



Applied data science in tourism (Interdisciplinary approaches, methodologies, and applications)

Roman Egger ISBN 978-3-030-88388-1, Springer Nature Switzerland, Switzerland, 2022, pp. 647, price

Omid Oshriyeh¹

Received: 21 November 2022 / Revised: 26 December 2022 / Accepted: 17 January 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Our modern culture is defined by an unheard-of capacity to generate and store mind-boggling volumes of data as well as by the ability to navigate through them in a way that allows us to extract relevant information, and thereby knowledge (Succi and Coveney 2019). Big Data's rise in social research presents both opportunities and difficulties (Tinati et al. 2014). As the use of big data is gaining more prominence in the fields of social sciences, some emerging challenges need to be addressed. As a result of the rapid evolution of big data analytics, new tools, methodologies, and algorithms are constantly being developed and applied to tourism and hospitality cases.

The problem however is the fact that only a handful of books have been authored on data science and big data research in the tourism and hospitality sector, which underscores the importance of the issue under discussion. For instance, in their book *Big Data and Innovation in Tourism, Travel, and Hospitality: Managerial Approaches, Techniques, and Applications*, Sigala, Rahimi, and Thelwall. (2019), try to demonstrate the practical implications of big data from a managerial perspective. In a book chapter entitled “*Big Data Analytics in Tourism*”, Łukasiewicz (2020) has provided an overview of the trends and the use of big data in tourism. Overall, it appears that while the application and implications of big data and data science in tourism have been covered by these books, there is still a need for an in-depth survey of tools, methods, and algorithms. Roman Egger and his co-authors have provided an attempt to fill this gap. The result of their work is entitled *Applied Data Science in Tourism*, a comprehensive look at the various methodological options and possibilities of applying data science to the interdisciplinary field of tourism.

✉ Omid Oshriyeh
oshriyeh@email.sc.edu

¹ University of South Carolina, Columbia, USA

The book comprises 26 chapters, which are organized into four main parts. The chapters in this book highlight up-to-date information on the application of data science to the tourism industry. Covering a variety of subjects such as network analysis, natural language processing, sentiment analysis, machine learning, and feature engineering. The content covers innovative tools and approaches that can be used in tourism and related sectors such as leisure, hospitality, and event management areas. In terms of research methodologies, case studies, and technological tools, this book offers an update on recent developments in artificial intelligence (AI), machine learning (ML), and data science.

After a thorough introduction of the related topics, concepts and definitions of the phenomenon known as data science, the authors present a comprehensive question-and-answer session with industry experts across the tourism and hospitality industry. This is one of the parts that make the book stand out, as they are the real-world industry leaders (not academics) who are sharing their related experiences of utilizing data science on a practical level in tourism and hospitality sector. Helping the reader hear the current trends and market needs before proceeding to the more detailed chapters on each method and topic.

Before talking about the methodology-focused parts of the book, it is necessary to highlight the authors' positionality on the theoretical and epistemological concerns. To be more precise, the focus of design science in information systems research appears to be moving away from IT artifacts and toward social systems, a development where theory input is offered for the scientific design process (Mazanec 2020). While some big data champions have gone as far as to describe the use of big data as "the end of theory", (Anderson 2008), arguing that research in the age of big data no longer needs theories; this book has a more conservative take on the use of big data. Egger and his co-authors insist on the need for coherent models, unified theories, and the need for casual explanations through theory. The chapter entitled "Epistemological Challenges" thoroughly discusses these concerns.

In the first part of the book, the need for interdisciplinary approaches is stressed. This part is entitled *Theoretical Fundamentals*, and discusses how domain knowledge, mathematics, and statistical methods interact with computer technologies to form what we refer to as data science, highlighting the need for a higher degree of interdisciplinarity across the field. Chapter 4 calls for ethical considerations in the new trends of data science, where data is readily available and may not be fully disclosed and unanimous.

In the second part, several chapters on machine learning are presented. In Chap. 6 the underlying idea behind machine learning is illustrated, while Chaps. 7–9 address more technical topics in machine learning including feature engineering, and various approaches in unsupervised and supervised machine learning are discussed, with examples presented for each of the topics to help readers gain a better understanding of how can they use tools such as clustering or dimensionality reduction in their research projects.

The third part of the book is mainly centered around natural language processing. Natural language processing is defined as the use of computational methods to learn, comprehend, and produce human language content (Hirschberg and Manning 2015). Improved Natural Language Processing techniques are a factor in sentiment

analysis's development (Kirilenko et al. 2018). Hence, Chap. 17 is an in-depth discussion regarding sentiment analysis. The following Chaps. (18–20) continue covering other commonly used techniques in text analysis such as topic modeling and entity matching. The application of knowledge graphs in a real-world tourism setting is the concluding chapter in the third and arguably the most technically complicated, yet practical section of the book.

The remaining Chaps. (21–26) mainly discuss other novel approaches and methods in which researchers can process their data using data science techniques. Ranging from visualizing networks and computing main measures through the use of network analysis, to predicting tourism growth and demands by conducting time series analysis. One of the more unique methods that have been featured in this book is the use of agent-based modeling, which can be extremely useful in the post-pandemic world and with COVID-19-related data.

Although the application of geographic information systems is not new to tourism, and researchers have long been using software and tools such as ArcGIS or QGIS in tourism geographics, Chap. 24 presents a distinctive perspective on spatial data analysis and GIS data. Data visualization is the subject of Chap. 25, which is one of the methods that can be applied to both micro and macro-level analyses. Finally, a list of useful software and tools is provided, many of which are free, open source, and easily accessible to readers.

Applied Data Science in Tourism can be an ideal book to teach students of tourism, hospitality, and leisure sciences on the subjects of data science. This is because the book is well-organized, each chapter starts with a set of learning objectives, and concludes with a summary of the material and its highlights. Also, chapters can be read and taught independently from each other, which can help pedagogical design in a potential big data course for graduate students or even senior undergrads. Another interesting feature of the book is the fact that while the concepts and methods may seem complicated at the first glance, the book makes use of parsimony and a simplified tune that makes it possible for those with no analytical and IT background to understand the topics with no prior knowledge.

Overall, *Applied Data Science in Tourism* is an incredibly helpful contribution to data science research in the fields of tourism and hospitality that is both easy to read and tremendously fascinating. This book is a valuable source of theoretical and methodological knowledge for both academics and industry practitioners of tourism and related fields such as hospitality, leisure, and event management. Especially considering the increasing demand for data analytics and the use of big data in service industries.

References

- Anderson C (2008) The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*. <https://www.wired.com/2008/06/pb-theory/> (Accessed 10th, November, 2022)
- Hirschberg J, Manning CD (2015) Advances in natural language processing. *Science* 349(6245):261–266. <https://doi.org/10.1126/science.aaa8685>

-
- Kirilenko AP, Stepchenkova SO, Kim H, Li X (2018) (Robert). Automated Sentiment Analysis in Tourism: Comparison of Approaches. *Journal of Travel Research*, 57(8), 1012–1025. <https://doi.org/10.1177/0047287517729757>
- Łukasiewicz K (2020) Big Data Analytics in Tourism: overview and Trends. Management in the era of Big Data. Auerbach Publications
- Mazanec JA (2020) Hidden theorizing in big data analytics: with a reference to tourism design research. *Annals of Tourism Research* 83:102931. <https://doi.org/10.1016/j.annals.2020.102931>
- Sigala M, Rahimi R, Thelwall M (eds) (2019) Big Data and Innovation in Tourism, Travel, and hospitality: managerial approaches, techniques, and applications. Springer. <https://doi.org/10.1007/978-981-13-6339-9>
- Succi S, Coveney PV (2019) Big data: the end of the scientific method? *Philosophical Trans Royal Soc A: Math Phys Eng Sci* 377(2142):20180145. <https://doi.org/10.1098/rsta.2018.0145>
- Tinati R, Halford S, Carr L, Pope C (2014) Big Data: Methodological Challenges and Approaches for Sociological Analysis. *Sociology* 48(4):663–681. <https://doi.org/10.1177/0038038513511561>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.