I.E., Seegers, D	Tried and tested: The impact of online hotel reviews on consumer consideration	TM	2009
13	Xiang, Z., Gretzel, U.	Role of social media in online travel information search	TM	2010
14	Ye, Q., Law, R., Gu, B., Chen, W.	The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings	CHB	2011

15	Ye, Q., Law, R., Gu, B.	The impact of online user reviews on hotel room sales	IJHM	2009
16	Zhang, Z., Ye, Q., Law, R., Li, Y.	The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews	IJHM	2010
17	Chevalier, J.A., Mayzlin, D.	The effect of word of mouth on sales: Online book reviews	JMR	2006
18	Dellarocas, C.	The digitization of word of mouth: Promise and challenges of online feedback mechanisms	MS	2003
19	Forman, C., Ghose, A., Wiesenfeld, B.,	Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets	ISR	2008
20	Zhu, F., Zhang, X	Impact of online consumer reviews on Sales: The moderating role of product and consumer characteristics	JM	2010

Supplementary file 2

Keywords	ANC	APY	Total	2019	2018	2017	2016	2015	2014	2013	2012	2011	2010	2009	2008	2007	2006
accommodation	1.2371	2017	8	1	3	2	0	1	1	0	0	0	0	0	0	0	0
airbnb	2.0539	2018.4545	22	13	6	3	0	0	0	0	0	0	0	0	0	0	0
animalia	0.5152	2018.3333	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
artificial intelligence	0.6172	2017	4	0	1	3	0	0	0	0	0	0	0	0	0	0	0
artificial neural network	4.4952	2018	4	3	0	0	0	1	0	0	0	0	0	0	0	0	0
attitude	1.2355	2017.5	4	1	2	0	0	1	0	0	0	0	0	0	0	0	0
authenticity	0.4065	2014.75	4	1	0	0	0	0	1	2	0	0	0	0	0	0	0
behavioral research	1.1578	2015.8333	6	0	1	2	2	0	0	0	1	0	0	0	0	0	0
big data	2.3596	2018.1622	37	18	11	6	0	2	0	0	0	0	0	0	0	0	0
big data analytics	0.3489	2018.6667	6	6	0	0	0	0	0	0	0	0	0	0	0	0	0
booking intentions	0.6257	2016.5	8	1	2	3	0	1	0	0	1	0	0	0	0	0	0
booking.com	0.8197	2017.25	8	0	4	3	0	1	0	0	0	0	0	0	0	0	0
branding	0.5158	2016.6667	9	0	5	1	1	0	1	1	0	0	0	0	0	0	0
business analytics	2.2667	2017.6667	6	2	2	1	0	1	0	0	0	0	0	0	0	0	0
china	0.8649	2016.5789	19	5	3	5	0	1	3	1	0	1	0	0	0	0	0
classification	0.745	2016.2857	7	1	2	2	0	0	0	1	1	0	0	0	0	0	0
cluster analysis	1.266	2017.75	4	1	3	0	0	0	0	0	0	0	0	0	0	0	0
communication	2.2763	2015	3	1	0	0	0	0	0	2	0	0	0	0	0	0	0
comparative study	3.5693	2018.25	4	2	2	0	0	0	0	0	0	0	0	0	0	0	0
competition	0.9843	2016.6667	9	1	3	2	2	0	0	0	0	1	0	0	0	0	0
conceptual framework	1.0104	2015.3333	3	0	0	1	1	0	0	1	0	0	0	0	0	0	0
conservation	0.5152	2018.3333	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
consumer behaviour	0.5992	2016.375	8	3	2	0	0	1	0	0	1	1	0	0	0	0	0

consumer-generated media	1.3298	2015.2	10	1	3	1	0	2	0	0	2	0	0	0	1	0	0
consumption behavior	2.6387	2016.5	10	3	1	2	1	1	1	0	0	1	0	0	0	0	0
content analysis	0.9718	2017.04	25	10	5	2	1	2	2	2	1	0	0	0	0	0	0
corporate social responsibility	1.2867	2016.6667	3	1	1	0	0	0	0	1	0	0	0	0	0	0	0
credibility	0.6998	2016.2857	7	2	1	1	0	1	1	0	1	0	0	0	0	0	0
cruise tourism	0.2956	2018.2	5	3	2	0	0	0	0	0	0	0	0	0	0	0	0
cultural differences	1.0082	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
cultural heritage	0.3968	2017	4	2	0	1	0	0	1	0	0	0	0	0	0	0	0
cultural tourism	0.5863	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
culture	1.0413	2017.5714	7	4	1	0	1	0	1	0	0	0	0	0	0	0	0
customer dissatisfaction	2.0716	2017.7143	7	3	1	1	2	0	0	0	0	0	0	0	0	0	0
customer feedback	0.6522	2015.3333	3	1	0	1	0	0	0	0	0	0	1	0	0	0	0
customer satisfaction	1.1899	2016.8627	51	17	7	11	4	6	1	1	1	1	2	0	0	0	0
dark tourism	0.0483	2018	4	0	4	0	0	0	0	0	0	0	0	0	0	0	0
data analytics	0.9777	2018.1667	6	2	4	0	0	0	0	0	0	0	0	0	0	0	0
data mining	1.0915	2016.931	29	10	2	10	0	3	0	3	0	1	0	0	0	0	0
decision making	2.2693	2017.6154	13	5	5	1	1	0	0	0	0	1	0	0	0	0	0
decision support systems	0.8895	2016.5	4	0	1	1	1	1	0	0	0	0	0	0	0	0	0
deep learning	0.5795	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
destination image	0.4222	2016.25	8	2	1	1	1	2	0	0	0	1	0	0	0	0	0
distrust	1.3235	2018.3333	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
e-tourism	1.0339	2014.5	4	1	0	0	1	0	0	0	1	1	0	0	0	0	0
ecotourism	0.8375	2016.8571	7	2	3	0	0	0	1	0	1	0	0	0	0	0	0
elaboration likelihood models	0.6759	2017.25	4	2	1	0	0	0	0	1	0	0	0	0	0	0	0
electronic commerce	1.8597	2016.1667	12	2	2	2	2	2	0	1	0	1	0	0	0	0	0
emotions	0.7413	2017.5	6	1	3	0	2	0	0	0	0	0	0	0	0	0	0
empirical analysis	1.8768	2016	7	0	2	2	1	1	0	0	0	1	0	0	0	0	0
environment	0.8816	2017.375	8	2	2	2	1	1	0	0	0	0	0	0	0	0	0
europe	1.2726	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
ewom	1.0867	2016.4048	84	20	18	10	11	8	4	5	2	5	0	0	1	0	0
experience	0.7103	2017.1667	6	1	2	1	1	1	0	0	0	0	0	0	0	0	0
experimental study	1.8029	2013.6667	3	0	0	1	0	0	0	1	0	1	0	0	0	0	0
expertise	1.4596	2015	3	1	0	0	0	1	0	0	0	1	0	0	0	0	0
facebook	1.0405	2016.3333	3	0	2	0	0	0	1	0	0	0	0	0	0	0	0
familiarity	1.6368	2017	3	2	0	0	0	1	0	0	0	0	0	0	0	0	0
feature extraction	0.1439	2014.5	4	0	0	2	0	0	0	1	0	1	0	0	0	0	0
forecasting	0.3177	2017	3	1	0	1	0	1	0	0	0	0	0	0	0	0	0

france	1.3588	2015	5	0	1	1	1	0	0	0	2	0	0	0	0	0	0
gender	0.3658	2016.5	4	2	0	1	0	0	0	0	0	1	0	0	0	0	0
genre analysis	0.2069	2018	5	2	1	2	0	0	0	0	0	0	0	0	0	0	0
gis	0.3133	2017	3	0	1	2	0	0	0	0	0	0	0	0	0	0	0
green practices	0.7574	2017.4	5	0	3	2	0	0	0	0	0	0	0	0	0	0	0
guest experiences	1.5347	2017.75	4	2	1	0	0	1	0	0	0	0	0	0	0	0	0
guest satisfaction	0.322	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
hedonic analysis	0.9062	2017.2222	9	4	1	0	3	0	0	1	0	0	0	0	0	0	0
helpfulness	0.5683	2017.3333	3	1	1	0	0	1	0	0	0	0	0	0	0	0	0
heritage tourism	0.3838	2018.2	5	2	3	0	0	0	0	0	0	0	0	0	0	0	0
heuristics	1.6431	2016.4	5	2	1	0	0	1	0	0	0	1	0	0	0	0	0
hong kong	1.1404	2015.2	5	1	0	1	1	0	0	1	0	1	0	0	0	0	0
hospitality	1.8546	2017.6364	33	9	10	10	1	3	0	0	0	0	0	0	0	0	0
hospitality services	0.5303	2016.3333	3	1	1	0	0	0	0	0	1	0	0	0	0	0	0
hotel attributes	0.9276	2016.1111	9	1	1	2	1	2	2	0	0	0	0	0	0	0	0
hotel image	0.4384	2014.6667	3	0	1	0	0	0	1	0	1	0	0	0	0	0	0
hotel location	1.8098	2018.3333	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
hotel performance	1.7117	2016.6364	11	1	2	3	3	1	1	0	0	0	0	0	0	0	0
hotel reviews	1.1454	2016.6	30	8	5	5	5	2	1	3	0	0	0	0	1	0	0
hotel star rating	0.7969	2016.0833	12	0	3	1	3	4	1	0	0	0	0	0	0	0	0
hotels	1.3198	2016.7714	105	26	26	20	11	5	5	3	4	2	2	1	0	0	0
industrial performance	2.4272	2016.25	4	0	1	1	0	2	0	0	0	0	0	0	0	0	0
information overloads	1.4352	2017.6667	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0
information systems	1.6556	2015.125	16	0	4	3	2	2	1	1	0	2	0	0	1	0	0
innovation	2.1606	2018.3333	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
international tourism	1.3442	2016	6	1	1	2	0	0	1	0	1	0	0	0	0	0	0
italy	0.4265	2017.6667	6	2	1	3	0	0	0	0	0	0	0	0	0	0	0
language	2.0594	2017.5714	7	2	2	3	0	0	0	0	0	0	0	0	0	0	0
leximancer	0.7045	2017.2857	7	3	1	1	0	1	1	0	0	0	0	0	0	0	0
life satisfaction	0.1846	2018	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
literature review	2.2386	2018	8	2	4	2	0	0	0	0	0	0	0	0	0	0	0
location	1.0927	2018	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
lodging	0.7972	2016.75	4	0	2	0	1	1	0	0	0	0	0	0	0	0	0
luxury hotels	0.5631	2014.6667	3	1	0	0	0	0	0	1	1	0	0	0	0	0	0
macau	0.309	2017.6667	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0
machine learning	2.4658	2018.4	10	5	5	0	0	0	0	0	0	0	0	0	0	0	0
management	1.4913	2017.4	5	1	2	1	0	1	0	0	0	0	0	0	0	0	0

management practice	2.1027	2017.25	4	1	1	0	2	0	0	0	0	0	0	0	0	0	0
management responses	0.8553	2017.4375	16	6	2	4	2	1	1	0	0	0	0	0	0	0	0
market conditions	2.9621	2018	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0
market segmentation	1.3096	2016.5	4	2	1	0	0	0	0	0	1	0	0	0	0	0	0
media role	2.4673	2016	6	1	1	0	1	2	0	1	0	0	0	0	0	0	0
methodology	0.3094	2016	5	0	2	0	0	2	1	0	0	0	0	0	0	0	0
modeling	2.127	2017	5	2	0	1	2	0	0	0	0	0	0	0	0	0	0
natural language processing	1.4074	2017.1429	7	1	1	4	0	1	0	0	0	0	0	0	0	0	0
netnography	0.3885	2016.0769	13	3	3	2	0	0	1	2	2	0	0	0	0	0	0
network analysis	0.7151	2018.3333	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
new york	2.5521	2016	5	1	0	1	0	2	1	0	0	0	0	0	0	0	0
numerical model	1.6548	2016.8	5	0	1	4	0	0	0	0	0	0	0	0	0	0	0
online complaints	1.2518	2014.6364	11	1	1	0	2	1	2	2	1	1	0	0	0	0	0
online hotel bookings	0.7537	2016.4286	7	3	0	0	2	0	1	0	1	0	0	0	0	0	0
online marketing	1.8147	2014.9545	22	3	4	0	2	4	0	5	2	1	0	0	1	0	0
online ratings	1.4153	2017.1786	28	10	4	6	3	3	0	0	2	0	0	0	0	0	0
online reputation	1.8779	2017.625	8	1	4	3	0	0	0	0	0	0	0	0	0	0	0
online reviews	1.1554	2016.8202	228	68	49	37	20	19	7	11	7	5	3	1	1	0	0
opinion mining	0.4425	2014.25	8	2	0	0	1	2	0	0	1	0	1	0	1	0	0
overtourism	0.5507	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
paris	1.2841	2015	3	0	0	1	1	0	0	0	1	0	0	0	0	0	0
peer-to-peer accommodation	0.4575	2019	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
perceived value	1.1864	2016.6667	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0
perception	1.1038	2017.35	20	6	6	4	1	2	0	0	0	1	0	0	0	0	0
performance	0.851	2018	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
performance assessment	2.2672	2015.4444	9	1	2	0	2	1	0	1	2	0	0	0	0	0	0
portugal	0.5711	2015.875	8	1	2	1	1	0	1	1	1	0	0	0	0	0	0
prediction	1.3465	2017.75	4	2	2	0	0	0	0	0	0	0	0	0	0	0	0
preference behavior	1.0486	2017	3	1	0	2	0	0	0	0	0	0	0	0	0	0	0
price dynamics	1.8093	2017	6	2	2	0	1	0	0	0	1	0	0	0	0	0	0
products and services	0.6457	2014	4	0	0	1	0	1	0	1	0	1	0	0	0	0	0
protected area	0.6439	2018	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0
psychology	0.7013	2017	3	0	1	2	0	0	0	0	0	0	0	0	0	0	0
public attitude	1.2561	2017.6667	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
public relations	0.5601	2016.5	4	0	0	4	0	0	0	0	0	0	0	0	0	0	0
qualitative analysis	1.5513	2016.9091	11	2	4	1	1	2	0	1	0	0	0	0	0	0	0
quality control	1.8499	2017	4	1	0	2	0	1	0	0	0	0	0	0	0	0	0

quality management	0.0487	2016	3	0	1	0	0	2	0	0	0	0	0	0	0	0	0
quality of service	0.4091	2015.6667	3	1	0	0	0	1	1	0	0	0	0	0	0	0	0
quantitative analysis	2.2479	2014.3333	3	0	0	0	1	1	0	0	1	0	0	0	0	0	0
ranking	2.9354	2017.3333	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0
rapport management	0.1402	2018	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
regression analysis	1.4846	2016.75	8	2	2	1	1	1	0	0	1	0	0	0	0	0	0
reputation management	1.134	2016.1304	23	5	1	4	5	3	2	2	0	1	0	0	0	0	0
research	1.8253	2017	8	3	3	0	0	0	1	0	1	0	0	0	0	0	0
response management	0.8018	2017.2	5	2	1	0	0	2	0	0	0	0	0	0	0	0	0
restaurant reviews	0.8035	2016.375	8	1	1	3	1	1	0	0	1	0	0	0	0	0	0
restaurants	1.4255	2015.7273	11	2	2	1	1	1	2	1	0	0	1	0	0	0	0
revenue management	0.5667	2016.6667	3	1	0	0	1	1	0	0	0	0	0	0	0	0	0
review helpfulness	1.7555	2017.2727	11	2	2	4	3	0	0	0	0	0	0	0	0	0	0
review quality	0.9015	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
review valence	0.8817	2017.4	5	2	0	2	0	1	0	0	0	0	0	0	0	0	0
reviewer expertise	1.7509	2017.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
revisit intention	0.6611	2016	5	0	2	1	0	2	0	0	0	0	0	0	0	0	0
revpar	2.0453	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
sales	0.9076	2015.2143	14	2	1	3	2	1	0	2	1	1	1	0	0	0	0
semantics	0.6208	2018.3333	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
sensitivity analysis	0.6536	2018	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
sentiment analysis	1.4459	2017.3095	42	10	12	10	6	2	0	1	1	0	0	0	0	0	0
service failure	2.3602	2016.25	4	0	2	0	2	0	0	0	0	0	0	0	0	0	0
service quality	1.1311	2017.35	20	5	7	4	1	1	1	1	0	0	0	0	0	0	0
service recovery	0.6831	2016.625	8	3	1	1	1	0	0	1	1	0	0	0	0	0	0
service sector	2.257	2015.4286	7	1	0	0	2	3	0	0	1	0	0	0	0	0	0
sharing economy	1.7518	2018.3529	17	10	3	4	0	0	0	0	0	0	0	0	0	0	0
smart tourism destination	2.3001	2016.75	4	0	1	2	0	1	0	0	0	0	0	0	0	0	0
social influence	1.4044	2016.3333	6	2	0	1	1	1	0	0	1	0	0	0	0	0	0
social media	1.355	2016.6633	98	17	28	14	13	12	3	6	5	0	0	0	0	0	0
social media analytics	1.8564	2017.8	10	3	4	2	0	1	0	0	0	0	0	0	0	0	0
social media marketing	1.3424	2017.5	4	2	0	0	2	0	0	0	0	0	0	0	0	0	0
software	1.7676	2016.4	5	0	2	1	0	1	1	0	0	0	0	0	0	0	0
source credibility	1.075	2016.1667	6	3	0	0	1	0	0	1	0	1	0	0	0	0	0
spain	0.6936	2016.7273	11	3	3	1	2	1	0	0	0	0	0	1	0	0	0
spatial analysis	0.7381	2018.2	5	1	4	0	0	0	0	0	0	0	0	0	0	0	0
stress	0.3227	2016	3	0	0	1	1	1	0	0	0	0	0	0	0	0	0

support vector machine	2.6155	2017.5	4	1	0	2	0	0	0	1	0	0	0	0	0	0	0
sustainability	0.5962	2017.8667	15	7	4	2	1	0	0	1	0	0	0	0	0	0	0
sustainable tourism	0.3343	2017	5	2	1	0	1	0	0	1	0	0	0	0	0	0	0
systematic review	1.7029	2017	4	1	0	2	0	1	0	0	0	0	0	0	0	0	0
technology acceptance model	0.4508	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
text analytics	1.1669	2017.6327	49	20	8	12	4	3	1	1	0	0	0	0	0	0	0
text classification	0.757	2015.75	4	1	1	0	0	0	0	2	0	0	0	0	0	0	0
thailand	0	2018.6667	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
topic modeling	0.5223	2018.2857	7	3	4	0	0	0	0	0	0	0	0	0	0	0	0
tourism	1.3336	2017.2241	58	15	22	6	5	3	3	1	3	0	0	0	0	0	0
tourism development	0.8812	2016.875	8	2	2	1	0	2	1	0	0	0	0	0	0	0	0
tourism economics	2.2052	2015.5556	9	1	3	0	1	2	0	1	0	0	0	0	1	0	0
tourism management	1.6121	2016.7647	34	10	8	6	4	1	0	1	1	2	0	0	1	0	0
tourism marketing	1.7477	2016.8125	16	4	5	0	1	4	1	1	0	0	0	0	0	0	0
tourism research	3.6481	2016.6667	3	1	1	0	0	0	0	1	0	0	0	0	0	0	0
tourist attraction	1.0979	2017.2308	13	2	5	3	1	1	1	0	0	0	0	0	0	0	0
tourist behavior	1.8131	2017	22	6	6	4	1	2	1	1	1	0	0	0	0	0	0
tourist destination	1.2429	2017.2093	43	16	12	5	2	3	0	1	1	3	0	0	0	0	0
tourist experience	0.3649	2017.7143	14	5	5	2	0	1	1	0	0	0	0	0	0	0	0
tourist perceptions	0.7512	2018	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
tourist satisfaction	0.1202	2018	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
travel	0.8536	2016.2	10	0	6	1	0	1	0	0	1	0	1	0	0	0	0
travel behavior	2.073	2016.381	21	4	5	2	4	2	0	3	0	1	0	0	0	0	0
travel blogs	0.8564	2016	3	0	0	1	1	1	0	0	0	0	0	0	0	0	0
travel reviews	0.5588	2016.875	8	3	1	0	2	0	2	0	0	0	0	0	0	0	0
travel websites	1.2115	2018	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0
tripadvisor	0.8669	2017.1772	79	25	22	11	7	5	3	2	2	1	0	1	0	0	0
trust	2.1957	2016.2727	11	4	2	0	1	1	0	1	1	1	0	0	0	0	0
trustworthiness	0.5511	2018	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
turkey	0.365	2017.25	4	1	1	1	0	1	0	0	0	0	0	0	0	0	0
twitter	1.04	2017.2	5	1	2	0	1	1	0	0	0	0	0	0	0	0	0
united kingdom	2.1768	2014.8333	6	1	1	0	0	1	0	1	2	0	0	0	0	0	0
united states	2.0448	2017.1429	14	4	4	3	0	1	1	0	1	0	0	0	0	0	0
urban area	2.2351	2017.3333	3	1	1	0	0	1	0	0	0	0	0	0	0	0	0
urban tourism	0.4936	2018	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0
usefulness	1.0344	2015.6667	3	0	0	1	2	0	0	0	0	0	0	0	0	0	0
user-generated content	0.8569	2017.0923	65	17	15	14	8	7	1	0	1	1	1	0	0	0	0

valence	1.1979	2016.6	5	2	0	1	1	0	0	0	1	0	0	0	0	0	0
value co-creation	1.2036	2016.25	4	0	1	2	0	0	0	1	0	0	0	0	0	0	0
value co-destruction	1.2118	2017.6667	3	0	3	0	0	0	0	0	0	0	0	0	0	0	0
value creation	2.0418	2017.4	5	0	3	2	0	0	0	0	0	0	0	0	0	0	0
visibility	1.4273	2016	3	0	0	0	3	0	0	0	0	0	0	0	0	0	0
web 2.0	1.8064	2013.7778	9	0	1	0	2	1	0	1	2	2	0	0	0	0	0
wildlife management	0.7156	2017	3	1	1	0	0	0	1	0	0	0	0	0	0	0	0
wildlife tourism	0.4185	2017.5	4	2	1	0	0	0	1	0	0	0	0	0	0	0	0
willingness to pay	1.4138	2018	4	2	2	0	0	0	0	0	0	0	0	0	0	0	0
wine tourism	0.9527	2018.6	5	3	2	0	0	0	0	0	0	0	0	0	0	0	0
word-of-mouth	1.0828	2015.8696	23	5	6	1	1	3	0	4	2	0	0	0	1	0	0

CHAPTER 3. COVID-19 AND CRUISE TOURISM PERCEPTION

Preface

This chapter of the thesis examines incidents early in the COVID-19 pandemic, when the *Diamond Princess* cruise ship became the center of the largest outbreak outside the original epicenter in China. This outbreak which left 712 passengers infected and 14 dead, followed by subsequent outbreaks affecting over one-third of the active ships in the cruise industry's global fleet, quickly became a crisis that captured public attention and dominated mainstream news and social media. This study investigates the perception of cruising during these outbreaks by analyzing the tweets on cruising using Natural Language Processing (NLP). The findings show a prevalent negative sentiment in most of the analyzed tweets, while the criticisms directed at the cruise industry were based on perceptions and stereotypes of the industry before the pandemic. This study provides insight into the concerns raised in these conversations and highlights the need for new business models outside the pre-pandemic mass-market model and to genuinely make cruising more environmentally friendly. The article based on this chapter is titled: "#CoronavirusCruise: Impact and implications of the COVID-19 outbreaks on the perception of cruise tourism" was published in 2022 in *Tourism Management Perspectives*, volume 41, article number 100948. The journal is indexed in Scopus and the Social Sciences Citation Index (SSCI) of the Journal Citation Reports (JCR). The paper was presented at the International Meeting on Business, Institutions, and the New Normal conference, which held in Madrid and online on March 1st-2nd, 2021. It was also presented at the ECO-SOS Doctorials held at the Faculty of Economics and Business of the Universitat Rovira i Virgili on 15th of December, 2021.

Introduction

The COVID-19 pandemic caused by the novel coronavirus SARS-CoV-2 has exceeded 5 million confirmed deaths at the time of writing (AP, 2021), causing widespread disruption in many sectors of the world economy. Tourism was one of the worst-hit sectors due to the closure of national borders and the reduction in global travel (Gössling et al., 2020). Cruise tourism in particular was severely affected because of COVID-19 outbreaks on several cruises early in the pandemic (Ito et al., 2020), which resulted in stranded ships due to port closures and the temporary ban of cruises in some countries (Gössling et al., 2020). Some returning passengers from cruises contributed to the spread of the virus in their home countries (Ito et al., 2020). For example, as at the end of April 2020, one in ten COVID-19 cases in Australia were attributed to a *Ruby Princess* cruise voyage (ABC, 2020), while between February and mid-March 2020, about 17% of the confirmed cases in the United States (US) were linked to returning cruisers (Moriarty et al., 2020).

Official and unofficial counts show that there were at least 3,908 confirmed COVID-19 cases and 111 confirmed deaths linked to over 102 COVID-19 outbreaks involving at least 124 cruise ships as at October 2020 (CDC, 2020; Miami Herald, 2020). Figure 3.1 shows that many cruise companies had outbreaks on their ships, including Carnival Corporation, Royal Caribbean, and Norwegian Cruise Line, the three biggest companies in the industry, which account for 80 percent of the industry's passenger capacity (Papathanassis, 2017). These cruise companies also suffered economic setbacks because of the pandemic, with their share prices falling off a cliff in the first

quarter of 2020 (see Figure 3.2). Cruises continue to record transmission of COVID-19 despite high vaccination rates among passengers and crew, with 1,359 reported confirmed cases in the US between June and October 2021 (CDC, 2021).

Figure 3.1: Number of COVID-19 cases by Cruise Company (Data source: Miami Herald last updated on 2020-10-02 (Miami Herald, 2020))

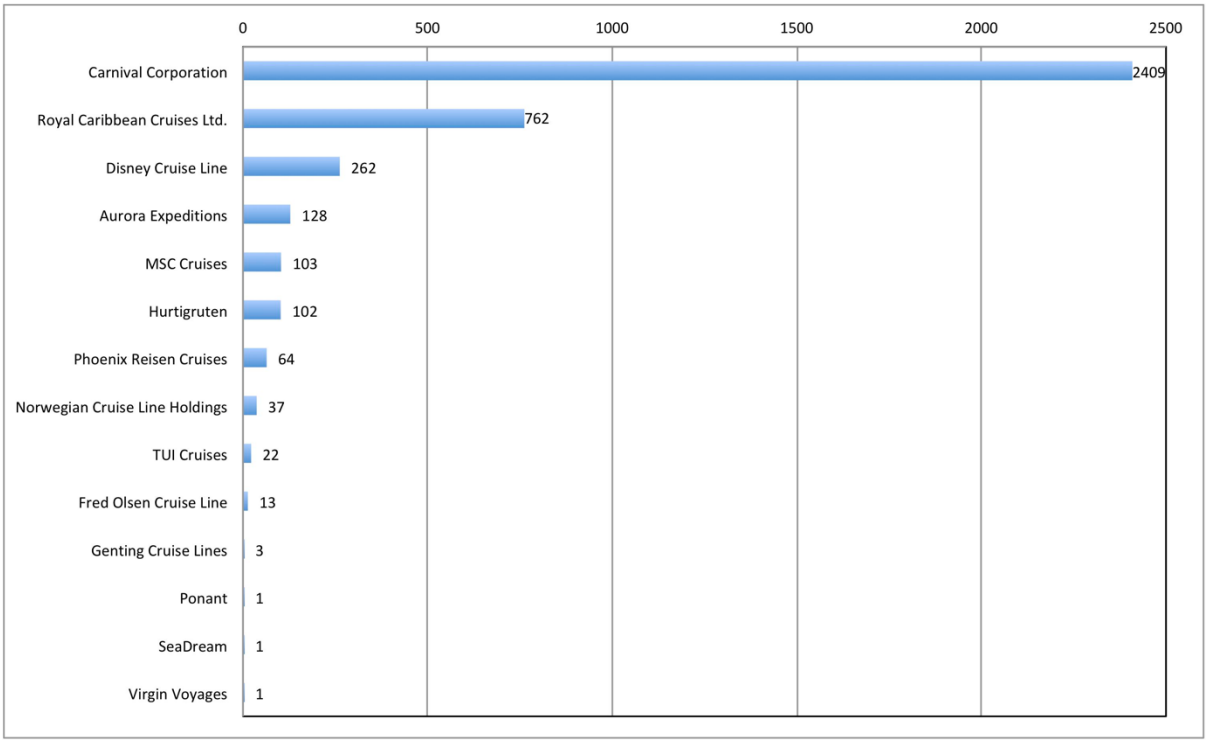
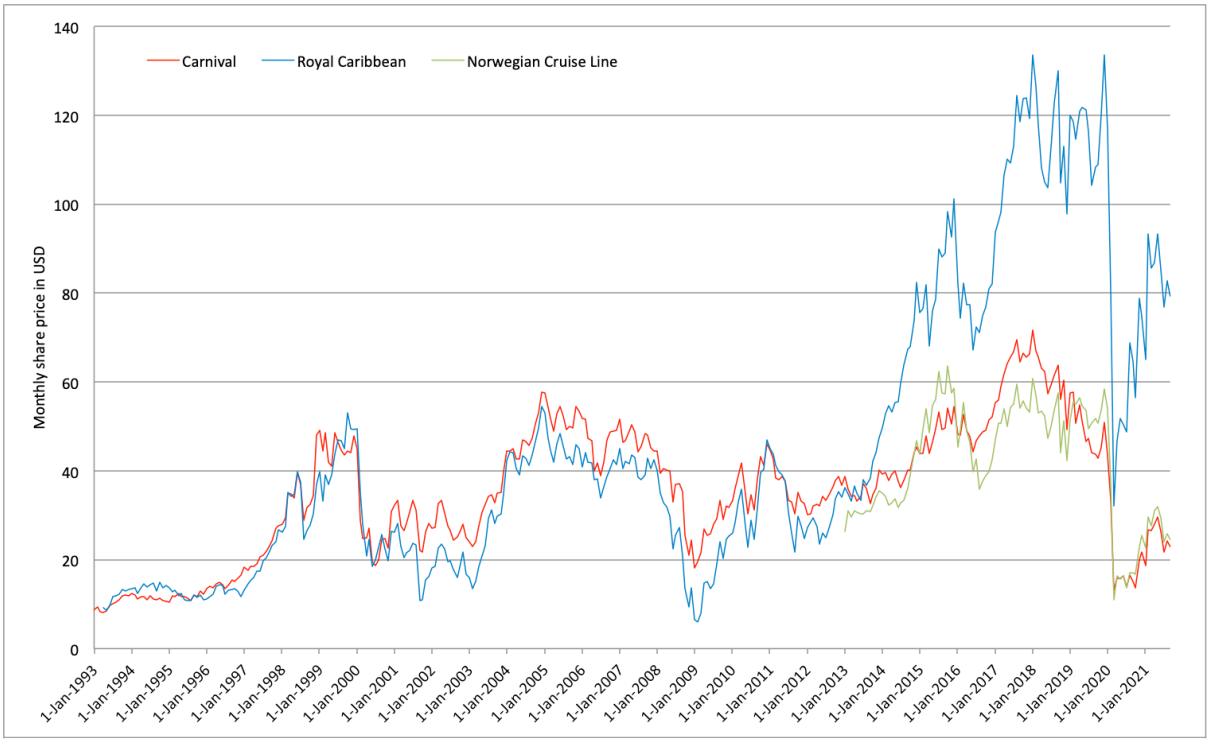


Figure 3.2: Share prices of the three largest cruise operators from January 1993 to September 2021 (Data source: Eikon, downloaded on 2021-09-07)



Cruising was the most mentioned sub-sector of tourism in global news in the early period of the pandemic as the media reported these COVID-19 outbreaks (Gössling et al., 2020) and some cruise industry executives have claimed that the industry was unfairly tarnished by the media in the reporting of the outbreaks (Financial Times, 2020; Washington Post, 2020). It is important to have an appraisal of the impact that the extensive coverage of the cruise COVID-19 outbreaks in the news and social media have had on people and how it has affected their perception of cruising. Such research is scarce given the novelty of the topic, but will help to evaluate the

purchase intention to cruise and the recovery of the cruise industry post-pandemic (Holland et al., 2021; Pan et al., 2021).

Related studies already published used questionnaire survey polls to gauge the outlook towards cruising and the willingness to cruise in the aftermath of the COVID-19 pandemic. For example, Holland et al. (2021) investigated the impact that the COVID-19 pandemic had on the perceived risk of cruising considering a list of 20 items, and found that the country of residence had a significant impact on risk perception of cruising and future intentions to cruise. Pan et al. (2021) tried to identify consumer perception of the cruise industry during the COVID-19 pandemic under the theoretical lens of leisure constraints and prospect theories. They found that travel constraints negatively influenced behavioral intention to travel with cruises, although perceived crisis management positively affects this behavioral intention. The survey instrument used in both studies is universally accepted for evaluating public opinion but has been criticized for treating public opinion as a quantitative distribution of opinions by disparate individuals having equal weight in society with a blind spot to the hierarchical and conversational nature of public opinion formation (Blumer, 1948). Social media conversations make up for these weak points, since they are hierarchical and conversational by nature (McGregor, 2019). Furthermore, surveys are appropriate to capture subjects' attitudes by directly asking respondent about their subjective experiences, perceptions, and attitudes about a topic, but have limited ability to observe actual behavior. Social media provides a wealth of information about actual user behavior, and not as limited or focused on specific questions (Couper, 2013). Research has also found a correlation between public opinion and sentiments expressed on Twitter (O'Connor et

al., 2010). Consequently, analysis of social media conversations can provide insight that can be used with survey results to provide a holistic view of the public opinion on an issue.

This study fills this research gap in developing a better understanding of the public perception towards cruising during the COVID-19 outbreaks on cruises in the early part of the COVID-19 pandemic, extending the research on this topic beyond the perceived risk of cruising or the intention to cruise. To do this, we analyze the relevant tweets using Natural Language Processing (NLP) methods. Based on this premise, this study intends to answer the research question: What insight can be derived from the public conversation on Twitter about cruising during the COVID-19 outbreaks on cruises?

The contribution of this study is to serve either as an empirical counterpart to the findings of the previous studies with survey data, as well as to uniquely show the hierarchical dimension of the public conversation with the influence of elites in driving awareness to issues in a way that surveys may not show. It is also hoped that the discussion of the implications of the findings will be a valuable addition to the academic literature on the prospects and outlook of the cruise industry after the COVID-19 pandemic.

Literature Review

The COVID-19 cruise outbreaks were a crisis event that we investigated using social media data. Hence, it is important to examine the literature on the use of social media in crisis communication in tourism and cruise tourism. As mentioned, there was extensive media coverage and social media mentions of cruising in the early period of the pandemic (Gössling et al., 2020). This wall-

to-wall media coverage usually affects public opinion, risk perception, and consumer behavior in a particular pattern. In the following sections, the theoretical foundation of this study is presented, which is based on the information integration theory (IIT) and the social amplification of risk framework (SARF) to explain risk perception and decision-making of tourists influenced by social media in times of crisis.

Role of social media during crises in tourism

One of the defining characteristics of a crisis is the mass generation of mostly negative comments and information (Coombs, 2018). Social media provides a platform for communicating decisions during a crisis and collecting feedback from the public (Sigala, 2011). The information integration theory (IIT) states that new information is added into the preliminary beliefs of people, affecting how attitudes and behavior are formed, and the value of information determined to be favorable or unfavorable regarding the events and subjects involved (Holland et al., 2021). This theory provides one of the theoretical frameworks for this study.

The amount of mass information generated by social media creates awareness and a forum for discussion during times of crisis (Sigala, 2011). This is because social media facilitates real-time interactivity, reciprocity, and instant reactions between users, hence it plays a central role when crisis events happen and in crisis management in tourism with potential effects on tourists' behavior (Schroeder et al., 2013; Sigala, 2011; Zeng & Gerritsen, 2014). International tourists have been found to have a high probability of turning to social media for information during periods of crisis (Schroeder et al., 2013). The electronic word-of-mouth (eWOM) shared on social

media by members of the public during a crisis is a curated heritage or collective memory that can be explored to learn about the crisis (Liu, 2009). This fact supports the suitability of using social media to explore crisis effects. Hence, tourism scholars have used social media data to investigate health-related crisis (Yu et al., 2020), political crisis (Luo & Zhai, 2017), natural disasters (Möller et al., 2018), climate crisis (Schweinsberg et al., 2020), service failures (Su et al., 2019), pest infestation (Liu et al., 2015), and terror attacks (Barbe et al., 2018).

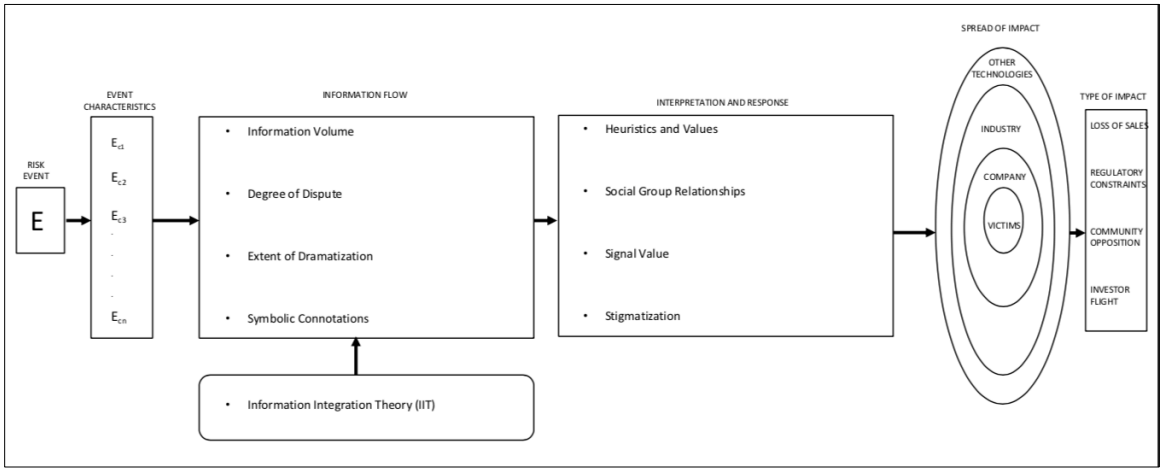
In cruise research, past literature on the role of social media during crises has focused on its impact on the corporate reputation of cruise companies. Ryschka et al. (2016) found that a prompt corporate response on social media and strong brand familiarity during a crisis led to positive perceptions of the cruise company. Penco et al. (2019) found that the "degree of anger" and prior corporate reputation during a crisis influenced the intention to take a future cruise. While Penco et al. (2018) found that social media may help cruise companies to communicate more effectively during a crisis.

Media coverage, public opinion, and risk perception

The news media influences public opinions through its news coverage (Gene Zucker, 1978). In the absence of direct personal experience, members of the public learn about risks from other people and from the media (Kasperson et al., 1988). Therefore, as the media informs the public about happenings beyond their immediate circle, it shapes the public's perception and understanding of risks (Rowe et al., 2000; Smith, 2005).

The social amplification of risk framework (SARF) is a seminal integrative and interdisciplinary risk perception framework that has been used to account for findings from a wide range of fields, from media to medical research (Kasperson et al., 2003; Y. Wang et al., 2021; Womack et al., 2020). We combined the SARF with the IIT to explore how risk perceptions during COVID-19 influenced potential tourist attitudes and behavior about cruising, as illustrated in Figure 3, which is an adaptation of the simplified representation of the SARF framework.

Figure 3.3: Proposed conceptual map of SARF and IIT adapted from Kasperson et al. (1988)



The SARF framework suggests that the social amplification of risk by the media increased its memorability and imaginability, which leads to increased risk perception (Kasperson et al., 1988). Four attributes of the information flow from the media influence the extent of social amplification. They include volume of information flow, disputability of the information, degree of dramatization, and symbolic connotations (Kasperson et al., 1988). It is worthy of note that high volumes of information flow about a risk attract public attention, mobilize latent fears about the risk, and trigger the recollection of previous failures or accidents (Kasperson et al., 1988;

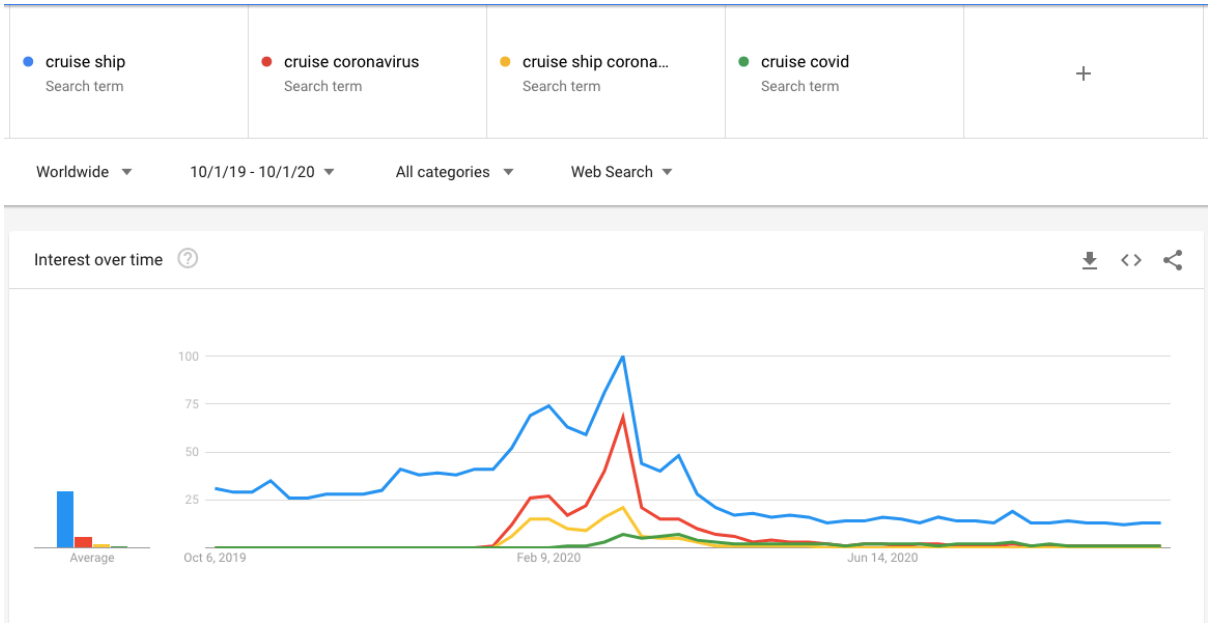
Renn, 1986). Degree of dispute deals with how much the facts about the event are disputed. Dramatization, usually in the form of sensational headlines, is an important attribute of information that increases perceived risk and memorability of the incident. While symbolic connotations deal with the specific terms used to convey information and how they may have other meanings and be interpreted by individuals and groups (Kasperson et al., 1988). Likewise, there are four attributes of the interpretation and response mechanism in the second stage of social amplification: heuristics and values which deals with the simplification of risk, social group relationships which deals with politicization and/or polarization, signal value on the seriousness of the risk, and stigmatization (Kasperson et al., 1988).

Method

Data collection

We analyzed the tweets from the start of the pandemic through the period of the early outbreaks on cruises, highlighted by an increased publicity on cruising, which was pinpointed using Google search trends. Google Trends has been demonstrated to show the popularity of a search query as an indicator of real public interest (D'Avanzo et al., 2017). Figure 3.4 shows that the worldwide search queries for selected relevant phrases like "cruise coronavirus" and "cruise ship coronavirus" rose from zero between January 19-25 and returned to zero around May 3-9, a period coinciding with the major outbreaks (Google Trends, 2020). "Cruise covid" rose from zero after February 11, the day the International Committee on Taxonomy of Viruses (ICTV) officially named the new virus SARS-CoV-2 and the disease it causes COVID-19 (WHO, 2020).

Figure 3.4: Google search trends for cruise ship, cruise coronavirus, cruise ship coronavirus, and cruise covid (Data source: Google Trends)



Therefore, we used the ninth version of the open-source COVID-19 dataset of tweet IDs by Banda et al. (2020). The data covers the period from 4th January to 10th May 2020 and includes our period of interest, to capture the tweets posted about the COVID-19 outbreaks on cruises at the time of the major outbreaks at the start of the pandemic when the outbreaks dominated news and social media. It was the time when the attention on the cruise COVID-19 outbreaks were highest, as indicated by excessive mentions of cruise and COVID-19, and pinpointed with Google Trends. This data was gathered by collecting all tweets on the novel coronavirus from the Twitter application programming interface (API) stream by filtering the API stream with relevant keywords connected to the pandemic as stated in the dataset publication (Banda et al., 2021). These keywords include, among others, “coronavirus”, “2019ncov”, “corona virus”, “COVD19”,

"CoronavirusPandemic", "COVID-19", "2019nCoV", "CoronaOutbreak", "WuhanVirus". The authors of the dataset provide a full and a clean version of the data; with the clean version comprising only original tweets without retweets. We used this clean version, which contained 66,538,356 tweet IDs because retweets contain the text of the original tweet and may disrupt our intended natural language processing tasks (Banda et al., 2021).

We rehydrated these tweet IDs to retrieve the full tweets that had not been deleted for analyses using TWARC version 1.8.3 (Summers et al., 2021) because deleted tweets are unretrievable. The rehydration process lasted about 186 hours. We successfully retrieved 60,483,491 full tweets (9.1% deleted), and the resulting JSONL (JavaScript Object Notation Lines) file was 263 gigabytes. The retrieved tweets were in over 64 languages, with the majority (57%) in English (34,431,646). Due to difficulties in translating such a huge number of languages, we restricted our analyses to only the tweets in English.

The required information (posting time, name of Twitter handle, text of tweet, number of retweets, and number of likes) from each tweet were parsed from the tweet json files and loaded into a Python Jupyter Notebook (Kluyver et al., 2016; Van Rossum & Drake, 2009) for analysis using Pandas DataFrames (McKinney, 2010). Preprocessing tasks like tokenization (breaking of words to smaller units), lemmatization (reduction of words to their root), and removal of unnecessary elements like stop words, URL links, and mentions (@user) were carried out. We also split and removed the hashtag signs (#) and the newline characters (\n) from the tweet text before analysis. The data preprocessing is a necessary step in order to clean the raw, noisy data

before advanced analytical processes. We searched for the occurrences of cruise(s), cruise line(s), cruise ship(s), and cruising and got 139,054 tweets. We randomly checked the obtained tweets manually to ensure that there were no false positives or irrelevant tweets in the data, as done in previous studies (Ainin et al., 2020) by exporting the tweets as a csv file and checking them in Microsoft Excel. This manual inspection revealed that there were 586 tweets about actor Tom Cruise halting a movie shoot because of the pandemic; which were then removed.

Data analysis

This study used sentiment analysis as the main analytical method to automatically classify the tweets into positive, neutral, and negative. Sentiment analysis is the automated process of opinion detection by using semantic relationships to determine the overall polarity of a text document as positive, neutral, or negative (Alaei et al., 2019; Feldman, 2013). The rapid growth of sentiment analysis has coincided with the growth of social media, which has provided enormous volumes of digital opinionated data (B. Liu, 2012). Therefore, social media sites like Twitter and Facebook are a centre of interest for sentiment analysis applications (Feldman, 2013). Social media offers tourists a platform to share their views, feelings, and sentiments about their experiences (Li et al., 2018). Sentiment analysis is one of the most common techniques for analyzing online text data on social media (Khong et al., 2021; Muritala et al., 2020) and at a relatively low cost (Neri et al., 2010). Sentiment analysis and opinion mining are used interchangeably (Feldman, 2013), although the use of the term “opinion mining” has been declining in the tourism and hospitality literature (Muritala et al., 2020).

However, because of its unstructured form and noisiness, social media data presents some challenges for automated sentiment analysis. Sentiment analysis algorithms are usually not equally effective across a variety of languages, hence, it is usual for researchers to restrict analysis to a single language, leading to loss of information (Li et al., 2018). The noisiness of social media with liberal use of slangs and lots of spelling, grammatical, and punctuation errors makes accurate analysis more difficult (Feldman, 2013). Sentiment analysis is also domain-dependent, with positive words in one domain, not positive in another (Pang & Lee, 2008). Identification of sarcasm also poses a challenge to sentiment analysis systems (Feldman, 2013) with Twitter exhibiting rich sarcasm (M. Zhang et al., 2016). Despite these challenges, sentiment analysis remains very useful for researchers and industry practitioners for analyzing large amounts of data compared to manual processing (Philander & Zhong, 2016).

Kietzmann et al.'s (2011) seven functional blocks of social media provide the conceptual framework for understanding the sentiment analysis results. Twitter is a conversation-based platform (Kaplan & Haenlein, 2011), hence, the relevant block for this study is the conversation block of the framework. The conversation block posits that to make sense of the conversation; the conversation velocity, i.e. the frequency of new conversations over time, and the change in how favorable or unfavorable the sentiment they contain needs to be analyzed in order to understand the conversation (Kietzmann et al., 2011). To make sense of the conversation on cruising during the COVID-19 outbreaks on cruise ships, the frequency of new conversations is the frequency of tweets on the topic over the analyzed period and the sentiment analysis score measures the change in sentiment of the tweets. This informed the decision to make a plot of

the daily frequency of cruise-related tweets and their average daily sentiment in Figure 6, to show how the frequency and sentiment of the conversation toward cruising changed during the analyzed period.

Apart from sentiment analysis, we construct a word cloud based on the 500 most frequent words and bigrams (two-word combination), and also present the three tweets with the highest amount of engagement. Tweets have unequal reach, and the engagement metrics on a tweet indicate the amount of people that saw and interacted with the tweet. We selected the most engaged tweets by checking the tweets with the highest number of likes, because viral tweets usually have more likes than retweets.

The word cloud of the most frequent words was constructed using WordCloud version 1.5.0 package in Python (Mueller, 2020). A simple ranking of the tweets based on their number of likes produced the most engaged tweets. Finally, the sentiment analysis was performed using VADER (Valence Aware Dictionary and sEntiment Reasoner) (Hutto & Gilbert, 2014) implementation in Python's NLTK (Natural Language Toolkit) (Bird et al., 2009). VADER is well suited for analyzing social media text since it handles many of its typical elements like acronyms, emojis, and slangs well. VADER assigns a score to the analyzed text based on the summation of valence scores of the words in a tweet between 1 (extreme positive sentiment) and -1 (extreme negative sentiment), while neutral sentiment spans 0.05 to -0.05 (Hutto & Gilbert, 2014). It is also important to do a manual validation of the sentiment classification results to confirm that the assigned sentiment scores reflect the content of a tweet. This validation effort showed that there

were some negative tweets given positive scores. These tweets were about reports of people who tested positive for the coronavirus, and "positive" as a positive word, shifted the sentiment weighting of these tweets to positive, even though this carries a negative sentiment in this context. We adjusted this by replacing the word "positive" with "infected" in these tweets, which made the algorithm to reflect the sentiment more accurately.

Results

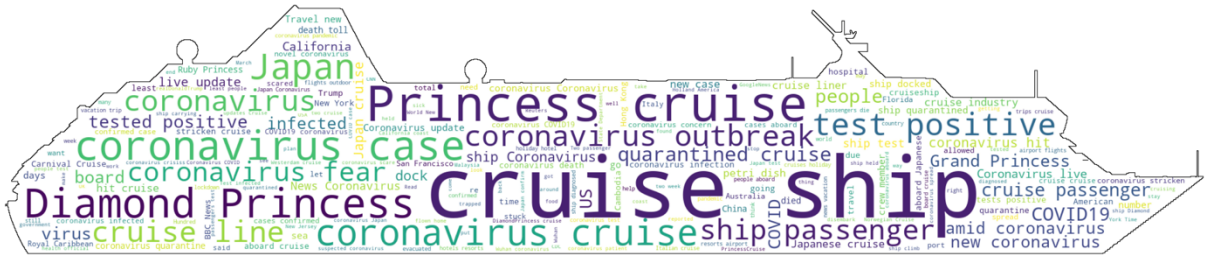
138,468 cruise-related tweets posted by 58,644 Twitter accounts were obtained. The mean and median number of tweets per account was 2.37 and 1 respectively, showing a likely preponderance of tweets from personal accounts, compared to news organizations or blogs that usually have multiple posts reporting the same information.

Word cloud

The word cloud of the 500 most frequent words and bigrams in the analyzed tweets is presented in Figure 3.5. Figure 3.5 shows the names of some of the cruise ships (*Diamond Princess, Grand Princess, Ruby Princess, Westerdam*), company/brand names (Carnival, Royal Caribbean, Norwegian Cruise (Line), Holland America), and places (Japan, US, China, Australia, Hong Kong, Italy, Cambodia, UK, Malaysia) that featured prominently in the Twitter conversation during the examined period. There are also many names related to the US like former president Donald Trump, the CDC, as well as American cities that were associated with some cruise COVID-19 outbreaks (California, San Francisco, Florida, New Jersey, New York). Petri dish was one of the most frequent bigrams visible on the word cloud because several tweets referred to cruise ships

as “petri dishes” or “floating petri dishes”, a derogatory cliché about cruise ships being a fertile platform for infectious diseases. Similar words or bigrams that reappear on the word cloud i.e. “test positive”, “tested positive”, and “tests positive” or different words that refer to the same entity like US, USA, and America are because the word cloud algorithm counts any change in spelling separately, and this shows the noisiness of social media data. We did not combine these similar words to avoid introducing any bias.

Figure 3.5: Word cloud of the 500 most frequent words and bigrams



Most-engaged tweets

The three tweets with the highest amount of engagement (retweets, quotes, and likes) were uncomplimentary remarks about the cruise industry by influential public personalities in the US with verified Twitter accounts and many followers (see Figure 3.6). The first was by Judd Apatow, an American film director with over 2.4 million followers, whose tweet had over 9,300 retweets and quote tweets and over 75,800 likes. The second was by Sheldon Whitehouse, an American senator with over 468,000 followers. His tweet, which also shared a *Washington Post* article, had over 10,400 retweets and quote tweets and over 22,600 likes. While the third was by Bill Maher, an American talk show host with over 11 million followers, who shared a video of an episode of

his talk show about the cruise industry. The tweet had more than a thousand retweets and quote tweets and over 4,600 likes, while the attached video was viewed over 234,500 times.

Figure 3.6: Top-three most engaged tweets (Source: Twitter)

Coincidentally, these three tweets were posted within a week from each other during a period when there were rumors that cruise companies would be included in the COVID-19 government bailout package following former President Trump's tweet on 12th March 2020, in which he described the cruise industry as a "great and important industry" that will be kept that way. These rumors turned out to be untrue but as these tweets show, there was opposition to the idea on the basis that cruise companies incorporate outside the US and sail under flags of convenience (FOC) to avoid paying taxes in the US and pollute the environment.

Sentiment analysis

The sentiment analysis result is presented on a dual-axis time series plot of the daily frequency of cruise-related tweets and the average daily sentiment of these tweets in Figure 3.7. The

highest volume of tweets per day were posted during the outbreak on the *Diamond Princess* with the two highest peaks in February 2020. The negative sentiment between January 26 and February 3 was due to the cancellation of cruises with Chinese port of calls and fears about the spreading novel coronavirus. However, from February 3rd with the outbreak on the *Diamond Princess* and subsequent cruise outbreaks, the negative sentiment was mainly because of the negative news stories on new confirmed cases, hospitalizations, passenger deaths, stranded ships, and people's reactions. The negative sentiment about any specific incident was usually persistent long after the news initially broke on Twitter because of repetition and round-the-clock coverage of the 24-hour news cycle, resulting in the predominantly negative sentiment throughout the analyzed period. The only time the sentiment crossed the neutral line was on the release of a CDC report on March 24 showing that the novel coronavirus survived in the *Diamond Princess* cabins for 17 days after passengers disembarked. The tweets about this report had a positive sentiment scores because the verb "survive" is a positive word.

Figure 3.7: Daily frequency and average daily sentiment of cruise-related tweets

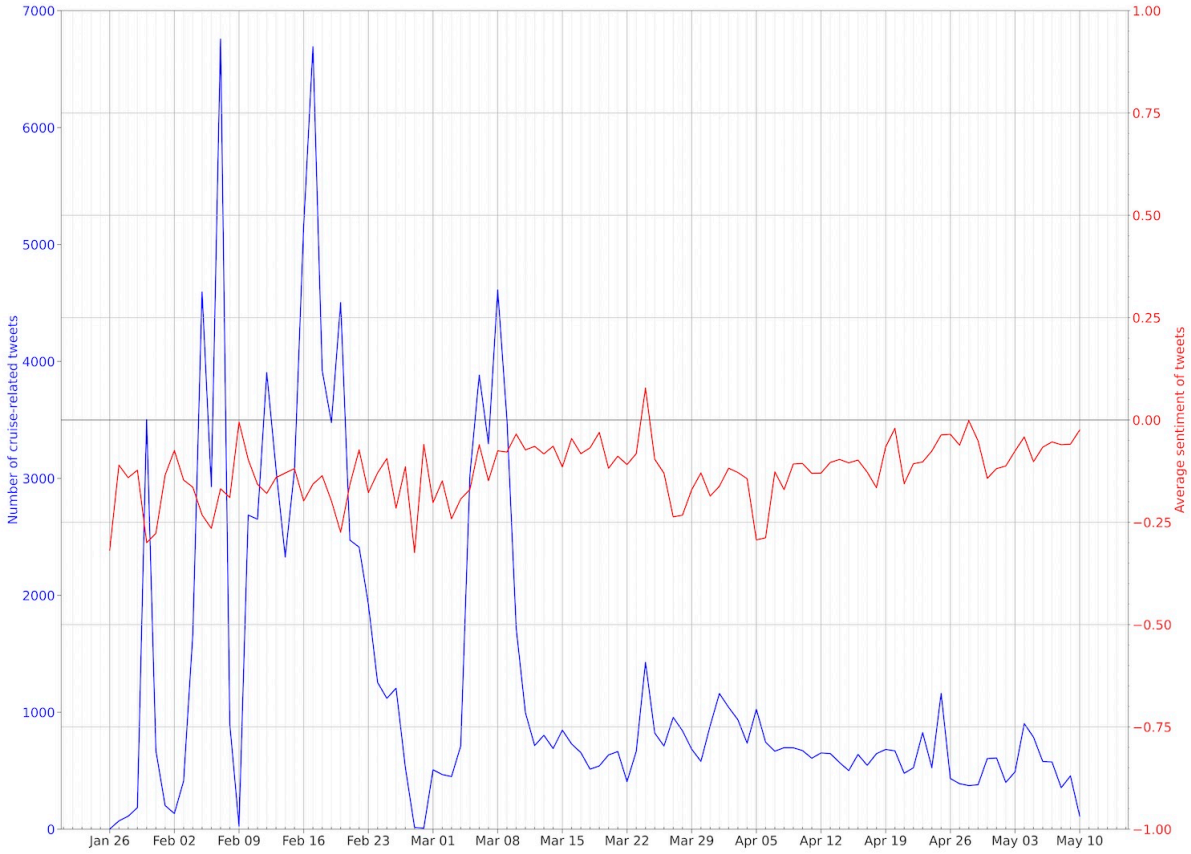
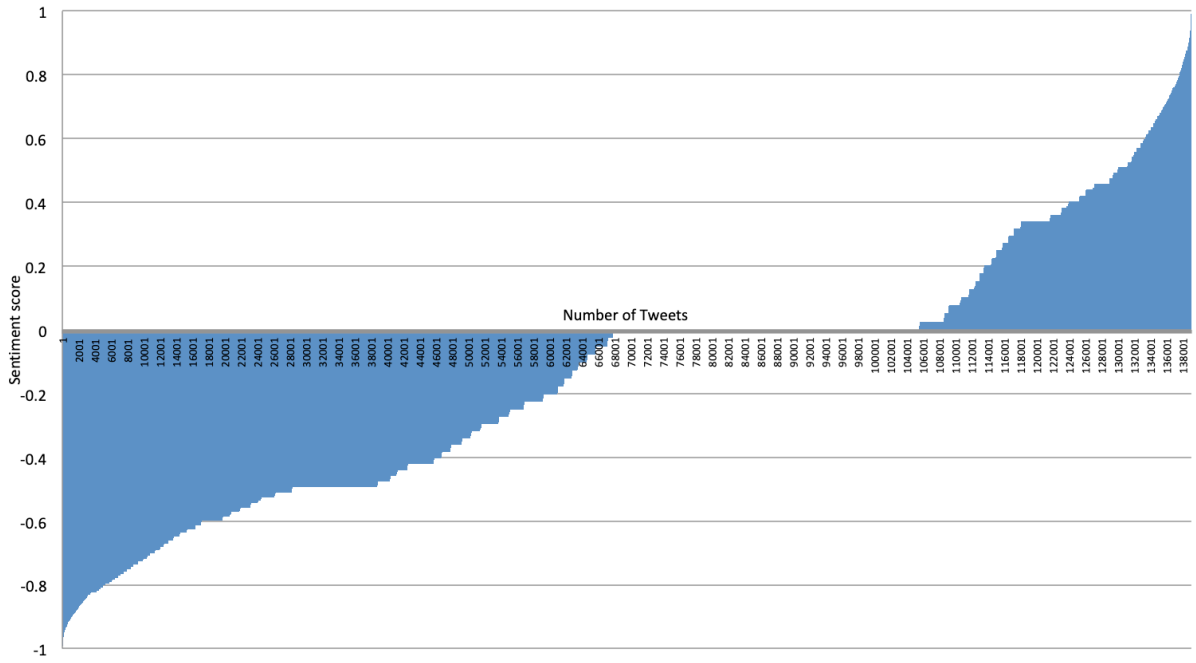


Figure 3.8 shows the proportion of positive, neutral, and negative tweets. 67,022 (48%) of the tweets in our dataset had a negative sentiment score, 41,636 (30%) had a neutral sentiment score, and 30,352 (22%) of the tweets had a positive sentiment score. A random sample of the positive, neutral, and negative tweets is presented in Tables 1 to 3, with the time they were posted and their sentiment scores. We do not include account details to preserve the privacy of the individuals that posted them. The tweets were randomly selected using random numbers generated by the RandArray function in Microsoft Excel.

Figure 3.8: Chart showing the proportion of positive, neutral, and negative tweets

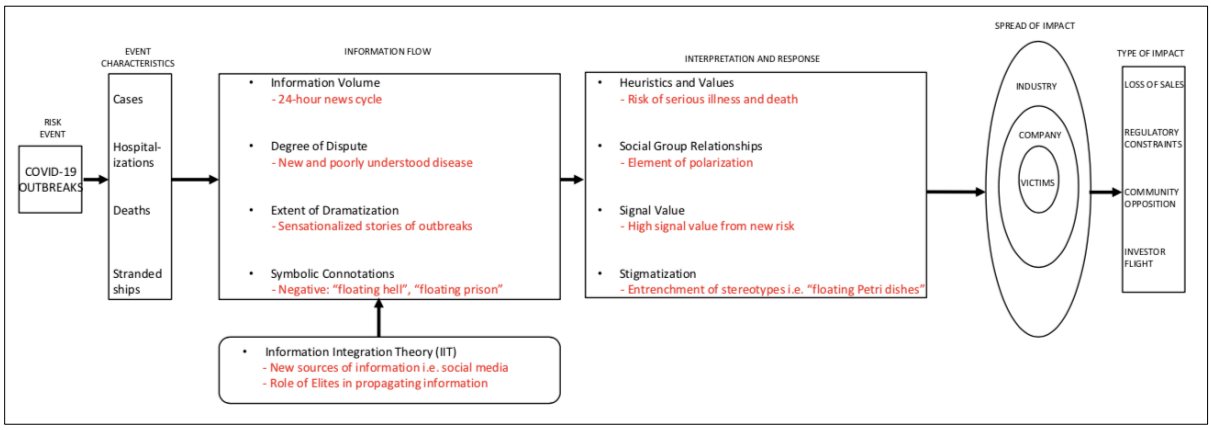


A random sample of positive, neutral, and negative tweets is provided in the Appendix. The positive tweets in Table 1 show that the algorithm had difficulty identifying sarcasm and had some positive-bias, i.e. tweets scored more positively than they were. For example, tweet numbers 3, 4, 5, 9, 17, and 20 are sarcastic negative tweets scored as positive. While tweet numbers 2, 7, 8, 10, 12, 13, 15, and 19 are not actually positive towards cruising. Genuinely positive tweet number 1 and 18 were about an *MS Westerdam* cruise that stayed virus free, which departed Hong Kong on February 1, 2020 and was prevented from docking in five countries before eventually docking successfully in Cambodia. The neutral and negative tweets in Tables 2 and 3 seem to have been scored more accurately.

Revised conceptual framework

A revised conceptual framework of the combination of the SARF and IIT is provided in Figure 3.9 based on the results of our analysis. The biggest change is that the ripple effects from the spread of impact do not have a "other technologies" dimension like in the original framework, with the effects stopping at the industry level. The entire industry suffered economic reversal although some companies had more outbreaks on their ships than others as shown in Figure 1. Aspects of the information flow, interpretation and response, and the information integration theory (IIT) that are noteworthy or new are highlighted in red. For example, the 24-hour news cycle under the information volume attribute and social media sources under the IIT are relatively recent phenomenon that have now been incorporated into the conceptual model.

Figure 3.9: Revised conceptual map of SARF and IIT for impact of COVID-19 outbreaks on the Cruise industry



Discussion and conclusions

This study investigated the public perception towards cruising during the COVID-19 outbreaks on cruises in the early part of the COVID-19 pandemic. To answer the research question on the insight that can be derived from the analyzed tweets during the cruise ship COVID-19 outbreaks, the results show that there was an overwhelming negative sentiment in a majority of the tweets. While this finding was expected, it is still an important result with implications for the cruise industry. This study is in line with the results of previous research that has shown the influence of media coverage on public opinion and risk perceptions (Gene Zucker, 1978; Rowe et al., 2000; Smith, 2005), supporting also the assumptions of SARF that suggests that the spreading of information that involves risk amplifies its perception. The media amplification of the cruise COVID-19 outbreaks had the four attributes of the information flow, which increase memorability and risk perception (Kasperson et al., 1988). The first attribute was are the massive volume of information flow on the incidents from the 24-hour news cycle. The second was the degree of dispute element from a fast-spreading contagious disease that was poorly understood at the time, spreading quickly in the vulnerable environment of a cruise ship. The sensational headlines about the increasing number of cases and deaths from the outbreaks provided the element of dramatization. The symbolic connotations were usually negative in these stories, with intrinsically negative words, such as hell or prison, used to refer to ships with outbreaks or under quarantine. Under the attributes of the interpretation and response phase, the simplification of the risk under the heuristics and values boiled down to the risk of disease and death. There was an element of politicization of social group relationships with the three most engaged tweets as

they were reactions to the rumor that cruise companies would be offered government bailout which started after the former president's tweet about the cruise industry on March 12. There was unmistakable high signal value from an unknown risk of a deadly disease and the stigmatization of the cruise industry was obvious in the recurrence of words like Petri dish in referring to the industry's ships. SARF also emphasizes the difficulty in changing risk perceptions once formed, causing perceptions that can endure far into the future (Kasperson et al., 1988). Hence, these factors underline the importance of this finding, as there is a possibility of an enduring perception of cruising as an undesirable or risky activity even after the threat of COVID-19 goes away. This finding is also consistent with the previous study, which found that cruiser and non-cruiser respondents had anxieties about cruising and were negative about taking a cruise in the future (Holland et al., 2021).

The results also show that criticisms directed towards the cruise industry usually referenced people's previous objections to the cruise industry on various issues such as infectious disease outbreaks before the pandemic, environmental impact and sustainability, sailing under flags of convenience (FOC) to avoid tax and laws, size of cruise ships, crime incidents, and crew working conditions. This finding is consistent with the assumptions of IIT, which highlights that new information is added and integrated into existing beliefs and knowledge, influencing people's attitudes and behavior. It also agrees with the literature on the impact of high volumes of information flow, which mobilizes latent fears and a recollection of previous issues, accidents, or failures (Kasperson et al., 1988). Thus, the high volumes of information through repeated stories and round-the-clock coverage of the COVID-19 outbreaks on cruise ships triggered the

recollection of previous concerns about cruising. Cruise researchers have discussed many of these issues at length before the pandemic (Klein, 2002, 2009, 2011, 2016b, 2016a; Papathanassis, 2016; Terry, 2017).

Our results contribute to previous studies showing the hierarchical dimension of the public conversation with the influence of elites, which is also exhibited on social media. The three most-engaged tweets by public personalities confirm this hierarchical nature, where elites shape public opinion as reported in the literature (Blumer, 1948). Tweets by public personalities can earn an outsized amount of engagement because of their societal prominence and a substantial number of followers. Furthermore, the fact that these tweets were by American personalities, in addition to several American-related entities appearing in the word cloud, shows that the analyzed tweets were very American-centric. This could be explained as evidence of strong American interest in cruising matters, although it could have also been affected by the fact that this study only analyzed English tweets. Besides, industry data shows the United States was the largest source of cruise passengers by far with 11.9 million passengers or approximately 40% of global cruise passengers in 2019 (CLIA, 2019, 2021). Overall, this shows the central importance of the American market to the cruise industry.

It is noteworthy that some of the criticized aspects of cruising reported in the results are a consequence of the mass-market cruise tourism business model. This business model uses cost reductions to enable mass-market cruises on increasingly larger ships (Vogel, 2017) which produce three times more greenhouse emissions than long haul planes (Lloret et al., 2021;

Mozuni & Jonas, 2016) and sail under FOCs, which frees the cruise lines from paying hefty taxes and many labor, environmental, and safety regulations (Terry, 2017), while externalizing environmental and social costs to the society (Klein, 2016b). Cost saving through FOC has also enabled the cruise industry to offer the affordable mass-market product to middle-class customers, and without it, ships would be smaller and cruises would be more expensive (Terry, 2017). However, despite these economic advantages, this business model has not been hugely successful without the subsidy afforded by the lucrative cruise onboard business (Vogel, 2017). An analysis of the finances of the biggest cruise operators between 2001 and 2014 shows that their total revenues are becoming less profitable (Vogel, 2017) and it can be observed that the share prices of the big three had either stagnated (Royal Caribbean and Norwegian Cruise Line) or was declining (Carnival) before the pandemic (see Figure 2). Cruise researchers have long questioned the increasing size of modern cruise ships (Klein, 2016; Papathanassis, 2016). For example, the world's largest cruise ship, Royal Caribbean's *Symphony of the Seas,* has 18 decks, measures 362 meters, and can carry almost 9,000 passengers and crew (CNN, 2018). Researchers have also linked the COVID-19 outbreaks with cruise ship size: "It is clear that cruise ships infected with COVID-19 are large ships" (Ito et al., 2020, p. 5).

The year 2020 started promisingly with the cruise industry preparing for a period of unprecedented boom, with 117 new ships on order by 2027 (Cruise Industry News, 2020). However, the COVID-19 pandemic has now reversed the industry's fortunes. But as the development of COVID-19 vaccines and antiviral drugs signals a path to the end of the pandemic; the cruise industry needs to do a lot of rethinking to make a successful recovery. The findings of

this study and that of previous research (Holland et al., 2021) show that the outbreaks of COVID-19 on cruises and the ensuing media coverage has hurt confidence in this form of tourism and could hamper future growth if the negative perception persists. Reports of full bookings for future cruises do not contradict these findings, since these reported bookings are usually not always new bookings. For instance, in the quarter ending on August 31, 2020, Carnival Corporation filings show that only 55% of these bookings were new bookings, while the rest were future cruise credits (FCC) re-bookings from previously canceled cruises during the pandemic (Carnival, 2020). The cruise industry has a reputation for discounting prices enough to stimulate demand to fill their fixed capacities after external shocks like 9/11 or the 2008 financial crisis (Vogel, 2017), but this line of action could further reduce profitability.

Therefore, the cruise industry needs a new paradigm away from the pre-pandemic mass-market model with diminishing profitability and whose consequences are stoking negative perceptions of cruising towards more sustainable, environmentally friendly, and profitable business models. This would not be a simple task for the cruise industry and could be further complicated by the huge debt taken on by the cruise companies to survive the pandemic. For example, Carnival has raised $23.6 billion through debt and equity so far during the pandemic (2021-04-23: Financial Times, 2021). However, it is difficult to see these challenges successfully resolved with the industry moving in the same direction as it was before the COVID-19 pandemic.

Theoretical implications

The theoretical contribution of this study is both the proposed conceptual model in Figure 3 and the revised conceptual model in Figure 9, combining SARF and IIT. Kasperson et al. (1988) provided a detailed framework, which is quite complex and seems highly suited to extreme events. For example, they listed social protest and disorder as part of the group responses while sabotage terrorism is listed as a potential impact. These types of outcomes can only be anticipated during extraordinary events or circumstances. They also provided a highly simplified framework, which is very skeletal. We have adapted the simplified framework and combined it with the IIT in order to provide new and useful contexts to the conceptual framework that can be directly adopted or adapted in future studies on risk perception. The results offer additional support to the assumptions of the SARF and IIT, on the influence of integrating new information on people's actual beliefs and the influence of media coverage on public opinion and risk perceptions.

However, the monitoring of changes in consumer sentiment expressed on social media is a continuous process, as perceptions can change as industry-related events happen. Cruise researchers can monitor these changes in future research and use them together with other established research methodologies, like surveys, to understand consumer behavior and intentions in the industry. For example, the video of a cruise ship crashing into a dock that went viral (see Appendix) and the release of a documentary film on the COVID-19 outbreak on the

Diamond Princess titled *The Last Cruise* on March 30, 2021 have also generated many social media comments on cruising in recent times.

Managerial implications

The practical contribution of this study is to provide insight into the public perception of cruising during the COVID-19 outbreaks on cruises using social media data. The rebuilding process after the pandemic provides an opportunity for the cruise industry to reinvent itself. One way the industry can achieve this is by doubling down on green credentials and observance of environmental regulations, and not simply by cost reductions and increasing ship size. These green efforts have to be genuine because consumers see through greenwashing and only genuine green behavior improves organizational reputation (de Jong et al., 2020). There should be an emphasis on cleaner and greener ships for new ship orders rather than on increasingly larger vessels. Smaller ships that are more environmentally friendly and offer more intimate experiences could be prioritized. Even though seniors remain the largest demographic of cruisers, the average age of cruise passengers has been falling steadily (Dowling & Weeden, 2017). The industry needs to maintain this trend by attracting younger new-to-cruise passengers with its green credentials when they have it. These young consumers among the millennials and Gen Z have a reputation for driving sustainability, environmental, and ethical consciousness (Choudhary, 2020; Deloitte, 2020; Yeoman, 2008). They are also over-represented on social media (Perrin, 2015) and hence, with a louder voice to express their discontent on sustainability issues on social media, which gets picked up by their peers. After facing a much sterner test with

COVID-19, cruise lines should now also endeavor to keep cruising safe from the routine outbreaks of other infectious diseases by preserving many of the COVID-19 health protocols like frequent handwashing, some measure of social distancing, increased ventilation, ultraviolet air filtration, additional medical facilities onboard, contactless apps for food-ordering, etc. In the aftermath of COVID-19, the cruise industry should vigorously look into alternative business models that would be sustainable and profitable in the long term. Finally, other areas of tourism can also benefit from monitoring the online conversations on social media when industry-related events happen.

Limitations

The restriction of the analyzed tweets to only those in English is an important limitation, since an analysis of tweets in all languages could have painted a more comprehensive picture. It is also possible some relevant tweets were not included in the analysis, as we based the data collection on the presence of keywords and some relevant tweets may not mention these keywords, especially when replying to a tweet that already provided context. As an example, there was a viral tweet (see Appendix), which replied to a news report about the planned resumption of cruises by Carnival with the dancing coffin meme photoshopped with a Carnival cruise ship in place of the coffin to imply danger of death.

References

ABC. (2020). How the coronavirus pandemic would look in Australia if Ruby Princess had never docked - ABC News. https://www.abc.net.au/news/2020-04-23/coronavirus-across-australia-if-ruby-princess-never-docked/12172314

Ainin, S., Feizollah, A., Anuar, N. B., & Abdullah, N. A. (2020). Sentiment analyses of multilingual tweets on halal tourism. Tourism Management Perspectives, 34. https://doi.org/10.1016/j.tmp.2020.100658

Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment Analysis in Tourism: Capitalizing on Big Data. Journal of Travel Research, 58(2), 175–191. https://doi.org/10.1177/0047287517747753

AP. (2021). COVID-19's global death toll tops 5 million in under 2 years. https://apnews.com/article/coronavirus-pandemic-africa-health-pandemics-infectious-diseases-838421260675827b7e735729b90ff95f

Bailey, S. (2020). https://t.co/31DtABZQof. https://twitter.com/samjabailey/status/1260045166674575360

Banda, J. M., Tekumalla, R., Wang, G., Yu, J., Liu, T., Ding, Y., Artemova, E., Tutubalina, E., & Chowell, G. (2021). A Large-Scale COVID-19 Twitter Chatter Dataset for Open Scientific Research—An International Collaboration. In Epidemiologia (Vol. 2, Issue 3). https://doi.org/10.3390/epidemiologia2030024

Banda, J. M., Tekumalla, R., Wang, G., Yu, J., Liu, T., Ding, Y., & Chowell, G. (2020). A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration, v.9. https://doi.org/10.5281/ZENODO.3819464

Barbe, D., Pennington-Gray, L., & Schroeder, A. (2018). Destinations' response to terrorism on Twitter. International Journal of Tourism Cities, 4(4), 495–512. https://doi.org/10.1108/IJTC-04-2018-0027

Bird, S., Loper, E., & Ewan, K. (2009). Natural Language Processing with Python. In O'Reilly Media Inc. http://www.nltk.org/book/

Blumer, H. (1948). Public Opinion and Public Opinion Polling. American Sociological Review, 13(5), 542–549. https://doi.org/10.2307/2087146

Carnival. (2020). SEC Filing | Carnival Corporation & plc. https://www.carnivalcorp.com/node/62011/html

CDC. (2020). Third Modification and Extension of No Sail Order and other Measures Related to Operations. https://www.cdc.gov/quarantine/pdf/CDC-NSO-Third-Extension-09-30-2020-p.pdf

CDC. (2021). Temporary Extension & Modification of Framework for Conditional Sailing Order (CSO). https://www.cdc.gov/quarantine/cruise/pdf/CDC-CSO-Extension-10-25-21-p.pdf

Choudhary, A. (2020). Generation Green is leading the sustainability agenda. https://www.capgemini.com/2020/08/generation-green-is-leading-the-sustainability-

agenda/

CLIA. (2019). 2019 State of the Industry. https://cruising.org/-/media/research-updates/research/clia-2019-state-of-the-industry-presentation-(1).ashx

CLIA. (2021). 2021 State of the Industry Outlook. https://cruising.org/-/media/research-updates/research/2021-state-of-the-cruise-industry_optimized.ashx

CNN. (2018). Symphony of the Seas: World's largest cruise ship sets sail | CNN Travel. https://edition.cnn.com/travel/article/symphony-of-the-seas-world-largest-cruise-ship/index.html

Coombs, W. T. (2018). Ongoing crisis communication planning, managing, and responding. SAGE Publications Inc.

Couper, M. P. (2013). Is the sky falling? new technology, changing media, and the future of surveys. Survey Research Methods, 7(3), 145–156.

Cruise Industry News. (2020). 117 New Cruise Ships On Forward-Looking Orderbook. https://www.cruiseindustrynews.com/cruise-news/22364-117-new-cruise-ships-on-forward-looking-orderbook.html

D'Avanzo, E., Pilato, G., & Lytras, M. (2017). Using Twitter sentiment and emotions analysis of Google Trends for decisions making. Program, 51(3), 322–350. https://doi.org/10.1108/PROG-02-2016-0015

de Jong, M. D. T., Huluba, G., & Beldad, A. D. (2020). Different Shades of Greenwashing:

Consumers' Reactions to Environmental Lies, Half-Lies, and Organizations Taking Credit for Following Legal Obligations. Journal of Business and Technical Communication, 34(1), 38–76. https://doi.org/10.1177/1050651919874105

Deloitte. (2020). Millennial Survey 2020. https://www2.deloitte.com/global/en/pages/about-deloitte/articles/millennialsurvey.html

Dowling, R., & Weeden, C. (2017). The world of cruising. In Cruise ship tourism (pp. 1–39). https://doi.org/10.1079/9781780646084.0001

Feldman, R. (2013). Techniques and applications for sentiment analysis: The main applications and challenges of one of the hottest research areas in computer science. Communications of the ACM, 56(4), 82–89. https://doi.org/10.1145/2436256.2436274

Financial Times. (2020). Coronavirus: is this the end of the line for cruise ships? https://www.ft.com/content/d8ff5129-6817-4a19-af02-1316f8defe52

Financial Times. (2021). The $24bn dash for cash that kept Carnival afloat. https://www.ft.com/content/033f6161-7fb7-4936-8f3b-d78afec94cb2

Gene Zucker, H. (1978). The Variable Nature of News Media Influence. Annals of the International Communication Association, 2(1), 225–240. https://doi.org/10.1080/23808985.1978.11923728

Google Trends. (2020). Google Trends. https://www.google.com/trends

Gössling, S., Scott, D., & Hall, C. M. (2020). Pandemics, tourism and global change: a rapid

assessment of COVID-19. Journal of Sustainable Tourism, 1–20. https://doi.org/10.1080/09669582.2020.1758708

Holland, J., Mazzarol, T., Soutar, G. N., Tapsall, S., & Elliott, W. A. (2021). Cruising through a pandemic: The impact of COVID-19 on intentions to cruise. Transportation Research Interdisciplinary Perspectives, 9. https://doi.org/10.1016/j.trip.2021.100328

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014, 216–225.

Ito, H., Hanaoka, S., & Kawasaki, T. (2020). The cruise industry and the COVID-19 outbreak. Transportation Research Interdisciplinary Perspectives, 5. https://doi.org/10.1016/j.trip.2020.100136

Jamie. (2020). jamie on Twitter: “Who the fuck is driving these cruise ships? https://t.co/0slDSqbXGq.” https://twitter.com/gnuman1979/status/1215799298111262722

Kaplan, A. M., & Haenlein, M. (2011). The early bird catches the news: Nine things you should know about micro-blogging. Business Horizons, 54(2), 105–113. https://doi.org/10.1016/j.bushor.2010.09.004

Kasperson, J.X., Kasperson, R. E., Pidgeon, N., & Slovic, P. (2003). The social amplification of risk: Assessing fifteen years of research and theory. In The Social Amplification of Risk.

https://doi.org/10.1017/CBO9780511550461.002

Kasperson, R.E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., Kasperson, J. X., & Ratick, S. (1988). The Social Amplification of Risk: A Conceptual Framework. Risk Analysis, 8(2), 177–187. https://doi.org/https://doi.org/10.1111/j.1539-6924.1988.tb01168.x

Khong, K. W., Teng, S., Butt, M. M., & Muritala, B. A. (2021). Analysing online reviews of restaurants in Malaysia: a novel approach to descriptive and predictive analytic. International Journal of Electronic Business, 16(4), 315–335. https://doi.org/10.1504/IJEB.2021.118491

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. Business Horizons, 54(3), 241–251. https://doi.org/10.1016/j.bushor.2011.01.005

Klein, R. A. (2002). Cruise Ship Blues: The Underside of the Cruise Ship Industry.

Klein, R. A. (2009). Getting a grip on cruise ship pollution. Friends of the Earth.

Klein, R. A. (2011). Responsible cruise tourism: Issues of cruise tourism and sustainability. Journal of Hospitality and Tourism Management, 18(1), 107–116. https://doi.org/10.1375/jhtm.18.1.107

Klein, R. A. (2016a). Crime at Sea: A Comparison of Crime on Carnival Cruise Lines, 2007--2011. In A. Papathanassis (Ed.), Cruise Business Development: Safety, Product Design and Human Capital (pp. 17–28). Springer International Publishing. https://doi.org/10.1007/978-3-319-

27353-2_2

Klein, R. A. (2016b). Safety, security, health and social responsibility. In Cruise Business Development: Safety, Product Design and Human Capital. https://doi.org/10.1007/978-3-319-27353-2_1

Kluyver, T., Ragan-kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., Willing, C., & Development Team, J. (2016). Jupyter Notebooks—a publishing format for reproducible computational workflows. Positioning and Power in Academic Publishing: Players, Agents and Agendas, 87–90. https://doi.org/10.3233/978-1-61499-649-1-87

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. Tourism Management, 68, 301–323. https://doi.org/10.1016/j.tourman.2018.03.009

Liu. (2009). Informing Design of Next Generation Social Media to Support Crisis-Related Grassroots Heritage. PhD Colloquium of the 6th International ISCRAM Conference, Gothenburg, Sweden, May.

Liu, B. (2012). Sentiment analysis and opinion mining. In Synthesis Lectures on Human Language Technologies (Vol. 5, Issue 1). https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Liu, Pennington-Gray, L., Donohoe, H., & Omodior, O. (2015). New York City bed bug crisis as framed by tourists on tripadvisor. Tourism Analysis, 20(2), 243–250. https://doi.org/10.3727/108354215X14265319207597

Lloret, J., Carreño, A., Carić, H., San, J., & Fleming, L. E. (2021). Environmental and human health impacts of cruise tourism: A review. Marine Pollution Bulletin, 173. https://doi.org/10.1016/j.marpolbul.2021.112979

Luo, Q., & Zhai, X. (2017). "I will never go to Hong Kong again!" How the secondary crisis communication of "Occupy Central" on Weibo shifted to a tourism boycott. Tourism Management, 62, 159–172. https://doi.org/10.1016/j.tourman.2017.04.007

McGregor, S. C. (2019). Social media as public opinion: How journalists use social media to represent public opinion. Journalism, 20(8), 1070–1086. https://doi.org/10.1177/1464884919845458

McKinney, W. (2010). Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference, 1697900(Scipy), 51–56. http://conference.scipy.org/proceedings/scipy2010/mckinney.html

Miami Herald. (2020). COVIDCRUISES MIAMI HERALD DATA. https://docs.google.com/spreadsheets/d/1KTJj94nWd88Gml98AbTZKYZejbEJ8q4Xa7oE6g5r9ec/edit#gid=0

Möller, C., Wang, J., & Nguyen, H. T. (2018). #Strongerthanwinston: Tourism and crisis communication through Facebook following tropical cyclones in Fiji. Tourism Management, 69, 272–284. https://doi.org/10.1016/j.tourman.2018.05.014

Moriarty, L., Plucinski, M., Marston, B., & et al. (2020). Public health responses to covid-19

outbreaks on cruise ships - Worldwide, February-March 2020. In MMWR Morb Mortal Wkly Rep (Vol. 69, Issue 12). https://doi.org/10.15585/MMWR.MM6912E3

Mozuni, M., & Jonas, W. (2016). Rise of Alternative Business Models in the Era of Expensive Megaships. In A. Papathanassis (Ed.), Cruise Business Development: Safety, Product Design and Human Capital (pp. 55–70). Springer International Publishing. https://doi.org/10.1007/978-3-319-27353-2_4

Mueller, A. (2020). WordCloud for Python documentation. https://amueller.github.io/word_cloud/index.html

Muritala, B. A., Sánchez-Rebull, M.-V., & Hernández-Lara, A.-B. (2020). A Bibliometric Analysis of Online Reviews Research in Tourism and Hospitality. Sustainability 2020, Vol. 12, Page 9977, 12(23), 9977. https://doi.org/10.3390/SU12239977

Neri, F., Geraci, P., & Camillo, F. (2010). Monitoring the Web Sentiment, The Italian Prime Minister's Case. 2010 International Conference on Advances in Social Networks Analysis and Mining, 432–434. https://doi.org/10.1109/ASONAM.2010.26

O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. ICWSM 2010 - Proceedings of the 4th International AAAI Conference on Weblogs and Social Media, 122–129.

Pan, T., Shu, F., Kitterlin-Lynch, M., & Beckman, E. (2021). Perceptions of cruise travel during the COVID-19 pandemic: Market recovery strategies for cruise businesses in North

America. Tourism Management, 85. https://doi.org/10.1016/j.tourman.2020.104275

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1–2), 1–135. https://doi.org/10.1561/1500000011

Papathanassis, A. (2016). 'Stranded at sea': Exploring passengers' reactions during incidents at Sea. In Cruise Business Development: Safety, Product Design and Human Capital. https://doi.org/10.1007/978-3-319-27353-2_3

Papathanassis, A. (2017). Cruise tourism management: state of the art. Tourism Review, 72(1), 104–119. https://doi.org/10.1108/TR-01-2017-0003

Penco, L., Profumo, G., & Remondino, M. (2018). Negative events and corporate communication strategies in the cruise industry. International Journal of Transport Economics, 45(3), 485–514. https://doi.org/10.19272/201806703007

Penco, L., Profumo, G., Remondino, M., & Bruzzi, C. (2019). Critical events in the tourism industry: factors affecting the future intention to take a cruise. International Journal of Contemporary Hospitality Management, 31(9), 3547–3566. https://doi.org/10.1108/IJCHM-05-2018-0423

Perrin, A. (2015). Social media usage. In Pew research center (Vol. 125).

Philander, K., & Zhong, Y. (2016). Twitter sentiment analysis: Capturing sentiment from integrated resort tweets. International Journal of Hospitality Management, 55, 16–24. https://doi.org/10.1016/j.ijhm.2016.02.001

Renn, O. (1986). Risk perception : a systematic review of concepts and research results.

Rowe, G., Frewer, L., & Sjöberg, L. (2000). Newspaper reporting of hazards in the UK and Sweden. Public Understanding of Science, 9(1), 59–78. https://doi.org/10.1088/0963-6625/9/1/304

Ryschka, A. M., Domke-Damonte, D. J., Keels, J. K., & Nagel, R. (2016). The Effect of Social Media on Reputation During a Crisis Event in the Cruise Line Industry. International Journal of Hospitality and Tourism Administration, 17(2), 198–221. https://doi.org/10.1080/15256480.2015.1130671

Schroeder, A., Pennington-Gray, L., Donohoe, H., & Kiousis, S. (2013). Using Social Media in Times of Crisis. Journal of Travel and Tourism Marketing, 30(1–2), 126–143. https://doi.org/10.1080/10548408.2013.751271

Schweinsberg, S., Darcy, S., & Beirman, D. (2020). 'Climate crisis' and 'bushfire disaster': Implications for tourism from the involvement of social media in the 2019–2020 Australian bushfires. Journal of Hospitality and Tourism Management, 43, 294–297. https://doi.org/10.1016/j.jhtm.2020.03.006

Sigala, M. (2011). Social Media and Crisis Management in Tourism: Applications and Implications for Research. J. Inf. Technol. Tour., 13, 269–283.

Smith, J. (2005). Dangerous news: Media decision making about climate change risk. Risk Analysis, 25(6), 1471–1482. https://doi.org/10.1111/j.1539-6924.2005.00693.x

Su, L., Stepchenkova, S., & Kirilenko, A. P. (2019). Online public response to a service failure incident: Implications for crisis communications. Tourism Management, 73, 1–12. https://doi.org/10.1016/j.tourman.2019.01.011

Summers, E., Brigadir, I., Kemenade, H. van, Hames, S., Binkley, P., tinafigueroa, Ruest, N., Walmir, Chudnov, D., recrm, celeste, Chosak, A., McCain, R. M., Milligan, I., Segerberg, A., Shahrokhian, D., Walsh, M., Lausen, L., Woodward, N., … Kerchner, D. (2021). DocNow/twarc. https://doi.org/10.5281/ZENODO.593575

Terry, W. C. (2017). Flags of convenience and the global cruise labour market. In Cruise ship tourism (pp. 72–85). https://doi.org/10.1079/9781780646084.0072

Twitter. (2020). More on restricted use cases – Twitter Developers. https://developer.twitter.com/en/developer-terms/more-on-restricted-use-cases

Van Rossum, G., & Drake, F. L. (2009). Python 3 Reference Manual.

Vogel, M. P. (2017). Economics of cruise shipping: the need for a new business model. In Cruise ship tourism (pp. 124–137). https://doi.org/10.1079/9781780646084.0124

Wang, Y., Zheng, L., & Zuo, J. (2021). Online rumor propagation of social media on NIMBY conflict: Temporal patterns, frameworks and rumor-mongers. Environmental Impact Assessment Review, 91. https://doi.org/10.1016/j.eiar.2021.106647

Washington Post. (2020). Cruise ships kept sailing after coronavirus was detected, exposing thousands. https://www.washingtonpost.com/graphics/2020/politics/cruise-ships-

coronavirus/

WHO. (2020). Naming the coronavirus disease (COVID-19) and the virus that causes it. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it

Womack, J. J., Anderson, L. N., & Ledford, C. J. W. (2020). Presence of Complex and Potentially Conflicting Information in Prenatal Mobile Apps. Health Promotion Practice, 21(2), 238–245. https://doi.org/10.1177/1524839918796216

Yeoman, I. (2008). Chapter 19 - Life on Mars — How Different Generations Perceive the Future of Holidays in Scotland. In I. B. T.-T. T. Yeoman (Ed.), Advances in Tourism Research (Vol. 16, pp. 261–272). Butterworth-Heinemann. https://doi.org/https://doi.org/10.1016/S1572-560X(08)00419-8

Yu, M., Li, Z., Yu, Z., He, J., & Zhou, J. (2020). Communication related health crisis on social media: a case of COVID-19 outbreak. Current Issues in Tourism. https://doi.org/10.1080/13683500.2020.1752632

Zeng, B., & Gerritsen, R. (2014). What do we know about social media in tourism? A review. Tourism Management Perspectives, 10, 27–36. https://doi.org/10.1016/j.tmp.2014.01.001

Zhang, M., Zhang, Y., & Fu, G. (2016). Tweet sarcasm detection using deep neural network. COLING 2016 - 26th International Conference on Computational Linguistics, Proceedings of COLING 2016: Technical Papers, 2449–2460.

Appendix

	Time	Positive tweets	Sentiment Score
1	15/02/2020 16:10	Love it when it all works out good... Wandering ship becomes best cruise ever despite coronavirus fears https://t.co/db7FOK1e2O	0.9062
2	27/03/2020 23:22	.@IHG It would be really nice to get a refund for my reservations in June, seeing as the cruise I booked last August is now canceled. Are people really still visiting Seattle? #Coronavirus sucks but you don't have to!	0.0763
3	15/03/2020 21:44	To all you good Christians out there please say a prayer for the CRUISE INDUSTRY. They need ALL OUR OUR THOUGHTS and SUPPORT. Also WALL STREET, SAUDI ARABIA, and PUTIN. They've all suffered so much. #trump #trumpspeech #trumpspandemic #coronavirus #trumpvirus	0.6633
4	14/04/2020 14:16	@i9998217 @CruiseIndustry I think they'd be unsuitable for those purposes since cruise ships have proven themselves to be petri dishes for #Covid19. They'll make wonderful reefs! ðŸ˜Ž	0.5719
5	07/02/2020 15:10	#pdknews Good news NYC/New Jersey, a cruise ship with 12 passengers who have the Coronavirus just pulled into port. https://t.co/R3dSYH0pEe	0.4404
6	12/04/2020 21:14	FYI followers this article offers a good explanation regarding FOC Flag Of Convenience Ships this is the Cruise Ship Industry this applies also to applies to 99% of ships sailing around Australia on our coastal shipping trade, the same for all ships taking our exports O/S https://t.co/zvLFp0u9jJ	0.6633
7	17/04/2020 14:31	I will never, ever go on a cruise ship again. you could not pay me enough to endanger my life on these floating petri dishes. https://t.co/4p5pibjhY2	0.0762
8	05/02/2020 7:40	@TIME Reason number 10946 not to vacation on a cruise ship. #coronavirus #ncov #Japan	0.0772
9	16/04/2020 5:12	@MollyJongFast But Molly, cruise ships are a wonderful place for Americans to congregate and spread love and cheer and Covid-19 and...	0.9042
10	24/03/2020 15:26	@SenTedCruz @JohnCornyn @RepRWilliams cruise lines do not pay US taxes nor are they US corporations. DO NOT BAIL THEM OUT. https://t.co/3ZBtcJ9Jdz	0.0762
11	15/02/2020 12:55	Be thankful this Saturday morning that you are not stuck on a cruise ship. #CoronaVirus #DiamondPrincess https://t.co/yVnRHhFpQ1	0.6641
12	09/04/2020 6:23	@JohnnyJet Ya know, I wand much interested in cruises because of norovirus. Definitely not interested anymore after one Coronavirus outbreak.	0.7964
13	25/03/2020 2:57	Plus my cruise line is only giving "cruise credits" not refunds for a future cruise they will jack the price up on. #UGH #COVIDIOT45 https://t.co/DuOLYpHTyU	0.0943
14	07/02/2020 5:34	@DisneyCruise Are you taking any precautions regarding coronavirus on your ships? We are scheduled to sail in a few weeks. TIA Our family has our Flu vaccines already ðŸ˜Š	0.347
15	14/04/2020 14:01	Long before Coronavirus, you couldn't have paid me to take a cruise. Hopefully a infected legacy of the pandemic will mean others won't have to pay again either https://t.co/otYaO58Vfg	0.4215
16	11/02/2020 0:35	As a reminder, we will not have an episode tomorrow, since Jordan is going on a cruise. Let's hope he doesn't get coronavirus!	0.4404
17	18/04/2020 0:33	@smh Coronavirus and Cruise ships a perfect match. https://t.co/rrXwX2Lwx2	0.5719
18	21/04/2020 5:19	The tale of The last cruise ship on Earth. "We became like a family - our guests and our crew together. The spirit has been beautiful." #GoodNews #WereAllInThisTogether #COVID19 https://t.co/4aioMAF37u https://t.co/jTA7O6bMTM	0.7964
19	02/04/2020 21:20	These "healthy" cruise ship passengers arriving in Ft. Lauderdale now have been exposed and are required to self-quarantine for 14 days once they get home. On the way home, however, they will potentially expose plenty of other travelers. #COVID19 #Coronavirus #Zaandam #Rotterdam https://t.co/ZVSVX8rqYK	0.2023
20	10/02/2020 17:03	At least the price of cruises are going down. And you wouldn't believe the deals on bat soup. Thanks #coronavirus https://t.co/mjwVveZxhI	0.4404

	Time	Neutral tweets	Sentiment Score
1	05/03/2020 18:16	Man, going to be tough refusing some of these cruise ship deals that will be popping up. #COVID19	0.0493
2	27/04/2020 1:34	Cleaning a floating petri dish: How is a cruise ship sanitized after a coronavirus outbreak? https://t.co/cyixhGrB38	0
3	14/02/2020 7:13	Well this seems like a very bad time to take a #Cruise #Coronavirus #COVID19 https://t.co/unVsTtCdxO	-0.0498
4	27/04/2020 4:11	Coronavirus watching you board a cruise ship https://t.co/9dnwf4Zol3 https://t.co/RcJ7CNUZto	0
5	05/03/2020 5:11	'Cluster' of 21 people on Princess Cruise ship heading toward Calif. show possible coronavirus symptoms: officials https://t.co/00uus5Jwme #FoxNews	0
6	05/02/2020 6:05	Quarantined on a cruise ship, food running out. #coronavirus https://t.co/vkAWZBmAIR	0
7	04/04/2020 19:16	I ain't cruising until 2022 https://t.co/NtMZa1mGjC	-0.0267
8	11/02/2020 6:00	Going on a cruise? Here's how the coronavirus will change your trip https://t.co/hYVCOx2Egt	0
9	27/04/2020 5:43	Coronavirus: How did Australia's Ruby Princess cruise debacle happen? https://t.co/5pFLnZ5WGX https://t.co/YSpCrXKFcK	0
10	11/02/2020 6:53	Japan might test everyone on cruise ship for coronavirus https://t.co/PMkvBFCpLq via @CBSNews	0
11	27/04/2020 2:39	Two things not to do in a pandemic... book a cruise, download a tracking app... https://t.co/keMd4UzJJ7	0
12	25/03/2020 19:29	The size of cruise ships is mad. This pandemic hopefully will result in a lot of rethinking, not just about cruise ships. but travel in general. https://t.co/v3KFQjCbmI	-0.0444
13	06/02/2020 15:33	Just saw an article about cruise ships being quarantined due to coronavirus and it's a great reminder that there's no reason to ever go on a cruise.	0.0344
14	26/04/2020 22:22	Can you imagine not being able to disembark from a cruise for more than 40 days? @fdilella spoke with @dan_domenech who has been stranded due to strict CDC guidelines, amid the Coronavirus pandemic. Here is his story: https://t.co/oGqsuAcq5N	0
15	26/04/2020 23:35	[Feed]Coronavirus journey: The 'last cruise ship on Earth' finally comes home https://t.co/3wH1wuYkzm	0
16	30/04/2020 3:43	Sixty-seven new cases of #Coronavirus were confirmed on the DiamondPrincess cruise ship, bring the total 285. The ship docked in Yokohama, Japan remains quarantined. https://t.co/BfncVKGIt0 #WednsdayWisdom	0
17	11/02/2020 7:15	American coronavirus patient describes "surreal" cruise experience - CBS News https://t.co/fase1Stk7T #cruise #travel	0
18	27/04/2020 2:48	@ABCWorldNews What about the guests or shall I say COVID-19 patients on that cruise ship? Hmmmm.	0
19	11/02/2020 6:24	How a cruise ship turns into a coronavirus breeding ground https://t.co/HfCyHZHgWy https://t.co/J4fkKCjbQd	0
20	05/02/2020 6:11	NBC News: 10 coronavirus cases confirmed from cruise ship quarantined in Japan. https://t.co/tjTmcV632f via @GoogleNews	0

	Time	Negative tweets	Sentiment Score
1	24/03/2020 11:09	Until cruise ships are entirely sustainable and powered by renewable energy they're utterly at odds with our climate change crisis, let alone the Coronavirus Cruises they've become. https://t.co/9dpR49UrDy	-0.6124
2	27/01/2020 23:29	My elderly mother is scheduled to depart this week on a 28 day cruise and ALL ports have cases of coronavirus. She called @hollandamerica to see if she could receive a refund or voucher. The answer was NO. #coronavirus #badcustomerservice	-0.4466

3	09/05/2020 22:42	imagine surviving coronavirus, taking a tropical cruise to celebrate, and then puking and shitting yourself to death in a "stateroom" the size of a closet because you got norovirus from the all you can eat dessert bar. just let the plague ship industry die. https://t.co/jYoeiVDF9z	-0.8176
4	10/02/2020 18:54	@billburr So ummm. Ya know that whole sinking Cruise Ships idea? I think now would be a good time to start.. with the Coronavirus and all.	-0.4404
5	05/05/2020 9:07	Make that forever. Cruises are an environmental catastrophe on every dimension, fiscally irresponsible operating under flags of convenience if not connivance and a health hazard even in the best of times. https://t.co/0TsLNQrhl2	-0.4767
6	23/04/2020 6:43	The cruise industry shouldn't recover. It's bad for the seas and the planet.	-0.3482
7	06/02/2020 1:20	You could not pay me enough money to go on a cruise. Coronavirus aside, these stories of ships losing power and everyone getting sick and waste all over the place. No way in hell. https://t.co/GOdXb95sFv	-0.9349
8	09/05/2020 17:20	.@NPR @KHNS_FM There were already outbreaks of illness on these vessels before coronavirus. Now with a number of outbreaks in assorted US states traced back to cruises, and their off-shore registration to avoid taxes, I'm not sorry to see the industry decline.	-0.5232
9	05/05/2020 6:11	what if we stop taking cruises altogether since they're gross af? https://t.co/4FIHUIGJFO	-0.6486
10	09/03/2020 23:51	Grand Princess Passengers horror on coronavirus cruise ship as people 'fight over rotten food' - Yahoo News Australia - That's a maritime emergency now they should be allowed to dock! https://t.co/BWNtSc1P0I	-0.8481
11	16/04/2020 15:49	Maybe I'm biased because going on cruise falls behind root canal for me, but I can't think of a worse, more disgusting collection of grossness than a cruise ship. Public restrooms are probably cleaner.	-0.7386
12	06/05/2020 11:51	@GeraldoRivera I've never been on a cruise.I always hated the idea of cruises especially since many of them were getting the norovirus. At this point , with the coronavirus, I am never ever going on a cruise ship.	-0.6369
13	13/04/2020 21:27	I have gone on cruises in the past, but you could not pay me to go on another one after this mess! #CruiseControl https://t.co/rUVEY91P2d	-0.5217
14	08/04/2020 7:14	#COVID19 BAN grotesque size of polluting cruise liners to save oceans. And now, they are floating hulks of the diseased and dying due to virus. Things turn into their opposites. https://t.co/RGbmVCChgX	-0.2808
15	16/04/2020 16:02	Someday we will all be saying, "Do you remember when we packed ourselves on disgusting petri dish cruise ships in rooms smaller than closets with no windows, where we got stomach viruses, drank cheap booze and ended up sea sick? What were we thinking? Glad that's over.	-0.1596
16	07/05/2020 14:04	Several cruise ships currently parked in our harbour, polluting the city with their tons of exhaust gases every day. On board: underpayed employees, threatened by a COVID19 outbreak, unable to get home. Despicable industry needs to change.	-0.6369
17	02/04/2020 18:38	Half of all #Covid_19 #coronavirus #pandemic cases in #Australia are linked to #cruiseships Floating petri dishes! Crews are not Australian, ships are not registered in Australia to avoid paying taxes. Get them out of our waters #scomo #ScottMorrison #LNP #auspol	-0.7274
18	04/05/2020 15:07	.@RichieFed Cruise ships were enormous floating Petri dishes before COVID19. Cruises are for the newlywed & nearly dead.	-0.6486
19	24/03/2020 3:46	Read this thread, then listen to @CrimeJunkiePod episode about the death at sea, and then you will understand why I will never take a cruise ever. If you're not dying of some infectious disease, then you're being pushed off a boat or sold into human trafficking and no one cares. https://t.co/nAg2y8c1XN	-0.4767
20	28/03/2020 12:41	NO MASS CRUISES!! THEY THROW THEIR SEWAGE IN THE OCEAN! THEY ARE NOT REGISTERED IN THE US SO THEY DONT HAVE TO ABIDE BY OUR ENVIRO REGS!! https://t.co/yUb2HWt3sq	-0.296

CHAPTER 4. STAYCATIONS DURING COVID-19

Preface

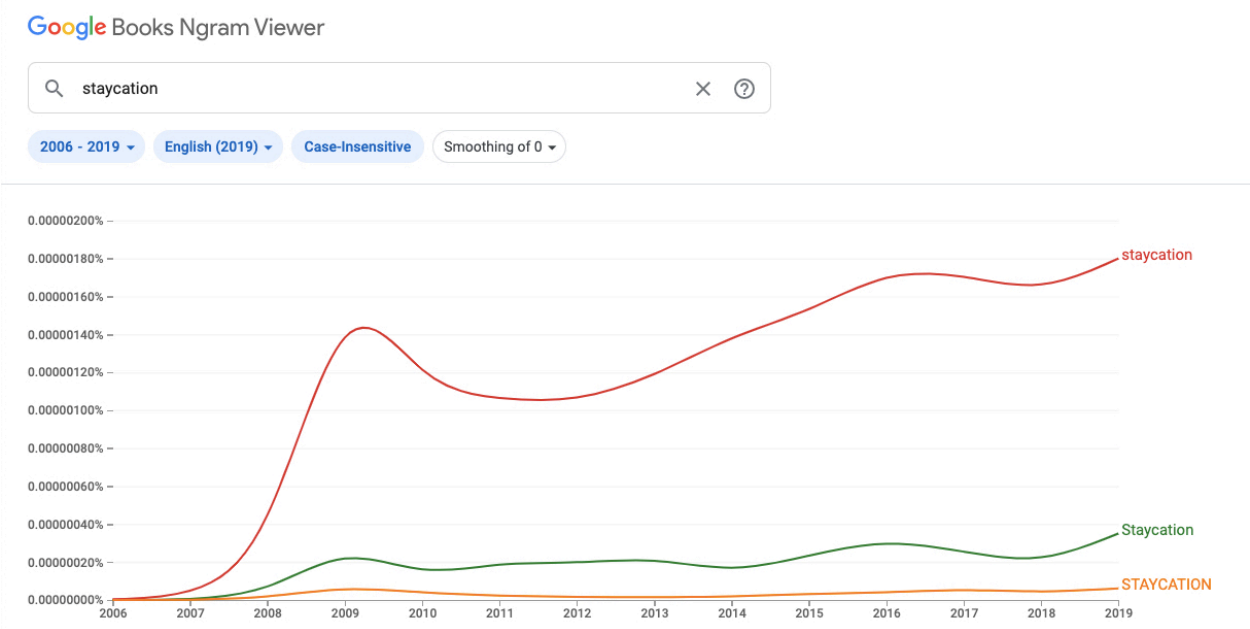
This chapter of the thesis examines the re-emergence of staycations to the fore during the COVID-19 pandemic, as many were forced to spend their vacations close to home. Staycations went mainstream during the 2008 financial crisis, and have now been further accelerated by the pandemic. This study investigated the growth and practice of staycations during the first two years of the pandemic by analyzing social media and internet search data using Latent Dirichlet Allocation (LDA) topic modeling and Google Trends analytics. Key findings point to a strong interest in spending staycations at hotels rather than at home, while the optimal LDA topic model produced 38 topics which were classified under four aggregate dimensions of antecedents, attributes, activities, and consequences of staycations. The findings provide critical insights to hotel managers and policymakers on boosting revenue through the practice, and the role of staycations in promoting tourism closer to home and sustainable tourism.

The article based on this chapter is titled: “COVID-19 staycations and the implications for hospitality and tourism” and was submitted to the *Journal of Leisure Research*.

Introduction

International tourism fell dramatically in the aftermath of the 2008 financial crisis as the global economy went into a deep recession with some regions suffering a decline of up to 18% in international arrivals (Smeral, 2010; UNWTO, 2009). As a result of the financial crisis and rising fuel prices, people had less discretionary income to spend on leisure travel and tourism; which led to the substitution of international leisure travel with domestic activities (Papatheodorou et al., 2010). Hence, consumers embraced tourism to places close to home and the word "staycation" received attention in the popular press, entered mainstream usage, and was added to the dictionary during this period (Frew & Winter, 2010; Merriam-Webster, n.d.). Tracing the evolution of staycation in the Google Books Ngram index shown in Figure 4.1 equally reflects this. This index contains the text corpora of millions of digitized books and printed sources from 1500 to 2019 in its latest version (Jean-Baptiste et al., 2011). A search in the index for staycation shows that the use of the word was flat until 2006, before growing rapidly and reaching its first peak in 2009 during the financial crisis. Afterwards, its prevalence initially fell slightly, before continuing to grow from 2012 onwards.

Figure 4.1: Use of staycation in Google Books Ngram index from 2006 to 2019 (Source: Google Books Ngram Viewer)



The hospitality and tourism industry has been hit hard by the COVID-19 pandemic (Gössling et al., 2020; Williams, 2021). Compared to the global financial crisis, the pandemic has had far devastating consequences for international travel and tourism. At different periods since the beginning of the COVID-19 pandemic, a common tool by various national governments to control the spread of the SARS-CoV-2 virus has been non-pharmaceutical intervention (NPI) measures such as curfews, stay-at-home orders, restriction of public gatherings, lockdowns, travel restrictions, and travel bans with cross-border travel restricted or constrained by tests, quarantine or vaccine requirements (Gössling et al., 2020; C. Michael Hall et al., 2020; Rice et al.,

2022). During the early part of the pandemic, international travel came to an almost complete halt with international arrivals plunging by 97% in April 2020 to levels not seen since the early 1990s and resulting in the loss of international tourism revenues by more than ten times those of the financial crisis (UNWTO, 2020). Italy was the first western country to order a nationwide lockdown on March 9, 2020 (Ren, 2020) and by the end of March 2020, over 90% of the world's population was under some form of international travel restriction resulting in a de facto pause in international tourism (Gössling et al., 2020).

The COVID-19 infection rate in different countries came in waves, and as the first wave of infection subsided with a decline in positive cases around June 2020, some control measures were relaxed that allowed some forms of tourism to resume (Bontempi, 2021; Collins-Kreiner & Ram, 2020). However, research showed that COVID-19 had already inspired a "pandemic travel fear" in people (Zheng et al., 2021). According to a United Kingdom (UK) national survey in August 2020, 68% of respondents cited fears of being stranded abroad and 62% cited uncertainty around COVID-19 as reasons for preferring a staycation in the UK rather than an international holiday (The Cumberland, 2020). After this relaxation of travel restrictions, attempts were made to restart tourism through the promotion of travel to in-country destinations, and domestic tourism showed some signs of recovery in many countries (C. Michael Hall et al., 2020). Several governments such as in Ireland, Iceland, Italy, Poland, Slovenia, Lithuania, South Korea, Macau, and Thailand started staycation initiatives to encourage visits to local destinations (Cvelbar et al., 2021; Wong et al., 2021).

Hence, staycations have boomed during the COVID-19 pandemic and have been part of the emergent customer behavior inspired by the pandemic (Wong et al., 2021). It deserves investigation by hospitality and tourism scholars because of its possible effects on the post-COVID-19 agenda of hotels, hospitality, and sustainable tourism (Pichierri et al., 2022). However, the literature on staycations is still sparse, especially in the context of the pandemic, with most of the studies published after 2008 (de Bloom et al., 2017; Heimtun, 2017; Molz, 2009; Sharma, 2009) and its definition remains unclear. There is confusion about its meaning, especially a conflation with domestic tourism. For example, the British Prime Minister's summer break in August 2020 in Scotland, hundreds of miles away from his place of usual residence and work in London was widely reported as a "staycation" because it was within the UK (Evening Standard, 2020). Hence, there is a gap in the literature to accurately define and examine the concept of staycations in light of the pandemic and also evaluate how it has evolved since 2008. Therefore, the objective of this study is to investigate staycations during the first two years of the COVID-19 pandemic by answering the following research questions:

RQ1. How have staycations changed since 2008, i.e. with regard to growth and new activities?

RQ2. What are the notable aspects of staycations in the first two years of the COVID-19 pandemic?

This study analyzes user-generated content (UGC) from Twitter using topic modeling to understand the discourse about staycations on social media and internet search data from Google to study the search interest and behavior about staycations. As a first step to answering

the research questions, this study explores the identifying elements of staycation and its differentiators from adjacent concepts like domestic tourism, based on the extant academic and industry literature, in order to propose a definition and highlight its benefits. The results of the study are interpreted using Construal Level Theory (CLT), described in the theoretical framework of this study. Finally, the implications of the insights for hospitality organizations and sustainable tourism are discussed.

Literature review

Staycation in the literature

Vacations (used interchangeably with holidays in the literature) provide a means for macro-recovery to help people recover from work and everyday stress (de Bloom et al., 2009). Empirical studies have shown that vacations contribute to the quality of life and well-being of many (Dolnicar et al., 2012; Gilbert & Abdullah, 2004). This is besides several other benefits associated with going on vacations. As McCabe (2009b, p. 683) stated: “it is clear that a holiday can have demonstrable impacts connected with many areas of current government policy on: health and well-being particularly in relation to the treatment of stress-related illnesses and disorders but also in a range of other potential treatments or holistic and alternative therapies.” Because of these benefits, going away for a holiday at least once a year has become an integral part of the postmodern society (Urry, 1988). The advent of inexpensive travel amplified this and contributed to the modern consumer culture of mass tourism in which to stay home is to be pitied (Urry, 1988). Mass tourism has made tourist destinations class-neutral where people of different

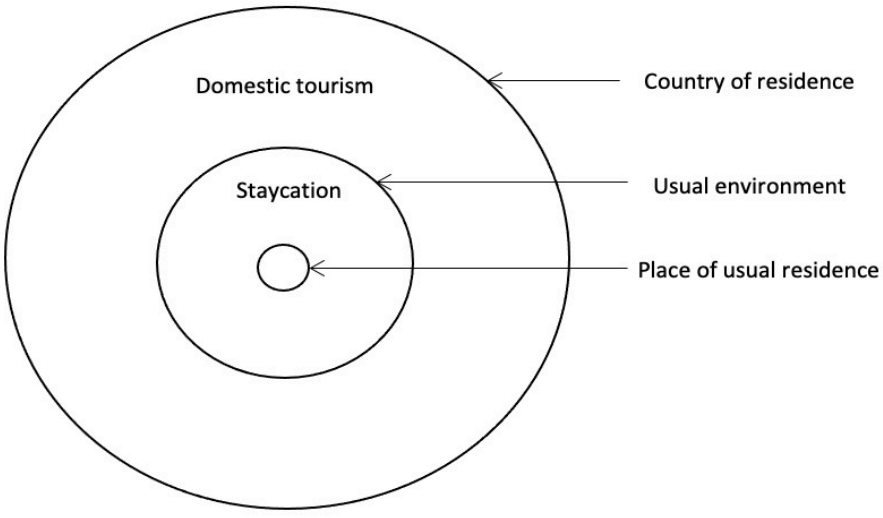
classes in the society can share the same type of holidays. While resulting in a reversal of the roles of the home and the destination, with the home taking the role as the primary place of displaying identity and status (McCabe, 2009a).

Staycation is a portmanteau expression derived from the combination of stay and vacation; and to "take a staycation" means to stay at home during a vacation rather than traveling to a destination, which is presumed to be the point of vacations (West, 2018). Sharma (2009) defined staycations as a neologism referring to "the activity of making a vacation out of staying at home." Molz (2009) calls it an invented term for describing vacationing at home. Some previous studies credited Terry Massey as the first person to use the word in 2003 (de Bloom et al., 2017; Hay, 2010; James et al., 2017), but the Merriam-Webster dictionary cites a much earlier source in the *Cincinnati Enquirer* newspaper from 1944: "Take a Stay-cation instead of a Va-cation, this year" (Merriam-Webster, n.d.). The use of the word went mainstream during the global financial crisis as soaring fuel prices pushed up airfares, room rates, and other miscellaneous expenses while the global economy was in a recession, which made going for a vacation prohibitively expensive (Fox, 2009; Molz, 2009). Interestingly, the first written usage from 1944 can also be ascribed to a period of crisis as it was a wartime admonition to conserve fuel during the Second World War (Merriam-Webster, n.d.).

Hay (2010) described staycations as just another name for domestic tourism. However, researchers like de Bloom et al. (2017) distinguished between domestic tourism and staycation and measured the difference in the subjective well-being of a vacation in one's home or domicile

and that of a domestic holiday in their study. This study agrees with the latter and proposes a definition of staycation based on the convention of the World Tourism Organization (UNWTO) for domestic tourism, place of usual residence, and usual environment. The UNWTO defines domestic tourism as activities of a resident visitor within the country as part of a domestic tourism trip or an outbound tourism trip. The place of usual residence as the geographical location a person lives. And the usual environment of a person as the geographical perimeter within which they conduct their regular life routines, which includes their place of usual residence, place of work or study, and all other places visited regularly or frequently, even if far from the usual residence or in another locality, except for vacation homes (UNWTO, 2010). Therefore, staycation could be described as tourism characteristic activities performed within one's usual environment, as represented in Figure 4.2. Hence, staycation is a subset of domestic tourism, but a domestic tourism trip might not be a staycation if outside one's usual environment.

Figure 4.2: Domestic tourism and staycation



Benefits of staycations

People have always spent some of their holidays in their immediate environment without taking a trip for various reasons, but this was not previously considered a vacation or type of tourism (Haukeland, 1990). The novelty is its treatment as a valid type of vacation and the neologism. Researchers like Heimtun (2017) have challenged simplistic notions of tourism and the stereotypical assumption that proper holidays must involve travel in a qualitative study of midlife single women's home holidays, arguing that it was possible to be a tourist in the home area. Lohmann (1996) found that subjects who traveled during their holidays felt a higher increase in recreation in the middle of their vacation, which fell rapidly after the vacation while the recreation effect of subjects that spent their holiday at home lasted longer after the holidays. In similar findings, de Bloom et al. (2017) found domestic travel to be higher in stimulating engagement in social activity and detachment from work and everyday stress but found no significant difference in the level of pleasure and hedonic well-being generated between domestic travel and a staycation.

Sharma (2009) criticized the manifestation of staycations in the United States (US) during the financial crisis and its gender, class, and race dimensions. In this study, she described this manifestation of staycations as an exercise in waiting out the crisis before real life resumes and everyone got moving again rather than as an opportunity to embrace stillness as an alternative to life in constant motion. Likewise, Molz (2009) also criticized the ambivalent attitude towards staying still in the media coverage and public conversation about staycations in the US during this

period as an abnormality to be endured rather than embraced and enjoyed. Staycations have been promoted as beneficial based on advantages due to less travel, cost savings, reduced stress, and a boost to local economies (Frew & Winter, 2010). According to Dodds (2012), staycations like slow travel highlights the opportunities for enjoyable holidays in one's backyard, which can be as good as that got by traveling for thousands of miles. Finally, a social shift from long to short travel distances from one's domicile could substantially reduce the greenhouse gas emission contributed by tourism and is more environmentally friendly (Gössling et al., 2010).

Construal Level Theory

The construal level theory (CLT) is the theory behind the notion of psychological distance. This theory from psychology proposes that human beings only directly experience the present, here, and now. The experience of all other objects beyond the immediate situation is done by forming abstract mental construals of these objects. Hence, psychological distance is a subjective experience that something is close or far away from the self in the present. Psychological distance is egocentric with the self as the reference point, while how an object may be removed from this reference are distance dimensions in time (temporal), space (spatial), social distance, and hypothetical distance (Trope & Liberman, 2010). Research from different domains has converged on the notion that transcending the present requires and is enabled by the human capacity for abstract mental representation (Trope & Liberman, 2010).

CLT postulates that we focus on the abstract at a higher level of construal and on the concrete at lower levels. The perceived distance influences decisions and behavior, i.e. when making choices

in psychologically distant situations, people focus on the central or global features of an object (Trope et al., 2007). Several studies in hospitality and tourism have used CLT and there has been a call for more practical applications of CLT in this field (Lindblom et al., 2020). This seminal theory has been employed to investigate how temporal distance and gender influences the assessment of hotel attributes (Kim et al., 2018), the effect of psychological distance or proximity via pictorial information on the evaluation of tourism products (Jia et al., 2021; Marlow & Dabbish, 2014), the relationship between the perception of psychological distance and length of tourist stay (Hateftabar, 2021), the effect of psychological distance on consumer response to destination advertisements (S. Wang & Lehto, 2020), and the influence of psychological distance on promotional messages in tourism (Kim et al., 2016).

Vacation travel is a beloved pastime and when it was not possible because of the pandemic, people settled for various activities within their usual environment. CLT offers a theoretical foundation for understanding the relationship between psychological distance and these activities and behaviors. Therefore, this study interpreted findings in light of this theory as people tried to create a simulacrum of their regular vacation while physically at or close to home through the lens of psychological distance. In addition, this theory has not been previously applied to staycations and is a gap in the literature that this study covers.

Methods

This research consists of two studies, the first is an analysis of Twitter UGC and the second based on analytics of Google Trends data. In the first study, topic modeling of tweets containing

staycation was conducted to get the major topics in the data. Topic modeling is a text analytical technique to extract the main features from a body of text that has been applied extensively in the hospitality and tourism literature (Muritala et al., 2020; Taecharungroj & Mathayomchan, 2021). The tweet IDs were obtained from a published COVID-19 Twitter dataset which covers the year 2020 (Banda et al., 2021). These tweet IDs were rehydrated to get the full tweets using Twarc (Summers et al., 2021). The dataset contained tweets in 65 languages, which we could not translate for our analysis. Therefore, only the tweets in English were rehydrated to save time and computing resources. 103,785,252 full tweets were successfully rehydrated. We loaded the required entries of each tweet with Pandas (McKinney, 2010) into a Python Jupyter Notebook (Kluyver et al., 2016). We then selected the tweets in which staycation appeared and performed painstaking data cleaning to reduce the noise in the data by removing irrelevant tweets, i.e. tweets about something else with the staycation hashtag attached. After data cleaning, we carried out pre-processing steps like removal of stop words, URL links, white space, newline characters, distracting single quotes, and other strange quirks in the text before lemmatizing the data. After these essential steps, topic modeling was applied to the tweets to identify the major topics or aspects of the online discourse about staycations using Latent Dirichlet Allocation (LDA) implemented using MALLET (version 2.0.8). LDA is an efficient topic modeling technique for extracting the hidden topics from large unlabeled textual data (Blei et al., 2003). MALLET is a Java-based program that uses sophisticated machine learning algorithms to enable faster processing and get a better quality of topic classification (McCallum, 2002).

In the second study, data from Google Trends is analyzed to obtain insights from the search behavior around staycations. Google Trends is an open platform that shows an unbiased sample of anonymized, categorized, and aggregated search queries of a keyword or topic on Google Search, the most popular search engine. It can measure the level of interest in a topic across the globe, in a specific geographical location, or over a specific period (Rogers, 2020). Analysis of internet search query results have been shown to reflect an interest in the real world (D'Avanzo et al., 2017). The data for this study was downloaded using Pytrends (version 4.7.3) (Hogue & DeWilde, 2020), an application programming interface (API) for Google Trends in Python (Van Rossum & Drake, 2009). We obtained the worldwide historical search data for "staycation" from 2004 (when Google started collecting the data) until December 2021, the interest by region, and the "top" and "rising" related search queries and topics for 2020 and 2021, to provide context around the search term. The interest by region shows the relative popularity of staycation searches across different locations. The Top related queries are the most popular terms that users who searched for staycation also searched for. The Rising related queries are terms with the biggest increase in search frequency over the period. The Top related topics are the most popular topics that users who searched for staycation also searched for. While the Rising related topics are topics with the biggest increase in search frequency.

Findings

Study 1: Topic Modeling

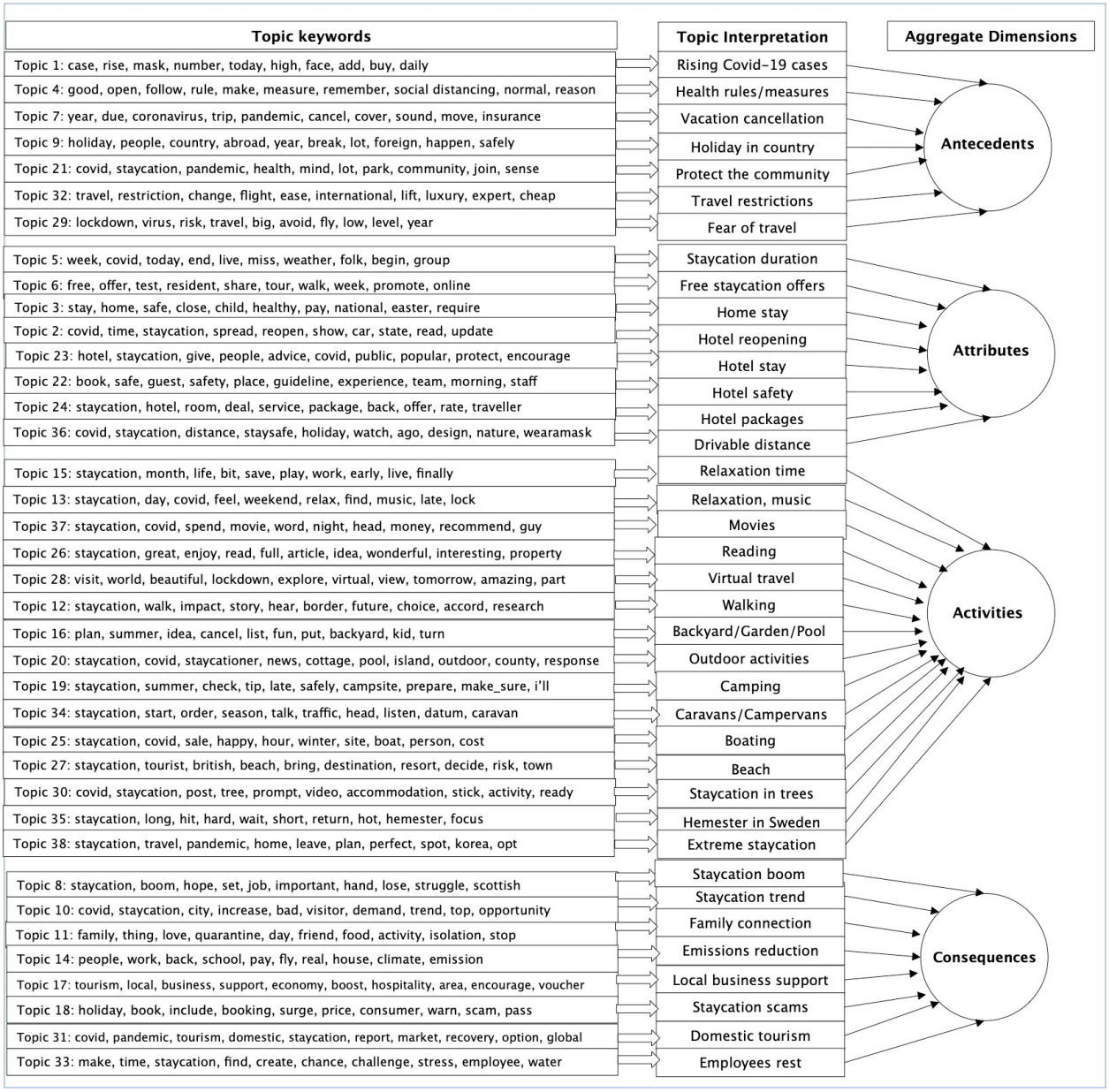
Topic modeling enabled the identification of the major topics in the Twitter UGC to primarily provide insight on the features of COVID-19 staycations in the second research question (RQ2) and also about new staycation-related activities in the first research question (RQ1). 7,729 tweets containing staycation posted by 6,059 unique users were obtained. The mean, median, and standard deviation of tweets per user were 1.276, 1, and 1.929. This indicates that most of the tweets were probably from personal accounts, and not repeated posts from the same account as typical with news media or blog accounts. Manual examination to identify irrelevant tweets like adverts was carried out, which were then removed (McKinney, 2010). 6,067 tweets were left after data cleaning. Before applying the MALLET algorithm to get the dominant topics in the staycation tweets, we obtain the optimal number of topics for the LDA topic model by calculating the topic coherence score for different numbers of topics and picking the number of topics with the highest topic coherence. The topic coherence score of a topic model is a measure of the quality of the topic model (Syed & Spruit, 2017) which computes the sum:

$$Coherence = \sum_{i<j} score(w_i, w_j)$$

Of the pair-wise scores of words $w_1,\ldots, w_n$ used to describe the topic, which are usually the top n words by frequency $p(w|k)$ (Pleplé, 2013). After computation, our optimal LDA model was the model with 38 topics with the highest topic coherence score of 0.413. The topic keywords, count

of tweets, and portion of each topic in the optimum LDA model are presented in Appendix. The interpretation of the topic model is presented in Figure 4.3. This interpretation was made based on the topic keywords and manual reading of the tweets assigned to each topic. Furthermore, we grouped the topics into aggregate dimensions in order to generate insights into the broad aspects of the tweets. These aggregate dimensions developed were grouped into the antecedents, attributes, activities, and consequences of staycations. The authors worked separately to group the topics into these aggregate dimensions during a first run, then came to agreement on the classification during a joint second run.

Figure 4.3: Topic Model Interpretation and Classification



The aggregate dimension on antecedents comprised of tweets on the causes of growth in staycations and is made up of seven topics as follows; the rise in the number of daily cases of COVID-19 (Topic 1), health rules or measures to control the spread of the virus (Topic 4),

cancellation of vacation trips (Topic 7), spending holiday in country (Topic 9), protecting the community (Topic 21), travel restrictions (Topic 32), and travel fear (Topic 29).

The aggregate dimension on attributes comprised of tweets on the characteristics of staycations and is made up of eight topics as follows; duration of staycations (Topic 5), staycation deals and hotel staycation packages (Topic 6 & 24), staying home (Topic 3), hotels reopening and encouraging staycations (Topic 2), staying at hotels (Topic 23), hotels mentioning their safety measures (Topic 22), staycation within a drivable distance (Topic 36).

The aggregate dimension on activities comprised of the activities people engaged in and is made up of 15 topics as follows; relaxation (Topics 13&15), watching movies (Topic 37), reading (Topic 26), virtual travel (Topic 28), walking (Topic 12), backyard activities (Topic16), outdoor activities/cottage stays/swimming pools (Topic 20), camping/caravans/campervans/road trips (Topics 19&34), boating (Topic 25), staying on the beach (Topic 27), Belgians staycationing in tents nested in trees (Topic 30), Hemester (Swedish name for staycation) (Topic 35), “Extreme staycation” in South Korea (Topic 38).

The aggregate dimension on consequences comprised of topics on the consequences of staycations and is made up of eight topics. These include tweets discussing the staycation trend or boom in staycations (Topics 8&10), family connection from staycations (Topic 11), reduction in carbon emissions (Topic 14), support for local businesses (Topic 17), warnings about staycation scams (Topic 18), rise in domestic tourism (Topic 31), and providing employees with an opportunity to rest (Topic 33).

Study 2: Google search trends

The Google search trends provided insights on the search behavior around staycations to answer the first question (RQ1) on how staycations have changed since 2008 in terms of growth and new activities as well as second research question (RQ2) on the features of COVID-19 staycations. The historical search query for staycation from 2004 till the end of 2021, showing the search interest in staycations over time, is presented in Figure 4. The data for Figure 4 is normalized with the period of maximum search interest indexed at 100, while all other data points are relative to this. Therefore, an index value of 50 signifies that the term is half as popular relative to the period of peak popularity. While a value of zero means there was not enough data. The figure shows that the frequency of the worldwide search for staycation took off from zero in 2008 during the financial crisis and has grown steadily year on year from 2011, with regular peaks around the middle of each year or summer in the Northern Hemisphere when many people take vacations. The search interest skyrocketed in the middle of 2020 because of the COVID-19 pandemic to over three times the level in 2019. Surprisingly, the search interest grew further by approximately 10% from the peak in 2020 to an all-time high in July 2021 although there were less travel restrictions in 2021 compared to 2020 and international tourism grew 4% in 2021 from 2020 (UNWTO, 2022).

Figure 4.4: Search interest for staycation from 2004 to 2021 (Data source: (Google Trends, 2022)

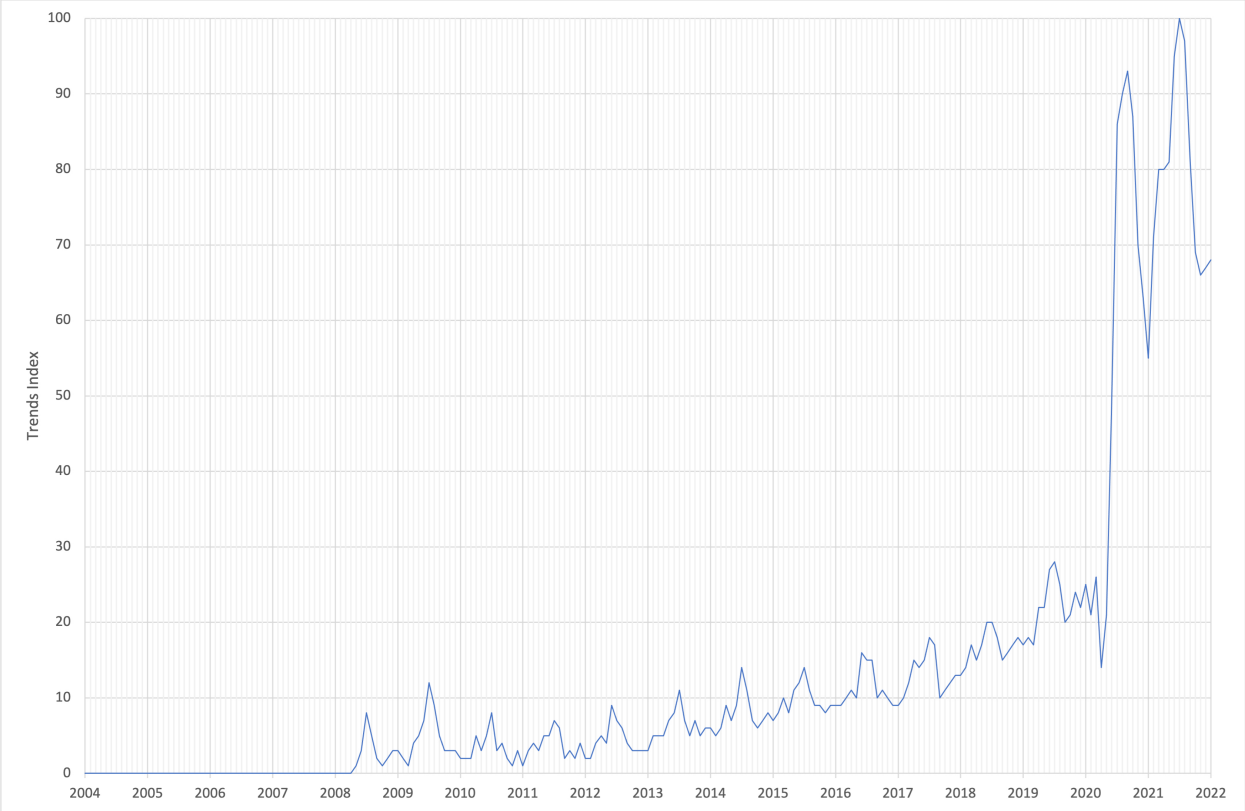


Table 4.1 shows the top 20 locations where searches for staycation were the most popular. The index values do not indicate the absolute query count but the relative proportion of staycation to all other queries in a location. The values are calculated so that the location with 100 is the place where staycation had the most popularity as a percentage of all searches, so a place with a value of 50 means staycation searches were half as popular as the place ranked first. Therefore, a smaller location where staycation-related queries are 50% of all searches in the location gets double the score of a bigger place where staycation-related queries are 25% of all searches. This explains why relatively smaller places like Hong Kong and Singapore top the list. This means that

the larger countries on the list, regardless of a low index score, had a significant amount of staycation-related queries relative to all other search queries. The spread of the locations across different continents shows that searches for staycation were a worldwide trend and not restricted to any region.

Table 4.1: Top locations for staycation searches

	Country	**Index**
1	Hong Kong	100
2	Singapore	86
3	Barbados	77
4	St. Lucia	70
5	Cayman Islands	52
6	Philippines	21
7	United Arab Emirates	14
8	Ireland	6
9	Malaysia	6
10	Qatar	5
11	Indonesia	4
12	France	3
13	United Kingdom	3
14	Kenya	2

15	Canada	2
16	United States	1
17	Australia	1
18	Belgium	1
19	Finland	1
20	Thailand	1

Tables 4.2 – 4.5 show the Top and Rising related queries and topics in 2020 and 2021. The Top related queries and topics (Tables 4.2 and 4.4) are on a relative scale with the most popular query or topic indexed at 100, while all other queries are relative to this. The Rising related queries and topics (Tables 4.3 and 4.5) have seen the biggest increase in search volume during the period and are presented with the percentage increase. Rising queries and topics labeled “Breakout” grew by over 5000% during the period. The dominant languages in these tables are English and Chinese (which is spoken in Hong Kong and Singapore, the top two locations for staycation searches).

Table 4.2: Top related queries in 2020 and 2021

	2020 Top Related queries	**Index**		**2021 Top Related queries**	**Index**
1	hotel	100	1	hotel staycation	100
2	hotel staycation	100	2	酒店 staycation (hotel)	59
3	singapore staycation	95	3	staycation singapore	57

4	staycation 2020	32	4	staycation 2021	41
5	singapore hotel staycation	28	5	staycation 優惠 (discount)	33
6	staycation 酒店 (hotel)	27	6	uk staycation	30
7	staycation deals	27	7	staycation hong kong	29
8	staycation uk	25	8	klook staycation	25
9	staycation ideas	22	9	staycation hotels	25
10	staycation hong kong	20	10	klook	24
11	sentosa	19	11	staycation deals	22
12	sentosa staycation	19	12	staycation 香港 (Hong Kong)	19
13	staycation meaning	17	13	staycation meaning	19
14	staycation 優惠 (discount)	17	14	staycation hk	15
15	香港 staycation (Hong Kong)	14	15	hotel staycation singapore	15
16	staycation in singapore	13	16	staycation ideas	14
17	staycation covid	12	17	stay	13
18	singapore staycation deals	12	18	tagaytay staycation	13
19	staycation 中文 (Chinese)	12	19	staycation near me	11
20	staycation dubai	12	20	rosewood staycation	11
21	staycation singapore sentosa	11	21	staycation package	11
22	staycation ireland	11	22	rosewood	10
23	staycation promotion	11	23	sentosa staycation	10

24	staycation singapore 2020	11	24	staycation jakarta	10
25	airbnb	10	25	airbnb staycation	10

Table 4.3: Rising related queries in 2020 and 2021

	2020 Rising Related queries	**Percentage**		**2021 Rising Related queries**	**Percentage**
1	staycation 酒店 (hotel)	Breakout	1	disney staycation cruise	Breakout
2	staycation 優惠 (discount)	Breakout	2	uk cruise staycation	Breakout
3	香港 staycation (Hong Kong)	Breakout	3	disney cruise uk	Breakout
4	staycation covid	Breakout	4	ontario staycation tax credit	Breakout
5	klook	Breakout	5	disney staycation cruise uk	Breakout
6	klook staycation	Breakout	6	staycation 優惠 8 月 (offer August)	Breakout
7	staycation approved hotels	Breakout	7	klook staycation 香港 (Hong Kong)	Breakout
8	staycation uk 2020	Breakout	8	disney hotel staycation	Breakout
9	stb staycation	Breakout	9	wm hotel	Breakout

10	staycation deals singapore 2020	Breakout	10	staycation ireland 2021	Breakout
11	staycation voucher singapore	Breakout	11	klook staycation hk	Breakout
12	murray staycation	Breakout	12	東 涌 喜來登 酒店 staycation (Sheraton Tung Chung Hotel)	Breakout
13	staycation singapore phase 2	Breakout	13	staycation lonavala	Breakout
14	upper house staycation	Breakout	14	the arca staycation	Breakout
15	staycation wales	Breakout	15	hyperair staycation	Breakout
16	staycation during covid	Breakout	16	staycation in hyderabad	Breakout
17	半島 staycation (peninsula)	Breakout	17	staycation 2021	2100%
18	hotels open for staycation	Breakout	18	柏 寧 酒店 staycation (Park Lane Hotel)	2100%
19	hotels approved for staycation	Breakout	19	大 澳 文物 酒店 staycation (Tai O Heritage Hotel)	2100%
20	staycation 意思 (meaning)	Breakout	20	staycation roma	2100%
21	staycation 10 月 (October)	Breakout	21	staycation semarang	2050%
22	staycation rebate	Breakout	22	staycation approved hotels manila	2000%

23	hilton singapore	Breakout	23	staycation uk 2021	1600%
24	rosewood staycation 優惠 (offer)	Breakout	24	staycation cruises	1400%
25	andaz staycation	Breakout	25	staycation tax credit	1100%

Table 4.4: Top related topics in 2020 and 2021

	2020 Top Related topics	**Index**		**2021 Top Related topics**	**Index**
1	Staycation	100	1	Staycation	100
2	Hotel	23	2	Hotel	22
3	Singapore	15	3	Singapore	8
4	Discounts and allowances	8	4	Discounts and allowances	8
5	Hong Kong	6	5	Hong Kong	6
6	Idea	4	6	Klook	3
7	Resort	4	7	Resort	3
8	Sentosa	3	8	Tagaytay	2
9	Swimming pool	3	9	Swimming pool	2
10	Ireland	2	10	Sentosa	2
11	Ireland	2	11	Vacation	1
12	Airbnb	2	12	Buffet	1

13	Vacation	2	13	Cruise ship	1
14	Dubai	2	14	Airbnb	1
15	Paris	2	15	Spa	1
16	Spa	1	16	Jakarta	1
17	Klook	1	17	Ritz-Carlton Hotel Company	1
18	Jakarta	1	18	Rosewood Hotels & Resorts	1
19	Marina Bay Sands Singapore	1	19	Voucher	1
20	Lodging	1	20	Ireland	1
21	Voucher	1	21	Hong Kong Disneyland	1
22	Rosewood Hotels & Resorts	1	22	Disney	1
23	Shangri-La Hotels and Resorts	1	23	Selangor	1
24	Conrad Hong Kong	1	24	Hong Kong Disneyland Hotel	1

Table 4.5: Rising related topics in 2020 and 2021

	2020 Rising Related topics	**Percentage**		**2021 Rising Related topics**	**Percentage**
1	Voucher	Breakout	1	Disney Cruise Line	Breakout
2	Rosewood Hotels & Resorts	Breakout	2	P&O Cruises	Breakout
3	Shangri-La Hotels and Resorts	Breakout	3	Selangor	1600%
4	Conrad Hong Kong	Breakout	4	Cruise ship	1300%

5	Rosewood Hong Kong	Breakout	5	Hong Kong Disneyland	1050%
6	Mandarin Oriental Hotel Group	Breakout	6	Sai Kung	700%
7	Kerry Hotel, Hong Kong	Breakout	7	香港东涌世茂喜来登酒店 (Sheraton Hong Kong Tung Chung Hotel)	650%
8	Resorts World Sentosa	Breakout	8	Goa	550%
9	Singapore Tourism Board	Breakout	9	Disney	500%
10	香港W酒店 (W Hotel Hong Kong)	Breakout	10	Tung Chung	500%
11	Express train	Breakout	11	The Walt Disney Company	450%
12	W Singapore - Sentosa Cove	Breakout	12	Hyatt Regency Hong Kong, Sha Tin	450%
13	Grand Hyatt Hong Kong	Breakout	13	Welsh language	450%
14	RUN HOTEL	Breakout	14	ALVA HOTEL BY ROYAL	400%
15	The Langham	Breakout	15	Hong Kong Gold Coast Hotel	400%
16	Four Seasons Hotel Hong Kong	Breakout	16	Hong Kong Disneyland Hotel	300%
17	The Upper House	Breakout	17	Sheraton Hotels and Resorts	300%
18	Andaz	Breakout	18	Tax credit	300%
19	Hyatt Centric Victoria Harbour Hong Kong	Breakout	19	Royal Caribbean International	300%
20	Sofitel Singapore Sentosa Resort & Spa	Breakout	20	Batangas	250%

21	Capella Singapore	Breakout	21	The Mira Hong Kong	250%
22	The Fullerton Hotel Singapore	Breakout	22	Klook	250%
23	Hyatt Centric	Breakout	23	K11 ARTUS	200%
24	Andaz Singapore - a concept by Hyatt	Breakout	24	Island Shangri-La, Hong Kong	200%
25	Holiday Cottage	Breakout	25	Buffet	200%

The most significant observation from Tables 4.2 – 4.5 was the strong search association between staycation and hotels. Hotel and staycation were the top two most popular related queries in both years. Hotel was also the most popular related topic in both years (after a subsequent staycation-related topic search). Users also searched for the names of specific hotels and hotel chains in association with staycation as seen in the Rising related lists in Tables 4.3 and 4.5. This shows strong search interest for hotels and other types of hospitality accommodation like resorts and Airbnb to spend staycations. Related searches for staycation deals, packages, discounts, allowances, vouchers, and tax credits show that people were interested in low-cost staycation opportunities. An element of people's search behavior was to add the name of the location, year, or month while searching on the internet for current staycation opportunities. Comparing both years of the pandemic, the Top lists in Tables 4.2 and 4.4 are similar with just the position of the terms changing. However, the Rising lists in Tables 4.3 and 4.5 show more difference as there was a rise in searches for "staycation cruises" i.e. cruises that sail close to home to no particular

or foreign destination before returning. The increase in search frequency pushed cruise ship into the Top related topics in 2021 on Table 4.4 and this search for staycation cruises was most popular in Singapore, UK, Hong Kong, and the US.

Discussion

This study investigated staycations during the COVID-19 pandemic with data from Twitter and Google Trends to provide insights on how staycations have changed since 2008 in terms of growth and new activities and the principal features of staycations during the first two years of the pandemic. The findings indicate that there has been a dramatic increase in internet search interest in staycations worldwide since 2008, including unprecedented growth during the pandemic, answering part of RQ1. Responding to RQ2, the most significant aspect of COVID-19 staycations, as shown by the findings, is a dominant association between staycations and hotels as demonstrated by four of the topics in the topic model and the internet search analysis. This indicates that many people had staycations at hotels during the pandemic based on the Twitter analysis or were interested in doing so based on the Google search analysis. To interpret this key finding using the Construal Level Theory (CLT), a principal benefit of vacation travel is the provision of psychological distance and freedom from everyday routines and bothersome chores (de Bloom et al., 2017). CLT contends that travel can be construed as freedom from routine, representing a high-level construal (Kim et al., 2016). The function of high-level construals is to enable people to mentally transcend the present by forming a representation of the central features of the object and projecting those representations onto distal situations. Since one of

the central features of travel tourism is staying at hotels or other hospitality accommodations, it is easy to build a mental construal of the travel process by staying in a hotel even when spatially close to home. Hence, staycation in a hotel may simulate travel abstractly and help to achieve psychological distance away from home, as everyday routines and chores are avoided while lodging in a hotel, leading to a more positive and satisfactory experience and a higher increase in recreation.

In the same vein, the numerous 5-star hotels and luxury hotel amenities like spa, swimming pool, and buffet in the related search tables which people searched for with staycation indicate a desire for something different by users from what is available at home. All these fit into the mold of people trying to create experiences that can generate high-level construal, as suggested by CLT. Safety measures by the hotels, such as increased cleaning regimes to encourage patronage for staycations, was an important topic in the context of the COVID-19 pandemic. Other studies during the pandemic have shown that these safety measures were of prime importance to attracting customers (Park & Lehto, 2021). The strong connection between staycation and hotels in this study represents a novel finding that was not present in the past literature.

Other key aspects of COVID-19 staycations, as posed in RQ2, include staycation-related Google searches for deals, packages, discounts, allowances, vouchers, and tax credits, including two topics in the topic model show that people were interested in staycation opportunities at a reduced cost. Benefits related to travel like breaking the routine and enjoying free time are facilitated when they are affordable, making it easy to form a mental construal of travel according

to CLT. These offers, deals or discounts can be divided into two; those provided by hotels and those implemented by governments. Hotel managers are encouraged to provide staycation packages for locals, which may compensate for part of the shortfall from international visitors during the ongoing pandemic. After the pandemic, this practice can remain useful, especially during off-season periods for tourism. Search users were also observed to add the current year, month, or place to their search queries. Hence, adding these time and place markers to the staycation package information on the hotel website is recommended as a good Search Engine Optimization (SEO) tactic, so that the hotel package shows up when users search like this.

Staycation offers through vouchers or tax credits can also be deployed by governments to stimulate demand and encourage people to visit local. Even after the pandemic, these tools remain useful to promote local tourism. These staycation vouchers can be a win-win for all parties as reported in the literature, the government, tourism operators, and the locals, while making the people that redeem such staycation vouchers to spend more money locally (Cvelbar et al., 2021). It could also be a tool to address overtourism by attracting local tourists away from the popular tourist destinations in the country while stimulating tourism demand and opening up areas that are not yet so popular. This does not amount to the government paying for people's holidays, but a proven strategic investment to promote local tourism. For example, Slovenia's €350 million staycation voucher scheme was estimated to earn up to 0.47 cents back on every euro spent with estimated multiplicative effects on the economy of up to €0.7 billion (Cvelbar et al., 2021).

Many staycation activities surfaced in the topic model such as relaxation, watching television and movies, visiting the shopping mall, outdoor activities, hiking, road trips, backyard activities, visiting the countryside, eating out at restaurants, visiting local attractions, lounging by swimming pools etc. were similar to the leisure activities reported in previous literature and confirms them (Heimtun, 2017; James et al., 2017; Sharma, 2009). According to CLT, these activities facilitate building a mental image, increase relaxation, and reduce the psychological distance to destinations and are associated with positive attitudes.

In response to the second part of RQ1, new types of activity surfaced in the findings that were not mentioned in the past staycation literature were virtual travel and staycation cruises. Virtual travel involves visiting unknown places through an online immersion process via the information and videos available on the internet. In more sophisticated implementations; this immersion is carried out through Virtual Reality (VR) headsets (Talwar et al., 2022). The COVID-19 pandemic has been tough for the cruise industry (Muritala et al., 2022). However, some parts of the industry opened up in 2021, as vaccination permitted the resumption of sailing in some places, leading to a breakout search for staycation cruises in 2021 for voyages to nowhere, especially in the UK as shown in the related search tables.

There were also activities or concepts that were unique to some locations in the topics, such as people in Belgium spending their staycations in tents constructed in trees. An activity named “extreme staycation” in South Korea, which involved the transformation of the home into a chosen popular destination like Bali, for example. Hemester is the Swedish word for staycation

and it was captured as a topic of its own from tweets about Swedes engaging in staycations. According to CLT, these activities facilitate building a mental image and reduce the psychological distance to destinations and are associated with positive attitudes (Marlow & Dabbish, 2014). These activities are also new findings as they were absent from the past literature on staycations.

Perhaps, the most important topic surfaced under the consequences of staycations is how it can help to reduce the carbon emissions due to tourism. As frequent extreme weather events in different parts of the world continues to highlight the reality of a changing climate (Stott, 2016), the role of tourism, which contributes at least eight percent of global greenhouse gas emissions will be increasingly scrutinized (Bigano et al., 2005; Lenzen et al., 2018). International tourism accounts for a disproportionate contribution to environmental emissions but constitutes only 16% of tourism trips (C.M. Hall et al., 2013). Flight-shaming is already a growing concept and has been found to be more pronounced against people flying for holidays than for work or to visit family and friends (Doran et al., 2021; Gössling et al., 2019). Furthermore, media reports detail government plans in different countries for green taxes and carbon pricing for flights, which may raise airfares (Abnett, 2021; Gatten & Gill, 2021). Therefore, it is not inconceivable for more people to go on fewer holidays abroad in the near future involving long haul travel for leisure for environmental concerns while embracing staycations and domestic tourism in its stead. Periods of crises have usually steered people towards taking holidays closer to home, such as during the Second World War, the global financial crisis or the current pandemic. In the light of the climate change emergency, which increasingly looks like a chronic crisis that will be confronted over

several years, it would not be surprising to have a similar outcome, especially if carbon taxes are integrated into travel costs.

Theoretical and practical implications

This study has made some conceptual and theoretical contributions. First, it has proposed a definition for staycation, which we hope is clearer based on the standing conventions of the UNWTO. This study has also used CLT to explain its findings, demonstrating its applicability and an avenue for further experimental research. To the knowledge of the authors, this study is also the first to analyze related web search queries and topics data in hospitality and tourism research. Many of the past research in this area using Google Trends data has been on forecasting studies. The interesting findings from this data show that this is a useful method for acquiring insight from Google-search activity on a global or local scale. Finally, this study also contributes to the literature on the impact of COVID-19 on hospitality and tourism.

Previous research has demonstrated the value of internet-generated data and UGC for hospitality and tourism businesses (Mendes-Filho et al., 2018). From a practical point of view, this study recommends hotel managers to use staycations as an opportunity to boost revenue, provide staycation deals or coupons, and share the details of such deals online with the location and time information specified. They are also recommended to provide information on the cleaning regime in their hotels as a central feature in light of the pandemic. Many governments have offered staycation vouchers and tax credits so far during this pandemic, and research has showed that they are beneficial to all parties concerned. These programs should therefore be sustained

and exploited in other countries to support and keep hospitality and tourism businesses afloat through the continuing headwinds caused by the pandemic.

Conclusion

This study has proposed a definition for staycation based on the standing convention of the UNWTO and extant literature. The most significant finding about the characteristics of the staycations during the COVID-19 pandemic was the uncovering of a dominant interest to spend staycations at hotels, which was interpreted using CLT as people trying to achieve a psychological distance away from home while being spatially close to home. New staycation-related activities like virtual travel and staycation cruises and specific practices in some locations were revealed in the findings.

At the time of writing in 2022, cross-border travel in many places still remains with COVID-19 related requirements for tests or vaccination. The current high inflation and fuel prices in several OECD countries may also impact on the ability to engage in international travel for tourism by consumers in the near term. Tourism experts have stated that they expect a recovery in international tourism to 2019 levels in 2024 or later (DeMicco et al., 2021; UNWTO, 2022). Research during the pandemic has also reported a desire by people to travel closer to home (ETC, 2022; Lee & Han, 2022). Therefore, while staycations and domestic tourism cannot make up for the revenue shortfall from international tourism, it is essential that in-country travel and staycation schemes be harnessed as much as possible to protect hospitality and tourism operators from bankruptcy before international tourism returns to pre-pandemic levels. In the

longer term, concerns about the environmental impact of tourism and climate change and a shift towards more sustainable forms of tourism may also encourage tourism closer to home. The COVID-19 pandemic has accelerated the growth and acceptance of staycations, and while people will continue to desire travel to distant tourist destinations, staycations are not a fad. Hence, it is important for government policy and businesses to adapt to this new reality.

Limitations and future research

This study has the following limitations. Social media data is noisy while containing useful information and it is possible that some valuable information was removed in the process of data cleaning. Also, the 38 topics surfaced by the LDA topic model were not the only topics present in the data, but the dominant topics, as the algorithm needs to stop at a particular threshold. If not, it is possible to have an extensive but weak model with many more topics. Furthermore, Google does not provide the absolute number of search queries for the Google Trends data, as all the data provided are normalized and relative to the maximum in each category and is therefore not suitable for robust statistical manipulation.

This study has interpreted the findings using CLT, but the method employed cannot directly prove the theory. Future experimental studies that can prove the theory are recommended using the aggregate dimensions of antecedents, attributes, activities, and consequences as research variables. Future studies exploring how to exploit staycations in particular locations to derive more economic benefits in terms of revenues and profits, as well as to produce cultural benefits for citizens increasing their knowledge about their local environment is recommended to expand

the research in this area. Studies into the market segments likely to be interested in staycations as well as surveys about people's staycation preferences are also recommended.

References

Abnett, K., 2021. Draft shows EU to propose aviation fuel tax in green policy push | Reuters [WWW Document]. Reuters. URL https://www.reuters.com/business/sustainable-business/draft-shows-eu-propose-aviation-fuel-tax-green-policy-push-2021-07-04/ (accessed 7.28.21).

Banda, J.M., Tekumalla, R., Wang, G., Yu, J., Liu, T., Ding, Y., Artemova, E., Tutubalina, E., Chowell, G., 2021. A Large-Scale COVID-19 Twitter Chatter Dataset for Open Scientific Research—An International Collaboration. Epidemiologia. https://doi.org/10.3390/epidemiologia2030024

Bigano, A., Goria, A., Hamilton, J.M., Tol, R.S.J., 2005. The Effect of Climate Change and Extreme Weather Events on Tourism. FEEM Work. Pap. No. 30.05, C. Res. Pap. No. 01.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent Dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022. https://doi.org/10.1016/b978-0-12-411519-4.00006-9

Bontempi, E., 2021. The europe second wave of COVID-19 infection and the Italy “strange” situation. Environ. Res. 193. https://doi.org/10.1016/j.envres.2020.110476

Chica, M., Hernández, J.M., Perc, M., 2022. Sustainability in tourism determined by an asymmetric game with mobility. J. Clean. Prod. 355. https://doi.org/10.1016/j.jclepro.2022.131662

Collins-Kreiner, N., Ram, Y., 2020. National tourism strategies during the Covid-19 pandemic.

Ann. Tour. Res. https://doi.org/10.1016/j.annals.2020.103076

Cvelbar, L.K., Farčnik, D., Ogorevc, M., 2021. Holidays for all: Staycation vouchers during covid-19. Ann. Tour. Res. Empir. Insights 2. https://doi.org/10.1016/j.annale.2021.100019

D'Avanzo, E., Pilato, G., Lytras, M., 2017. Using Twitter sentiment and emotions analysis of Google Trends for decisions making. Program 51, 322–350. https://doi.org/10.1108/PROG-02-2016-0015

de Bloom, J., Kompier, M., Geurts, S., de Weerth, C., Taris, T., Sonnentag, S., 2009. Do we recover from vacation? Meta-analysis of vacation effects on health and well-being. J. Occup. Health 51, 13–25. https://doi.org/10.1539/joh.K8004

de Bloom, J., Nawijn, J., Geurts, S., Kinnunen, U., Korpela, K., 2017. Holiday travel, staycations, and subjective well-being. J. Sustain. Tour. 25, 573–588. https://doi.org/10.1080/09669582.2016.1229323

DeMicco, F., Cetron, M., Davies, O., Guzman, J., 2021. COVID-19 Impact on the Future of Hospitality and Travel. J. Hosp. Tour. Res. 45, 911–914. https://doi.org/10.1177/10963480211000822

Dodds, R., 2012. Questioning Slow As Sustainable. Tour. Recreat. Res. 37, 81–83. https://doi.org/10.1080/02508281.2012.11081691

Dolnicar, S., Yanamandram, V., Cliff, K., 2012. The contribution of vacations to quality of life. Ann. Tour. Res. 39, 59–83. https://doi.org/10.1016/j.annals.2011.04.015

Doran, R., Pallesen, S., Böhm, G., Ogunbode, C.A., 2021. When and why do people experience flight shame? Ann. Tour. Res. 103254. https://doi.org/https://doi.org/10.1016/j.annals.2021.103254

ETC, 2022. Sustainable Travel in an Era of Disruption: Impact of COVID-19 on Sustainable Tourism Attitudes [WWW Document]. URL https://etc-corporate.org/reports/sustainable-travel-in-an-era-of-disruption-impact-of-covid-19-on-sustainable-tourism-attitudes/ (accessed 4.7.22).

Evening Standard, 2020. Boris Johnson "going to Scotland for two-week staycation" [WWW Document]. URL https://www.standard.co.uk/news/uk/johnson-scotland-two-week-staycation-a4520176.html (accessed 6.24.21).

Fox, S., 2009. Vacation or Staycation? Neumann Bus. Rev. 1–7.

Frew, E., Winter, C., 2010. Tourist Response to Climate Change: Regional and Metropolitan Diversity. Tour. Rev. Int.

Gatten, E., Gill, O., 2021. Cars and flights to be hit with green taxes [WWW Document]. Dly. Telegr. URL https://www.telegraph.co.uk/environment/2021/07/14/cars-flights-hit-green-taxes/ (accessed 7.28.21).

Gilbert, D., Abdullah, J., 2004. Holidaytaking and the sense of well-being. Ann. Tour. Res. 31, 103–121. https://doi.org/10.1016/j.annals.2003.06.001

Google Trends, 2022. Google Trends [WWW Document]. URL https://www.google.com/trends

(accessed 3.5.21).

Gössling, S., Hall, C.M., Peeters, P., Scott, D., 2010. The Future of Tourism: Can Tourism Growth and Climate Policy be Reconciled? A Mitigation Perspective. Tour. Recreat. Res. 35, 119–130. https://doi.org/10.1080/02508281.2010.11081628

Gössling, S., Hanna, P., Higham, J., Cohen, S., Hopkins, D., 2019. Can we fly less? Evaluating the 'necessity' of air travel. J. Air Transp. Manag. 81, 101722. https://doi.org/https://doi.org/10.1016/j.jairtraman.2019.101722

Gössling, S., Scott, D., Hall, C.M., 2020. Pandemics, tourism and global change: a rapid assessment of COVID-19. J. Sustain. Tour. 1–20. https://doi.org/10.1080/09669582.2020.1758708

Hall, C.M., Scott, D., Gössling, S., 2020. Pandemics, transformations and tourism: be careful what you wish for. Tour. Geogr. 22, 577–598. https://doi.org/10.1080/14616688.2020.1759131

Hall, C.M., Scott, D., Gössling, S., 2013. The Primacy of Climate Change for Sustainable International Tourism. Sustain. Dev. 21, 112–121. https://doi.org/10.1002/sd.1562

Hateftabar, F., 2021. The impact of psychological distance on tourists' length of stay: Survival analysis. J. Hosp. Tour. Manag. 46, 1–11. https://doi.org/10.1016/j.jhtm.2020.11.006

Haukeland, J.V., 1990. Non-travelers. The flip side of motivation. Ann. Tour. Res. 17, 172–184. https://doi.org/10.1016/0160-7383(90)90082-3

Hay, B., 2010. What's in a Name: A Review of Popular New Words to Describe Holidays - a Clever Marketing Ploy or a Pointless Waste of Time?, in: In: CAUTHE 2010: Tourism and Hospitality: Challenge the Limits. Hobart, Tas.: University of Tasmania. School of Management, 2010: [1776]-[1778].

Heimtun, B., 2017. Home holidays as real holidays? Midlife single women's experiences, Women and Travel: Historical and Contemporary Perspectives. https://doi.org/10.1201/9781315365879

Hogue, J., DeWilde, B., 2020. Pytrends · PyPI [WWW Document]. URL https://pypi.org/project/pytrends/ (accessed 12.21.20).

James, A., Ravichandran, S., Chuang, N.-K., Bolden, E., 2017. Using Lifestyle Analysis to Develop Lodging Packages for Staycation Travelers: An Exploratory Study. J. Qual. Assur. Hosp. Tour. 18, 387–415. https://doi.org/10.1080/1528008X.2016.1250240

Jean-Baptiste, M., Kui, S.Y., Presser, A.A., Adrian, V., K., G.M., null, null, P., P.J., Dale, H., Dan, C., Peter, N., Jon, O., Steven, P., A., N.M., Lieberman, A.E., 2011. Quantitative Analysis of Culture Using Millions of Digitized Books. Science (80-.). 331, 176–182. https://doi.org/10.1126/science.1199644

Jia, Y., Ouyang, J., Guo, Q., 2021. When rich pictorial information backfires: The interactive effects of pictures and psychological distance on evaluations of tourism products. Tour. Manag. 85. https://doi.org/10.1016/j.tourman.2021.104315

Kim, E.L., Tanford, S., 2021. The windfall gain effect: Using a surprise discount to stimulate add-on purchases. Int. J. Hosp. Manag. 95, 102918. https://doi.org/10.1016/J.IJHM.2021.102918

Kim, J., Kim, P.B., Kim, J.-E., 2018. Impacts of temporal and gender difference on hotel selection process. J. Hosp. Mark. Manag. 27, 711–732. https://doi.org/10.1080/19368623.2018.1438325

Kim, J., Kim, P.B., Kim, J.-E., Magnini, V.P., 2016. Application of Construal-Level Theory to Promotional Strategies in the Hotel Industry. J. Travel Res. 55, 340–352. https://doi.org/10.1177/0047287514550097

Kluyver, T., Ragan-kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., Willing, C., Development Team, J., 2016. Jupyter Notebooks—a publishing format for reproducible computational workflows. Position. Power Acad. Publ. Play. Agents Agendas 87–90. https://doi.org/10.3233/978-1-61499-649-1-87

Lee, S., Han, H., 2022. A new tourism paradigm and changes in domestic tourism for married Koreans in their 30s and 40s. Curr. Issues Tour. 1–13. https://doi.org/10.1080/13683500.2022.2054402

Lenzen, M., Sun, Y.-Y., Faturay, F., Ting, Y.-P., Geschke, A., Malik, A., 2018. The carbon footprint of global tourism. Nat. Clim. Chang. 8, 522–528. https://doi.org/10.1038/s41558-018-0141-x

Lindblom, J., Vogt, C., Andereck, K., 2020. Construal Level Theory as a Framework for Navigating Community Contexts in Tourism Planning. Tour. Plan. Dev. https://doi.org/10.1080/21568316.2020.1855238

Lohmann, M., 1996. You'll better stay at home? Studies on the recreational effects of holidays and holiday tourism. Tour. Rev. 51, 39–44. https://doi.org/10.1108/eb058228

Marlow, J., Dabbish, L., 2014. When is a picture not worth a thousand words? the psychological effects of mediated exposure to a remote location. Comput. Human Behav. https://doi.org/10.1016/j.chb.2013.08.002

McCabe, S., 2009a. Who needs a holiday? Evaluating social tourism. Ann. Tour. Res. 36, 667–688. https://doi.org/10.1016/j.annals.2009.06.005

McCabe, S., 2009b. The tourist experience and everyday life., in: The Tourist as a Metaphor of the Social World. pp. 61–75. https://doi.org/10.1079/9780851996066.0061

McCallum, A.K., 2002. MALLET: A Machine Learning for Language Toolkit.

McKinney, W., 2010. Data Structures for Statistical Computing in Python. Proc. 9th Python Sci. Conf. 1697900, 51–56.

Mendes-Filho, L., Mills, A.M., Tan, F.B., Milne, S., 2018. Empowering the traveler: an examination of the impact of user-generated content on travel planning. J. Travel Tour. Mark. 35, 425–436. https://doi.org/10.1080/10548408.2017.1358237

Merriam-Webster, n.d. Staycation [WWW Document]. URL https://www.merriam-

webster.com/words-at-play/staycation-date-meaning (accessed 7.1.21).

Molz, J.G., 2009. Representing pace in tourism mobilities: Staycations, slow travel and the amazing race. J. Tour. Cult. Chang. 7, 270–286. https://doi.org/10.1080/14766820903464242

Muritala, B.A., Hernández-Lara, A.-B., Sánchez-Rebull, M.-V., Perera-Lluna, A., 2022. #CoronavirusCruise: Impact and implications of the COVID-19 outbreaks on the perception of cruise tourism. Tour. Manag. Perspect. 100948. https://doi.org/https://doi.org/10.1016/j.tmp.2022.100948

Muritala, B.A., Sánchez-Rebull, M.-V., Hernández-Lara, A.-B., 2020. A Bibliometric Analysis of Online Reviews Research in Tourism and Hospitality. Sustain. 2020, Vol. 12, Page 9977 12, 9977. https://doi.org/10.3390/SU12239977

Papatheodorou, A., Rosselló, J., Xiao, H., 2010. Global economic crisis and tourism: Consequences and perspectives. J. Travel Res. 49, 39–45. https://doi.org/10.1177/0047287509355327

Park, S., Lehto, X., 2021. Understanding the opaque priority of safety measures and hotel customer choices after the COVID-19 pandemic: an application of discrete choice analysis. J. Travel Tour. Mark. 38, 653–665. https://doi.org/10.1080/10548408.2021.1985038

Pichierri, M., Petruzzellis, L., Passaro, P., 2022. Investigating staycation intention: the influence of risk aversion, community attachment and perceived control during the pandemic. Curr.

Issues Tour. 1–7. https://doi.org/10.1080/13683500.2022.2069551

Pleplé, Q., 2013. Topic Coherence To Evaluate Topic Models.

Ren, X., 2020. Pandemic and lockdown: a territorial approach to COVID-19 in China, Italy and the United States. Eurasian Geogr. Econ. 61, 423–434. https://doi.org/10.1080/15387216.2020.1762103

Rice, W.L., Reigner, N., Freeman, S., Newman, P., Mateer, T.J., Lawhon, B., Taff, B.D., 2022. The impact of protective masks on outdoor recreation crowding norms during a pandemic. J. Leis. Res. 53, 340–356. https://doi.org/10.1080/00222216.2021.1981791

Rogers, S., 2020. What is Google Trends data — and what does it mean? [WWW Document]. Google News Lab. URL https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8 (accessed 12.21.20).

Sharma, S., 2009. The Great American Staycation and the Risk of Stillness. M/C J. 12. https://doi.org/10.5204/mcj.122

Smeral, E., 2010. Impacts of the world recession and economic crisis on tourism: Forecasts and potential risks. J. Travel Res. 49, 31–38. https://doi.org/10.1177/0047287509353192

Stott, P., 2016. How climate change affects extreme weather events. Science (80-.). 352, 1517 LP – 1518. https://doi.org/10.1126/science.aaf7271

Summers, E., Brigadir, I., Kemenade, H. van, Hames, S., Binkley, P., tinafigueroa, Ruest, N., Walmir, Chudnov, D., recrm, celeste, Chosak, A., McCain, R.M., Milligan, I., Segerberg, A.,

Shahrokhian, D., Walsh, M., Lausen, L., Woodward, N., Münch, F.V., haruna, Ramaswami, A., Hereñú, D., Milajevs, D., Elwert, F., Westerling, K., rongpenl, Costa, S., Shawn, Kerchner, D., 2021. DocNow/twarc. https://doi.org/10.5281/ZENODO.593575

Syed, S., Spruit, M., 2017. Full-Text or abstract? Examining topic coherence scores using latent dirichlet allocation, in: Proceedings - 2017 International Conference on Data Science and Advanced Analytics, DSAA 2017. pp. 165–174. https://doi.org/10.1109/DSAA.2017.61

Talwar, S., Kaur, P., Nunkoo, R., Dhir, A., 2022. Digitalization and sustainability: virtual reality tourism in a post pandemic world. J. Sustain. Tour. https://doi.org/10.1080/09669582.2022.2029870

The Cumberland, 2020. Boost for hospitality sector as 83% of Brits opt for UK staycation [WWW Document]. URL https://www.cumberland.co.uk/news/staycation-survey (accessed 6.22.21).

Trope, Y., Liberman, N., 2010. Construal-Level Theory of Psychological Distance. Psychol. Rev. 117, 440–463. https://doi.org/10.1037/a0018963

Trope, Y., Liberman, N., Wakslak, C., 2007. Construal levels and psychological distance: Effects on representation, prediction, evaluation, and behavior. J. Consum. Psychol. https://doi.org/10.1016/S1057-7408(07)70013-X

UNWTO, 2022. Tourism Grows 4% in 2021 but Remains Far Below Pre-Pandemic Levels [WWW Document]. URL https://www.unwto.org/news/tourism-grows-4-in-2021-but-remains-far-

below-pre-pandemic-levels (accessed 1.20.22).

UNWTO, 2020. Impact assessment of the COVID-19 outbreak on international tourism [WWW Document]. URL https://www.unwto.org/impact-assessment-of-the-covid-19-outbreak-on-international-tourism (accessed 6.14.21).

UNWTO, 2010. International Recommendations for Tourism Statistics 2008. https://doi.org/10.18356/05265168-en

UNWTO, 2009. Tourism Highlights 2009 Edition [WWW Document]. URL https://www.e-unwto.org/doi/pdf/10.18111/9789284413591 (accessed 6.9.21).

Urry, J., 1988. Cultural Change and Contemporary Holiday-Making. Theory, Cult. Soc. 5, 35–55. https://doi.org/10.1177/026327688005001003

Van Rossum, G., Drake, F.L., 2009. Python 3 Reference Manual. Scotts Valley, CA.

Wang, S., Lehto, X., 2020. The Interplay of Travelers' Psychological Distance, Language Abstraction, and Message Appeal Type in Social Media Advertising. J. Travel Res. 59, 1430–1446. https://doi.org/10.1177/0047287519880000

West, J., 2018. Neologisms: Semiotic deconstruction of the new words "lizardy," "staycation," and "wannarexia" as peircean indexes of culture, Empirical Research on Semiotics and Visual Rhetoric. https://doi.org/10.4018/978-1-5225-5622-0.ch009

Williams, C.C., 2021. Impacts of the coronavirus pandemic on Europe's tourism industry: Addressing tourism enterprises and workers in the undeclared economy. Int. J. Tour. Res.

23, 79–88. https://doi.org/10.1002/jtr.2395

Wong, I.A., Lin, Z., Kou, I.E., 2021. Restoring hope and optimism through staycation programs: an application of psychological capital theory. J. Sustain. Tour. https://doi.org/10.1080/09669582.2021.1970172

Xu, S., Sun, K., Yang, B., Zhao, L., Wang, B., Zhao, W., Wang, Z., Su, M., 2021. Can public participation in haze governance be guided by government？ –Evidence from large-scale social media content data mining. J. Clean. Prod. 318. https://doi.org/10.1016/j.jclepro.2021.128401

Zhang, J., Zhang, Y., 2020. Assessing the low-carbon tourism in the tourism-based urban destinations. J. Clean. Prod. 276. https://doi.org/10.1016/j.jclepro.2020.124303

Zheng, D., Luo, Q., Ritchie, B.W., 2021. Afraid to travel after COVID-19? Self-protection, coping and resilience against pandemic 'travel fear.' Tour. Manag. 83. https://doi.org/10.1016/j.tourman.2020.104261

Appendix

Topic	Topic Keywords	Count	Portion
1	case, rise, mask, number, today, high, face, add, buy, daily	266	0.0438
2	covid, time, staycation, spread, reopen, show, car, state, read, update	143	0.0236
3	stay, home, safe, close, child, healthy, pay, national, easter, require	246	0.0405
4	good, open, follow, rule, make, measure, remember, social distancing, normal, reason	208	0.0343
5	week, covid, today, end, live, miss, weather, folk, begin, group	191	0.0315
6	free, offer, test, resident, share, tour, walk, week, promote, online	183	0.0302
7	year, due, coronavirus, trip, pandemic, cancel, cover, sound, move, insurance	179	0.0295
8	staycation, boom, hope, set, job, important, hand, lose, struggle, scottish	184	0.0303
9	holiday, people, country, abroad, year, break, lot, foreign, happen, safely	191	0.0315
10	covid, staycation, city, increase, bad, visitor, demand, trend, top, opportunity	155	0.0255
11	family, thing, love, quarantine, day, friend, food, activity, isolation, stop	190	0.0313
12	staycation, walk, impact, story, hear, border, future, choice, accord, research	134	0.0221
13	staycation, day, covid, feel, weekend, relax, find, music, late, lock	117	0.0193
14	people, work, back, school, pay, fly, real, house, climate, emission	186	0.0307
15	staycation, month, life, bit, save, play, work, early, live, finally	153	0.0252
16	plan, summer, idea, cancel, list, fun, put, backyard, kid, turn	182	0.03
17	tourism, local, business, support, economy, boost, hospitality, area, encourage, voucher	228	0.0376
18	holiday, book, include, booking, surge, price, consumer, warn, scam, pass	133	0.0219
19	staycation, summer, check, tip, late, safely, campsite, prepare, make sure, i'll	144	0.0237
20	staycation, covid, staycationer, news, cottage, pool, island, outdoor, county, response	152	0.0251
21	covid, staycation, pandemic, health, mind, lot, park, community, join, sense	117	0.0193
22	book, safe, guest, safety, place, guideline, experience, team, morning, staff	179	0.0295
23	hotel, staycation, give, people, advice, covid, public, popular, protect, encourage	138	0.0227
24	staycation, hotel, room, deal, service, package, back, offer, rate, traveller	138	0.0227
25	staycation, covid, sale, happy, hour, winter, site, boat, person, cost	91	0.015
26	staycation, great, enjoy, read, full, article, idea, wonderful, interesting, property	139	0.0229
27	staycation, tourist, british, beach, bring, destination, resort, decide, risk, town	148	0.0244
28	visit, world, beautiful, lockdown, explore, virtual, view, tomorrow, amazing, part	165	0.0272
29	lockdown, virus, risk, travel, big, avoid, fly, low, level, year	149	0.0246
30	covid, staycation, post, tree, prompt, video, accommodation, stick, activity, ready	101	0.0166

31	covid, pandemic, tourism, domestic, staycation, report, market, recovery, option, global	126	0.0208
32	travel, restriction, change, flight, ease, international, lift, luxury, expert, cheap	185	0.0305
33	make, time, staycation, find, create, chance, challenge, stress, employee, water	142	0.0234
34	staycation, start, order, season, talk, traffic, head, listen, datum, caravan	99	0.0163
35	staycation, long, hit, hard, wait, short, return, hot, hemester, focus	166	0.0274
36	covid, staycation, distance, staysafe, holiday, watch, ago, design, nature, wearamask	139	0.0229
37	staycation, covid, spend, movie, word, night, head, money, recommend, guy	125	0.0206
38	staycation, travel, pandemic, home, leave, plan, perfect, spot, korea, opt	155	0.0255

CHAPTER 5. CONCLUSIONS

Summary of main results and conclusions

The aim of this doctoral thesis was to contribute originally to the body of knowledge in the research area on electronic word-of-mouth (eWOM) and customer opinions in hospitality and tourism. To conclude this thesis, this final chapter presents the main conclusions from the studies carried out in each chapter.

The first publication presented in the second chapter of the thesis made a critical evaluation of the previous literature in online reviews in hospitality and tourism with a bibliometric analysis and presented the current state-of-the research in this area. The study uses a rigorous methodological approach, starting with a comprehensive keyword search and systematic refining process, resulting in 1440 journal articles in the English language. Subsequently, the title and abstracts of each of these articles were read carefully to select 632 relevant publications. The bibliographic data of the relevant publications was obtained from the Scopus database, followed by co-citation and co-word analysis. The breakdown of the general literature trends showed that the literature in this area was fast-growing, especially in recent years and that the studies based on electronic word-of-mouth (eWOM) and online reviews were being published in the top hospitality and tourism journals, which answered the first research question posed by this study i.e. “What are the general trends in this research area?” The co-citation analysis provided the answer to the second research question on finding the foundational literature of the research area and revealed the foundational publications in this study area with the top 20 most co-cited studies, which have also been well-cited. These 20 studies were divided into three clusters, with

the clusters showing an alignment with the time of publication. The first cluster containing the earliest foundational literature published between 2003 and 2010 dealt with the effect of electronic word-of-mouth (eWOM) on sales and were written from outside the hospitality and tourism domain from adjacent areas on online reviews in fields such as marketing, management, and information systems. The next cluster comprising papers published between 2008 and 2011 were the most co-cited papers out of the 20 and marked the period when hospitality and tourism scholars started research into eWOM based on online reviews from hotels, restaurants, and tourism services. The third cluster, with relatively newer publications between 2013 and 2016, addressed the motivations, values, and usefulness of online reviews.

The co-word analysis answered the third and fourth research questions, namely: “What are the major research themes and how the research interests have evolved in the research area?” The co-word analysis revealed the conceptual structure and research themes underlying the publications, showing the dimensions of their network structure and the thematic evolution of the research area. The most frequent keywords were online reviews, social media, eWOM, UGC, and online ratings. The central domains in the network structure were based around hotels and hospitality, tourism, restaurants, customer satisfaction, analytical techniques, Airbnb, sustainability, and tourist experience. The evolution of the research area showed the hot and emerging themes and topics, subtle changes in terminology over time, and the progression in the type of data analytical techniques employed in the research area. The conclusion of the study was a verdict on the maturity of the research area on online reviews and eWOM in hospitality and tourism with a rapidly increasing volume of publications built upon a reliable base of seminal

research, applied to highly varied areas, using increasingly sophisticated analytical methods such as artificial neural networks and deep learning. Therefore, this study contributes importantly to scholars in this field to understand the current state of knowledge amidst the overwhelming growth in the literature, providing a glimpse of future research growth areas. It also provides managers with a window to understanding the concept of online reviews, which is now commonplace and which they confront every day in the tourism and hospitality business.

The COVID-19 pandemic is a significant macro-phenomenon which has affected all parts of the hospitality and tourism industry and consequently provided an important research opportunity which inspired the next two academic outputs to provide novel and meaningful contributions to knowledge by helping to understand the impacts of the pandemic and employing electronic word-of-mouth (eWOM) from social media as recommended in the first publication (bibliometric literature review). Therefore, the second publication in the third chapter of the thesis made a crucial examination of the effects of the outbreaks of COVID-19 during several cruise voyages on the perception of cruising and cruise tourism. Cruising was probably the most affected area of tourism and hospitality during the early part of the pandemic, with the COVID-19 outbreaks on the *Diamond Princess* and other ships generating tremendous media interest. Hence, it was important to investigate the impact of this media coverage and derive insight from the public conversation on Twitter about the perception of cruise tourism, which was designated as the research question of this study. The study used electronic word-of-mouth from Twitter and analyzed 139,054 collected tweets about cruising in English using Natural Language Processing (NLP) techniques. The main analytical method was sentiment analysis, which automatically

classified the tweets into positive, negative, and neutral groups. The results answered the research question and found a prevalent negative sentiment in most of the analyzed tweets. More importantly, it was observed that criticisms directed at the cruise industry in the tweets were often based on perceptions and stereotypes about the industry before the pandemic on issues such as frequent outbreaks of infectious diseases like the Norovirus, the environmental impact of cruising, sailing under flags of convenience to avoid tax and laws, increasing size of cruise ships, etc. which were all connected to the mass-market cruise business model. It was also feared that this negative sentiment could persist long afterwards based on the postulation of the social amplification of risk theory (SARF). The conclusion of the study was a declaration that the cruise tourism industry needed to make an effort to become more sustainable and environmentally friendly with a profitable business model different from the main pre-pandemic mass-market model. Furthermore, the preservation of some of the health protocols introduced during the pandemic to help prevent the outbreaks of other infectious diseases on cruises after the pandemic was recommended. These findings are expected to make an important contribution to the cruise industry, as it reopens after the worst periods of the pandemic.

The final study in the fourth chapter of the thesis investigated the rise of staycations as a result of the COVID-19 pandemic forcing many people to spend their vacations at or close to home. Similar to the previous study, this study also used electronic word-of-mouth from Twitter. It analyzed 7,729 unique tweets about staycation using NLP techniques. The main analytical method was Latent Dirichlet Allocation (LDA) topic modeling, which automatically clusters the tweets into its main topics. In addition, internet search data from Google Trends analytics

provided insights and context on the historical search behavior on staycations and the related search queries and topics. The results answered the research questions about how staycations have changed since 2008, as well as the notable aspects of staycations during the first two years of the COVID-19 pandemic. The key findings provide useful insights and contributions to hotel managers and policymakers and found out that the search interest on staycations had been increasing since it first went mainstream after the 2008 financial crisis. The findings also showed a strong interest in spending staycations at hotels rather than at home. The discussion of the results highlighted how to harness staycations to boost revenue and promote sustainable tourism to places close to home.

This thesis was presented as a compendium of three articles, with two published and one currently under review. The connecting thread for the studies in this compendium is a better understanding of the knowledge area of electronic word-of-mouth (eWOM) and customer opinions in hospitality and tourism from the review of the literature. This is followed by the use of electronic word-of-mouth (eWOM) from social media as the principal source of data to understand some impacts of the COVID-19 pandemic on the hospitality and tourism industry. This doctoral thesis fills a gap in the literature to outline the current state of the literature on online reviews in hospitality and tourism. Furthermore, in the context of the COVID-19 pandemic, it fills many gaps in the literature by helping to understand the perception of one of the most affected segments of tourism by the pandemic, as well as the re-emergence and accelerated growth of staycations because of the pandemic. It has contributed to the knowledge about online

reviews, electronic word-of-mouth (eWOM) in hospitality and tourism, as demonstrated by the peer-reviewed publications that were achieved from the studies.

Contribution and novelty of the thesis

Academic and theoretical contributions

This thesis has made some important academic and theoretical contributions. The first study was able to show the conceptual structure and thematic evolution of online reviews in tourism and hospitality, an individual or researcher new to this research area can quickly get a bird's-eye view of the seminal publications and how the research themes have evolved to date. These findings are useful for scholars in this field to understand the current state of knowledge amidst the overwhelming growth in the literature. Another contribution for the academia is the thematic evolution chart made with a new methodology showing the hot and emergent topics. They offer a glimpse of future research growth areas toward the hot and emergent topics and can be used in future bibliometric studies in any research area.

The theoretical contribution of the second study is both the proposed conceptual model and the revised conceptual model, which combined the social amplification of risk framework (SARF) and the information integration theory (IIT). This conceptual framework can be directly adopted or adapted in future studies on risk perception, the influence of integrating new information on people's actual beliefs, and the influence of media coverage on public opinion.

To further the conceptual background and the theory, the third study has proposed a definition for staycation based on the conventions of the UNWTO. This study has also used construal level

theory (CLT) to explain its findings, demonstrating its applicability and an avenue for further experimental research.

Practical contributions

This study has made the following practical contributions. The first study has provided a quick and practical introduction for general readers, practitioners and business owners interested in learning about online reviews and offers a window for understanding an ubiquitous concept in the tourism and hospitality industry.

The second study provided insight into the negative public perception of cruising generated because of the COVID-19 outbreaks on cruise ships, revealing how the sentiment was connected in part to the perception of the cruise industry before the pandemic. The study also provided important recommendations on the rebuilding process of the cruise industry as it recovers from the effects of the pandemic by seeking alternative business models, doubling down on green credentials, strict observance of environmental regulations, emphasis on cleaner and greener ships for new ship orders with more intimate experiences onboard, and attraction of younger new-to-cruise passengers. Cruise lines were also encouraged to keep cruising safe from the routine outbreaks of other infectious diseases by preserving some of the COVID-19 health protocols like frequent handwashing, some measure of social distancing, increased ventilation, ultraviolet air filtration, additional medical facilities onboard, contactless apps for food-ordering, etc.

The third study revealed a hitherto unknown strong connection to spend staycations at hotels, and recommends that hotel managers use staycations as an opportunity to boost revenue, provide staycation deals or coupons, and share the details of such deals online with the location and time information specified. They are also recommended to provide information on the cleaning regime in their hotels as a central feature considering the elevation of the importance of cleaning because of the pandemic. Governments and policy-makers are also advised to maintain or introduce staycation vouchers and tax credits as practiced during the pandemic, as research has showed the enormous benefits of such programs to all parties; hospitality and tourism businesses, tourist areas, the government, and local tourists.

Limitations

While many of the limitations of this work have been mentioned in each chapter, it is important to reiterate the most salient ones, in addition to a few new limitations that have been identified. The identification and use of only journal articles from the Scopus database for the bibliometric literature analysis is a limitation that could have resulted in leaving out some relevant publications indexed by Web of Science, but not in Scopus or other types of academic literature apart from journal articles. While the bibliometric literature review was published in 2020, the study was carried out in 2019 before the COVID-19 pandemic. Hence, there could have been a few sharp changes in the literature in the research area because of the pandemic in the ensuing time.

In the studies investigating COVID-19 and its impact on cruise tourism and on staycations, the analysis was restricted to only tweets in English. While there was a reflection of views from different parts of the world in the analyzed tweets, an analysis of tweets in all languages containing the selected keywords could have provided a more global and comprehensive picture on the topics at hand. In addition, the process of removing the unwanted aspects and noise in the social media data through data cleaning could have eliminated some useful and valuable data accidentally.

Suggestions for future research

Electronic word-of-mouth (eWOM), user generated content (UGC), and customer opinions on social media and online review websites in the hospitality and tourism industry continue to grow at an exponential rate. As a result, there are several useful lines of future research anticipated in this research area. Social media data possesses some advantages over traditional survey methods for evaluating public or customer opinions because surveys treat these opinions as a quantitative distribution by individuals with equal weight in the society. eWOM and UGC from social media reflect the hierarchical and conversational nature of public opinion and captures subjects' attitudes through the observation of actual behavior, rather than the limited or focused questions of a survey. Consequently, analysis of social media conversations can provide insight that can be used with other traditional research methods such as surveys to provide a holistic view of the public or customer opinions in hospitality and tourism, and a wide variety of interesting research questions can be crafted, including on real-time events such as competitions

or festivals. In addition, analysis of eWOM and UGC data from new social media sites such as TikTok, which is currently the fastest growing social media platform regarding the hospitality and tourism industry, is still an under-researched area in the literature because of the relative novelty of the platform.

The bibliometric survey of the literature in the first publication found the application of eWOM and UGC in increasingly niche areas of tourism, such as wine tourism, heritage tourism, cruise tourism, dark tourism, etc. Hence, this was part of the motivation to embark on the study to investigate the perception of cruise tourism when the COVID-19 outbreaks on cruise ships started grabbing the headlines in the early part of the pandemic. Therefore, future lines of research into similar niche areas of tourism can also benefit from monitoring the online conversations on social media or online reviews when relevant events happen. For example, there was an increased conversation on social media about space tourism after two billionaires, Richard Branson and Jeff Bezos, both went on space flights in 2021. Hence, studies looking into these types of salient topics can be done with eWOM data. Finally, the use of the latest development in natural language algorithms, such as foundation artificial intelligence language models, may enable more accurate text analysis in the future.

This doctoral thesis opens an interesting path in the literature, and its methods and approaches could be useful to explore both mainstream and niche areas of the hospitality and tourism industry in order to develop a more holistic understanding of the industry with the help of information provided through eWOM and UGC sources.

APPENDICES

Appendix 1: A Bibliometric Analysis of Online Reviews Research in Tourism and Hospitality

sustainability

MDPI

Review

A Bibliometric Analysis of Online Reviews Research in Tourism and Hospitality

Babajide Abubakr Muritala *, Maria-Victoria Sánchez-Rebull and Ana-Beatriz Hernández-Lara

Departament de Gestió d'Empreses, Facultat d'Economia i Empresa, Universitat Rovira i Virgili, 43204 Reus, Spain; mariavictoria.sanchez@urv.cat (M.-V.S.-R.); anabeatriz.hernandez@urv.cat (A.-B.H.-L.)
* Correspondence: babajide.muritala@urv.cat

Received: 5 November 2020; Accepted: 26 November 2020; Published: 29 November 2020

Abstract: This paper reviews the literature on online reviews in tourism and hospitality, and presents the current state of research in the area. A bibliometric approach was used to analyze 632 journal articles on online reviews in tourism and hospitality from 2005 to 2019 from the Scopus Database. This study identifies the most prolific journals, foundational works, and major research themes in the research area. In addition, we analyzed some dimensions of their network structure and the thematic evolution of the research area. The bibliometric method is quantitative and objective, and we carry out an analysis of the area based on citations and keywords. Researchers and business managers can gain useful insights on the current state of the art in this area. There have been only a few literature reviews tracking the growth in this research area, and even fewer using bibliometric methods or science maps. Therefore, this work provides an updated review of this fast-growing area with a bibliometric approach to highlight the recent developments with the aid of science maps, and shows the thematic network structure and evolution with an innovative visualization.

Keywords: online reviews; literature review; VOSviewer; Ucinet; citation analysis; thematic evolution; bibliometric analysis; co-word analysis; conceptual structure; science maps

1. Introduction

Online reviews, also referred to as electronic word of mouth (eWOM) or user-generated content (UGC) [1,2], are all similar concepts with minor differences. eWOM is all electronic communications between producers and consumers and between consumers themselves: emails, websites, consumer review sites, blogs, virtual communities, chat rooms, newsgroups, and instant messaging [3]. UGC is data generated online by consumers, e.g., text (consumer reviews and blogs) and picture data [4].

Tourism and hospitality relies very much on word of mouth among consumers [5]. Before the Internet, this word of mouth was usually obtained from friends and acquaintances, but has now moved to the Internet, with consumers overwhelmingly consulting the opinions of fellow consumers online [6,7]. Their decision-making and travel experience is shaped by the reviews they read, and this has been a significant development in the industry [8].

Consumers use online reviews at every stage of the travel process—pre-trip, during trip, and post-trip—to engage in information search, share experiences, and give feedback [5]. Companies use it for customer engagement, to build online presence, get consumers' opinions, influence booking intentions, and earn revenue [9–11]. With its usefulness to both producers and consumers, it is no surprise that online reviews are a highly popular topic in tourism and hospitality for both researchers and managers [12].

As a result, there have been many studies on online reviews [13], but only a few keeping track of the literature. These works presented the state of the research at their time of publication, providing a

now-obsolete image of the field because of the astonishing speed at which the literature is growing. Therefore, recent studies on the application of online reviews in tourism and hospitality have not been integrated with earlier studies to understand the present state of research and current research trends. Many of these previous reviews have also been systematic [1,10] and none of them have used bibliometric methods or science maps [14] to visually show the development, conceptual structure, and thematic evolution of online reviews in tourism and hospitality.

Hence, the aim of this paper is to integrate and provide an organized summary of the existing research by identifying the foundational and seminal works and the old and new areas of the prior scholarship in order to build knowledge, gain an understanding, and show the future direction in this research area. We achieve this by means of bibliometric analytical techniques that enable us to deconstruct the main anchors and evolution of the research area. We also contribute an innovative way to show the thematic evolution of the key research themes. We draw four research questions based on this research objective to achieve this goal:

RQ1. What are the general trends in this research area?
RQ2. What is the foundational literature?
RQ3. What are the major research themes?
RQ4. How have the research interests evolved?

2. Literature Review

Even though research on online reviews in tourism and hospitality is fast-growing [14], it has a relatively brief history [15]. This is because notable online opinion platforms for customers rose to prominence in the early 2000s [12]. A few attempts have been made to review the literature and Table 1 presents 16 previous studies that we identified that reviewed the literature on online reviews in tourism and hospitality directly or indirectly.

Table 1 shows that these studies were published in reputable journals, with the International Journal of Contemporary Hospitality Management, Journal of Travel and Tourism Marketing, and International Journal of Hospitality Management having five, three, and two publications, respectively. The remaining journals—Tourism Management, Current Issues in Tourism, Journal of Hospitality Marketing and Management, Tourism Management Perspectives, International Journal of Hospitality and Tourism Administration, and Hospitality and Society—had one publication each.

An analysis of these previous reviews reveals that there are differences in their depth of focus. While some focused directly on online reviews in tourism and hospitality, many considered a larger domain such as social media or big data in tourism and hospitality, with online reviews treated as a subset of this larger domain.

Law et al. [16] had the biggest scope and explored information and communication technology (ICT) in tourism and hospitality, with online reviews included under ICT. This is followed in scope by Mariani et al. [17], who investigated business intelligence and big data in tourism and hospitality. Some studies reviewed research focused on types of big data in tourism and hospitality, including UGC such as online reviews, device data such as GPS (Global Positioning System) and Wi-Fi data, and transaction data such as web search data [4,14]. Other studies examined the role of social media in hospitality and tourism by reviewing the influence of online consumer review platforms, social networking sites, Internet forums, and other online communities [10,15,18–20]. Two studies examined eWOM relating only to hotels [1,21], whereas the remaining studies reviewed online reviews/UGC/eWOM in tourism and hospitality [2,12,13,22].

Table 1. Previous literature reviews.

Authors	Scope	Journal	Databases	Method	Sample Size	Years Covered
Leung et al. (2013)	SM [1] in T&H [2]	JTTM	SD [3], EBSCOHost, and GS [4]	Systematic	44	2007–2011
Serra Cantallops and Salvi (2014)	eWOM in Hotels	IJHM	6 selected journals	Systematic	28	2007–2011
Zeng and Gerritsen (2014)	SM in T&H	TMP	WoS [5], EBSCOHost, and GS	Bibliometric	279	2007–2013
Law et al. (2014)	ICT in T&H	IJCHM	SD, EBSCOHost, Emerald, and Sage	Content analysis	107	2009–2013
Schuckert et al. (2015)	Online reviews in T&H	JTTM	SD, EBSCOHost, and GS	Content analysis	50	2004–2013
Lu and Stepchenkova (2015)	UGC in T&H	JHMM	GS, WoS, top 5 tourism, and top 5 hospitality journals	Systematic	122	2001–2013
Chen and Law (2016)	eWOM in T&H	IJHTA	SD, EBSCOHost	Systematic	43	2008–2013
Kwok et al., (2017)	Online reviews in T&H	IJCHM	Top 7 T&H journals	Systematic	67	2000–2015
Sotiriadis (2017)	SM in T&H	IJCHM	SD, GS	Systematic	146	2009–2016
Bore et al. (2017)	eWOM in Hotels	HS	SD, EBSCOHost, Emerald, and SpringerLink	Systematic	45	2000–2015
Leung et al. (2017)	SM in T&H	IJHM	Top 8 business and top 8 T&H journals in WoS	Bibliometric	406	2007–2016
Li et al. (2018)	BD [6] in Tourism	TM	WoS, SD, Sage, Emerald, SpringerLink, Wiley Online Library, and GS	Systematic	165	2007–2017
Mariani et al. (2018)	BD in T&H	IJCHM	Scopus and WoS	Systematic	173	2000–2016
Lu et al. (2018)	SM in T&H	JTTM	SD, EBSCOHost, GS, and top 7 T&H journals	Systematic	105	2004–2014
Centobelli and Ndou (2019)	BD in Tourism	CIT	Scopus	Bibliometric	109	1990–2017
Nusair et al. (2019)	SM in T&H	IJCHM	SD and EBSCOHost	Bibliometric	439	2002–2016

[1] Social media; [2] Tourism and hospitality; [3] Science Direct; [4] Google Scholar; [5] Web of Science; [6] Big Data.

The previous works can be classified into three groups regarding the academic databases from which they selected their samples. The first group, which comprises most of the studies (11), used a keyword search of well-known databases such as Scopus and Google Scholar [4,10,13–18,21–23]. The second group concentrated on selected top journals such as *Tourism Management*, *Journal of Travel Research*, etc. [1,12,19]. The third group used both selected top journals and databases to get their samples [2,20].

Most of the studies considered only journal articles, except those of Li et al. [4], which used articles and conference papers; Centobelli and Ndou [14], which used articles, conference papers, and book chapters; and Zeng and Gerritsen [15], which used a lot of grey literature such as research degree theses, electronic articles, books, and reports in addition to articles and conference papers.

The literature reviews in this field started in 2013, and the year of publication explained the number of citations received by the documents, with the older publications having more citations than the newer ones, apart from some exceptions. Five of the 16 studies reviewed literature from 2007 up until their time of publication [1,4,10,15,19]. Three were from 2000 [12,17,21], two each started in 2004 [13,20] and 2009 [16,18], and one each in 2001 [2], 2002 [23] and 2008 [22]. Even though Centobelli and Ndou [14] searched for literature from 1990 to 2017, they reviewed documents published from 2011 onwards. In summary, the earliest time of publication of the reviewed literature was from the year 2000, with most of the studies considering literature from 2007 onwards. The last year considered in the most recent reviews was 2017 [4,14].

Regarding method, the previous reviews showed a preference for systematic reviews, with 10 of them adopting this approach [1,2,4,10,12,17,18,20–22]. Two applied content analysis [13,16] and four used a bibliometric approach [14,15,19,23]. The bibliometric reviews examined more documents, with three bibliometric reviews having the biggest sample sizes [15,19,23].

We employ the bibliometric analytical method for this study since it provides the tools to answer our research questions on the foundations and themes of this research area. The bibliometric method provides an advantage of objectivity and quantifiability, and helps to avoid subjective biases. It also helps to provide validation for findings that scholars had intuitively inferred in earlier studies and is more informative [24,25], which justifies the suitability of this methodological approach. Bibliometrics is the use of mathematical and statistical methods to quantify and analyze the bibliographic information of publications [26]. This bibliographic information enables researchers to make linkages between authors or papers [27]. Thomson Reuters Web of Science (WoS) and Elsevier's Scopus are the traditional databases for carrying out bibliometric studies since they are reliable sources of citation data [24].

Many bibliometric studies often utilize science maps [28]. The ability to visualize bibliometric networks with science maps was an important technical development in bibliometrics [24]. Science maps, also known as bibliometric maps or knowledge maps, help to reveal the conceptual, intellectual, or social structure of a field [28]. It is a spatial representation of the interrelationship between research elements such as authors or citations that facilitates the understanding of the structure and developments in a field [29]. There are different software tools for constructing science maps and Cobo et al. [28] made a comparative study of these tools. There are also different approaches to extracting a bibliometric network depending on the selected unit of analysis; e.g., authors, documents, journals, cited references, or keywords [28].

Science maps were used in just two of the four previous bibliometric reviews. Centobelli and Ndou [14] used a citation network, while Leung et al. [19] used a combination of co-citation and co-word analyses to make up for the weaknesses in a single method, and to reveal the theoretical foundation and thematic evolution of their area of study. This study follows in the steps of Leung et al. [19] with this combination for the same reason.

Co-citation analysis is a study of cited documents to get the frequency of citation of two earlier documents together. Hence, a co-citation link is a link between two documents cited together by a later document. When a set of documents is frequently co-cited together in a certain area, this may show

that the co-cited documents contain important concepts that peers have recognized. Thus, a co-citation network analysis can identify the core literature in a particular area of study [30].

Co-word or co-occurrence analysis measures the frequency of co-occurrence of two keywords in the same literature [31]. It relies on the assumption that the co-occurrence of keywords shows a non-trivial relationship [32]. When sets of keywords are frequently used together by different authors in a certain area, this may show that these keywords have a significant relationship within the research area [31]. This bibliometric method directly extracts the major themes of a research area and the linkages between these themes based on the co-occurrence of word pairs without relying on a priori definitions [33]. Hence, they help to reveal the conceptual structure and major research themes of the area [28,32].

The strength of a co-citation link is derived from the number of citing authors who cite two earlier works together. Therefore, co-citation is not a permanent relationship and the co-citation network pattern changes with time. The same applies to co-word analysis, which changes with a change in vocabulary co-occurrences in a field. Therefore, co-citation and co-word patterns in a field will change as the field evolves with a change in the interests and intellectual patterns [30,32].

From the review of the literature, it is obvious that the previous literature review studies have mainly been systematic reviews, and the few bibliometric studies covered a bigger scope, namely, social media [15,19,23] and big data [14]. Hence, their emphasis on online reviews was only peripheral. Similarly, the two previous bibliometric studies that used science maps focused on social media [19] and big data [14]. In summary, there have been no studies to the best of our knowledge providing the foundational literature and major research themes in the maturing research area of online reviews in tourism and hospitality. In addition, our study also covers the recent publications in a fast-growing field, as most of the documents reviewed in this work were published in the last two years. Finally, the thematic evolution chart and the analysis of some dimensions of its network structure is a novel contribution of this work that allows a better understanding of the development and state of the art in the field.

3. Method

3.1. Data Collection

We obtained the publications considered for this study from the Scopus database through a systematic process as shown in Figure 1. Scopus was chosen because it is a recognized index with a wide coverage of peer-reviewed publications and provides reliable bibliographic data [24,34]. Several search queries with relevant keywords were used (see Supplementary File S1) to search in the title, abstract, and keywords of the publications in the database from 2005 to 2019. The year 2005 was chosen as the start date because most of the previous reviews reported the first publications in this area to be from 2007 or later. Two years were added to this to ensure that all relevant documents were found. We filtered the results to include only publications in English and journal articles, resulting in 1440 documents. Their titles and abstracts were reviewed, and we selected 632 relevant publications. We selected theoretical and empirical articles on the application of online reviews in tourism and hospitality, with an emphasis on studies that used online review data and excluded technical papers focused on highlighting the performance and accuracy of algorithmic models for prediction, classification, recommendation engines, or detection of fake reviews.

Figure 1. Systematic process to select literature.

3.2. Data Analysis

To answer the first research question posed by this study, the general literature trends are presented. Co-citation analysis was done to determine the foundational literatures of the study area, which are the most co-cited studies by the selected publications, and answers the second research question. A co-word analysis was also carried out because it enables the determination of the conceptual structure and research themes of the study area and helps to answer the third and fourth research questions. VOSviewer software (version 1.6.13) [35] and Ucinet 6.0 [36] were used to conduct the bibliometric analysis. Data cleaning, an important preprocessing step [28], was performed before data analysis by using a thesaurus file. The thesaurus file was used to combine variants of the same term or related concepts, such as "eWOM," "e-wom," "electronic word of mouth," etc., which were all merged into "eWOM." The thesaurus file was also used to ignore irrelevant words that did not provide any useful information, e.g., "article." The first step in the analysis was to produce a chart showing the growth in the number of published articles per year to show the growth in the literature. Next was a study of the most important journals, followed by the co-citation analysis to identify the core documents. Finally, the co-word analysis was carried out to identify the major research themes, with a study of the crucial dimensions of the keyword network structure and the evolution of the research interests.

4. Results

4.1. General Literature Trends

Figure 2 shows the growth in the number of publications within the study period. There was steady growth from 2006 to 2013 and a slight drop from 2013 to 2014, followed by rapid growth from 2014 onwards.

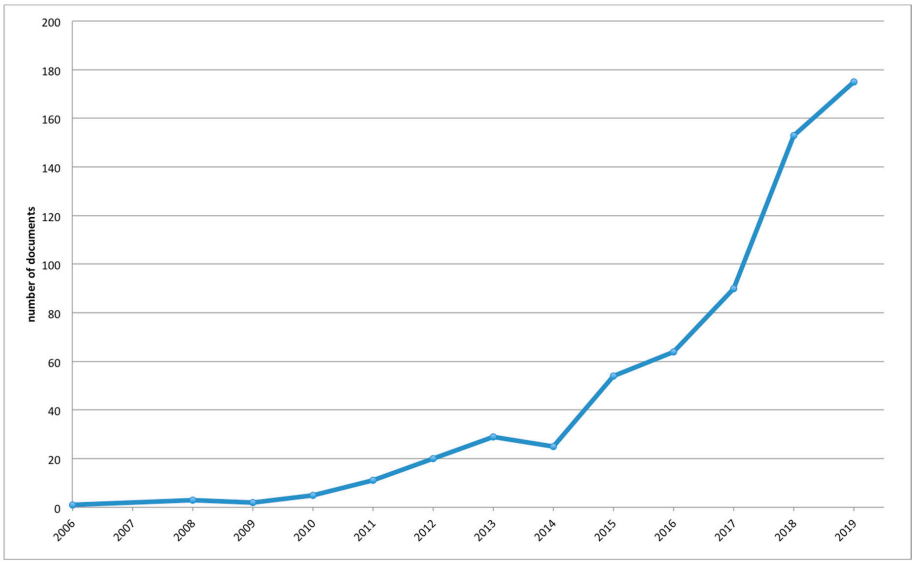


Figure 2. Growth of literature over time.

Table 2 shows the most prolific journals, using a threshold of six or more publications. The list includes many reputable tourism and hospitality journals, with most of them in the first quartile in the 2018 Scopus CiteScore under the Tourism, Leisure and Hospitality Management category. The 632 articles were from 202 journals.

Table 2. List of most prolific journals.

	Journal	Total Articles	Total Citations	Quartile
1	*Tourism Management*	56	3855	Q1
2	*International Journal of Hospitality Management*	53	2644	Q1
3	*International Journal of Contemporary Hospitality Management*	42	721	Q1
4	*Journal of Hospitality Marketing and Management*	25	675	Q1
5	*Current Issues in Tourism*	17	98	Q1
6	*Journal of Travel and Tourism Marketing*	17	885	Q1
7	*Sustainability (Switzerland)*	14	40	Q1 *
8	*Journal of Travel Research*	13	557	Q1
9	*Tourism Management Perspectives*	13	154	Q1
10	*Information Technology and Tourism*	12	70	Q1 *
11	*Journal of Hospitality and Tourism Technology*	12	165	Q1
12	*Cornell Hospitality Quarterly*	11	641	Q1
13	*Journal of Vacation Marketing*	9	150	Q1
14	*Journal of Hospitality and Tourism Research*	8	166	Q1
15	*Tourism Analysis*	8	15	Q3
16	*Asia Pacific Journal of Tourism Research*	7	115	Q1
17	*Journal of Business Research*	7	153	Q1 *

Table 2. *Cont.*

	Journal	Total Articles	Total Citations	Quartile
18	*Journal of China Tourism Research*	7	15	Q2
19	*Anatolia*	6	38	Q2 *
20	*Annals of Tourism Research*	6	131	Q1
21	*E Review of Tourism Research*	6	8	Q4
22	*International Journal of Culture Tourism and Hospitality Research*	6	26	Q2
23	*International Journal of Hospitality and Tourism Administration*	6	37	Q3
24	*Journal of Destination Marketing and Management*	6	165	Q1

Tourism Management and *International Journal of Hospitality Management* had the most publications; with both of them accounting for 17% of all publications. *Tourism Management* published articles with more impact, which have accumulated more citations than the others with about 26% of the total citations of all selected documents.

4.2. Foundational Literature

The co-citation network map was made to identify the foundational literature. To obtain this map, co-citation was selected as the type of analysis, cited references as the unit of analysis, and fractional counting as the counting method, as recommended [37]. Given the high number of cited references of over 30,000, an adequate minimum threshold of citations was mandatory for clarity. To obtain the top 20 most co-cited articles, as done in previous bibliometric studies [38], we set the threshold at 29 citations. The normalization method was association strength and the visualization weight was citations.

Figure 3 shows the co-citation network map. The proximity of two papers, the thickness of the lines connecting them, and their size shows the strength of their co-citation links [39]. These documents in the nodes are the 20 most influential or seminal papers in this area at this time. It is important to point out that even though these papers are not necessarily the most frequently cited papers in terms of bulk citations; they have been highly co-cited by the selected documents and hence represent the foundational literature.

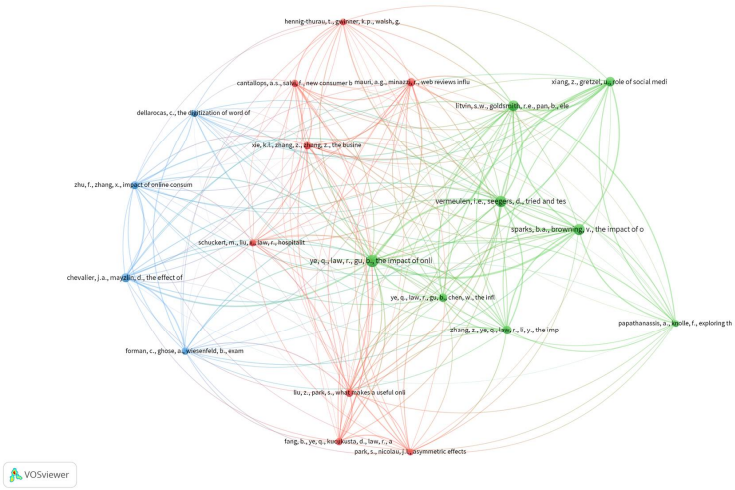


Figure 3. Co-citation network map.

Tourism Management and *International Journal of Hospitality Management* are the dominant sources of these influential papers, with seven and five publications, respectively, while other journals have one each of the remaining eight publications. These papers are in three clusters and there is a uniform pattern with the papers in each cluster. The blue cluster has four papers and are the earliest publications of the 20, published between 2003 and 2010. They deal with the effect of eWOM on sales and were written from outside the tourism and hospitality domain, specifically marketing, information systems, and management science [40–43].

The red cluster consists of eight papers. These are the relatively newer documents, and seven of them are the most recent documents of the 20, published between 2013 and 2016. They can be divided into three subgroups: The first two include literature reviews [1,13] and a paper on the motivations to write a review [44], whereas the remaining five papers are on the consequences of reviews, e.g., value and usefulness of online reviews [45–47], the impact of hotel reviews on consumer purchase intention and service expectations [48], and hotel performance [7].

The green cluster is also made up of eight papers and has larger nodes, which implies that they are the most co-cited papers on the map. They were published from 2008 to 2011 and can be situated in time between the early blue cluster and the recent red cluster. The papers in this mid-period were all written from a tourism and hospitality perspective, in contrast to the blue cluster. We can divide this cluster into two subgroups. The first is made up of influential conceptual papers on the role of social media in online travel information search [49], eWOM in hospitality and tourism [3], and an exploration of the adoption and processing of online reviews [50]. Although similar to some papers in the red cluster [7,48], the second subgroup comprises papers on the consequences of online reviews on different aspects of the sales process, such as consumer consideration [51], hotel booking and sales [9,52], and popularity of restaurants [53]. We find these similar papers in the red and green clusters relatively close on the visualization map.

4.3. Main Research Themes

The co-word network map was made to identify the major research themes. To obtain this map, co-occurrence was selected as the type of analysis, all keywords as the unit of analysis, and fractional counting as the counting method, and the thesaurus file was uploaded. The total number of keywords was 2028, which were too many to fit on a chart. Therefore, a threshold of three occurrences was set, which 219 keywords met. The normalization method was fractionalization and the visualization weight was occurrences.

Figure 4 shows the main keywords that have been used in online reviews literature (full list in Supplementary File S1). The proximity of two terms and the thickness of the lines connecting them show how frequently they co-occurred as keywords, and the size of a node is determined by the frequency of its occurrence as a keyword.

The most frequent keywords, such as “online reviews,” “social media,” “eWOM,” “UGC,” and “online ratings,” had the biggest sizes. To avoid labels from overlapping, VOSviewer does not display labels close to a bigger label in the static image of the map, which prevents some keyword labels from showing. Hence, these oft-recurring keywords that cut across the entire field adding no extra information were removed from the visualization to simplify it and reveal the structure of the field spatially [33].

We calculated centrality measures to determine the precise structure of this network and identify the keywords that occupy critical positions within it. There are different centrality measures, all of which identify the most central elements in a network. Using Ucinet 6.0, degree, betweenness, and eigenvector centrality were calculated. These measures determine the prominence of a keyword in the network and existing conceptual overlap between them, however, some differences may also emerge from their use depending on the network configuration. Thus, considering all of them provides a more complete picture not just about the keywords that occupy highly central positions, but also

of the influence of these keywords on others [54] The top 10 central keywords for each measure are presented in Table 3, as done in previous studies [55].

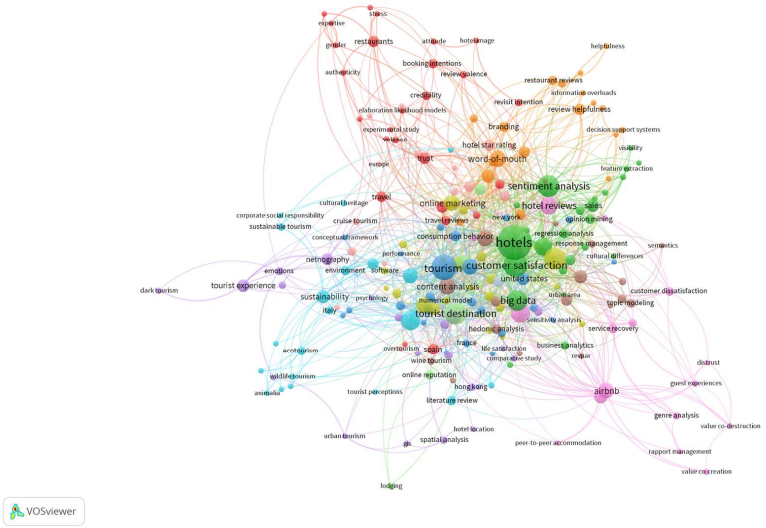


Figure 4. Co-word network map.

Table 3. Top 10 keywords with high centrality.

Keyword	Degree Centrality	Keyword	Betweenness Centrality	Keyword	Eigenvector Centrality
Hotels	96	Hotels	3100.50	Hotels	1.00
Tourism	53	Tourism	1453.66	Customer satisfaction	0.65
Customer satisfaction	45	Tourist destination	1240.57	Text analytics	0.61
Text analytics	43	Tourism management	1166.91	Tourism	0.60
Tourist destination	41	Customer satisfaction	1023.74	Sentiment analysis	0.51
Sentiment analysis	40	Hospitality	865.59	Big data	0.47
Tourism management	34	Text analytics	835.86	Hospitality	0.44
Big data	33	Sentiment analysis	767.95	Online ratings	0.42
Hospitality	33	Perception	644.48	Data mining	0.36
Data mining	28	Tourist behavior	584.82	Tourist destination	0.36

Hotels/hospitality (hotels, hotel reviews, hotel star rating, etc.) and tourism (tourism, tourist destination, tourism management, travel, travel reviews, wine tourism, cruise tourism, etc.) are the prominent industry domains and occupy a central position on the map, as expected according to all three centrality measures. Restaurants (restaurants, authenticity, booking intentions, etc.) was also an expected theme because of its relevance in the hospitality industry.

Customer satisfaction is also central, and highly connected on the map as a concept that has been investigated extensively using online reviews. Big data and text analytics techniques (sentiment analysis, topic modeling, etc.), and traditional research and statistical methodologies (regression

analysis, sensitivity analysis, content analysis, etc.) also occupy central positions as the processing tools for online reviews. It is worth mentioning that the high values of these keywords in terms of betweenness centrality underline not just their central position in the network, but also their influence on other keywords close to them.

Airbnb (Airbnb, sharing economy, peer-to-peer accommodation, guest experiences, value co-creation and co-destruction, etc.), sustainability (sustainability, sustainable tourism, ecotourism, overtourism, environment, wildlife tourism, corporate social responsibility, etc.), and tourist experience (emotions, dark tourism, heritage tourism, etc.) are also themes that can be identified on the map, showing the underlying structure of the research area.

4.4. Evolution of Research Interests

VOSviewer provides an overlay visualization of a network diagram, which shows the nodes on a network map with a color gradation based on the scores calculated by the software for either the average publication year (APY), average number of citations, or average normalized citations (ANC). To show the impact of the keywords over time in the selected literature in an easy-to-understand chart, the calculated figures from VOSviewer were taken and the keyword impact, represented by the ANC, was plotted against publication year (APY).

Figure 5 shows the wide range of keywords that have been used in the selected documents and the evolution of research interests in online reviews in tourism and hospitality in a timeline. ANC is the average number of normalized citations received by the documents that the keyword represents. The normalized citation count is the number of citations of the document divided by the average number of all citations of all documents published in the same year it was published, included in VOSviewer data [39]. ANC was used to compensate for the fact that citation counts favor older publications. APY is the average year of publication of the documents that the keyword represents. We placed the number of occurrences of the keywords in brackets in front of the keywords.

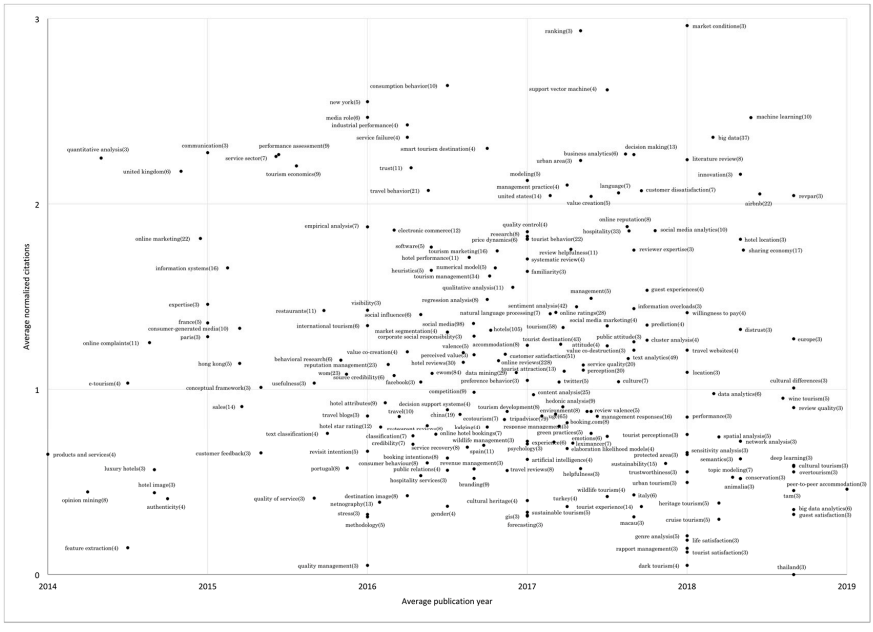


Figure 5. Evolution timeline of keywords.

This chart originally contained the 219 keywords from the co-word network map after choosing a threshold of three occurrences, with this threshold eliminating many keywords from earlier years. However, placing all 219 keywords on the chart, including those with very high ANC values and very low APY values, made it difficult to visualize most of the keywords. Therefore, five keywords with outlying values were excluded to get a clearer plot, with axes 0–3 for ANC and 2014–2019 for APY. We present these excluded keywords in Table 4. In addition, we present the full data of all 219 keywords in Supplementary File S2, i.e., ANC, APY, total number of occurrences, and the breakdown of occurrences by year.

Table 4. Outlier values removed from Figure 4.

Keyword	Occurrences	APY	ANC
Artificial neural network	4	2018.00	4.50
Tourism research	3	2016.67	3.65
Comparative study	4	2018.25	3.57
Experimental study	3	2013.67	1.80
Web 2.0	9	2013.78	1.81

Many of the keywords from the co-word analysis are visible in Figure 5, including those that could not be seen on the co-word network map to avoid overlapping. They include keywords about the industry sectors, consumer behavior, management processes to achieve management objectives, online review platforms, research methodologies, and country/city names. The closer to the right side of the plot the keyword appears, the more recent its occurrences in the literature are, and the closer to the top of the plot, the better cited (normalized citations) the articles are that used the keyword. Keywords found to the bottom right of the plot are emergent in the literature, while those at the top right are recent and well cited or hot topics [56].

The applications of artificial intelligence (AI) or machine learning techniques to the big data of online reviews are prominent themes in the hot topic area, e.g., "big data," "machine learning," "artificial neural network," "support vector machine," and "business analytics." Other keywords related to this theme, such as "deep learning," "text analytics," "natural language processing," "data analytics," "big data analytics," and "AI," are found at lower values of ANC, but still in the recent areas. This highlights their current relevance in the literature. The Airbnb and sharing economy theme is also in the hot topic area.

In the emergent area, online reviews are increasingly applied to investigate different niche areas of tourism, such as wine tourism, heritage tourism, cultural tourism, urban tourism, cruise tourism, wildlife tourism, and dark tourism. Sustainability themes such as sustainability, ecotourism, overtourism, environment, green practices, and sustainable tourism can also be found in this emergent area.

Figure 5 also shows a subtle evolution in terminology for some terms. For example, the average year of usage of "opinion mining" was 2014.3 (six of its eight appearances were between 2008 and 2016), and it seems to have now been replaced by terms such as "data mining" (2016.9), "text analytics" (2017.6), and "data analytics" (2018.2) in the tourism and hospitality literature. In the same vein, we can also observe an evolution in techniques. For example, feature extraction, a machine learning process used in text analysis to extract the unique features of a document [57], had an average year of 2014.5, and it seems to have now been eclipsed in the tourism and hospitality literature by topic modeling (2018.3), a method for carrying out feature extraction [57], which suggests that researchers now frequently use this technique for feature extraction. Another way the use of keywords has changed is the gradual change with time from the use of broader keywords to more specific sub-techniques in the literature as the research in the subfield deepens, for example, from a broad keyword like "AI" (2017) to "machine learning" (2018.4), a subset of AI, and then to "artificial neural networks" (2018) and "deep learning" (2018.67), which are both subsets of machine learning.

5. Discussion and Conclusions

This study used a bibliometric approach to review the research on online reviews in tourism and hospitality by analyzing 632 relevant documents selected from the Scopus database published between 2005 and 2019. The study applied co-citation and co-word analyses, including an innovative visualization of the co-word network to explore the theoretical foundations, network structure, and thematic evolution in this research area. It has successfully carried out an update, reviewed more studies, and revealed the theoretical foundations of the field with the use of science maps and shown the evolution of the major themes in the knowledge area. It has also provided an important quantitative complement with bibliometric methods to the previous literature reviews to enhance the understanding of this fast-growing knowledge area and has validated previous findings, such as the central research themes in online reviews.

Online review research is well established in the tourism and hospitality literature, with the top journals publishing literature in this research area. The earliest foundations of this research area was anchored to the adjacent body of knowledge on online reviews from other disciplines such as marketing and management, conducted on electronic commerce websites such as Amazon and eBay, before extensive research by tourism and hospitality scholars based on hotel and tourism reviews grew. Besides influential conceptual papers and literature reviews that integrate the knowledge base of the area, most of the foundational papers focused on the consequences, value, usefulness, and effect of online reviews on sales and financial performance, customer booking intention, and service expectations. These publications were important because they answered the fundamental question of whether online reviews were relevant to tourism and hospitality, because it is only when they are relevant that attention needs to be paid to them. Therefore, these papers played a crucial role as foundational works since they tackled the important theoretical concepts as well as the relevance, consequences, and challenges of eWOM. Hence, these documents were often co-cited by many publications in this research area to establish a theoretical base upon which to build their own research.

The co-word analysis confirms the research that has been carried out in the past in the traditional cornerstones of the tourism and hospitality industry: hotels, travel and tourism, and restaurants, as established in previous works [13,22], with customer satisfaction as a central concept and sentiment analysis as a predominant mode of analysis [4]. These themes not only occupy central positions but also maintain close relationships among themselves.

Big data and the application of machine learning are impacting virtually every industry [58]. The tourism and hospitality industry is not an exception, with the application of machine learning techniques such as artificial neural networks, support vector machines, and deep learning to natural language processing tasks, especially sentiment analysis. Many of these techniques were developed in the 1970s and 1980s, but big data and increased computer-processing capacity has resulted in their revival and wide application in different fields, including natural language processing [58,59]. We expect the application of these techniques to grow in the text processing of online reviews, especially in the deployment of novel sub-techniques as they are developed, as has happened in the past, i.e., in novel means of implementing neural networks and multilayer neural network (deep learning) models.

There is also an expansion of the application of online reviews beyond the traditional cornerstones of tourism and hospitality research into other themes, such as Airbnb, sustainability, and tourist experience. Airbnb's emergence with a unique accommodation business model has disrupted the tourism accommodation sector. The platform's celebration in March 2019 of reaching a milestone of 500 million guests, from about 10 million in June 2012, is evidence of its explosive growth [60,61]. Guest reviews are a feature of each Airbnb listing, and researchers have used these reviews to investigate guest experiences [62], host experiences [63], and other research perspectives, and is an active area of research in the application of online reviews to tourism and hospitality.

The theme of tourist experience, connected with emotions, heritage, and dark tourism, confirms the increasing demand for new forms of tourism and unique experiences, characterized by a departure from mass tourism [64] and an increase in demand for experiential tourism [65], which these niche

areas of tourism provide. The provision of memorable tourist experiences is directly related to a business's ability to generate revenue [66], and destination managers and even countries can leverage this. For example, Airbnb reported a five-fold increase in heritage travel globally since 2014 on its platform [67,68]. The application of online reviews to investigate these niche areas of tourism is likely to continue to grow, as there have been papers using online reviews to investigate other niche areas such as polar tourism [69], cycling tourism [70], plantation tours [71], etc. among the reviewed papers.

Tourism accounts for about 8% of global greenhouse gas emissions [72], and sustainability has become an important policy issue in tourism and hospitality [73]. Hence, "sustainability," "sustainable tourism," "ecotourism," and other related keywords are an important emergent area and reviews have been used to investigate guest reactions to hotel sustainability measures and the impact on guest experiences [74].

In conclusion, the research on online reviews and eWOM is reaching maturity with a rapidly increasing volume of new publications built on a solid base of seminal research and applied to highly varied areas, while increasingly sophisticated analytical methods are employed for analysis. This literature review has shown the foundational literature of this research area, recent growing areas of application in Airbnb and tourist experience, and in niche tourism areas such as wine tourism, dark tourism, etc. It has also revealed the latest analytical techniques being used, such as artificial neural networks and deep learning, and the hot and emergent topic areas, such as machine learning, big data analytics, and sustainability themes.

These findings are useful for scholars in this field to understand the current state of knowledge amidst the overwhelming growth in the literature and for general readers interested in learning about online reviews, and also provide practitioners and business owners a window to understanding a ubiquitous concept that they confront every day in the tourism and hospitality industry. The findings also provide a glimpse of future research growth areas toward the hot and emergent topics and the application of new machine learning sub-techniques using deep learning and neural networks for text processing. The thematic evolution chart, which highlights the hot and emergent topics, can be used in future bibliometric studies in any research area.

Limitations, Theoretical Implications, and Future Research

A limitation of this work is that authors sometime add arbitrary keywords to their papers that do not truly reflect the content of the paper, which could have affected the veracity of the co-word network map and thematic evolution timeline chart. The use of only the Scopus database is another limitation of this study that could have possibly resulted in leaving out a few publications indexed by WoS, but not in Scopus. It was not possible to combine both databases for this study because VOSviewer only allows the use of bibliographic data from either database, but not a combination of both. However, there is a significant overlap between both databases [34] and we recommend future work that can conduct a bibliometric analysis on a combination of both databases.

Our results have important implications for theory and future research. They show the conceptual structure and thematic evolution of this research area, and an individual or researcher new to this research area can quickly obtain a bird's-eye view of the seminal publications and how the research themes have evolved to date. However, research is a continuous process and as mentioned, the co-citation network map, co-word network map, and, by extension, the thematic evolution chart are not static but change with time. Future studies are thereby needed to track the changes that take place in the conceptual structure and major research themes of this research area with time. In addition, in light of the COVID-19 pandemic, studies employing online reviews and eWOM from various sources such as social media are suggested to understand the various effects of the pandemic.

Supplementary Materials: The following are available online at http://www.mdpi.com/2071-1050/12/23/9977/s1.

Author Contributions: Conceptualization, B.A.M., M.-V.S.-R., and A.-B.H.-L.; methodology, B.A.M., M.-V.S.-R., and A.-B.H.-L.; formal analysis, B.A.M.; writing—original draft preparation, B.A.M.; writing—review and editing, B.A.M., M.-V.S.-R., and A.-B.H.-L.; visualization, B.A.M.; supervision, M.-V.S.-R. and A.-B.H.-L. All authors have read and agreed to the published version of the manuscript.

Funding: This project received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 713679 and from the Universitat Rovira i Virgili (URV).

Acknowledgments: We appreciate the constructive recommendations made by the three anonymous reviewers.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Serra Cantallops, A.; Salvi, F. New consumer behavior: A review of research on eWOM and hotels. *Int. J. Hosp. Manag.* **2014**, *36*, 41–51. [CrossRef]
2. Lu, W.; Stepchenkova, S. User-Generated Content as a Research Mode in Tourism and Hospitality Applications: Topics, Methods, and Software. *J. Hosp. Mark. Manag.* **2015**, *24*, 119–154. [CrossRef]
3. Litvin, S.W.; Goldsmith, R.E.; Pan, B. Electronic word-of-mouth in hospitality and tourism management. *Tour. Manag.* **2008**, *29*, 458–468. [CrossRef]
4. Li, J.; Xu, L.; Tang, L.; Wang, S.; Li, L. Big data in tourism research: A literature review. *Tour. Manag.* **2018**, *68*, 301–323. [CrossRef]
5. Fotis, J.; Buhalis, D.; Rossides, N. *Social Media Use and Impact during the Holiday Travel Planning Process BT—Information and Communication Technologies in Tourism 2012*; Fuchs, M., Ricci, F., Cantoni, L., Eds.; Springer Vienna: Vienna, Austria, 2012; pp. 13–24.
6. Wang, Y.; Yu, Q.; Fesenmaier, D.R. Defining the virtual tourist community: Implications for tourism marketing. *Tour. Manag.* **2002**, *23*, 407–417. [CrossRef]
7. Xie, K.L.; Zhang, Z.; Zhang, Z. The business value of online consumer reviews and management response to hotel performance. *Int. J. Hosp. Manag.* **2014**, *43*, 1–12. [CrossRef]
8. Sparks, B.A.; Perkins, H.E.; Buckley, R. Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tour. Manag.* **2013**, *39*, 1–9. [CrossRef]
9. Sparks, B.A.; Browning, V. The impact of online reviews on hotel booking intentions and perception of trust. *Tour. Manag.* **2011**, *32*, 1310–1323. [CrossRef]
10. Leung, D.; Law, R.; van Hoof, H.; Buhalis, D. Social Media in Tourism and Hospitality: A Literature Review. *J. Travel Tour. Mark.* **2013**, *30*, 3–22. [CrossRef]
11. He, W.; Tian, X.; Tao, R.; Zhang, W.; Yan, G.; Akula, V. Application of social media analytics: A case of analyzing online hotel reviews. *Online Inf. Rev.* **2017**, *41*, 921–935. [CrossRef]
12. Kwok, L.; Xie, K.L.; Richards, T. Thematic framework of online review research: A systematic analysis of contemporary literature on seven major hospitality and tourism journals. *Int. J. Contemp. Hosp. Manag.* **2017**, *29*, 307–354. [CrossRef]
13. Schuckert, M.; Liu, X.; Law, R. Hospitality and Tourism Online Reviews: Recent Trends and Future Directions. *J. Travel Tour. Mark.* **2015**, *32*, 608–621. [CrossRef]
14. Centobelli, P.; Ndou, V. Managing customer knowledge through the use of big data analytics in tourism research. *Curr. Issues Tour.* **2019**, *22*, 1862–1882. [CrossRef]
15. Zeng, B.; Gerritsen, R. What do we know about social media in tourism? A review. *Tour. Manag. Perspect.* **2014**, *10*, 27–36. [CrossRef]
16. Law, R.; Buhalis, D.; Cobanoglu, C. Progress on information and communication technologies in hospitality and tourism. *Int. J. Contemp. Hosp. Manag.* **2014**, *26*, 727–750. [CrossRef]
17. Mariani, M.; Baggio, R.; Fuchs, M.; Höepken, W. Business intelligence and big data in hospitality and tourism: A systematic literature review. *Int. J. Contemp. Hosp. Manag.* **2018**, *30*, 3514–3554. [CrossRef]
18. Sotiriadis, M.D. Sharing tourism experiences in social media: A literature review and a set of suggested business strategies. *Int. J. Contemp. Hosp. Manag.* **2017**, *29*, 179–225. [CrossRef]
19. Leung, X.Y.; Sun, J.; Bai, B. Bibliometrics of social media research: A co-citation and co-word analysis. *Int. J. Hosp. Manag.* **2017**, *66*, 35–45. [CrossRef]

20. Lu, Y.; Chen, Z.; Law, R. Mapping the progress of social media research in hospitality and tourism management from 2004 to 2014. *J. Travel Tour. Mark.* **2018**, *35*, 102–118. [CrossRef]
21. Bore, I.; Rutherford, C.; Glasgow, S.; Aheri, B.; Antony, J. A systematic literature review on eWOM in the hotel industry: Current trends and suggestions for future research. *Hosp. Soc.* **2017**, *7*, 63–85. [CrossRef]
22. Chen, Y.-F.; Law, R. A Review of Research on Electronic Word-of-Mouth in Hospitality and Tourism Management. *Int. J. Hosp. Tour. Adm.* **2016**, *17*, 347–372. [CrossRef]
23. Nusair, K.; Butt, I.; Nikhashemi, S.R. A bibliometric analysis of social media in hospitality and tourism research. *Int. J. Contemp. Hosp. Manag.* **2019**, *31*, 2691–2719. [CrossRef]
24. Mingers, J.; Leydesdorff, L. A review of theory and practice in scientometrics. *Eur. J. Oper. Res.* **2015**, *246*, 1–19. [CrossRef]
25. Nerur, S.P.; Rasheed, A.A.; Natarajan, V. The intellectual structure of the strategic management field: An author co-citation analysis. *Strateg. Manag. J.* **2008**, *29*, 319–336. [CrossRef]
26. Pritchard, A. Statistical bibliography or bibliometrics? *J. Doc.* **1969**, *25*, 348–349.
27. Garfield, E.; Malin, V.M.; Small, H. Citation Data as Science Indicators. *Towar. A Metr. Sci.* **1983**, *6*, 580.
28. Cobo, M.J.; López-Herrera, A.G.; Herrera-Viedma, E.; Herrera, F. Science mapping software tools: Review, analysis, and cooperative study among tools. *J. Am. Soc. Inf. Sci. Technol.* **2011**, *62*, 1382–1402. [CrossRef]
29. Small, H. Visualizing science by citation mapping. *J. Am. Soc. Inf. Sci.* **1999**, *50*, 799–813. [CrossRef]
30. Small, H. Co-citation in the scientific literature: A new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* **1973**, *24*, 265–269. [CrossRef]
31. Whittaker, J. Creativity and Conformity in Science: Titles, Keywords and Co-word Analysis. *Soc. Stud. Sci.* **1989**, *19*, 473–496. [CrossRef]
32. Callon, M.; Courtial, J.P.; Laville, F. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics* **1991**, *22*, 155–205. [CrossRef]
33. He, Q. Knowledge Discovery Through Co-Word Analysis. *Libr. Trends* **1999**, *48*, 133–159.
34. Martín-Martín, A.; Orduna-Malea, E.; Thelwall, M.; Delgado López-Cózar, E. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *J. Informetr.* **2018**, *12*, 1160–1177. [CrossRef]
35. van Eck, N.J.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [CrossRef]
36. Borgatti, S.P.; Everett, M.G.; Freeman, L.C. *UCINET 6 for Windows: Software for Social Network Analysis*; Analytic Technologies: Harvard, MA, USA, 2002.
37. van Eck, N.J.; Waltman, L. *Visualizing Bibliometric Networks*; Springer: Cham, Switzerland, 2014; ISBN 9783319103778.
38. Liu, Z.; Yin, Y.; Liu, W.; Dunford, M. Visualizing the intellectual structure and evolution of innovation systems research: A bibliometric analysis. *Scientometrics* **2015**, *103*, 135–158. [CrossRef]
39. van Eck, N.J.; Waltman, L. *VOSviewer Manual*; Univeristeit Leiden: Leiden, The Netherlands, 2013.
40. Chevalier, J.A.; Mayzlin, D. The effect of word of mouth on sales: Online book reviews. *J. Mark. Res.* **2006**, *43*, 345–354. [CrossRef]
41. Dellarocas, C. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Manag. Sci.* **2003**, *49*, 1407–1424. [CrossRef]
42. Forman, C.; Ghose, A.; Wiesenfeld, B. Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Inf. Syst. Res.* **2008**, *19*, 291–313. [CrossRef]
43. Zhu, F.; Zhang, X. Impact of online consumer reviews on Sales: The moderating role of product and consumer characteristics. *J. Mark.* **2010**, *74*, 133–148. [CrossRef]
44. Hennig-Thurau, T.; Gwinner, K.P.; Walsh, G.; Gremler, D.D. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *J. Interact. Mark.* **2004**, *18*, 38–52. [CrossRef]
45. Fang, B.; Ye, Q.; Kucukusta, D.; Law, R. Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tour. Manag.* **2016**, *52*, 498–506. [CrossRef]
46. Liu, Z.; Park, S. What makes a useful online review? Implication for travel product websites. *Tour. Manag.* **2015**, *47*, 140–151. [CrossRef]

47. Park, S.; Nicolau, J.L. Asymmetric effects of online consumer reviews. *Ann. Tour. Res.* **2015**, *50*, 67–83. [CrossRef]
48. Mauri, A.G.; Minazzi, R. Web reviews influence on expectations and purchasing intentions of hotel potential customers. *Int. J. Hosp. Manag.* **2013**, *34*, 99–107. [CrossRef]
49. Xiang, Z.; Gretzel, U. Role of social media in online travel information search. *Tour. Manag.* **2010**, *31*, 179–188. [CrossRef]
50. Papathanassis, A.; Knolle, F. Exploring the adoption and processing of online holiday reviews: A grounded theory approach. *Tour. Manag.* **2011**, *32*, 215–224. [CrossRef]
51. Vermeulen, I.E.; Seegers, D. Tried and tested: The impact of online hotel reviews on consumer consideration. *Tour. Manag.* **2009**, *30*, 123–127. [CrossRef]
52. Ye, Q.; Law, R.; Gu, B.; Chen, W. The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Comput. Hum. Behav.* **2011**, *27*, 634–639. [CrossRef]
53. Zhang, Z.; Ye, Q.; Law, R.; Li, Y. The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *Int. J. Hosp. Manag.* **2010**, *29*, 694–700. [CrossRef]
54. Valente, T.; Coronges, K.; Lakon, C.; Costenbader, E. How Correlated Are Network Centrality Measures? *Connect. (Tor)* **2008**, *28*, 16–26.
55. Hu, C.-P.; Hu, J.-M.; Deng, S.-L.; Liu, Y. A co-word analysis of library and information science in China. *Scientometrics* **2013**, *97*, 369–382. [CrossRef]
56. Ávila-Robinson, A.; Wakabayashi, N. Changes in the structures and directions of destination management and marketing research: A bibliometric mapping study, 2005–2016. *J. Destin. Mark. Manag.* **2018**, *10*, 101–111. [CrossRef]
57. Liu, B. *Sentiment Analysis and Opinion Mining*; Morgan & Claypool Publishers: San Rafael, CA, USA, 2012; Volume 5.
58. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
59. Burges, C.J.C. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [CrossRef]
60. Guttentag, D. Airbnb: Disruptive innovation and the rise of an informal tourism accommodation sector. *Curr. Issues Tour.* **2015**, *18*, 1192–1217. [CrossRef]
61. Airbnb Airbnb Celebrates Half A Billion Guest Arrivals. Available online: https://news.airbnb.com/airbnb-celebrates-half-a-billion-guest-arrivals/ (accessed on 21 December 2019).
62. Cheng, M.; Jin, X. What do Airbnb users care about? An analysis of online review comments. *Int. J. Hosp. Manag.* **2019**, *76*, 58–70. [CrossRef]
63. Cheng, M.; Zhang, G. When Western hosts meet Eastern guests: Airbnb hosts' experience with Chinese outbound tourists. *Ann. Tour. Res.* **2019**, *75*, 288–303. [CrossRef]
64. Stamboulis, Y.; Skayannis, P. Innovation strategies and technology for experience-based tourism. *Tour. Manag.* **2003**, *24*, 35–43. [CrossRef]
65. Smith, W.L. Experiential tourism around the world and at home: Definitions and standards. *Int. J. Serv. Stand.* **2006**, *2*, 1–14. [CrossRef]
66. Kim, J.-H.; Ritchie, J.R.B.; McCormick, B. Development of a scale to measure memorable tourism experiences. *J. Travel Res.* **2012**, *51*, 12–25. [CrossRef]
67. Airbnb Heritage Travel on the Rise: Airbnb and 23andMe Team Up to Make it Even Easier. Available online: https://news.airbnb.com/heritage-travel-on-the-rise/ (accessed on 19 December 2019).
68. Yearofreturn.com Year Of Return—Celebrating 400 years of African Resilience. Available online: https://www.yearofreturn.com/ (accessed on 19 December 2019).
69. Heimtun, B. Emotions and affects at work on Northern Lights tours. *Hosp. Soc.* **2016**, *6*, 223–241. [CrossRef]
70. Chiu, W.; Leng, H.K. Let's go cycling: An analysis of tourists' experience on online user-generated content. *Int. J. Tour. Cities* **2017**, *3*, 30–42. [CrossRef]
71. Carter, P.L. Where are the enslaved?: Tripadvisor and the narrative landscapes of southern plantation museums. *J. Herit. Tour.* **2016**, *11*, 235–249. [CrossRef]
72. Lenzen, M.; Sun, Y.-Y.; Faturay, F.; Ting, Y.-P.; Geschke, A.; Malik, A. The carbon footprint of global tourism. *Nat. Clim. Chang.* **2018**, *8*, 522–528. [CrossRef]

73. Saarinen, J. Traditions of sustainability in tourism studies. *Ann. Tour. Res.* **2006**, *33*, 1121–1140. [CrossRef]
74. Gerdt, S.-O.; Wagner, E.; Schewe, G. The relationship between sustainability and customer satisfaction in hospitality: An explorative investigation using eWOM as a data source. *Tour. Manag.* **2019**, *74*, 155–172. [CrossRef]

Appendix 2: #CoronavirusCruise: Impact and implications of the COVID-19 outbreaks on the perception of cruise tourism

Tourism Management Perspectives 41 (2022) 100948

Contents lists available at ScienceDirect

Tourism Management Perspectives

journal homepage: www.elsevier.com/locate/tmp

#CoronavirusCruise: Impact and implications of the COVID-19 outbreaks on the perception of cruise tourism

Babajide Abubakr Muritala [a,*], Ana-Beatriz Hernández-Lara [a], Maria-Victoria Sánchez-Rebull [a], Alexandre Perera-Lluna [b,c,d]

[a] Departament de Gestió d'Empreses, Facultat d'Economia i Empresa, Universitat Rovira i Virgili, 43204 Reus, Tarragona, Spain
[b] B2SLab, Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial, Universitat Politècnica de Catalunya, Barcelona, Spain
[c] Networking Biomedical Research Centre, Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain
[d] Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain

ARTICLE INFO

Keywords:
Coronavirus
Crisis communication
Risk perception
Big data
eWOM
SARF
Information integration theory
Sentiment analysis

ABSTRACT

Early in the COVID-19 pandemic, the *Diamond Princess* became the center of the largest outbreak outside the original epicenter in China. This outbreak which left 712 passengers infected and 14 dead, followed by subsequent outbreaks affecting over one-third of the active ships in the cruise industry's global fleet, quickly became a crisis that captured public attention and dominated mainstream news and social media. This study investigates the perception of cruising during these outbreaks by analyzing the tweets on cruising using Natural Language Processing (NLP). The findings show a prevalent negative sentiment in most of the analyzed tweets, while the criticisms directed at the cruise industry were based on perceptions and stereotypes of the industry before the pandemic. The study provides insight into the concerns raised in these conversations and highlights the need for new business models outside the pre-pandemic mass-market model and to genuinely make cruising more environmentally friendly.

1. Introduction

The COVID-19 pandemic caused by the novel coronavirus SARS-CoV-2 has exceeded 5 million confirmed deaths at the time of writing (AP, 2021), causing widespread disruption in many sectors of the world economy. Tourism was one of the worst-hit sectors due to the closure of national borders and the reduction in global travel (Gössling, Scott, & Hall, 2020). Cruise tourism in particular was severely affected because of COVID-19 outbreaks on several cruises early in the pandemic (Ito, Hanaoka, & Kawasaki, 2020), which resulted in stranded ships due to port closures and the temporary ban of cruises in some countries (Gössling et al., 2020). Some returning passengers from cruises contributed to the spread of the virus in their home countries (Ito et al., 2020). For example, as at the end of April 2020, one in ten COVID-19 cases in Australia were attributed to a *Ruby Princess* cruise voyage (ABC, 2020), while between February and mid-March 2020, about 17% of the confirmed cases in the United States (US) were linked to returning cruisers (Moriarty, Plucinski, Marston, et al., 2020).

Official and unofficial counts show that there were at least 3908 confirmed COVID-19 cases and 111 confirmed deaths linked to over 102 COVID-19 outbreaks involving at least 124 cruise ships as at October 2020 (CDC, 2020; Miami Herald, 2020). Fig. 1 shows that many cruise companies had outbreaks on their ships, including Carnival Corporation, Royal Caribbean, and Norwegian Cruise Line, the three biggest companies in the industry, which account for 80% of the industry's passenger capacity (Papathanassis, 2017). These cruise companies also suffered economic setbacks because of the pandemic, with their share prices falling off a cliff in the first quarter of 2020 (see Fig. 2). Cruises continue to record transmission of COVID-19 despite high vaccination rates among passengers and crew, with 1359 reported confirmed cases in the US between June and October 2021 (CDC, 2021).

Cruising was the most mentioned sub-sector of tourism in global news in the early period of the pandemic as the media reported these COVID-19 outbreaks (Gössling et al., 2020) and some cruise industry executives have claimed that the industry was unfairly tarnished by the media in the reporting of the outbreaks (Financial Times, 2020; Washington Post, 2020). It is important to have an appraisal of the impact that the extensive coverage of the cruise COVID-19 outbreaks in the news

* Corresponding author.
E-mail addresses: babajide.muritala@urv.cat (B.A. Muritala), anabeatriz.hernandez@urv.cat (A.-B. Hernández-Lara), mariavictoria.sanchez@urv.cat (M.-V. Sánchez-Rebull), alexandre.perera@upc.edu (A. Perera-Lluna).

https://doi.org/10.1016/j.tmp.2022.100948
Received 7 December 2020; Received in revised form 23 December 2021; Accepted 6 February 2022
Available online 10 February 2022

and social media have had on people and how it has affected their perception of cruising. Such research is scarce given the novelty of the topic, but will help to evaluate the purchase intention to cruise and the recovery of the cruise industry post-pandemic (Holland, Mazzarol, Soutar, Tapsall, & Elliott, 2021; Pan, Shu, Kitterlin-Lynch, & Beckman, 2021).

Related studies already published used questionnaire survey polls to gauge the outlook towards cruising and the willingness to cruise in the aftermath of the COVID-19 pandemic. For example, Holland et al. (2021) investigated the impact that the COVID-19 pandemic had on the perceived risk of cruising considering a list of 20 items, and found that the country of residence had a significant impact on risk perception of cruising and future intentions to cruise. Pan et al. (2021) tried to identify consumer perception of the cruise industry during the COVID-19 pandemic under the theoretical lens of leisure constraints and prospect theories. They found that travel constraints negatively influenced behavioral intention to travel with cruises, although perceived crisis management positively affects this behavioral intention. The survey instrument used in both studies is universally accepted for evaluating public opinion but has been criticized for treating public opinion as a quantitative distribution of opinions by disparate individuals having equal weight in society with a blind spot to the hierarchical and conversational nature of public opinion formation (Blumer, 1948). Social media conversations make up for these weak points, since they are hierarchical and conversational by nature (McGregor, 2019). Furthermore, surveys are appropriate to capture subjects' attitudes by directly asking respondent about their subjective experiences, perceptions, and attitudes about a topic, but have limited ability to observe actual behavior. Social media provides a wealth of information about actual user behavior, and not as limited or focused on specific questions (Couper, 2013). Research has also found a correlation between public opinion and sentiments expressed on Twitter (O'Connor, Balasubramanyan, Routledge, & Smith, 2010). Consequently, analysis of social media conversations can provide insight that can be used with survey results to provide a holistic view of the public opinion on an issue.

This study fills this research gap in developing a better understanding of the public perception towards cruising during the COVID-19 outbreaks on cruises in the early part of the COVID-19 pandemic, extending the research on this topic beyond the perceived risk of cruising or the intention to cruise. To do this, we analyze the relevant tweets using Natural Language Processing (NLP) methods. Based on this premise, this study intends to answer the research question: What insight can be derived from the public conversation on Twitter about cruising during the COVID-19 outbreaks on cruises?

The contribution of this study is to serve either as an empirical counterpart to the findings of the previous studies with survey data, as well as to uniquely show the hierarchical dimension of the public conversation with the influence of elites in driving awareness to issues in a way that surveys may not show. It is also hoped that the discussion of the implications of the findings will be a valuable addition to the academic literature on the prospects and outlook of the cruise industry after the COVID-19 pandemic.

2. Literature review

The COVID-19 cruise outbreaks were a crisis event that we investigated using social media data. Hence, it is important to examine the literature on the use of social media in crisis communication in tourism and cruise tourism. As mentioned, there was extensive media coverage and social media mentions of cruising in the early period of the pandemic (Gössling et al., 2020). This wall-to-wall media coverage usually affects public opinion, risk perception, and consumer behavior in a particular pattern. In the following sections, the theoretical foundation of this study is presented, which is based on the information integration theory (IIT) and the social amplification of risk framework (SARF) to explain risk perception and decision-making of tourists influenced by social media in times of crisis.

2.1. Role of social media during crises in tourism

One of the defining characteristics of a crisis is the mass generation of mostly negative comments and information (Coombs, 2018). Social media provides a platform for communicating decisions during a crisis and collecting feedback from the public (Sigala, 2011). The information

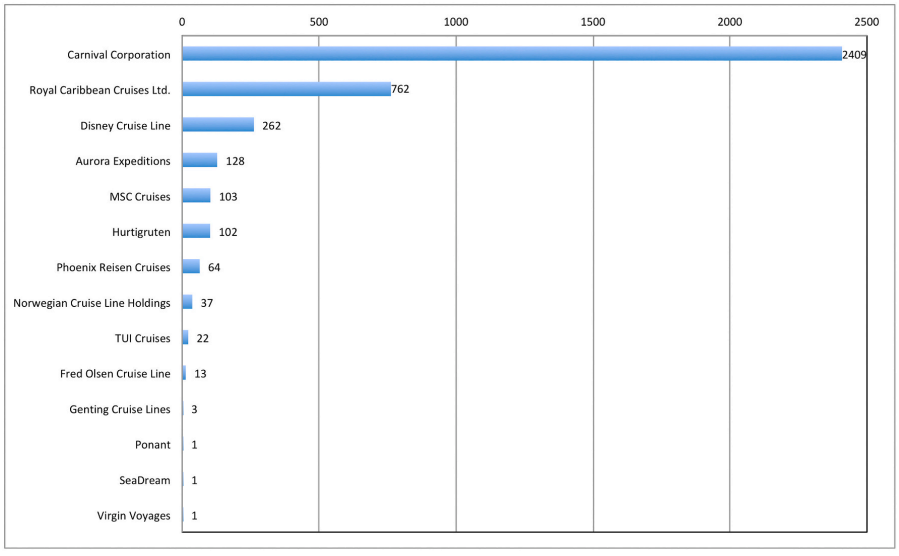


Fig. 1. Number of COVID-19 cases by Cruise Company (Data source: Miami Herald last updated on 2020-10-02 (Miami Herald, 2020)).

integration theory (IIT) states that new information is added into the preliminary beliefs of people, affecting how attitudes and behavior are formed, and the value of information determined to be favorable or unfavorable regarding the events and subjects involved (Anderson, 1981). This theory provides one of the theoretical frameworks for this study.

The amount of mass information generated by social media creates awareness and a forum for discussion during times of crisis (Sigala, 2011). This is because social media facilitates real-time interactivity, reciprocity, and instant reactions between users, hence it plays a central role when crisis events happen and in crisis management in tourism with potential effects on tourists' behavior (Schroeder, Pennington-Gray, Donohoe, & Kiousis, 2013; Sigala, 2011; Zeng & Gerritsen, 2014). International tourists have been found to have a high probability of turning to social media for information during periods of crisis (Schroeder et al., 2013). The electronic word-of-mouth (eWOM) shared on social media by members of the public during a crisis is a curated heritage or collective memory that can be explored to learn about the crisis (Liu., 2009). This fact supports the suitability of using social media to explore crisis effects. Hence, tourism scholars have used social media data to investigate health-related crisis (Yu, Li, Yu, He, & Zhou, 2020), political crisis (Luo & Zhai, 2017), natural disasters (Möller, Wang, & Nguyen, 2018), climate crisis (Schweinsberg, Darcy, & Beirman, 2020), service failures (Su, Stepchenkova, & Kirilenko, 2019), pest infestation (Liu, Pennington-Gray, Donohoe, & Omodior, 2015), and terror attacks (Barbe, Pennington-Gray, & Schroeder, 2018).

In cruise research, past literature on the role of social media during crises has focused on its impact on the corporate reputation of cruise companies. Ryschka, Domke-Damonte, Keels, and Nagel (2016) found that a prompt corporate response on social media and strong brand familiarity during a crisis led to positive perceptions of the cruise company. Penco, Profumo, Remondino, and Bruzzi (2019) found that the "degree of anger" and prior corporate reputation during a crisis influenced the intention to take a future cruise. While Penco, Profumo, and Remondino (2018) found that social media may help cruise companies to communicate more effectively during a crisis.

2.2. Media coverage, public opinion, and risk perception

The news media influences public opinions through its news coverage (Gene Zucker, 1978). In the absence of direct personal experience, members of the public learn about risks from other people and from the media (Kasperson et al., 1988). Therefore, as the media informs the public about happenings beyond their immediate circle, it shapes the public's perception and understanding of risks (Rowe, Frewer, & Sjöberg, 2000; Smith, 2005).

The social amplification of risk framework (SARF) is a seminal integrative and interdisciplinary risk perception framework that has been used to account for findings from a wide range of fields, from media to medical research (Kasperson, Kasperson, Pidgeon, & Slovic, 2003; Wang, Zheng, & Zuo, 2021; Womack, Anderson, & Ledford, 2020). We combined the SARF with the IIT to explore how risk perceptions during COVID-19 influenced potential tourist attitudes and behavior about cruising, as illustrated in Fig. 3, which is an adaptation of the simplified representation of the SARF framework.

The SARF framework suggests that the social amplification of risk by the media increased its memorability and imaginability, which leads to increased risk perception (Kasperson et al., 1988). Four attributes of the information flow from the media influence the extent of social amplification. They include volume of information flow, disputability of the information, degree of dramatization, and symbolic connotations (Kasperson et al., 1988). It is worthy of note that high volumes of information flow about a risk attract public attention, mobilize latent fears about the risk, and trigger the recollection of previous failures or accidents (Kasperson et al., 1988; Renn, 1986). Degree of dispute deals with how much the facts about the event are disputed. Dramatization, usually in the form of sensational headlines, is an important attribute of information that increases perceived risk and memorability of the incident. While symbolic connotations deal with the specific terms used to convey information and how they may have other meanings and be interpreted by individuals and groups (Kasperson et al., 1988). Likewise, there are four attributes of the interpretation and response mechanism in the second stage of social amplification: heuristics and values which deals with the simplification of risk, social group relationships which deals with politicization and/or polarization, signal value on the seriousness

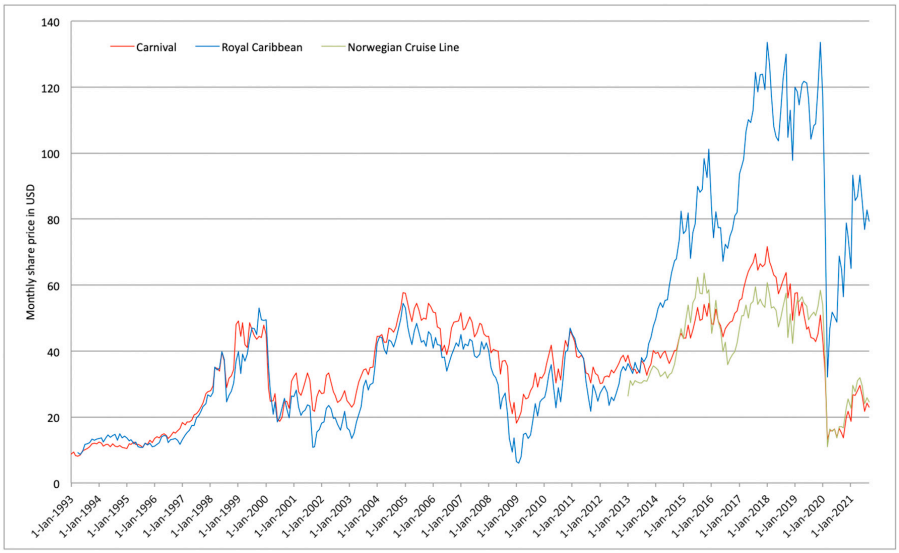


Fig. 2. Share prices of the three largest cruise operators from January 1993 to September 2021 (Data source: Eikon, downloaded on 2021-09-07).

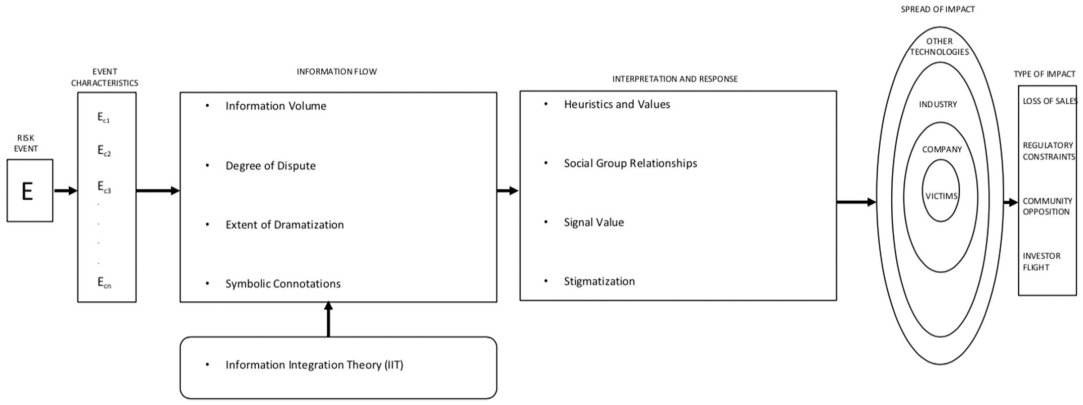


Fig. 3. Proposed conceptual map of SARF and IIT adapted from Kasperson et al. (1988).

of the risk, and stigmatization (Kasperson et al., 1988).

3. Method

3.1. Data collection

We analyzed the tweets from the start of the pandemic through the period of the early outbreaks on cruises, highlighted by an increased publicity on cruising, which was pinpointed using Google search trends. Google Trends has been demonstrated to show the popularity of a search query as an indicator of real public interest (D'Avanzo, Pilato, & Lytras, 2017). Fig. 4 shows that the worldwide search queries for selected relevant phrases like "cruise coronavirus" and "cruise ship coronavirus" rose from zero between January 19–25 and returned to zero around May 3–9,[1] a period coinciding with the major outbreaks (Google Trends, 2020). "Cruise covid" rose from zero after February 11, the day the International Committee on Taxonomy of Viruses (ICTV) officially named the new virus SARS-CoV-2 and the disease it causes COVID-19 (WHO, 2020).

Therefore, we used the ninth version of the open-source COVID-19 dataset of tweet IDs by Banda et al. (2020). The data covers the period from 4th January to 10th May 2020 and includes our period of interest, to capture the tweets posted about the COVID-19 outbreaks on cruises at the time of the major outbreaks at the start of the pandemic when the outbreaks dominated news and social media. It was the time when the attention on the cruise COVID-19 outbreaks were highest, as indicated by excessive mentions of cruise and COVID-19, and pinpointed with Google Trends. This data was gathered by collecting all tweets on the novel coronavirus from the Twitter application programming interface (API) stream by filtering the API stream with relevant keywords connected to the pandemic as stated in the dataset publication (Banda et al., 2021). These keywords include, among others, "coronavirus", "2019ncov", "corona virus", "COVD19", "CoronavirusPandemic", "COVID-19", "2019nCoV", "CoronaOutbreak", "WuhanVirus". The authors of the dataset provide a full and a clean version of the data; with the clean version comprising only original tweets without retweets. We used this clean version, which contained 66,538,356 tweet IDs because retweets contain the text of the original tweet and may disrupt our intended natural language processing tasks (Banda et al., 2021).

We rehydrated these tweet IDs to retrieve the full tweets that had not been deleted for analyses using TWARC version 1.8.3 (Summers et al., 2021) because deleted tweets are unretrievable. The rehydration process lasted about 186 h. We successfully retrieved 60,483,491 full tweets (9.1% deleted), and the resulting JSONL (JavaScript Object Notation Lines) file was 263 gigabytes. The retrieved tweets were in over 64 languages, with the majority (57%) in English (34,431,646). Due to difficulties in translating such a huge number of languages, we restricted our analyses to only the tweets in English.

The required information (posting time, name of Twitter handle, text of tweet, number of retweets, and number of likes) from each tweet were parsed from the tweet json files and loaded into a Python Jupyter Notebook (Kluyver et al., 2016; Van Rossum & Drake, 2009) for analysis using Pandas DataFrames (McKinney, 2010). Preprocessing tasks like tokenization (breaking of words to smaller units), lemmatization (reduction of words to their root), and removal of unnecessary elements like stop words, URL links, and mentions (@user) were carried out. We also split and removed the hashtag signs (#) and the newline characters (\n) from the tweet text before analysis. The data preprocessing is a necessary step in order to clean the raw, noisy data before advanced analytical processes. We searched for the occurrences of cruise(s), cruise line(s), cruise ship(s), and cruising and got 139,054 tweets. We randomly checked the obtained tweets manually to ensure that there were no false positives or irrelevant tweets in the data, as done in previous studies (Ainin, Feizollah, Anuar, & Abdullah, 2020) by exporting the tweets as a csv file and checking them in Microsoft Excel. This manual inspection revealed that there were 586 tweets about actor Tom Cruise halting a movie shoot because of the pandemic; which were then removed.

3.2. Data analysis

This study used sentiment analysis as the main analytical method to automatically classify the tweets into positive, neutral, and negative. Sentiment analysis is the automated process of opinion detection by using semantic relationships to determine the overall polarity of a text document as positive, neutral, or negative (Alaei, Becken, & Stantic, 2019; Feldman, 2013). The rapid growth of sentiment analysis has coincided with the growth of social media, which has provided enormous volumes of digital opionionated data (Liu, 2012). Therefore, social media sites like Twitter and Facebook are a centre of interest for sentiment analysis applications (Feldman, 2013). Social media offers tourists a platform to share their views, feelings, and sentiments about their experiences (Li, Xu, Tang, Wang, & Li, 2018). Sentiment analysis is one of the most common techniques for analyzing online text data on social media (Khong, Teng, Butt, & Muritala, 2021; Muritala, Sánchez-Rebull,

[1] The unit of Google Trends data for the past year is given in 6-day intervals

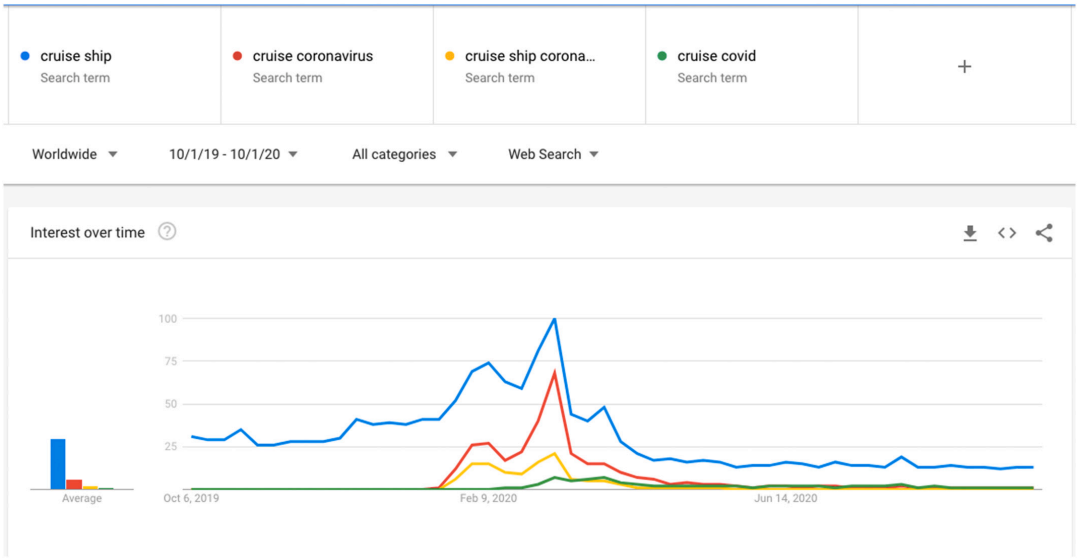


Fig. 4. Google search trends for cruise ship, cruise coronavirus, cruise ship coronavirus, and cruise covid (Data source: Google Trends).

& Hernández-Lara, 2020) and at a relatively low cost (Neri, Geraci, & Camillo, 2010). Sentiment analysis and opinion mining are used interchangeably (Feldman, 2013), although the use of the term "opinion mining" has been declining in the tourism and hospitality literature (Muritala et al., 2020).

However, because of its unstructured form and noisiness, social media data presents some challenges for automated sentiment analysis. Sentiment analysis algorithms are usually not equally effective across a variety of languages, hence, it is usual for researchers to restrict analysis to a single language, leading to loss of information (Li et al., 2018). The noisiness of social media with liberal use of slangs and lots of spelling, grammatical, and punctuation errors makes accurate analysis more difficult (Feldman, 2013). Sentiment analysis is also domain-dependent, with positive words in one domain, not positive in another (Pang & Lee, 2008). Identification of sarcasm also poses a challenge to sentiment analysis systems (Feldman, 2013) with Twitter exhibiting rich sarcasm (Zhang, Zhang, & Fu, 2016). Despite these challenges, sentiment analysis remains very useful for researchers and industry practitioners for analyzing large amounts of data compared to manual processing (Philander & Zhong, 2016).

Kietzmann, Hermkens, McCarthy, and Silvestre's (2011) seven functional blocks of social media provide the conceptual framework for understanding the sentiment analysis results. Twitter is a conversation-based platform (Kaplan & Haenlein, 2011), hence, the relevant block for this study is the conversation block of the framework. The conversation block posits that to make sense of the conversation; the conversation velocity, i.e. the frequency of new conversations over time, and the change in how favorable or unfavorable the sentiment they contain needs to be analyzed in order to understand the conversation (Kietzmann et al., 2011). To make sense of the conversation on cruising during the COVID-19 outbreaks on cruise ships, the frequency of new conversations is the frequency of tweets on the topic over the analyzed period and the sentiment analysis score measures the change in sentiment of the tweets. This informed the decision to make a plot of the daily frequency of cruise-related tweets and their average daily sentiment in Fig. 6, to show how the frequency and sentiment of the conversation towards cruising changed during the analyzed period.

Apart from sentiment analysis, we construct a word cloud based on the 500 most frequent words and bigrams (two-word combination), and also present the three tweets with the highest amount of engagement. Tweets have unequal reach, and the engagement metrics on a tweet indicate the amount of people that saw and interacted with the tweet. We selected the most engaged tweets by checking the tweets with the highest number of likes, because viral tweets usually have more likes than retweets.

The word cloud of the most frequent words was constructed using WordCloud version 1.5.0 package in Python (Mueller, 2020). A simple ranking of the tweets based on their number of likes produced the most engaged tweets. Finally, the sentiment analysis was performed using VADER (Valence Aware Dictionary and sEntiment Reasoner) (Hutto & Gilbert, 2014) implementation in Python's NLTK (Natural Language Toolkit) (Bird, Loper, & Ewan, 2009). VADER is well suited for analyzing social media text since it handles many of its typical elements like acronyms, emojis, and slangs well. VADER assigns a score to the analyzed text based on the summation of valence scores of the words in a tweet between 1 (extreme positive sentiment) and − 1 (extreme negative sentiment), while neutral sentiment spans 0.05 to −0.05 (Hutto & Gilbert, 2014). It is also important to do a manual validation of the sentiment classification results to confirm that the assigned sentiment scores reflect the content of a tweet. This validation effort showed that there were some negative tweets given positive scores. These tweets were about reports of people who tested positive for the coronavirus, and "positive" as a positive word, shifted the sentiment weighting of these tweets to positive, even though this carries a negative sentiment in this context. We adjusted this by replacing the word "positive" with "infected" in these tweets, which made the algorithm to reflect the sentiment more accurately.

4. Results

138,468 cruise-related tweets posted by 58,644 Twitter accounts were obtained. The mean and median number of tweets per account was 2.37 and 1 respectively, showing a likely preponderance of tweets from personal accounts, compared to news organizations or blogs that usually

have multiple posts reporting the same information.

4.1. Word cloud

The word cloud of the 500 most frequent words and bigrams in the analyzed tweets is presented in Fig. 5. Fig. 5 shows the names of some of the cruise ships (*Diamond Princess, Grand Princess, Ruby Princess, Westerdam*), company/brand names (Carnival, Royal Caribbean, Norwegian Cruise (Line), Holland America), and places (Japan, US, China, Australia, Hong Kong, Italy, Cambodia, UK, Malaysia) that featured prominently in the Twitter conversation during the examined period. There are also many names related to the US like former president Donald Trump, the CDC, as well as American cities that were associated with some cruise COVID-19 outbreaks (California, San Francisco, Florida, New Jersey, New York). Petri dish was one of the most frequent bigrams visible on the word cloud because several tweets referred to cruise ships as "petri dishes" or "floating petri dishes", a derogatory cliché about cruise ships being a fertile platform for infectious diseases. Similar words or bigrams that reappear on the word cloud i.e. "test positive", "tested positive", and "tests positive" or different words that refer to the same entity like US, USA, and America are because the word cloud algorithm counts any change in spelling separately, and this shows the noisiness of social media data. We did not combine these similar words to avoid introducing any bias.

4.2. Most-engaged tweets

The three tweets with the highest amount of engagement (retweets, quotes, and likes) were uncomplimentary remarks about the cruise industry by influential public personalities in the US with verified Twitter accounts and many followers (see Fig. 6). The first was by Judd Apatow, an American film director with over 2.4 million followers, whose tweet had over 9300 retweets and quote tweets and over 75,800 likes. The second was by Sheldon Whitehouse, an American senator with over 468,000 followers. His tweet, which also shared a *Washington Post* article, had over 10,400 retweets and quote tweets and over 22,600 likes. While the third was by Bill Maher, an American talk show host with over 11 million followers, who shared a video of an episode of his talk show about the cruise industry. The tweet had more than a thousand retweets and quote tweets and over 4600 likes, while the attached video was viewed over 234,500 times.

Coincidentally, these three tweets were posted within a week from each other during a period when there were rumors that cruise companies would be included in the COVID-19 government bailout package following former President Trump's tweet on 12th March 2020, in which he described the cruise industry as a "great and important industry" that will be kept that way. These rumors turned out to be untrue but as these tweets show, there was opposition to the idea on the basis that cruise companies incorporate outside the US and sail under flags of convenience (FOC)[2] to avoid paying taxes in the US and pollute the environment.

4.3. Sentiment analysis

The sentiment analysis result is presented on a dual-axis time series plot of the daily frequency of cruise-related tweets and the average daily sentiment of these tweets in Fig. 7. The highest volume of tweets per day were posted during the outbreak on the *Diamond Princess* with the two highest peaks in February 2020. The negative sentiment between January 26 and February 3 was due to the cancellation of cruises with Chinese port of calls and fears about the spreading novel coronavirus. However, from February 3rd with the outbreak on the *Diamond Princess* and subsequent cruise outbreaks, the negative sentiment was mainly because of the negative news stories on new confirmed cases, hospitalizations, passenger deaths, stranded ships, and people's reactions. The negative sentiment about any specific incident was usually persistent long after the news initially broke on Twitter because of repetition and round-the-clock coverage of the 24-h news cycle, resulting in the predominantly negative sentiment throughout the analyzed period. The only time the sentiment crossed the neutral line was on the release of a CDC report on March 24 showing that the novel coronavirus survived in the *Diamond Princess* cabins for 17 days after passengers disembarked. The tweets about this report had a positive sentiment scores because the verb "survive" is a positive word.

Fig. 8 shows the proportion of positive, neutral, and negative tweets. 67,022 (48%) of the tweets in our dataset had a negative sentiment score, 41,636 (30%) had a neutral sentiment score, and 30,352 (22%) of the tweets had a positive sentiment score. A random sample of the positive, neutral, and negative tweets is presented in Tables 1 to 3, with the time they were posted and their sentiment scores. We do not include account details to preserve the privacy of the individuals that posted them. The tweets were randomly selected using random numbers generated by the RandArray function in Microsoft Excel.

The positive tweets in Table 1 show that the algorithm had difficulty identifying sarcasm and had some positive-bias, i.e. tweets scored more positively than they were. For example, tweet numbers 3, 4, 5, 9, 17, and 20 are sarcastic negative tweets scored as positive. While tweet numbers 2, 7, 8, 10, 12, 13, 15, and 19 are not actually positive towards cruising. Genuinely positive tweet number 1 and 18 were about an *MS Westerdam* cruise that stayed virus free, which departed Hong Kong on February 1, 2020 and was prevented from docking in five countries before eventually docking successfully in Cambodia. The neutral and negative tweets in Tables 2 and 3 seem to have been scored more accurately.

4.4. Revised conceptual framework

A revised conceptual framework of the combination of the SARF and IIT is provided in Fig. 9 based on the results of our analysis. The biggest change is that the ripple effects from the spread of impact do not have a "other technologies" dimension like in the original framework, with the effects stopping at the industry level. The entire industry suffered economic reversal although some companies had more outbreaks on their ships than others as shown in Fig. 1. Aspects of the information flow, interpretation and response, and the information integration theory (IIT) that are noteworthy or new are highlighted in red. For example, the 24-h news cycle under the information volume attribute and social media sources under the IIT are relatively recent phenomenon that have now been incorporated into the conceptual model.

5. Discussion and conclusions

This study investigated the public perception towards cruising during the COVID-19 outbreaks on cruises in the early part of the COVID-19 pandemic. To answer the research question on the insight that can be derived from the analyzed tweets during the cruise ship COVID-19 outbreaks, the results show that there was an overwhelming negative sentiment in a majority of the tweets. While this finding was expected, it is still an important result with implications for the cruise industry. This study is in line with the results of previous research that has shown the influence of media coverage on public opinion and risk perceptions (Gene Zucker, 1978; Rowe et al., 2000; Smith, 2005), supporting also the assumptions of SARF that suggests that the spreading of information that involves risk amplifies its perception. The media amplification of the cruise COVID-19 outbreaks had the four attributes of the information flow, which increase memorability and risk perception (Kasperson et al., 1988). The first attribute was are the massive volume of information flow on the incidents from the 24-h news cycle. The second was the degree of dispute element from a fast-spreading contagious disease that

[2] Registration of a ship in another country to avoid restrictive labor and tax regulations in the shipowner's country.

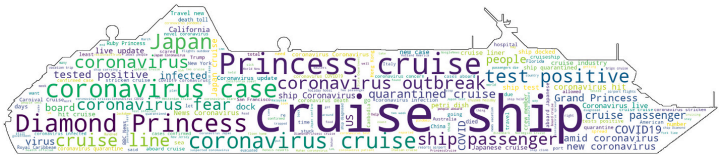


Fig. 5. Word cloud of the 500 most frequent words and bigrams.

Fig. 6. Top-three most engaged tweets (Source: Twitter, 2020).

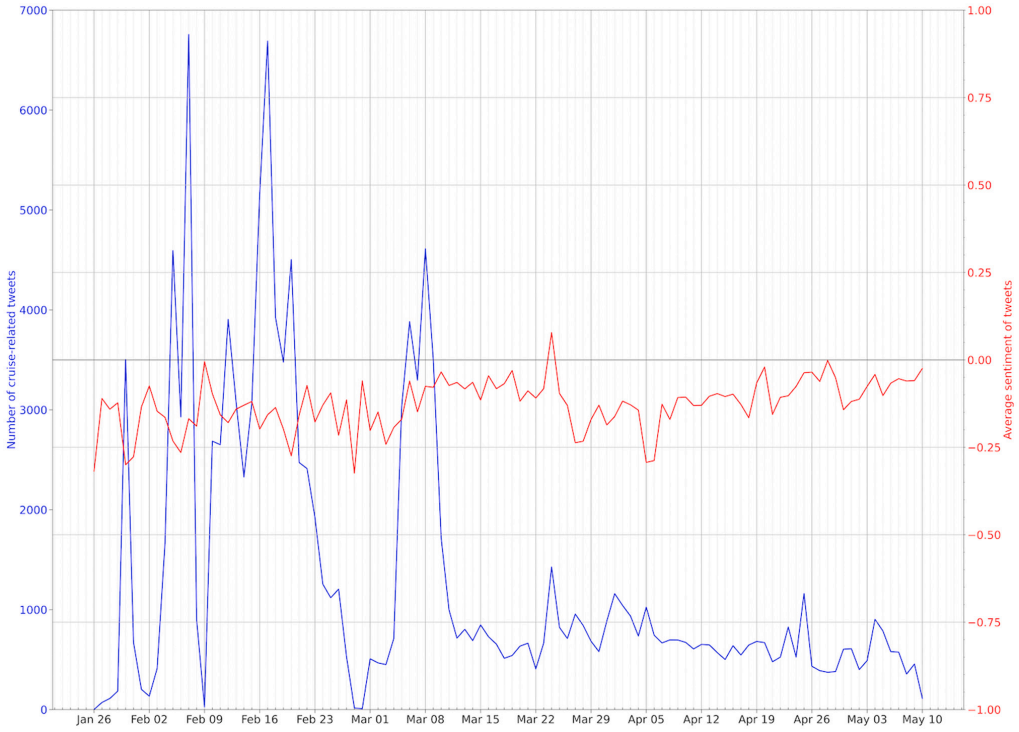


Fig. 7. Daily frequency and average daily sentiment of cruise-related tweets.

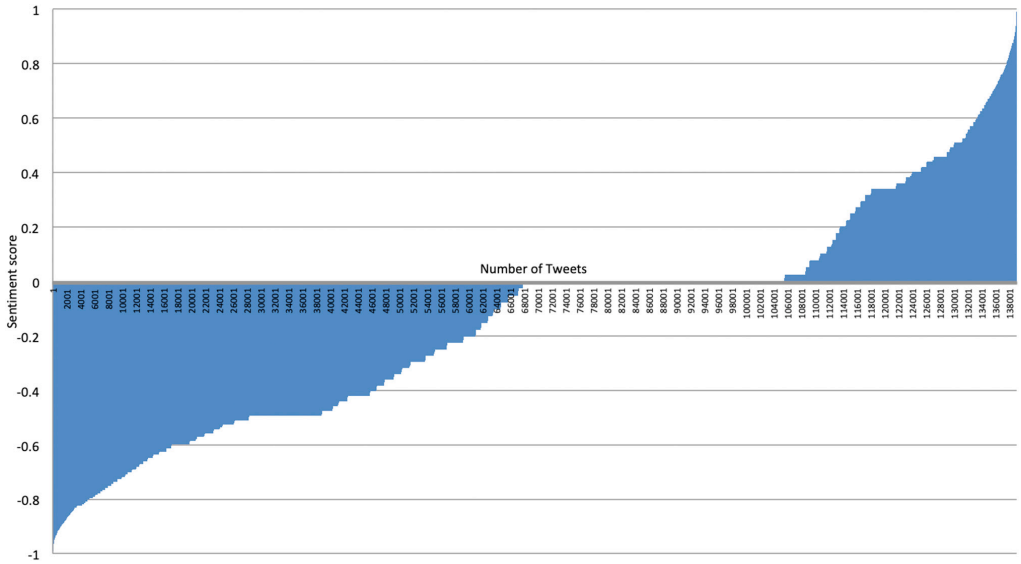


Fig. 8. Chart showing the proportion of positive, neutral, and negative tweets.

was poorly understood at the time, spreading quickly in the vulnerable environment of a cruise ship. The sensational headlines about the increasing number of cases and deaths from the outbreaks provided the element of dramatization. The symbolic connotations were usually negative in these stories, with intrinsically negative words, such as hell or prison, used to refer to ships with outbreaks or under quarantine. Under the attributes of the interpretation and response phase, the simplification of the risk under the heuristics and values boiled down to the risk of disease and death. There was an element of politicization of social group relationships with the three most engaged tweets as they were reactions to the rumor that cruise companies would be offered government bailout which started after the former president's tweet about the cruise industry on March 12. There was unmistakable high signal value from an unknown risk of a deadly disease and the stigmatization of the cruise industry was obvious in the recurrence of words like Petri dish in referring to the industry's ships. SARF also emphasizes the difficulty in changing risk perceptions once formed, causing perceptions that can endure far into the future (Kasperson et al., 1988). Hence, these factors underline the importance of this finding, as there is a possibility of an enduring perception of cruising as an undesirable or risky activity even after the threat of COVID-19 goes away. This finding is also consistent with the previous study, which found that cruiser and non-cruiser respondents had anxieties about cruising and were negative about taking a cruise in the future (Holland et al., 2021).

The results also show that criticisms directed towards the cruise industry usually referenced people's previous objections to the cruise industry on various issues such as infectious disease outbreaks before the pandemic, environmental impact and sustainability, sailing under flags of convenience (FOC) to avoid tax and laws, size of cruise ships, crime incidents, and crew working conditions. This finding is consistent with the assumptions of IIT, which highlights that new information is added and integrated into existing beliefs and knowledge, influencing people's attitudes and behavior. It also agrees with the literature on the impact of high volumes of information flow, which mobilizes latent fears and a recollection of previous issues, accidents, or failures (Kasperson et al., 1988). Thus, the high volumes of information through repeated stories and round-the-clock coverage of the COVID-19 outbreaks on cruise ships triggered the recollection of previous concerns about cruising. Cruise researchers have discussed many of these issues at length before the pandemic (Klein, 2002, 2009, 2011, 2016a, 2016b; Papathanassis, 2016; Terry, 2017).

Our results contribute to previous studies showing the hierarchical dimension of the public conversation with the influence of elites, which is also exhibited on social media. The three most-engaged tweets by public personalities confirm this hierarchical nature, where elites shape public opinion as reported in the literature (Blumer, 1948). Tweets by public personalities can earn an outsized amount of engagement because of their societal prominence and a substantial number of followers. Furthermore, the fact that these tweets were by American personalities, in addition to several American-related entities appearing in the word cloud, shows that the analyzed tweets were very American-centric. This could be explained as evidence of strong American interest in cruising matters, although it could have also been affected by the fact that this study only analyzed English tweets. Besides, industry data shows the United States was the largest source of cruise passengers by far with 11.9 million passengers or approximately 40% of global cruise passengers in 2019 (CLIA, 2019, 2021). Overall, this shows the central importance of the American market to the cruise industry.

It is noteworthy that some of the criticized aspects of cruising reported in the results are a consequence of the mass-market cruise tourism business model. This business model uses cost reductions to enable mass-market cruises on increasingly larger ships (Vogel, 2017) which produce three times more greenhouse emissions than long haul planes (Lloret, Carreño, Carić, San, & Fleming, 2021; Mozuni & Jonas, 2016) and sail under FOCs, which frees the cruise lines from paying hefty taxes and many labor, environmental, and safety regulations (Terry, 2017), while externalizing environmental and social costs to the society (Klein, 2016b). Cost saving through FOC has also enabled the cruise industry to offer the affordable mass-market product to middle-class customers, and without it, ships would be smaller and cruises would be more expensive (Terry, 2017). However, despite these economic advantages, this business model has not been hugely successful without the subsidy afforded by the lucrative cruise onboard business (Vogel, 2017). An analysis of the finances of the biggest cruise operators

Table 1
Random sample of positive tweets.

	Time	Tweet	Sentiment Score
1	2/15/2020 16:10	Love it when it all works out good… Wandering ship becomes best cruise ever despite coronavirus fears https://t.co/db7FOK1e2O	0.9062
2	3/27/2020 23:22	.@IHG It would be really nice to get a refund for my reservations in June, seeing as the cruise I booked last August is now canceled. Are people really still visiting Seattle? #Coronavirus sucks but you don't have to!	0.0763
3	3/15/2020 21:44	To all you good Christians out there please say a prayer for the CRUISE INDUSTRY. They need ALL OUR OUR THOUGHTS and SUPPORT. Also WALL STREET, SAUDI ARABIA, and PUTIN. They've all suffered so much. #trump #trumpspeech #trumpspandemic #coronavirus #trumpvirus	0.6633
4	4/14/2020 14:16	@i9998217 @CruiseIndustry I think they'd be unsuitable for those purposes since cruise ships have proven themselves to be petri dishes for #Covid19. They'll make wonderful reefs! ðŸŽ	0.5719
5	2/7/2020 15:10	#pdknews Good news NYC/New Jersey, a cruise ship with 12 passengers who have the Coronavirus just pulled into port. https://t.co/R3dSYH0pEe	0.4404
6	4/12/2020 21:14	FYI followers this article offers a good explanation regarding FOC Flag Of Convenience Ships this is the Cruise Ship Industry this applies also to applies to 99% of ships sailing around Australia on our coastal shipping trade, the same for all ships taking our exports O/S https://t.co/zvLFp0u9jJ	0.6633
7	4/17/2020 14:31	I will never, ever go on a cruise ship again. You could not pay me enough to endanger my life on these floating petri dishes. https://t.co/4p5pibjhY2	0.0762
8	2/5/2020 07:40	@TIME Reason number 10946 not to vacation on a cruise ship. #coronavirus #ncov #Japan	0.0772
9	4/16/2020 05:12	@MollyJongFast But Molly, cruise ships are a wonderful place for Americans to congregate and spread love and cheer and Covid-19 and...	0.9042
10	3/24/2020 15:26	@SenTedCruz @JohnCornyn @RepRWilliams cruise lines do not pay US taxes nor are they US corporations. DO NOT BAIL THEM OUT. https://t.co/3ZBtcJ9Jdz	0.0762
11	2/15/2020 12:55	Be thankful this Saturday morning that you are not stuck on a cruise ship. #CoronaVirus #DiamondPrincess https://t.co/yVnRHhFpQ1	0.6641
12	4/9/2020 06:23	@JohnnyJet Ya know, I wand much interested in cruises because of norovirus. Definitely not interested anymore after one Coronavirus outbreak.	0.7964
13	3/25/2020 02:57	Plus my cruise line is only giving "cruise credits" not refunds for a future cruise they will jack the price up on. #UGH #COVIDIOT45 https://t.co/DuOLYpHTyU	0.0943
14	2/7/2020 05:34	@DisneyCruise Are you taking any precautions regarding coronavirus on your ships? We are scheduled to sail in a few weeks. TIA Our family has our Flu vaccines already ðŸŽ	0.347
15	4/14/2020 14:01	Long before Coronavirus, you couldn't have paid me to take a cruise. Hopefully a infected legacy of the pandemic will mean others won't have to pay again either https://t.co/otYaO58Vfg	0.4215
16	2/11/2020 00:35	As a reminder, we will not have an episode tomorrow, since Jordan is going on a cruise. Let's hope he doesn't get coronavirus!	0.4404
17	4/18/2020 00:33	@smh Coronavirus and Cruise ships a perfect match. https://t.co/rrXwX2Lwx2	0.5719
18	4/21/2020 05:19	The tale of The last cruise ship on Earth. "We became like a family - our guests and our crew together. The spirit has been beautiful."	0.7964
19	4/2/2020 21:20	These "healthy" cruise ship passengers arriving in Ft. Lauderdale now have been exposed and are required to self-quarantine for 14 days once they get home. On the way home, however, they will potentially expose plenty of other travelers. #COVID19 #Coronavirus #Zaandam #Rotterdam https://t.co/ZVSVX8rqYK	0.2023
20	2/10/2020 17:03	At least the price of cruises are going down. And you wouldn't believe the deals on bat soup. Thanks #coronavirus https://t.co/mjwVveZxhI	0.4404

between 2001 and 2014 shows that their total revenues are becoming less profitable (Vogel, 2017) and it can be observed that the share prices of the big three had either stagnated (Royal Caribbean and Norwegian Cruise Line) or was declining (Carnival) before the pandemic (see Fig. 2). Cruise researchers have long questioned the increasing size of modern cruise ships (Klein, 2016a, 2016b; Papathanassis, 2016). For example, the world's largest cruise ship, Royal Caribbean's *Symphony of the Seas,* has 18 decks, measures 362 m, and can carry almost 9000 passengers and crew (CNN, 2018). Researchers have also linked the COVID-19 outbreaks with cruise ship size: "It is clear that cruise ships infected with COVID-19 are large ships" (Ito et al., 2020, p. 5).

The year 2020 started promisingly with the cruise industry preparing for a period of unprecedented boom, with 117 new ships on order by 2027 (Cruise Industry News, 2020). However, the COVID-19 pandemic has now reversed the industry's fortunes. But as the development of COVID-19 vaccines and antiviral drugs signals a path to the end of the pandemic; the cruise industry needs to do a lot of rethinking to make a successful recovery. The findings of this study and that of previous research (Holland et al., 2021) show that the outbreaks of COVID-19 on cruises and the ensuing media coverage has hurt confidence in this form of tourism and could hamper future growth if the negative perception persists. Reports of full bookings for future cruises do not contradict these findings, since these reported bookings are usually not always new bookings. For instance, in the quarter ending on August 31, 2020, Carnival Corporation filings show that only 55% of these bookings were new bookings, while the rest were future cruise credits (FCC) re-bookings from previously canceled cruises during the pandemic (Carnival, 2020). The cruise industry has a reputation for discounting prices enough to stimulate demand to fill their fixed capacities after external shocks like 9/11 or the 2008 financial crisis (Vogel, 2017), but this line of action could further reduce profitability.

Therefore, the cruise industry needs a new paradigm away from the pre-pandemic mass-market model with diminishing profitability and whose consequences are stoking negative perceptions of cruising towards more sustainable, environmentally friendly, and profitable business models. This would not be a simple task for the cruise industry and could be further complicated by the huge debt taken on by the cruise companies to survive the pandemic. For example, Carnival has raised $23.6 billion through debt and equity so far during the pandemic (2021-04-23: Financial Times, 2021). However, it is difficult to see these challenges successfully resolved with the industry moving in the same direction as it was before the COVID-19 pandemic.

5.1. Theoretical implications

The theoretical contribution of this study is both the proposed conceptual model in Fig. 3 and the revised conceptual model in Fig. 9, combining SARF and IIT. Kasperson et al. (1988) provided a detailed framework, which is quite complex and seems highly suited to extreme events. For example, they listed social protest and disorder as part of the

Table 2
Random sample of neutral tweets.

	Time	Tweet	Sentiment Score
1	3/5/2020 18:16	Man, going to be tough refusing some of these cruise ship deals that will be popping up. #COVID19	0.0493
2	4/27/2020 01:34	Cleaning a floating petri dish: How is a cruise ship sanitized after a coronavirus outbreak? https://t.co/cyixhGrB38	0
3	2/14/2020 07:13	Well this seems like a very bad time to take a #Cruise #Coronavirus #COVID19 https://t.co/unVsTtCdxO	−0.0498
4	4/27/2020 04:11	Coronavirus watching you board a cruise ship https://t.co/9dnwf4Zol3 https://t.co/RcJ7CNUZto	0
5	3/5/2020 05:11	'Cluster' of 21 people on Princess Cruise ship heading towards Calif. show possible coronavirus symptoms: officials https://t.co/00uus5Jwme #FoxNews	0
6	2/5/2020 06:05	Quarantined on a cruise ship, food running out. #coronavirus https://t.co/vkAWZBmAIR	0
7	4/4/2020 19:16	I ain't cruising until 2022 https://t.co/NtMZa1mGjC	−0.0267
8	2/11/2020 06:00	Going on a cruise? Here's how the coronavirus will change your trip https://t.co/hYVCOx2Egt	0
9	4/27/2020 05:43	Coronavirus: How did Australia's Ruby Princess cruise debacle happen? https://t.co/5pFLnZ5WGX https://t.co/YSpCrXKFcK	0
10	2/11/2020 06:53	Japan might test everyone on cruise ship for coronavirus https://t.co/PMkvBFCpLq via @CBSNews	0
11	4/27/2020 02:39	Two things not to do in a pandemic… book a cruise, download a tracking app… https://t.co/keMd4UzJJ7	0
12	3/25/2020 19:29	The size of cruise ships is mad. This pandemic hopefully will result in a lot of rethinking, not just about cruise ships. But travel in general. https://t.co/v3KFQjCbmI	−0.0444
13	2/6/2020 15:33	Just saw an article about cruise ships being quarantined due to coronavirus and it's a great reminder that there's no reason to ever go on a cruise.	0.0344
14	4/26/2020 22:22	Can you imagine not being able to disembark from a cruise for more than 40 days? @fdilella spoke with @dan_domenech who has been stranded due to strict CDC guidelines, amid the Coronavirus pandemic. Here is his story: https://t.co/oGqsuAcq5N	0
15	4/26/2020 23:35	[Feed]Coronavirus journey: The 'last cruise ship on Earth' finally comes home https://t.co/3wH1wuYkzm	0
16	4/30/2020 03:43	Sixty-seven new cases of #Coronavirus were confirmed on the DiamondPrincess cruise ship, bring the total 285. The ship docked in Yokohama, Japan remains quarantined. https://t.co/BfncVKGIt0 #WednsdayWisdom	0
17	2/11/2020 07:15	American coronavirus patient describes "surreal" cruise experience - CBS News https://t.co/fase1Stk7T #cruise #travel	0
18	4/27/2020 02:48	@ABCWorldNews What about the guests or shall I say COVID-19 patients on that cruise ship? Hmmmm.	0
19	2/11/2020 06:24	How a cruise ship turns into a coronavirus breeding ground https://t.co/HfCyHZHgWy https://t.co/J4fkKCjbQd	0
20	2/5/2020 06:11	NBC News: 10 coronavirus cases confirmed from cruise ship quarantined in Japan. https://t.co/tjTmcV632f via @GoogleNews	0

Table 3
Random sample of negative tweets.

	Time	Tweet	Sentiment Score
1	3/24/2020 11:09	Until cruise ships are entirely sustainable and powered by renewable energy they're utterly at odds with our climate change crisis, let alone the Coronavirus Cruises they've become. https://t.co/9dpR49UrDy	−0.6124
2	1/27/2020 23:29	My elderly mother is scheduled to depart this week on a 28 day cruise and ALL ports have cases of coronavirus. She called @hollandamerica to see if she could receive a refund or voucher. The answer was NO. #coronavirus #badcustomerservice	−0.4466
3	5/9/2020 22:42	imagine surviving coronavirus, taking a tropical cruise to celebrate, and then puking and shitting yourself to death in a "stateroom" the size of a closet because you got norovirus from the all you can eat dessert bar. Just let the plague ship industry die. https://t.co/jYoeiVDF9z	−0.8176
4	2/10/2020 18:54	@billburr So ummm. Ya know that whole sinking Cruise Ships idea? I think now would be a good time to start.. with the Coronavirus and all.	−0.4404
5	5/5/2020 09:07	Make that forever. Cruises are an environmental catastrophe on every dimension, fiscally irresponsible operating under flags of convenience if not connivance and a health hazard even in the best of times. https://t.co/0TsLNQrhl2	−0.4767
6	4/23/2020 06:43	The cruise industry shouldn't recover. It's bad for the seas and the planet.	−0.3482
7	2/6/2020 01:20	You could not pay me enough money to go on a cruise. Coronavirus aside, these stories of ships losing power and everyone getting sick and waste all over the place. No way in hell. https://t.co/GOdXb95sFv	−0.9349
8	5/9/2020 17:20	.@NPR @KHNS_FM There were already outbreaks of illness on these vessels before coronavirus. Now with a number of outbreaks in assorted US states traced back to cruises, and their off-shore registration to avoid taxes, I'm not sorry to see the industry decline.	−0.5232
9	5/5/2020 06:11	what if we stop taking cruises altogether since they're gross af? https://t.co/4FIHUIGJFO	−0.6486
10	3/9/2020 23:51	Grand Princess Passengers horror on coronavirus cruise ship as people 'fight over rotten food' - Yahoo News Australia - That's a maritime emergency now they should be allowed to dock! https://t.co/BWNtSc1P0I	−0.8481
11	4/16/2020 15:49	Maybe I'm biased because going on cruise falls behind root canal for me, but I can't think of a worse, more disgusting collection of grossness than a cruise ship. Public restrooms are probably cleaner.	−0.7386
12	5/6/2020 11:51	@GeraldoRivera I've never been on a cruise.I always hated the idea of cruises especially since many of them were getting the norovirus. At this point, with the coronavirus, I am never ever going on a cruise ship.	−0.6369
13	4/13/2020 21:27	I have gone on cruises in the past, but you could not pay me to go on another one after this mess! #CruiseControl https://t.co/rUVEY91P2d	−0.5217
14	4/8/2020 07:14	#COVID19 BAN grotesque size of polluting cruise liners to save oceans. And now, they are floating hulks of the diseased and dying due to virus. Things turn into their opposites. https://t.co/RGbmVCChgX	−0.2808
15	4/16/2020 16:02	Someday we will all be saying, "Do you remember when we packed ourselves on disgusting petri dish cruise ships in rooms smaller than closets with no windows, where we got stomach viruses, drank cheap booze and	−0.1596

(*continued on next page*)

group responses while sabotage terrorism is listed as a potential impact. These types of outcomes can only be anticipated during extraordinary events or circumstances. They also provided a highly simplified framework, which is very skeletal. We have adapted the simplified framework and combined it with the IIT in order to provide new and useful contexts

Table 3 (*continued*)

	Time	Tweet	Sentiment Score
		ended up sea sick? What were we thinking? Glad that's over".	
16	5/7/2020 14:04	Serveral cruise ships currently parked in our harbour, polluting the city with their tons of exhaust gases every day. On board: underpayed employees, threatened by a COVID19 outbreak, unable to get home. Despicable industry needs to change.	−0.6369
17	4/2/2020 18:38	Half of all #Covid_19 #coronavirus #pandemic cases in #Australia are linked to #cruiseships Floating petri dishes! Crews are not Australian, ships are not registered in Australia to avoid paying taxes. Get them out of our waters #scomo #ScottMorrison #LNP #auspol	−0.7274
18	5/4/2020 15:07	@RichieFed Cruise ships were enormous floating Petri dishes before COVID19. Cruises are for the newlywed & nearly dead.	−0.6486
19	3/24/ 2020 03:46	Read this thread, then listen to @CrimeJunkiePod episode about the death at sea, and then you will understand why I will never take a cruise ever. If you're not dying of some infectious disease, then you're being pushed off a boat or sold into human trafficking and no one cares. https://t.co/nAg2y8c1XN	−0.4767
20	3/28/ 2020 12:41	NO MASS CRUISES!! THEY THROW THEIR SEWAGE IN THE OCEAN! THEY ARE NOT REGISTERED IN THE US SO THEY DONT HAVE TO ABIDE BY OUR ENVIRO REGS!! https://t.co/yUb2HWt3sq	−0.296

to the conceptual framework that can be directly adopted or adapted in future studies on risk perception. The results offer additional support to the assumptions of the SARF and IIT, on the influence of integrating new information on people's actual beliefs and the influence of media coverage on public opinion and risk perceptions.

However, the monitoring of changes in consumer sentiment expressed on social media is a continuous process, as perceptions can change as industry-related events happen. Cruise researchers can monitor these changes in future research and use them together with other established research methodologies, like surveys, to understand consumer behavior and intentions in the industry. For example, the video of a cruise ship crashing into a dock that went viral (see Appendix) and the release of a documentary film on the COVID-19 outbreak on the *Diamond Princess* titled *The Last Cruise* on March 30, 2021 have also generated many social media comments on cruising in recent times.

5.2. Managerial implications

The practical contribution of this study is to provide insight into the public perception of cruising during the COVID-19 outbreaks on cruises using social media data. The rebuilding process after the pandemic provides an opportunity for the cruise industry to reinvent itself. One way the industry can achieve this is by doubling down on green credentials and observance of environmental regulations, and not simply by cost reductions and increasing ship size. These green efforts have to be genuine because consumers see through greenwashing and only genuine green behavior improves organizational reputation (de Jong, Huluba, & Beldad, 2020). There should be an emphasis on cleaner and greener ships for new ship orders rather than on increasingly larger vessels. Smaller ships that are more environmentally friendly and offer more intimate experiences could be prioritized. Even though seniors remain the largest demographic of cruisers, the average age of cruise passengers has been falling steadily (Dowling & Weeden, 2017). The industry needs to maintain this trend by attracting younger new-to-cruise passengers with its green credentials when they have it. These young consumers among the millennials and Gen Z have a reputation for driving sustainability, environmental, and ethical consciousness (Choudhary, 2020; Deloitte, 2020; Yeoman, 2008). They are also over-represented on social media (Perrin, 2015) and hence, with a louder voice to express their discontent on sustainability issues on social media, which gets picked up by their peers. After facing a much sterner test with COVID-19, cruise lines should now also endeavor to keep cruising safe from the routine outbreaks of other infectious diseases by preserving many of the COVID-19 health protocols like frequent handwashing, some measure of social distancing, increased ventilation, ultraviolet air filtration, additional medical facilities onboard, contactless apps for food-ordering, etc. In the aftermath of COVID-19, the cruise industry should vigorously look into alternative business models that would be sustainable and profitable in the long term. Finally, other areas of tourism can also benefit from monitoring the online conversations on social media when industry-related events happen.

5.3. Limitations

The restriction of the analyzed tweets to only those in English is an important limitation, since an analysis of tweets in all languages could have painted a more comprehensive picture. It is also possible some relevant tweets were not included in the analysis, as we based the data collection on the presence of keywords and some relevant tweets may not mention these keywords, especially when replying to a tweet that already provided context. As an example, there was a viral tweet (see Appendix), which replied to a news report about the planned

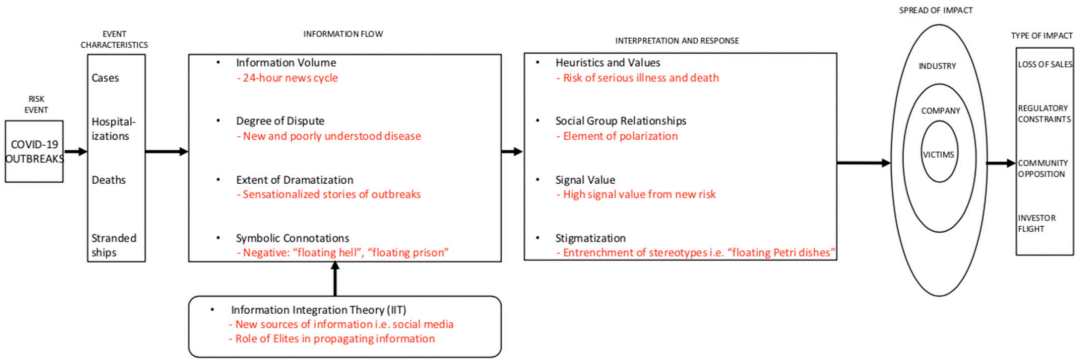


Fig. 9. Revised conceptual map of SARF and IIT for impact of COVID-19 outbreaks on the Cruise industry.

resumption of cruises by Carnival with the dancing coffin meme photoshopped with a Carnival cruise ship in place of the coffin to imply danger of death.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 713679 and from the Universitat Rovira i Virgili (URV).

Availability of data and material (data transparency)

The dataset of Twitter IDs used in this study is available online at https://doi.org/10.5281/ZENODO.3819464

CRediT authorship contribution statement

Babajide Abubakr Muritala: Conceptualization, Methodology, Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Ana-Beatriz Hernández-Lara:** Conceptualization, Writing – review & editing. **Maria-Victoria Sánchez-Rebull:** Conceptualization, Writing – review & editing. **Alexandre Perera-Lluna:** Conceptualization, Data curation, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgments

We thank Ed Summers of DocNow (docnow.io) and Paul R. Pival (distlib.blogs.com) for their help in obtaining the data. We are also grateful to the Editor and the anonymous reviewers for their excellent recommendations.

Appendix A. Appendix

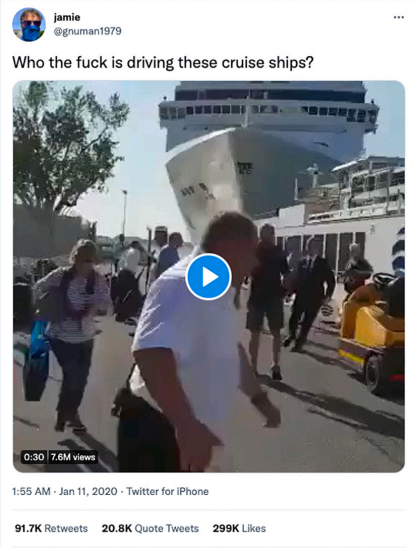


Fig. A1. Viral tweet of cruise ship crashing into a dock in Venice (Source: Twitter (Jamie, 2020)).

Fig. A2. Dancing coffin meme with cruise ship (Source: Twitter (Bailey, 2020 .

References

ABC. (2020). How the coronavirus pandemic would look in Australia if Ruby Princess had never docked - ABC News. https://www.abc.net.au/news/2020-04-23/coronavirus-across-australia-if-ruby-princess-never-docked/12172314.

Ainin, S., Feizollah, A., Anuar, N. B., & Abdullah, N. A. (2020). Sentiment analyses of multilingual tweets on halal tourism. *Tourism Management Perspectives, 34*. https://doi.org/10.1016/j.tmp.2020.100658

Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: capitalizing on big data. *Journal of Travel Research, 58*(2), 175–191. https://doi.org/10.1177/0047287517747753

Anderson, Norman H. (1981). Foundations of Information Integration Theory. *Academic Press*. https://philpapers.org/rec/ANDFOI-2/.

AP. (2021). COVID-19's global death toll tops 5 million in under 2 years. https://apnews.com/article/coronavirus-pandemic-africa-health-pandemics-infectious-diseases-838421260675827b7e735729b90ff95f.

Bailey, S. (2020). *https://t.co/31DtABZQof*.

Banda, J. M., Tekumalla, R., Wang, G., Yu, J., Liu, T., Ding, Y., … Chowell, G. (2021). A large-scale COVID-19 twitter chatter dataset for open scientific research—An international collaboration. *Epidemiologia, 2*(3). https://doi.org/10.3390/epidemiologia2030024

Banda, J. M., Tekumalla, R., Wang, G., Yu, J., Liu, T., Ding, Y., & Chowell, G. (2020). *A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration* (vol. 9). https://doi.org/10.5281/ZENODO.3819464

Barbe, D., Pennington-Gray, L., & Schroeder, A. (2018). Destinations' response to terrorism on twitter. *International Journal of Tourism Cities, 4*(4), 495–512. https://doi.org/10.1108/IJTC-04-2018-0027

Bird, S., Loper, E., & Ewan, K. (2009). *Natural language processing with Python*. O'Reilly Media Inc. http://www.nltk.org/book/.

Blumer, H. (1948). Public opinion and public opinion polling. *American Sociological Review, 13*(5), 542–549. https://doi.org/10.2307/2087146

Carnival. (2020). SEC Filing|Carnival Corporation & plc. https://www.carnivalcorp.com/node/62011/html.

CDC. (2020). Third modification and extension of No Sail Order and other measures related to operations. https://www.cdc.gov/quarantine/pdf/CDC-NSO-Third-Extension-09-30-2020-p.pdf.

CDC. (2021). Temporary extension & modification of framework for Conditional Sailing Order (CSO). https://www.cdc.gov/quarantine/cruise/pdf/CDC-CSO-Extension-10-25-21-p.pdf.

Choudhary, A. (2020). Generation green is leading the sustainability agenda. https://www.capgemini.com/2020/08/generation-green-is-leading-the-sustainability-agenda/.

CLIA. (2019). 2019 state of the industry. https://cruising.org/-/media/research-updates/research/clia-2019-state-of-the-industry-presentation-(1).ashx.

CLIA. (2021). 2021 state of the industry outlook. https://cruising.org/-/media/research-updates/research/2021-state-of-the-cruise-industry_optimized.ashx.

CNN. (2018). Symphony of the Seas: World's largest cruise ship sets sail|CNN Travel. https://edition.cnn.com/travel/article/symphony-of-the-seas-world-largest-cruise-ship/index.html.

Coombs, W. T. (2018). *Ongoing crisis communication planning, managing, and responding*. SAGE Publications Inc.

Couper, M. P. (2013). Is the sky falling? New technology, changing media, and the future of surveys. *Survey Research Methods, 7*(3), 145–156.

Cruise Industry News. (2020). 117 new cruise ships on forward-looking orderbook. https://www.cruiseindustrynews.com/cruise-news/22364-117-new-cruise-ships-on-forward-looking-orderbook.html.

D'Avanzo, E., Pilato, G., & Lytras, M. (2017). Using Twitter sentiment and emotions analysis of Google Trends for decisions making. *Program, 51*(3), 322–350. https://doi.org/10.1108/PROG-02-2016-0015

Deloitte. (2020). Millennial survey 2020. https://www2.deloitte.com/global/en/pages/about-deloitte/articles/millennialsurvey.html.

Dowling, R., & Weeden, C. (2017). The world of cruising. In *Cruise ship tourism* (pp. 1–39). https://doi.org/10.1079/9781780646084.0001

Feldman, R. (2013). Techniques and applications for sentiment analysis: The main applications and challenges of one of the hottest research areas in computer science. *Communications of the ACM, 56*(4), 82–89. https://doi.org/10.1145/2436256.2436274

Financial Times. (2020). Coronavirus: Is this the end of the line for cruise ships?. https://www.ft.com/content/d8ff5129-6817-4a19-af02-1316f8defe52.

Financial Times. (2021). The $24bn dash for cash that kept carnival afloat. https://www.ft.com/content/033f6161-7fb7-4936-8f3b-d78afec94cb2.

Gene Zucker, H. (1978). The variable nature of news media influence. *Annals of the International Communication Association, 2*(1), 225–240. https://doi.org/10.1080/23808985.1978.11923728

Google Trends. (2020). Google Trends. https://www.google.com/trends.

Gössling, S., Scott, D., & Hall, C. M. (2020). Pandemics, tourism and global change: A rapid assessment of COVID-19. *Journal of Sustainable Tourism, 1–20*. https://doi.org/10.1080/09669582.2020.1758708

Herald, M. (2020). COVIDCRUISES MIAMI HERALD DATA. https://docs.google.com/spreadsheets/d/1KTJj94nWd88Gml98AbTZKYZejbEJ8q4Xa7oE6g5r9ec/edit#gid=0.

Holland, J., Mazzarol, T., Soutar, G. N., Tapsall, S., & Elliott, W. A. (2021). Cruising through a pandemic: The impact of COVID-19 on intentions to cruise. *Transportation Research Interdisciplinary Perspectives, 9*. https://doi.org/10.1016/j.trip.2021.100328

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the 8th international conference on Weblogs and Social Media, ICWSM 2014* (pp. 216–225).

Ito, H., Hanaoka, S., & Kawasaki, T. (2020). The cruise industry and the COVID-19 outbreak. *Transportation Research Interdisciplinary Perspectives, 5*. https://doi.org/10.1016/j.trip.2020.100136

Jamie. (2020). *jamie on Twitter: "Who the fuck is driving these cruise ships*? https://t.co/0slDSqbXGq". https://twitter.com/gnuman1979/status/1215799298111262722.

de Jong, M. D. T., Huluba, G., & Beldad, A. D. (2020). Different shades of greenwashing: Consumers' reactions to environmental lies, half-lies, and organizations taking credit for following legal obligations. *Journal of Business and Technical Communication, 34*(1), 38–76. https://doi.org/10.1177/1050651919874105

Kaplan, A. M., & Haenlein, M. (2011). The early bird catches the news: Nine things you should know about micro-blogging. *Business Horizons, 54*(2), 105–113. https://doi.org/10.1016/j.bushor.2010.09.004

Kasperson, J. X., Kasperson, R. E., Pidgeon, N., & Slovic, P. (2003). The social amplification of risk: Assessing fifteen years of research and theory. In *The social amplification of risk*. https://doi.org/10.1017/CBO9780511550461.002

Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., … Ratick, S. (1988). The social amplification of risk: A conceptual framework. *Risk Analysis, 8*(2), 177–187. https://doi.org/10.1111/j.1539-6924.1988.tb01168.x

Khong, K. W., Teng, S., Butt, M. M., & Muritala, B. A. (2021). Analysing online reviews of restaurants in Malaysia: A novel approach to descriptive and predictive analytic. *International Journal of Electronic Business, 16*(4), 315–335. https://doi.org/10.1504/IJEB.2021.118491

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons, 54*(3), 241–251. https://doi.org/10.1016/j.bushor.2011.01.005

Klein, R. A. (2002). *Cruise ship blues: The underside of the cruise ship industry*.

Klein, R. A. (2009). *Getting a grip on cruise ship pollution*. Friends of the Earth.

Klein, R. A. (2011). Responsible cruise tourism: Issues of cruise tourism and sustainability. *Journal of Hospitality and Tourism Management, 18*(1), 107–116. https://doi.org/10.1375/jhtm.18.1.107

Klein, R. A. (2016a). Crime at Sea: A Comparison of Crime on Carnival Cruise Lines, 2007–2011. In A. Papathanassis (Ed.), *Cruise business development: Safety, product design and human capital* (pp. 17–28). Springer International Publishing. https://doi.org/10.1007/978-3-319-27353-2_2.

Klein, R. A. (2016b). Safety, security, health and social responsibility. In *Cruise business development: Safety, product design and human capital*. https://doi.org/10.1007/978-3-319-27353-2_1

Kluyver, T., Ragan-kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., Willing, C., & Development Team, J. (2016). *Jupyter Notebooks—a publishing format for reproducible computational workflows* (vol. 87–90). Positioning and Power in Academic Publishing: Players, Agents and Agendas. https://doi.org/10.3233/978-1-61499-649-1-87

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management, 68*, 301–323. https://doi.org/10.1016/j.tourman.2018.03.009

Liu, B. (2012). Sentiment analysis and opinion mining. In *Synthesis lectures on human language technologies*. https://doi.org/10.2200/S00416ED1V01Y201204HLT016 (Vol. 5, Issue 1).

Liu, Pennington-Gray, L., Donohoe, H., & Omodior, O. (2015). New York City bed bug crisis as framed by tourists on tripadvisor. *Tourism Analysis, 20*(2), 243–250. https://doi.org/10.3727/108354215X14265319207597

Liu.. (2009). Informing Design of Next Generation Social Media to support crisis-related grassroots heritage. In *PhD Colloquium of the 6th International ISCRAM Conference*. Gothenburg, Sweden, May.

Lloret, J., Carreño, A., Carić, H., San, J., & Fleming, L. E. (2021). Environmental and human health impacts of cruise tourism: A review. *Marine Pollution Bulletin, 173*. https://doi.org/10.1016/j.marpolbul.2021.112979

Luo, Q., & Zhai, X. (2017). "I will never go to Hong Kong again!" How the secondary crisis communication of "Occupy Central" on Weibo shifted to a tourism boycott. *Tourism Management, 62*, 159–172. https://doi.org/10.1016/j.tourman.2017.04.007

McGregor, S. C. (2019). Social media as public opinion: How journalists use social media to represent public opinion. *Journalism, 20*(8), 1070–1086. https://doi.org/10.1177/1464884919845458

McKinney, W. (2010). Data structures for statistical computing in Python. In *Proceedings of the 9th Python in science conference, 1697900(Scipy)* (pp. 51–56). http://conference.scipy.org/proceedings/scipy2010/mckinney.html.

Möller, C., Wang, J., & Nguyen, H. T. (2018). #Strongerthanwinston: Tourism and crisis communication through Facebook following tropical cyclones in Fiji. *Tourism Management, 69*, 272–284. https://doi.org/10.1016/j.tourman.2018.05.014

Moriarty, L., Plucinski, M., Marston, B., et al. (2020). Public health responses to covid-19 outbreaks on cruise ships - Worldwide, February-March 2020. In *MMWR Morb Mortal Wkly Rep*. https://doi.org/10.15585/MMWR.MM6912E3 (Vol. 69, Issue 12).

Mozuni, M., & Jonas, W. (2016). Rise of alternative business models in the era of expensive megaships. In A. Papathanassis (Ed.), *Cruise business development: Safety, product design and human capital* (pp. 55–70). Springer International Publishing. https://doi.org/10.1007/978-3-319-27353-2_4.

Mueller, A. (2020). WordCloud for Python documentation. https://amueller.github.io/word_cloud/index.html.

Muritala, B. A., Sánchez-Rebull, M.-V., & Hernández-Lara, A.-B. (2020). A bibliometric analysis of online reviews research in tourism and hospitality. *Sustainability 2020, 2*, 9977. https://doi.org/10.3390/SU12239977

Neri, F., Geraci, P., & Camillo, F. (2010). Monitoring the web sentiment, the Italian prime Minister's case. In *2010 international conference on advances in social networks analysis and mining* (pp. 432–434). https://doi.org/10.1109/ASONAM.2010.26

O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. In *ICWSM 2010 - proceedings of the 4th international AAAI conference on weblogs and social media* (pp. 122–129).

Pan, T., Shu, F., Kitterlin-Lynch, M., & Beckman, E. (2021). Perceptions of cruise travel during the COVID-19 pandemic: Market recovery strategies for cruise businesses in North America. *Tourism Management, 85*. https://doi.org/10.1016/j.tourman.2020.104275

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1–135. https://doi.org/10.1561/1500000011

Papathanassis, A. (2016). 'Stranded at sea': Exploring passengers' reactions during incidents at sea. In *Cruise business development: safety, product design and human capital*. https://doi.org/10.1007/978-3-319-27353-2_3

Papathanassis, A. (2017). Cruise tourism management: State of the art. *Tourism Review, 72*(1), 104–119. https://doi.org/10.1108/TR-01-2017-0003

Penco, L., Profumo, G., & Remondino, M. (2018). Negative events and corporate communication strategies in the cruise industry. *International Journal of Transport Economics, 45*(3), 485–514. https://doi.org/10.19272/201806703007

Penco, L., Profumo, G., Remondino, M., & Bruzzi, C. (2019). Critical events in the tourism industry: Factors affecting the future intention to take a cruise. *International Journal of Contemporary Hospitality Management, 31*(9), 3547–3566. https://doi.org/10.1108/IJCHM-05-2018-0423

Perrin, A. (2015). Social media usage. In , *Vol. 125. Pew research center*.

Philander, K., & Zhong, Y. (2016). Twitter sentiment analysis: Capturing sentiment from integrated resort tweets. *International Journal of Hospitality Management, 55*, 16–24. https://doi.org/10.1016/j.ijhm.2016.02.001

Renn, O. (1986). *Risk perception : A systematic review of concepts and research results*.

Rowe, G., Frewer, L., & Sjöberg, L. (2000). Newspaper reporting of hazards in the UK and Sweden. *Public Understanding of Science, 9*(1), 59–78. https://doi.org/10.1088/0963-6625/9/1/304

Ryschka, A. M., Domke-Damonte, D. J., Keels, J. K., & Nagel, R. (2016). The effect of social media on reputation during a crisis event in the cruise line industry. *International Journal of Hospitality and Tourism Administration, 17*(2), 198–221. https://doi.org/10.1080/15256480.2015.1130671

Schroeder, A., Pennington-Gray, L., Donohoe, H., & Kiousis, S. (2013). Using social Media in Times of crisis. *Journal of Travel & Tourism Marketing, 30*(1–2), 126–143. https://doi.org/10.1080/10548408.2013.751271

Schweinsberg, S., Darcy, S., & Beirman, D. (2020). 'Climate crisis' and 'bushfire disaster': Implications for tourism from the involvement of social media in the 2019–2020 Australian bushfires. *Journal of Hospitality and Tourism Management, 43*, 294–297. https://doi.org/10.1016/j.jhtm.2020.03.006

Sigala, M. (2011). Social media and crisis management in tourism: Applications and implications for research. *Information Technology Tourism, 13*, 269–283.

Smith, J. (2005). Dangerous news: Media decision making about climate change risk. *Risk Analysis, 25*(6), 1471–1482. https://doi.org/10.1111/j.1539-6924.2005.00693.x

Su, L., Stepchenkova, S., & Kirilenko, A. P. (2019). Online public response to a service failure incident: Implications for crisis communications. *Tourism Management, 73*, 1–12. https://doi.org/10.1016/j.tourman.2019.01.011

Summers, E., Brigadir, I., van Kemenade, H., Hames, S., Binkley, P., tinafigueroa, … Kerchner, D. (2021). *DocNow/twarc*. https://doi.org/10.5281/ZENODO.593575

Terry, W. C. (2017). Flags of convenience and the global cruise labour market. In *Cruise ship tourism* (pp. 72–85). https://doi.org/10.1079/9781780646084.0072

Twitter. (2020). More on restricted use cases – Twitter Developers. https://developer.twitter.com/en/developer-terms/more-on-restricted-use-cases.

Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*.

Vogel, M. P. (2017). Economics of cruise shipping: The need for a new business model. In *Cruise ship tourism* (pp. 124–137). https://doi.org/10.1079/9781780646084.0124

Wang, Y., Zheng, L., & Zuo, J. (2021). Online rumor propagation of social media on NIMBY conflict: Temporal patterns, frameworks and rumor-mongers. *Environmental Impact Assessment Review, 91*. https://doi.org/10.1016/j.eiar.2021.106647

Washington Post. (2020). Cruise ships kept sailing after coronavirus was detected, exposing thousands. https://www.washingtonpost.com/graphics/2020/politics/cruise-ships-coronavirus/.

WHO. (2020). Naming the coronavirus disease (COVID-19) and the virus that causes it. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it.

Womack, J. J., Anderson, L. N., & Ledford, C. J. W. (2020). Presence of complex and potentially conflicting information in prenatal Mobile apps. *Health Promotion Practice, 21*(2), 238–245. https://doi.org/10.1177/1524839918796216

Yeoman, I. (2008). Chapter 19 - life on Mars — How different generations perceive the future of holidays in Scotland. In I. B. T.-T. T. Yeoman (Ed.), *Vol. 16. Advances in tourism research* (pp. 261–272). Butterworth-Heinemann. https://doi.org/10.1016/S1572-560X(08)00419-8.

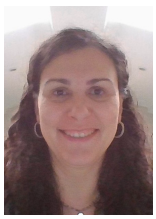
Yu, M., Li, Z., Yu, Z., He, J., & Zhou, J. (2020). Communication related health crisis on social media: A case of COVID-19 outbreak. *Current Issues in Tourism*. https://doi.org/10.1080/13683500.2020.1752632

Zeng, B., & Gerritsen, R. (2014). What do we know about social media in tourism? A review. *Tourism Management Perspectives, 10*, 27–36. https://doi.org/10.1016/j.tmp.2014.01.001

Zhang, M., Zhang, Y., & Fu, G. (2016). Tweet sarcasm detection using deep neural network. In *COLING 2016 - 26th international conference on computational linguistics, proceedings of COLING 2016: Technical papers* (pp. 2449–2460).

Babajide Abubakr Muritala is a doctoral candidate and Martí i Franquès COFUND fellow at the Department of Business Management, Universitat Rovira i Virgili (URV), Spain. He has an MBA from the University of Nottingham (Malaysia Campus). His research interests are in electronic word-of-mouth (eWOM), online reviews, big data, data mining/analytics, and their applications in tourism and hospitality.

Ana Beatriz Hernández-Lara is an Associate Professor and member of the Rectory team, as Teaching and Research staff Assistant at Universitat Rovira i Virgili (URV), Spain. She obtained her PhD at Pablo de Olavide University (Spain) in 2007 on the influence of corporate governance on innovation. Her research lines are related to corporate governance, innovation, tourism and e-learning. She currently participates in different research groups and projects in the field of social and organizational studies. Her research has been published in reputable journals like *Computers in Human Behavior*, *Behavior and Information Technology*, *Management Decision*, *Current Issues in Tourism*, *Tourism Management Perspectives*, among others.



Maria Victoria Sánchez-Rebull is an Associate Professor and Vice Dean at the Faculty of Economics and Business at the Universitat Rovira i Virgili (URV), Spain. She obtained her PhD at this University in 2002 on Activity Based Costing (ABC) in the hotel sector. Her research lines are related to cost management, internationalization and tourism. She currently participates in different research groups and projects in the field of finance and accounting research. Her research has been published in reputable journals like *Cornell Hospitality Quarterly*, *Total Quality Management & Business Excellence*, *Tourism and Hospitality Management*, *International Journal of Lean Six Sigma*, *Health Policy*, among others.

Alexandre Perera Lluna holds a degree in Physics (UB), Electronic Engineering (UB), PhD in Physical Sciences (UB), postdoctoral fellow at Texas A&M University, and Ramon y Cajal Fellow. He is currently tenured at the Polytechnic University of Catalonia (UPC) and also has a double affiliation as researcher of the Institut de Recerca de Sant Joan de Déu. He has published more than 65 papers in peer-reviewed journals, five patents, and more than 60 contributions to national and international conferences. He was co-founder of a technology-based company or spin-off (2001 eSense Systems SL), exheus.com and Vincer.IA, with applications of pattern recognition algorithms in the food industry, health and bioengineering. He is currently the coordinator of the B2Slab research group (http://b2slab.upc.edu) and director of the Biomedical Engineering Research Center (CREB). His research covers artificial intelligence algorithms, multivariate statistics, machine learning applied to bioinformatics, bioengineering, and clinical data science.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 713679 and from the Universitat Rovira i Virgili (URV).

UNIVERSITAT
ROVIRA i VIRGILI