

Connecting Big Data and Service Quality Evaluation – Developing a Service Quality Map of the Austrian Hospitality Industry through the Application of Big Data

Julia Beck, Margarita Danilenko, Laura Sperber, Brenda Wiersma,
Roman Egger*

Abstract

The impact of online reviews on guests, hotel owners and other parties, is growing in importance. In reference to online reviews, service quality plays a crucial role in hotel differentiation and influencing the choice of accommodation made by travellers. Thus, online reviews represent a valuable source of information about perceived service, that has not been fully exploited yet. This research paper attempts to look more closely at this extensive body of data. The authors have conceptualized a tool that assists governmental institutions, DMOs and investors in decision making. This tool accumulates intelligent data and provides a comprehensive overview of the Austrian hospitality industry and its service quality standards. It allows the user to conduct specific queries on how a certain dimension of service quality is perceived. The results can be either visualised on a density map or extracted as a structured .csv file for further analysis.

Keywords: *Big data; online reviews; service quality; density map; HolidayCheck.*

1. Introduction and research goals

The web 2.0 in the 21st century has demonstrated tremendous growth and its applications empower numerous people to generate, share and accumulate information (Sigala, 2008). This information is referred to as user generated content (UGC). This phenomenon affects not only the everyday lives of Internet users, but has a booming effect on a variety of industries throughout the globe, with the hospitality sector being no exception. The two-way exchange of digital content results

* Prof. Dr. Roman Egger is Head of Tourism Research Department at Salzburg University of Applied Sciences , Austria. The other authors are research associates who belong to the same Department. Email: roman.egger@fh-salzburg.

in a vast amount of online information about the hospitality and tourism sector and its services (Sigala, 2008).

According to Dong, McCarthy, O'Mahony, Schaal, & Smyth (2012) among various segments of UGC, there has been an explosion in the growth of user reviews. Nowadays, one-third of travel-related purchasers visit message boards, forums or an online community before the transaction, as they believe it would facilitate their decision making process (Complete, Inc (2007), as cited in Qiang, Law, Gu, & Chen, 2011). One of the main factors which consumers value most highly, that would influence the decision making process when it comes to the choice of accommodation, is service quality. However, according to Marine-Roig and Clave (2015), UGC data in tourism is not only an important source of information for prospective travellers, but also for destination management/marketing organisations (DMO), governmental bodies and other stakeholders, as it consists of freely expressed opinions. This in turn represents a whole new perspective on the decision making process and suggests a shift in potential implications for the individual traveller, or even on a regional or national scale for a country. Due to the growth of UGC, the use of big data technologies is needed to manipulate and further interpret it (Marine-Roig & Clave, 2015).

Tapping into the power of big data, this research paper proposes the development of a tool – a service quality map (SQM) - that compresses and processes large volume of user reviews in a visualised and structured way. The essence of this project lies within the strength of market intelligence about rated services, as a hotel stay, i.e. service quality indicators associated with it, is subject to evaluation. Those indicators create the opportunity to display the evaluated service in a new way (Litvin, Goldsmith & Pan, 2008). The Austrian hospitality sector serves as a research field due to the country's advanced involvement with the tourism industry. The research incorporates DMOs, governmental bodies and investors as a main target group expected to engage with the prototype that would serve them as a decision making tool depicting ratings in a visualised way. The map provides a visualised overview of how travellers perceive certain (service) quality aspects. Thus, the aim of the study is also to illustrate how big data can be further used in thorough service quality examinations. For this purpose, the designed tool allows one to extract data as .csv file, which can then be further interpreted and analysed using external software.

2. Literature review

In 2011, many scholars working in the area of technological research, hailed the term big data as the future buzzword. This paper refers to the trending framework of using large scaled information and data sets (Weinberg, Davis, & Berger 2013; Lohr, 2013; Harris, 2013). Taking into consideration that big data is still a relatively new phenomenon, one cannot offer a fully complete interpretation of the term

(Weinberg et al., 2013; Harris, 2013). Therefore, the authors of this paper outline a general definition of the big data phenomenon: big data is a large volume of complex, unstructured digital data generated through a variety of sources, typically requiring new database software tools that are able to handle this type of data in a timely manner - the definition derived from the key articles on the topic (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh, & Byers, 2011; Harris, 2013; Kim, Trimi, & Chung, 2014; Davenport, 2013). Furthermore, big data is often referred to as the three “Vs”: volume, velocity, and variety (Harris, 2013; Kim et al., 2014; McAfee & Brynjolfsson, 2012; Davenport, Barth & Bean, 2012; Russom, 2011; Jackson, 2014).

This research paper presents various arguments and expectations that are to some extent, sceptical but overall anticipates a wide range of applications for the use of big data and predicts a positive future for its use (Fox and Do, 2013; Rouen, 2012).

2.1 The use of big data in tourism

The travel industry is characterised by an enormous volume of structured data. Every reservation, hotel stay, flight or train ticket can be seen as a data trail. But the industry also has to deal with a massive amount of unstructured data, accumulated through social networks, online reviews and ratings, blogs, heap of mobile applications, and much more (Davenport, 2013; SOCAP Int., 2013). Furthermore, this relatively new type of data is collected faster and faster every day (Davenport, 2013). This change creates a challenge for the tourism industry, as it has to find new methods to analyse all these amounts of unstructured data and strategically make use of it. Intelligently arranged large-scale data can create undiscovered opportunities, which enable the tourism industry to improve the service quality that consequently can further enhance the customer experience (Davenport, 2013; SOCAP Int., 2013). The following below presents some progressive examples of big data application for the benefit of tourism industry.

One such example is the development of a data warehouse to support decision making in the Chinese tourism industry. For the purpose of a platform development data is retrieved from government websites, official statistics and tourism enterprises. For future projects, one suggests that further data be incorporated from online travel agencies (Qiao, Zhang, Li & Zhu, 2014).

Similarly, area search study by Marine-Roig and Clave (2015) focused on the collection of UGC from travel blogs and review websites in Catalonia, Spain. Researchers suggest innovative methods that would allow “semi-automatic downloading, arranging, cleaning, debugging, and analysing large scale travel blog and online travel reviews data” (Marine-Roig & Clave, 2015, p. 4), as opposed to traditional way of data collection “by hand”. Thus, this case provides useful insights and inspiration for the current research. In conclusion, the research by Marine-Roig and Clave (2015) demonstrates how big

data technologies can be helpful for branding and promotion, when properly utilized by National Tourism Organisations (NTOs) and DMOs.

Further examples of big data application in tourism include research by Schlegel (2015), Irudeen and Samaraweera (2013), or Xiang, Schwartz, Gerdes and Uysal (2015), which are essential to familiarize oneself with. These examples demonstrate that big data can assist DMOs as well as other institutions in their decision-making process (Davenport, 2013; SOCAP Int., 2013; Qiao et al., 2014; Marine-Roig & Clave, 2015; Schlegel, 2015; Irudeen & Samaraweera, 2013; Xiang et al., 2015). However, there are no clear examples of how big data can facilitate an examination of (perceived) service quality for the sake of business and operational decision making of DMOs and other stakeholders. As provision of excellent service quality is regarded as a prerequisite for gaining competitiveness and achieving long-term success on the market, measurement of service quality has been an important research subject over the years, e.g. (Atilgan, Akinci and Aksoy (2003); Attallah (2015); Lau Pei Mey, Khatibi Akbar and Yong Gun Fie (2006)). However, the scope of these studies is usually limited to a certain target group(s) or location, which thus fails to provide a broad overview. As proven by the research in question, big data can solve this problem, which allows for coverage of the service quality landscape of the entire country, namely Austria. Service quality is one of the key exploration topics in this study, which is further discussed in the next sub-chapter.

2.2 Service quality

Parasuraman, Zeithaml and Berry (1985, p.41) describe quality as “an elusive and indistinct construct”. When attributed to service, it gets even more intricate. Thus, it is often more difficult to evaluate the quality of service than the quality of physical goods. Evaluation is rather complex and consists of several stages: (a) pre-consumption, which is the selection among alternatives, (b) consumption, explained as the comparison of experiences with expectations, and (c) post-consumption, the comparison of experience with expectations formed during both the pre-consumption and consumption stages (Reisinger in Kandampully, Mok, & Sparks, 2001).

One of the most famous conceptualisations of service quality was done by Parasuraman, Zeithaml, and Berry (1985). It later evolved into the *SERVQUAL* scale (Parasuraman, et al., 1988), which is still widely used for measuring service quality. The following five dimensions of service quality, reassessed and partially reformulated, form the basis of *SERVQUAL*: tangibles, reliability, responsiveness, assurance, and empathy (Parasuraman et al., 1988). Nevertheless, various studies tend to further elaborate on the concept and introduce own measuring scales fully or partially based on the work of Parasuraman et al. (1988), e.g. *SERVPERF* by Cronin and Taylor (1994), *Normed Quality models* by Teas (1993), etc.

For the research in question, it is important to distinguish a broad set of service quality indicators in order to further justify the structure of the final product, the SQM. The following table provides a summary of different scales and sets of service quality determinants used by other researchers in an attempt to evaluate service quality within the hospitality industry.

Table 1. Overview of the service quality indicators (Sources: Briggs et al., 2006, p. 1011; Benitez et al., 2007, p. 546; Zafar, et al., 2007, p. 69; BechSerrat, 2011, p. 281)

Briggs, Sutherland, and Drummond (2006)	Benitez, Martin, and Roman (2007)	Zafar, Sadiq, Matthew, and Mohammad (2007)	Bech Serrat (2011)
Personal service	Reception desk	Reception: <ul style="list-style-type: none"> • Handling of check-in and check-out • Efficiency • Friendliness 	Environmental dimension: <ul style="list-style-type: none"> • Homogeneity of customer / shared behaviour • Atmosphere • Codes of appearance
Value of money	Room cleaning		
Friendliness/warmth	Room maintenance	Room: <ul style="list-style-type: none"> • Value for money • Appearance • Furnishing • Cleanliness • Functioning of facilities • Bathroom • Overall impression 	Human dimension: <ul style="list-style-type: none"> • Flexibility and versatility • Courtesy • Positive attitude • Uniformity of behaviour • Prompt response
Attention to detail			
High standards	Main restaurants (breakfast/dinner) and A-la-Carte: <ul style="list-style-type: none"> • Food offer/ cuisine • Service 	Restaurant: <ul style="list-style-type: none"> • Quality of food • Price • Services • Overall impression 	Restaurant: <ul style="list-style-type: none"> • Quality of food • Price • Services • Overall impression

Briggs, Sutherland, and Drummond (2006)	Benitez, Martin, and Roman (2007)	Zafar, Sadiq, Matthew, and Mohammad (2007)	Bech Serrat (2011)
Uniqueness	Bar service (day / evening)	Conference Facilities: <ul style="list-style-type: none"> • Comfort • Lighting • Temperature • Services 	
Natural approach			
Tangibles	Cleaning pool and garden area	Attitudes: <ul style="list-style-type: none"> • Friendliness • Helpfulness • Attentiveness • Efficiency • Discretion 	
Efficiency / professionalism	Entertainment and shows		

The list of service quality indicators can be continued further, as only a limited number of studies on the topic have been done so far. More importantly, the above mentioned indicators significantly coincide with the service quality indicators used by Holiday Check. And while it is not possible to claim whether Holiday Check is basing its choice on any theoretical research, it can still be argued that this choice is smart and the set of indicators resonates with the end users. It is a well-proven fact that travellers actively and regularly share their experiences with the Holiday Check community. Furthermore, according to Dong et al. (2012), millions of consumers are now using platforms such as Trip Advisor, Hotels.com, Amazon, Best Buy and others to contribute their opinions and experiences. In fact, the recent growth of online reviews has reached such a level that Dong et al. (2012) refer to it as a “unique service” between consumers, where there is a mutual exchange and regular provision of information.

Similarly, the aim of the research in question is to tap into the power of user reviews by incorporating data from Holiday Check and to ascertain what other beneficial implications exist on a higher level. Further details about the platform, its service quality indicators and how they are used for this research can be found in the next chapter.

3. Methodology

The application of big data is the main focus of this research. Big data allows the creation and manipulation of large data sets that can be rather easily aggregated, examined, and connected in a comprehensive manner. Moreover, it enables the cross-referencing of large data sets in order to achieve a clearer picture or even a new perspective on the research subject (Boyd & Crawford, 2012).

This section of the paper will explain the applied methods that lead to the end product. The following research questions are posed:

RQ₁: How can big data be used to develop a service quality map of the Austrian hospitality?

RQ₂: How can the use of big data further facilitate the examination of perceived service quality?

To develop the prototype, a total of eight workshops were conducted. The development process can be subdivided in the following phases: (1) planning (RQ₁), (2) data collection and processing (RQ₁,RQ₂). The first four workshops were dedicated to the planning phase, whereas the other four workshops were held to improve the prototype by testing it thoroughly and establishing a structure for the output in a .csv file.

3.1 Planning

The utilisation of workshops for the development of the prototype is comparable to Agile Project Management (APM) workflow. According to Hass (2007, p.8), “APM is a highly iterative and incremental process, where developers and project stakeholders actively work together to understand the domain, identify what needs to be built, and prioritize functionality”. This process is recommended when managing the development of software (Hass, 2007). By means of the workshops, the project team, application developers and the advanced stage end users remained in close contact for the realisation of the tool as dictated by the APM.

During the first four workshops, the core idea as well as the final product was conceptualised. After a comprehensive evaluation of different rating platforms, the decision to use Holiday Check was made, based on the wide range of accommodation attributes to be rated by the users. Holiday Check registers approximately 25 million visitors monthly and has to date, around eleven million ratings. Within the Austrian Hospitality sector, around 24,000 accommodations are listed (Holiday Check, 2015). Those user-generated ratings serve as a foundation for this prototype.

3.2 Data collection and processing

To collect the data from Holiday Check, a technique called web scraping was used. Web scraping is a process whereby content or information from a website is retrieved for another purpose (Bonneville, 2015; Cleaves, 2015). To scrape the large amount of data from the interface of Holiday Check, Node.js was used. Many rating platforms for hotels consist of ratings from travellers using a combination of star ranking and text evaluation (Banerjee & Chua, 2016). However, at this point it is important to mention that this research does not focus on narrative evaluation, but rather on scale ratings from one to six. The aim is to collect scale user reviews complemented with additional information. Thus, the scraped data then comprises

of(1) general hotel information, (2) socio-demographic information of travellers, and (3) ratings of various service quality determinants.

After the web scraping, the data needs to get structured in order to discover patterns. According to Witten and Frank (2005), this process can be referred to as data mining. The prototype organises the data, so it can be visualised through a density map. As defined by the geographic information system dictionary of ESRI Inc. (ESRI Inc., 2016), a density map is “a quantitative, thematic map on which dots of the same size are randomly placed in proportion to a numeric attribute associated with an area. Dot density maps convey the intensity of an attribute”. With regards to this paper, the theme of the map is the service quality of the Austrian hospitality sector. The numerical attributes are the scale ratings of the reviewers and the intensity of the attributes are shown by colour. The structure of the scraped data is reflected in the selection menu, providing the possibility to select various data items and thus generate a query. The menu is illustrated in the upper-side of Figure 1. The output of the queries that the viewer can conduct, is not only processed visually, but can also be extracted in a .csv file.

4. Results and discussion

The main outcome of this research is a prototype of a SQM of the Austrian hospitality industry. The aim of the application is to show how a specific type of tourist perceives different services within the hospitality industry in Austria or a certain area within Austria. It is a unique tool that should assist governmental institutions, DMOs and investors in decision making. Beneficiaries of this tool can either examine the data directly on the SQM or through the extracted data set. Figure 1 shows the interface of the developed prototype and how data filtered through a query is visualised on the map. For better understanding of the visualised result, the QR code leads to a video of a test version.

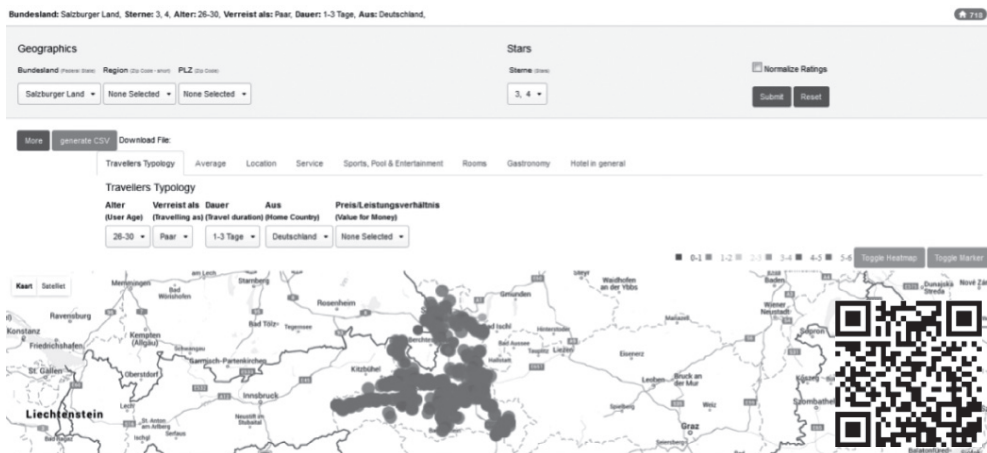


Fig 1. Illustration of the developed prototype (Source: Own Illustration)

4.1 Analysis and results: Service quality map

The purpose of this tool is to give governmental institutions and DMOs an insight into travellers' perceptions of service quality. With the SQM, the beneficiaries can view at a glance the perceived quality of the hospitality sector in a certain area in Austria. The displayed data is a compressed overview of the UGC, which is scraped from Holiday Check. Furthermore, the users can gain valuable data by asking themselves critical questions and creating hypothesis about possible relationships between various service quality attributes. Therefore, the user has the possibility to conduct specific queries by choosing different filters and attributes.

As it can be seen in Figure 1, the filters are drop down boxes where a selection can be made. After selecting a preferred area in Austria (federal state or zip code) and the official star classification for hotels, the socio-demographics of travellers can be specified. The travellers' typology consists of information about their origins, age group, duration of stay and mode of travelling (alone, with family, etc.) The last step is to select the service quality indicators one is interested in ((a) location, (b) service, (c) sports, pool & entertainment; (d) rooms, (e) gastronomy and (f) hotel in general) as well as a corresponding subcategory.

As a result, the data appears visually on the density map. Users can choose either a marker map or a heat map. The marker map shows different coloured dots, where each dot represents one hotel and gives information about the number of reviews and the average overall rating. This average overall rating is connected to a colour scheme with six categories, where red represents the lowest rating and blue respectively the highest. The heat map displays the density of the number of hotels and the average rating. Another functional capability of the prototype is the aggregation of the data in a .csv file. This data set can then be further analysed to gain deeper insights into possible relationships or differences.

4.2 Analysis and results: Examination of perceived service quality

With regards to RQ₂, by means of the above-mentioned techniques and procedures, an exemplary query was conducted and a .csv file was gathered. The extracted data set contains information about the ratings as well as socio-demographic information about the reviewers. For further analysis, the data set was processed with IBM SPSS Statistics.

In order to test usability of the extracted data, the authors formulated ten hypotheses. According to Bortz (2005), hypotheses refer to theory-derived statements or conclusions. However, for the purpose of this study, the following hypotheses were not based on any theoretical background, but rather were freely formulated, as the main goal was to test the functionality of the extracted data. To avoid redundancy only three out of ten hypotheses will be explained in detail:

Table 2. Hypotheses (Source: Own data)

H1	Travellers within the age range of 19-35 years perceive the quality of the condition of hotels in Salzburg differently than travellers within the age range of 36-50 years.
H2	There is a relationship between the perceived friendliness and the perceived competencies for hotels in Salzburg.
H3	There is a relationship between family friendliness and the quality of child-care or playground for hotels in Salzburg and Vienna.

The data to build the prototype was retrieved from Holiday Check on June 30, 2015, which means that ratings up to this date were included in the research. The sample size for the statistical analysis (according to H1, the relevant data was extracted from our tool into a .csv file) comprised of 113,494 reviews, with a total of 84,364 for the federal state of Salzburg and 29,130 for the federal state of Vienna. The majority of reviews were written for hotels in the 3* (Salzburg 28.4% and Vienna 32.9%) and 4* (64.2% and 53.7% respectively) segment. Furthermore, 66.9% of the people, who had rated an accommodation in Salzburg, came from Germany, 27.8% from Austria and 2.8% from Switzerland. Due to the small size of the remaining segments, the rest of the reviewers were then combined into regional groups based on their origins, for example, Benelux, Southern Europe, Overseas countries, etc.

Generally speaking, the data set can be subdivided into three main categories: (1) hotel information (location by zip code, title, star classification, etc.); (2) travellers' information (origin, age, travelling habits, etc.); (3) ratings of service quality parameters (condition of the hotel/rooms/restaurant, staff friendliness, etc.).

The reviewers on Holiday Check have the possibility to rate the quality on the scale from one to six "suns". Therefore, the data sets are based on ordinal scaled data that is characterized by non-normal distribution and hence requires application of non-parametric tests (Field, 2013). As explained by Bortz (2005), Spearman's correlation coefficient is used for studying the relationship between the two ordinal scaled variables. To assess two independent ordinal scaled samples on the central tendency, the Mann-Whitney U test is applied (Bortz, 2005). Kendall's Tau B is another non-parametric test, identifying correlation between ordinal scaled variables that can be regarded as a substitute to Spearman's Rho test (Field, 2013). However, Kendall's Tau B is recommended when it is not possible to distinguish between the dependent and independent variable (Acton, Miller, Fullerton, and Maltby, 2009).

Table 3. Results of statistical tests (Source: Own data)

	Spearman's Rho, coeff.	Kendall's Tau B, coeff.	Mann-Whitney U	Significance	Results
H1: Age range 19-35 and 36-50: Quality of the condition of the hotels in Salzburg	0.009	---	✓	0.071	Reject
H2: Relationship between perceived friendliness and handling of complaints	0.671	0.650	---	0.000**	Support
H3: Relationship between perceived quality of family friendliness and child care or playground (Salzburg and Vienna)	0.525	0.497	---	0.000**	Support

H_1 is rejected, which means that there is no difference in how travellers within the age range of 19-35 and within the age range of 36-50 rate the quality of the hotels' condition in Salzburg. H_2 and H_3 are supported at the 0.01 significance level. Spearman's Rho and Kendall's Tau B often result in similar values. However, it is advised to interpret the lower one when discrepancies occur (Field, 2013). Thus, there is a relationship between perceived friendliness and perceived competencies (0.650) and between perceived family friendliness and perceived quality of child care or playground (0.497).

The above explained and tested hypotheses demonstrate that big data, and more specifically in this case, the derivations from the prototype data set, can be used to examine the service quality of the Austrian hospitality industry as well as to explore possible relationships and regularities in it.

5. Conclusion and implications

The research demonstrates that it is possible to develop a fully functioning prototype, based on the integration of UGC, which due to the amount of data scraped, can be classified as big data. The SQM does not only represent available data from a rating platform in a new structured way, it also provides a new possibility for

statistical analysis and interpretation through the function of .csv file download. The key implication designated to the project is the facilitation of the decision making process for governmental bodies, DMOs and investors. Therefore, when a decision concerning investments in touristic infrastructure, education or other branches has to be weighted, it can be partially based upon the knowledge extracted from the prototype. The variety of service quality indicators as well as the function to conduct individual queries are the main characteristics of the prototype, allowing for the response to specific questions. In the end, addressing certain shortcomings revealed by the prototype allow for overall service quality improvements and consequently, higher competitiveness and long-term success of a destination.

Within the framework of future research, the authors propose that data be combined from several rating platforms and integrate official statistical data from the Statistik Austria (2016) platform. When considering the combination from different sources, it should be kept in mind that this step requires adjustments of the data base model, e.g. unification of the model in accordance with various data sources. Additionally, further development of the SQM should focus on the possibility to include real-time data. This would allow potential users to base their decisions on the latest data. This is relevant as the overall effectiveness of marketing, processes and strategic planning can be improved by astute use of large amounts of data (Schlegel, 2015). In order to achieve even higher functionality, it is recommended to integrate a statistical output directly into the prototype. This would lead to a more comprehensive understanding of the data, as the user could directly interpret the results without resorting to external software while having at the same time, the visualised output. For further development of the prototype, it is advisable to embark on collaborations with prospective end users. Should users accept the SQM, the prototype could then be released on the market. Otherwise, further development stages would be guided by the remaining phases of the Agile Life cycle.

As UGC is an unstoppable online force and is constantly growing and generating data, one assumes that the creation of the prototype of service quality map is only a starting point for further research and development of the concept.

References

- n.a., 2014, Sep. 4. TripAdvisor study reveals key factors driving hotel engagement. [online] Available at: http://hotelmkt.com/index.php/content/article/tripadvisor_study_reveals_key_factors_driving_hotel_engagement#sthash.xPrzfRco.dpuf [Accessed 20 January 2016].
- Acton, C., Miller, R., Fullerton, D., & Maltby, J., 2009. *SPSS Statistics for Social Scientists*, 2nd Ed. UK: Palgrave Macmillan.
- Atilgan, E., Akinci, S., & Aksoy, S. (2003). Mapping service quality in the tourism industry. *Managing Service Quality: An International Journal*, 13(5), 412-422.

- Attallah, N. F., 2015. Evaluation of perceived service quality provided by tourism establishments in Egypt. *Tourism and Hospitality Research*, 15(3), pp. 149-160.
- Banerjee, S. & Chua, A. Y. (2016). In search of patterns among travellers' hotel ratings in TripAdvisor. *Tourism Management*, Vol. 53, pp.123-131.
- Bonneville, E.,2015, April 8. Web Scraping With Node.js [online]. Available at: <https://www.smashingmagazine.com/2015/04/web-scraping-with-nodejs/> [Accessed 19 January 2016].
- Bortz, J., 2005. *Statistik für Human- und Sozialwissenschaftler*, 6.Auflage. Heildelberg: Springer.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*. 15 (5), pp. 662-679.
- Cleaves, C., 2015, July 22. Web Scraping: Everything You Wanted to Know [online]. Available at: <<http://resources.distilnetworks.com/h/i/111901208-web-scraping-everything-you-wanted-to-know-but-were-afraid-to-ask/181642>> [Accessed 19 January 2016].
- Davenport, T.H., Barth, P., & Bean, R., 2012, July 30. How 'Big Data' Is Different. *MIT Sloan Management Review*. [online] Available at: <http://sloanreview.mit.edu/article/how-big-data-is-different/>, [Accessed 10 January 2016].
- Davenport, T.H., 2013. At the Big Data Crossroads: turning towards a smarter travel experience. *Amadeus*, [online] Available at: www.bigdata.amadeus.com/assets/pdf/Amadeus_Big_Data.pdf [Accessed 10 January 2015]
- Dong, R., McCarthy, K., O'Mahony, M.P., Schaal, M., & Smyth, B., 2012. Towards an intelligent reviewer's assistant: recommending topics to help users to write better product reviews. In Duarte, C. et al (eds.), *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*, pp. 159-168. New York: ACM
- ESRI Inc., 2016. GIS Dictionary. Dot density map. [online] Available at: <http://support.esri.com/en/knowledgebase/GISDictionary/term/dot%20density%20map> [Accessed 20 February 2016]
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*, 4th Ed. London: SAGE Publications Ltd.
- Fox, S. & Do, T. (2013). Getting real about Big Data: applying critical realism to analyse Big Data hype. *International Journal of Managing Projects in Business*, Vol. 6, 4, pp.739 – 760.
- Harris, C. (2013). Diving Into Big Data. *Canadian Underwriter*, Vol. 80 (2), pp. 28-36.

- Hass, K. B. (2007). The blending of traditional and agile project management. *PM world today*, 9(5), 1-8.
- HolidayCheck, 2015. Über uns. Das Unternehmen. [online] Available at: <http://www.holidaycheck.at/ueberuns> [Accessed 27 June 2015].
- Irudeen, R. & Samaraweera, S. (2013). Big data solution for Sri Lankan development: A case study from travel and tourism. *Advances in ICT for Emerging Regions (ICTer), 2013 International Conference on*, pp. 207-216. IEEE.
- Jackson, S. (2014). Prediction, explanation and big(ger) data: a middle way to measuring and modelling the perceived success of a volunteer tourism sustainability campaign based on 'nudging'. *Current Issues in Tourism*, DOI: 10.1080/13683500.2014.898616.
- Kim, G.-H., Trimi, S., & Chung, J.-H. (2014). Big-Data applications in the Government sector. *Communications of the ACM*, 57 (3), pp. 78-85.
- Lau Pei Mey, Khatibi Akbar, A., Yong Gun Fie, D. (2006). Measuring Service Quality and Customer Satisfaction of the Hotels in Malaysia: Malaysian, Asian and Non-Asian Hotel Guests. *Journal of Hospitality and Tourism Management*, 13(2), pp.144-160
- Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, 29(3), 458-468.
- Lohr, S., 2013, February 1. The origins of 'big data': An etymological detective story. *The New York Times*. [online] Available at: http://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/?_r=0 [Accessed 20 January 2016].
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A.H., 2011. Big Data: The Next Frontier for Innovation, Competition, and Productivity. *McKinsey Global Institute Report*. [online] Available at: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation [Accessed 10 January 2016].
- Marine-Roig, E. & Clave, S.A. (2015). A Method for Analysing Large-Scale UGC Data for Tourism: Application to the Case of Catalonia. In Tussyadiah, I and Inversini, A., eds. 2015. *Information and Communication Technologies in Tourism 2015. Proceedings of the International Conference in Lugano, Switzerland, February 3-6, 2015*. Cham: Springer. pp.3-18.
- McAfee, A. & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, 90 (10), pp. 60-68.

- Qiang, Y., Law, R., Gu, B. & Chen, W. (2011). The influence of user-generated content on travel behaviour: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior*, 27, pp.634-639.
- Qiao, X., Zhang, L., Li, N. & Zhu, W. (2014). Constructing a Data Warehouse Based Decision Support Platform for China Tourism Industry. *Information and Communication Technologies in Tourism 2014*, pp.883 - 894.
- Rouen, E., 2012. Big Data won't solve your company's problems. *Fortune*. [online] Available at: <http://fortune.com/2012/03/19/big-data-wont-solve-your-companys-problems/> [Accessed 10 January 2016].
- Russom, P., 2011. Big Data Analytics. *TDWI Best Practices Report*, [online] Available at: <http://public.dhe.ibm.com/common/ssi/ecm/en/iml14293usen/IML14293USEN.PDF> [Accessed 10 January 2016].
- Schlegel, V., 2015. The mobile revolution in tourism: what are and will be the factors of success?, Lugano, *Enter Conference2015*, [online] Available at: <https://www.youtube.com/watch?v=6yo1K71pQ4> [Accessed 19 June 2015].
- Sigala, M. (2008). Web 2.0, social marketing strategies and distribution channels for city destinations: enhancing the participatory role of travellers and exploiting their collective intelligence. In M. Gascó-Hernández, T. Torres-Coronas (Eds.), *Information communication technologies and city marketing: digital opportunities for cities around the world*. IDEA Publishing.
- SOCAP International, 2013. Now Arriving: Big Data In The Hospitality, Travel, And Tourism Sector. [online] Available at: https://cdn.fsbx.com/hphotos-xap1/v/t59.2708-21/10656875_491040414368048_1625910113_n.pdf/BIG-DATA-IN-THE-HOSPITALITY-TRAVEL-AND-TOURISM.pdf?oh=36c87a9d98109ed51077da2ec517a1fe&oe=54B5BD43&dl=1 [Accessed 10 January 2016].
- Statistik Austria, 2016. Bundesanstalt Statistik. [online] Available at: http://www.statistik-austria.at/web_de/statistiken/index.html [Accessed 18 January 2016].
- Weinberg, B.D., Davis, L., & Berger, P.D. (2013). Perspectives on Big Data. *Journal of Marketing Analytics*, Vol. 1, 4, pp. 187 - 201.
- Witten, I. H., & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*, 2nd Ed. San Francisco: Morgan Kaufmann.
- Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, 44, pp. 120-130.