Big Data in Tourism

Karampatsou Maria

SCHOOL OF ECONOMICS, BUSINESS ADMINISTRATION & LEGAL STUDIES

A thesis submitted for the degree of

Master of Science (MSc) in Hospitality and Tourism Management

January 2018
Thessaloniki – Greece
Student Name: Maria Karampatsou
SID: 1109150017
Supervisor: Prof. Eleni Mavragani

I hereby declare that the work submitted is mine and that where I have made use of another’s work, I have attributed the source(s) according to the Regulations set in the Student’s Handbook.

January 2018
Thessaloniki - Greece
Abstract

This dissertation was written as part of the MSc in Hospitality and Tourism Management at the International Hellenic University.

Over the last years, the Internet brought a tremendous change in the way people interact with each other, do business and consume services (Rob Law, 2014). At the same time the use of smartphones, PCs and laptops has allowed the creation and communication of data that include every possible form of information: from digital pictures, GPS signals and climate information to personal data of individuals such as names, addresses, posts on social networking websites, bank details and medical information. In 2010, the Economist published a report titled “Data, Data Everywhere”. Its author Kenneth Cukier writes: “...the world contains an unimaginably vast amount of digital information which is getting ever vaster more rapidly... The effect is being felt everywhere, from business to science, from governments to the arts.” (Cukier, 2010)

This vast amount of Data, widely known as “Big Data” is nowadays one of the key factors that can drive innovation and growth, affecting all sectors of the economy (Monica Bulger, 2014), among them the tourism industry. This research is dedicated in examining the role of Big Data phenomenon in modern tourism industry. Its methodology is based on a combination of quantitative and qualitative data gathered in the form of questionnaires and in-depth face-to-face and phone interviews. The research is also underpinned by a rich collection of literature references and case studies. Going a step beyond this, the analysis is concluding by suggesting ways of better implementation of Big Data on behalf of tourism companies, with ultimate goal the creation of a more efficient and tailor-made travel experience.

At this point, I would like to acknowledge the people that helped me and contributed significantly in the successful completion of my dissertation. Initially, I would like to thank my supervisor, Ms Eleni Mavragani, for her help and trust she showed to me during my thesis. Also, I am deeply indebted to my family, who encouraged and motivated me during my studies.

Keywords: Big Data, Big Data analytics, tourism industry, innovation

Karampatsou Maria
14/01/2018
Table of Contents

ABSTRACT ........................................................................................................................................... III

INTRODUCTION ..................................................................................................................................... 6

LITERATURE REVIEW .......................................................................................................................... 8

HISTORY AND EVOLUTION OF BIG DATA ......................................................................................... 8

DEFINITION OF BIG DATA ................................................................................................................... 9

SOURCES OF BIG DATA .................................................................................................................... 13

BIG DATA’S VALUE PROPOSITION FOR TOURISM INDUSTRY ....................................................... 15

   Personalization of Services .............................................................................................................. 15

   Delivering New and Competitive Products and Services ................................................................. 17

   Better Decision Support – Internal Operations .............................................................................. 18

   Revenue Management – Improving Pricing Strategy .................................................................... 20

BIG DATA CHALLENGES ..................................................................................................................... 22

   Privacy and Security ......................................................................................................................... 22

   Data Ownership ............................................................................................................................... 24

   Cost/Operational Expenditures ....................................................................................................... 25

   Data skilled personnel ..................................................................................................................... 25

METHODOLOGY .................................................................................................................................... 1

PRIMARY RESEARCH RESULTS ......................................................................................................... 3

   QUANTITATIVE RESEARCH – RESULTS AND ANALYSIS ............................................................. 3

   QUALITATIVE RESEARCH – PRESENTATION OF INTERVIEWS .................................................... 11

DISCUSSION AND CONCLUSION ...................................................................................................... 16

LIMITATION AND FURTHER RESEARCH ......................................................................................... 23

BIBLIOGRAPHY .................................................................................................................................... 25

APPENDIX ........................................................................................................................................... 32

   QUESTIONNAIRE ............................................................................................................................ 32

   DESCRIPTIVE STATISTICS ............................................................................................................. 32

   RELATIVE FREQUENCY ................................................................................................................. 32

   INTERVIEWS ................................................................................................................................... 32
List of Tables

Table 1: UNECE Big Data Classification ......................................................................................... 14

List of Figures

Figure 1: Data Analytics .................................................................................................................. 10
Figure 2: The 5Vs of Big Data-Volatility, Variety, Velocity, Veracity and Volume ......................... 12
Figure 3: Hipmunk-Agony Index .................................................................................................. 18
Figure 4: Kayak – Flight Price Forecasting Index ......................................................................... 21
Figure 5: Skills and Competences of Data Scientists ..................................................................... 27
Figure 6: Sector of Business Operations ....................................................................................... 3
Figure 7: Base of Business Operations .......................................................................................... 4
Figure 8: Number of company’s employees ................................................................................. 5
Figure 9: Sources of Data employed by the company ................................................................. 5
Figure 10: Big Data’s level of use in Company’s business Strategy .............................................. 6
Figure 11: Big Data as Competitive Advantage ............................................................................ 6
Figure 12: Benefits of Big Data ...................................................................................................... 7
Figure 13: Challenges of Big Data ................................................................................................ 8
Figure 14: Importance of Data Skilled Personnel ......................................................................... 8
Figure 15: Employment of data skilled personnel ......................................................................... 9
Figure 16: Improvement in Big Data analytics Performance ...................................................... 9
Figure 17: Interest for further information about Big Data ........................................................ 10
Figure 18: Data Governance Disciplines .................................................................................... 19
Figure 19: DIKW Hierarchy - The Knowledge Pyramid ............................................................. 21
Introduction

One of the key characteristics of digital age is the exponential growth of Data. We live in a world, where with every online action we take, a digital trail is generated. The number of Data produced by Internet users nowadays is growing unabated and IT professionals estimate 4300% increase in annual data generation by 2020 (Reddy, 2016). This phenomenon, widely known as “Big Data”, meaning the vast amount of heterogeneous data being produced every second of the day, is becoming unquestionably the new frontier for growth and innovation as well as the tool for driving knowledge and competition.

No wonder Big Data is drawing the attention not only of academics and researchers but also of well-known companies and leading organizations. At its 45th session in 2014, the UN Statistical Commission officially recognized that “Big Data constitute a source of information that cannot be ignored” (C. L. Hammer, 2017). Accordingly, Gartner Inc, research and advisory company, in its report “Top 10 Strategic Technology Trends for 2017” points out that by 2018 at least 200 of the world’s largest companies are expected to exploit intelligent applications and leverage Big Data to improve customer experience (Panetta, 2016). At the same time, James Mundie, Senior Advisor of the CEO of Microsoft stresses out that “Data are becoming the new raw material of business.” (Cukier, 2010).

From this Big Data deluge, the tourism industry couldn’t remain unaffected. Big Data brings undoubtedly great opportunities to tourism companies, by providing insights to customer preferences and empowering them to ameliorate the whole travel experience. However, as every coin has two sides, Big Data benefits often go hand-in-hand with significant challenges that require proper action (Akerkar, 2012).

The purpose of this dissertation is to explore the reshaping role of Big Data in modern tourism industry. It addresses every tourism company and organization which wishes to acquire a better understanding of what Big Data actually is, how it can be used for boosting innovation and business growth and what are the possible risks that come along with its implementation. However, since Big Data affects all sectors of industry (Monica Bulger, 2014), the dissertation holds an interest for every person who wants to deepen its knowledge about this phenomenon, concentrating mostly on the Management aspect of it.
The study begins with the presentation and analysis of the literature review; particular emphasis at this point is given on the Big Data benefits and challenges, underpinned by distinctive case studies of well-known tourism companies from around the world. It proceeds with the presentation of primary research results, meaning a set of data gathered through quantitative and qualitative research with particular focus on Greek tourism companies. The study concludes by posing some key questions concerning the role of Big Data as a tool for gaining competitive advantage in tourism industry and by presenting the Big Data trends for the years to come.
Literature Review

Aim of this chapter is to provide an in-depth analysis of the existing literature relating to Big Data. It begins with a short reference in the history and evolution of data storage and analysis and then proceeds to the definition of Big Data and its sources. The chapter closes with an elaborate examination of the key benefits and challenges that arise from Big Data implementation.

History and evolution of Big Data

Data storage and analysis as a way of gaining deeper knowledge on business activities and organizing an efficient business strategy is, by all means, not a new idea; in fact, it is as old as decision making itself (Group, 2013).

The ability to store and analyze information has known a gradual evolution: the abacus in Babylon was the first device ever constructed in order to perform calculations. The Library of Alexandria is considered to be the largest collection of data of ancient times and the Antikythera Mechanism the earliest discovered mechanical computer. In 1663 John Graunt was the first who gathered and analyzed data relating to the rate of mortality in London, caused by the bubonic plague that was ravaging Europe at that time. His book “Natural and Political Observations Made upon the Bills of Mortality”, was the first statistical analysis of data ever recorded, establishing John Graunt as the pioneer of the field of statistics (Marr, 2015).

Data, in modern form, first appear in 1889 when Herman Hollerith, the father of modern automatic computation, invented a computing system for storing and processing information, with the aim to organize census data (Cruz, 2011). In the landmark year 1989, the World Wide Web was invented by Tim Berners-Lee, a British computer scientist. In the same year, Howard Dresner defined the term “Business Intelligence” as "concepts and methods to improve business decision making by using fact-based support systems". The years that followed, till 2005, are known as Analytics 1.0. and characterized by relatively small data sources, stored mostly in enterprise warehouses.

1 https://www.winshuttle.com/big-data-timeline/
and the data analysis was mainly descriptive (C. L. Hammer, 2017). At that time, very few organizations focused exclusively on data analytics; for the majority, data analysis was marginal to their business strategy.

With the rise of personal computing, the production of online data started accelerating rapidly and data began pouring from any-where and any-time. It was in 2005 when Roger Mougala-s, the current director of market research at O'Reilly Media, first coined the term “Big Data” defining it as “the wide range of large data sets almost impossible to manage and process using traditional data management tools—due to their size, but also their complexity”. Next year, in 2006, Yahoo created the open-source Hadoop, as a way of storing and processing data. The period 2005-2012, known as Analytics 2.0, is characterized by the use of Big Data and analytics mostly by online firms like Google and Yahoo (Thomas H. Davenport, 2013). Contrary to Analytics 1.0, in Analytics 2.0 data are very often externally sourced, large and usually unstructured. The flow of data is fast and it creates the need for an equally fast storage and analysis.

Nowadays, Big Data is among the top Business Intelligence trends. Analytics 3.0 era has begun and has as most important trait the accessibility of all and not only the online firms, in the data-driven economy. Any company in any industry can leverage Big Data, in order to shape its strategy and gain a competitive advantage against its business rivals.

**Definition of Big Data**

Although there is no official definition of “Big Data”, the majority of the now existing terms converge to the one originally given by Roger Mougala-s. Thus, they are most commonly used to describe “datasets so large and complex that could not be captured, managed and processed by general computers within an acceptable scope” (Baggio, 2016). Key parameter of all Big Data terms is the massive volume of both structured and unstructured data that are generated through the Internet use and are difficult to become warehoused and further analyzed by simply using traditional software techniques.
Big Data however in its raw form doesn’t have a true value; it is simply a vast number of heterogeneous data scattered across the internet. In order to confer real insights, it requires further processing and analysis. Hence, Big Data entails **Big Data analytics**, meaning the science of finding patterns and reaching to conclusions by using algorithms and mechanical processes. Respectively, Big Data technologies are replacing traditional database software and single servers, which are not capable any more of handling the volume and speed of Big Data. A new generation of tools like the open-source software Hadoop is used for transforming the raw data into valuable insights (Group, 2013). Thanks to these, businesses are now able to transcend from descriptive analytics - that answer the question “What happened?” and are traditionally based on charts, graphs and tables (Gartner, n.d.) - to predictive (Liran Einav, 2013) and prescriptive analytics that answer the questions “What will happen?” and “How can we make it happen?” (Gartner, n.d.). The new frontier of analytics provides businesses with a better understanding of the future and allows the optimization of business operations by maximizing the core values and mitigating any possible risks.

![Figure 1: Data Analytics](source:Siodmok A., 2017. From best practice to next practice)
Characteristics of Big Data

Doug Laney, analyst at research and advisory company Gartner, was the first who defined three of the most important characteristics of Big Data; in a 2001 METAGroup research publication titled “3D data management: Controlling data volume, variety and velocity”, he introduced the widely known 3Vs Model of Big Data (3V stands for Volume, Variety and Velocity) (Laney, 2001). In 2012 researchers and practitioners updated this definition by adding one more “V” for “Veracity”; in 2014 it was extended into 5Vs Model for “Value” (H. Song, 2017) and today it is reported as the 7Vs Model namely Volume, Variety, Veracity, Velocity, Variability, Visualization and Value of Data (U. Sivarajah, 2016) (Demunter, C., 2017).

**Volume:** One of the main characteristics of Big Data refers to its *large scale*: from terabytes to zettabytes, masses of data sets are generated every second of the day from various sources, like smartphones and digital devices (G. Bello-Orgaza, 2015) (Monica Bulger, 2014). For example, well-known companies like Facebook generate over 500 terabytes of data, forcing the development of new data mining techniques (U. Sivarajah, 2016).

**Variety:** Characterized by a *wide range* of data types, Big Data are found and captured in various forms like text, video, image content, noise, audio and sensor data; their forms are not only diverse but also dissimilar and of different quality level (H. Song, 2017).

**Veracity:** This characteristic stands for the *accuracy* and *correctness* of information (G. Bello-Orgaza, 2015). Researches are using veracity for referring to the untrustworthiness, doubts, imprecision, messiness inherent in Big Data (U. Sivarajah, 2016) (Monica Bulger, 2014).

**Velocity:** In today’s digital world, the rise of the number of digital devices goes hand in hand with a *high influx rate* of large data sets, that either generate new data or are added up to the already existing ones. This great speed of Big Data creates the need for equally fast methods of collection and analysis of the information (G. Bello-Orgaza, 2015) (H. Song, 2017).
**Variability:** This term is used to describe the data whose meaning is *constantly* and *rapidly changing*. For example, in Tweeter the meaning of a word can be a totally different even in the same tweets (U. Sivarajah, 2016).

**Visualization:** Visualization of data is about representing key information and knowledge effectively by using various visual formats such as a pictorial or graphical layout. For example eBay is using “Tableau” - a Big Data visualization tool - for transforming large and complex datasets into clear depictions, facilitating thus the employees and decision makers to grasp complex information and come to fast and accurate conclusions (U. Sivarajah, 2016).

**Value:** Last pillar of Big Data – but equally important – is Value. As Big Data is in fact generated by Internet users, it constitutes *valuable source of information* concerning end users’ needs and preferences. The value that could be extracted from its further processing could lead to useful business information and it is often described as the “backbone” of the internet economy (U. Sivarajah, 2016).

---

**Figure 2:** The 5Vs of Big Data-Volatility, Variety, Velocity, Veracity and Volume

Source: C. L. Hammer, 2017
Sources of Big Data

Unlike statistical ("made") data that are created for specific purposes, Big Data is actually a byproduct “found” in social networks, business and administrative systems and the Internet of Things (IoT). United Nations Economic Commission for Europe (UNECE), classifies Big Data sources into three main categories (UNECE, 2013) (Table 1):

- **Social Networks** (human – sourced information): This term refers to the online platforms, where people record their experiences, share pictures, post reviews and exchange messages, building thus social relations. Facebook, Twitter, Instagram, LinkedIn are based on human-sourced information that is digitalized and stored. Data belonging in this category does not follow a specified format; on the contrary it is most commonly ungoverned and unstructured (C. L. Hammer, 2017) (H. Song, 2017).

- **Traditional Business systems** (process-mediated data): According to UNECE, this category includes all processes set by a firm in order to record and monitor its operations and strategy, such as registering a client, taking an order, proceeding to a commercial transaction etc. These process-mediated data, including administrative data produced by public agencies, are usually well structured; that means it is easy to input, query, store and analyze in relational database systems (H. Song, 2017).

- **Internet of Things** (machine – generated data): Sensors that are embedded in our computing devices for recording and measuring situations in the physical world (e.g. weather, traffic and GPS sensors) as well as internet connectivity and web portals, like Booking.com or Airbnb.com create a vast amount of machine – generated data. These data are commonly well-structured and suitable for further computer processing, however due to their vast size and fast speed, traditional methods of analysis cannot be applied. (Demunter, 2017)
### Social Networks (human-sourced information)

- 1100. Social Networks: Facebook, Twitter, Tumblr etc.
- 1200. Blogs and comments
- 1300. Personal documents
- 1400. Pictures: Instagram, Flickr, Picasa etc.
- 1500. Videos: Youtube etc.
- 1600. Internet searches
- 1800. User-generated maps
- 1700. Mobile data content: text messages
- 1900. E-Mail

### Traditional Business systems (process-mediated data)

- 21. Data produced by Public Agencies
  - 2110. Medical records
- 22. Data produced by businesses
  - 2210. Commercial transactions
  - 2220. Banking/stock records
  - 2230. E-commerce
  - 2240. Credit cards

### Internet of Things (machine-generated data)

- 31. Data from sensors
  - 311. Fixed sensors
    - 3111. Home automation
    - 3112. Weather/pollution sensors
    - 3113. Traffic sensors/webcam
    - 3114. Scientific sensors
    - 3115. Security/surveillance videos/images
  - 312. Mobile sensors (tracking)
    - 3121. Mobile phone location
    - 3122. Cars
    - 3123. Satellite images
- 32. Data from computer systems
  - 3210. Logs
  - 3220. Web logs

---

**Table 1: UNECE Big Data Classification**

Source: UNECE, 2013.
Big Data’s Value Proposition for Tourism Industry

The disruptive power of Big Data can result to significant advantages for all early adopters. By focusing on travel and tourism industry, the research presents and analyses the most important Big Data benefits, underpinned by successful case studies from around the world.

Personalization of Services

One key attribute of Big Data is that it enables the personalization of company’s services. According to Habegger et al., “personalization is the process of adapting the output of a system to a user’s context and profile”. In this vein, a personalized service aims at providing the traveler with the most suitable solution based on his needs, obtaining at the same time access to the minimum amount of his personal data. As primary source of information may be used the customer profile characteristics, such as interests, preferences, membership in groups, academic and professional background, geographic location etc. (B. Habegger, 2014). Particularly nowadays the personalization of tourism services is constantly gaining in importance; the modern traveler is attracted by flexible and up to date services, rating “the ability to value me” and “the ability to understand my needs” as the most important prerequisites of customer loyalty (Sigala, 2016) . To that end, academics and tourism professionals emphasize on the need for transition from traditional CRM, which focuses solely on handling the customer transactions, to social CRM or CRM 2.0, which considers the traveler as a value co-creator and aims at achieving personalization of all services provided (Sigala, 2016) (Werner Kunz, 2016).

Big Data tools enable this strategic shift since it allows a deeper knowledge and insight in traveler’s needs and preferences, setting thus the foundations for a more holistic approach in the design of a customer centric strategy. Thanks to Big Data, tourism companies are nowadays able to process and analyze structured as well as unstructured data; the latter contain a great amount of multifarious and useful information about travelers that until now was inaccessible to companies. Furthermore, Big Data paves the way in the gathering of customer’s information by using various sources, such as social
media, blogs and e-commerce transactions, limiting thus any invalid and inaccurate information; those data are being processed and finally combined in one single user profile (B. Habegger, 2014).

Well-known online firms are already employing Big Data tools for achieving personalization of services. For example LinkedIn.com suggests future employers based on the user’s professional background; Google.com is using the search history to personalize user’s searches; and Facebook.com suggests potential friends based on existing contacts and groups in user’s profile (B. Habegger, 2014). The latter extended Big Data use also in personalized advertising; in particular Facebook’s tool called “Custom Audiences” allows tourism providers to use their list of customers and target them directly with advertisements posted on Facebook website (Group, 2013). In this case, travelers only receive advertisements that are most interested in, destinations that they wished to visit, without being overwhelmed by information that has no meaning for them. Particularly interesting is the discovery made by the commerce marketing company Criteo, according to which, when a personalized advertisement is posted there are eight times more chances that the user will select and buy the specific product or service offered (Group, 2013). It becomes apparent, that with customer’s profile analysis, a tourism provider can anticipate personal needs and offer for example a price affordable hotel room to a specific customer, boosting thus the sales and increasing satisfaction and loyalty.

A representative example of a tourism provider who successfully implemented Big Data analytics for enhancing its personalization strategy is MGM Resorts International. The company deployed successfully the aforementioned Facebook’s Big Data tools and as a result company’s revenues have grown 300% in a three-year period (Hertzfeld, 2014). The springboard of MGM personalized services is customer’s enrollment to company’s “M Life program”: once the customer’s data are registered into program’s database, MGM Resorts is able to track customer’s movements in all of its 19 properties and place the rights adds in Facebook. According to Nick Mattera, social media manager of MGM Resorts International “if travelers start looking at room rates on MGM.com and then leave to compare prices on a competitor’s site MGM can drive them back to its site with a relevant Facebook offer” (Kuchinskas, 2014) (Hertzfeld, 2014).
Delivering New and Competitive Products and Services

One more important benefit of Big Data is their enhancing role in the delivery of new and competitive products and services (Group, 2013). Thanks to predictive and prescriptive analytics, companies can nowadays achieve greater insights into markets and find unique and creative ways to respond rapidly to the emerging opportunities (U. Sivarajah, 2016). Innovative online travel companies have already implemented Big Data into their business strategy in order to refine their products and differentiate themselves from their competitors. Representative examples are considered to be Amadeus, which is one of largest distributors of leisure packages and technology providers and Hipmunk, the online travel website. More explicitly:

Since the beginning of its foundation, Amadeus incorporated Big Data analytics in its business strategy. Amadeus Data Centre, with more than 49 Petabytes of storage, 55,000 transactions per second and 3.9 million net bookings per day owns a vast number of data relating to its customer needs and preferences, likes and dislikes (Amadeus, n.d.). The company uses this information for designing and implementing innovative tools, such as “Amadeus Featured Results” and “Amadeus Extreme Search”.

In particular “Amadeus Featured Results” is a revolutionary data-based technology that combines the four most relevant flights’ itineraries – fastest, lowest priced, most popular and sponsored, in which the customer might be interested in (Amadeus, n.d.). Aim of this application is to offer the best possible choice, increasing thus customer satisfaction and loyalty. Supplementary to that, “Amadeus Extreme Search” (Amadeus, n.d.) enables the traveler to enter his preferences concerning the length of time for the trip, the number of passengers, the budget and desired minimum temperature of the destination and to receive recommendations for possible destinations.

Respectively, the online travel website Hipmunk is known for delivering innovative functionalities in travel industry. The site uses Big Data analytics to create customer-centered applications. Rather than solely ranking services based on of their ratings and prices, Hipmunk introduces the “Agony Index” (Figure 3), which is used in ranking flights combining multiple criteria such as price, duration of flight and number of stops (Group, 2013), (Delventhal, 2015). In April 2017, Hipmunk presented two more data-based tools: “Basic Coach” fares, that provides the travelers with call-outs portraying the trade-offs
for booking a low cost flight, and “*No Frills Fight*”, that informs travelers about any flight peculiarities (Hipmunk, 2017).

![Hipmunk Flight Comparison](image)

**Figure 3: Hipmunk-Agon Index**  
Source: Hipmunk, 2017

It becomes apparent that both travel companies are employing Big Data analytics for improving their products and services, transforming traveler’s experience and gaining a competitive advantage over their business rivals.

**Better Decision Support – Internal Operations**

Big Data can also ameliorate the *internal operations and decision – making* in a business. Business managers are employing it as a powerful tool for speeding up decisions that lead to an improved and customer-focused approach, bringing thus more value to their customers.

*British airways* is a representative example of an air travel provider, who makes extensive use of all data collected from its passengers. In the era of economy travel, the company implements Big Data analytics not only to *segment* its customers better but also to come up with *critical* and *real-time operational decisions*. In 2012 the company assembled all the information available from its website as well as the Executive Club
loyalty program and created the “Know Me” program (British, n.d.) (IPSOS, 2015), whose main characteristics are (Group, 2013):

- **Personal recognition.** Know me Program recognizes the travelers who are loyal to British Airways and expresses appreciation with targeted offers and benefits. The program uses Google Image Search for helping company’s staff to recognize million-mile fliers when they enter airport and to provide tailor-made service.

- **Service excellence and recovery.** The program tracks the services offered to its passengers; if for example a passenger experienced a flight delay, the staff can apologize in person and regain customer’s trust with an offer.

- **Offers that inspire and motivate.** The purpose is to provide the best possible offer according to the needs and wants of each passenger. For example, the Program can detect in which cases passengers choose a window seat (e.g. for shorthaul flights) and aisle seat (e.g. for longhaul flights) and it can repeat the same pattern automatically in the next booking of the passenger.

Thus British airlines concentrates on building a one-to-one relationship with its passengers, increasing their satisfaction and loyalty to the company (MOROZ, 2017).

Another important example of Big Data’s role in the improvement of internal operations and decision-making refers to the company’s energy consumption and environmental profile. Since November 2012 two InterContinental Hotels in San Francisco, *InterContinental Mark Hopkins (ICMH)* and *LEED Gold Certified InterContinental San Francisco (ICSF)* are collaborating with *Stem*, a startup company that uses Big Data analytics for energy management. The company’s software gathers and analyzes data from sources like weather forecasts, electricity rates and hotel’s energy consumption, with the aim to reduce energy costs and carbon footprint for both the hotels. The first results are very positive and the hotel chain is expected to benefit from Stem’s analytics by reducing energy expenditures up to 10% -15%. (Dragosavac, 2015) (Group, 2013) (Stem, n.d.).
Revenue Management – Improving Pricing Strategy

Without doubt, price is one of most critical decision factors for travelers when they are booking an air flight or a hotel room. Revenue management is the process by which a company sets the optimal price for its products and services, making them more appealing to potential customers. Tourism providers, such as Marriott International and Kayak.com, have since long time understood the importance of improving their revenue management capabilities and have turned into Big Data analytics for assistance and guidance.

Marriott International is a prime example of a hotel chain that used Big Data for predicting the optimal price of its rooms. Its strategy focuses on improving the related algorithms for a faster and more accurate data analysis, making the revenue management accessible over the Internet, as well as extending it in all levels of hotel chain’s operations such as the restaurants, catering and meeting areas (Group, 2013). Furthermore, Marriot’s data scientists are not restricted only to the analysis of internal data for predicting demand; they are combining also customers’ data with unstructured and semi-structured data gathered from external sources such as weather predictions and event schedules (Simpson, n.d.). This strategy allows Marriott International to optimize its product and services and adjust hotel’s price accordingly (Dragosavac, 2015). In an industry as competitive as hospitality, price differentiation - meaning displaying different room rates for every customer – is crucial since it offers a significant competitive advantage and equals to a high RevPar (Revenue Per Available Room).

The travel search site Kayak.com is also utilizing Big Data tools effectively for ensuring a consistent price optimization. Algorithms, analytical models and machine learning are tracking and analyzing competitors’ prices, allowing the company to build a competitive pricing strategy. In 2013 Kayak.com has also launched a new predictive model, the Flight Price Forecasting (Figure 4). This tool is used in forecasting whether the price of a specific flight will go up or down in the seven days (Charriez, 2013) (Group, 2013) allowing the travel search site to offer better-targeted products and services and in the long-term to drive new revenue streams and improve its revenue management.
Concluding is evident that the scope and applicability of Big Data phenomenon offers a dynamic perspective to travel companies, bringing along Big Opportunities and empowering organizations at all possible levels such as strategic, organizational and operational (U. Sivarajah, 2016).

Figure 4: Kayak – Flight Price Forecasting Index
Source: Charriez, A., 2013
**Big Data Challenges**

So far it has been proved that Big Data is a tool that, when it is appropriately managed, has the potential to provide deep insights into the tourism market thus generating new knowledge and setting the foundations for innovation and differentiation for tourism firms (Jukić et al., 2015) (U. Sivarajah, 2016). However, as every coin has two sides, Big Data brings not only Big Opportunities but also *Big Challenges* (U. Sivarajah, 2016). Researchers refer to a great deal number of risks and challenges that tourism companies are confronted with when employing Big Data in their business strategy.

According to researchers and academics, the broad challenges of Big Data can be divided into three main categories (U. Sivarajah, 2016):

- **Data challenges** that derive from the Big Data characteristics (e.g. data volume, variety, velocity, veracity, volatility, quality, discovery and dogmatism).
- **Process challenges** that refer to the technical aspects of capturing data, transforming and integrating to the computing system, selecting the right model for analysis and extracting the desired results.
- **Management challenges** that focus on issues such privacy and security, data ownership, cost/operational expenditures and lack of data skilled personnel.

Although all three categories are of equal importance, this study focuses mostly on the management related challenges.

**Big Data Management Challenges**

**Privacy and Security**

Privacy and security issues has always belonged to the major Big Data concerns, posing a challenge for both companies and legislators.

In the Big Data industry, it is often encountered the phenomenon that *private customers’ data* are being passed from one firm to another, without the owner’s knowledge and consent. It all start with customers providing personal data to companies, which then pass it to data aggregators, which on their turn may transfer it or even sell it to any interested party, such as polling companies, researches or government agencies. Although that certain benefits can derive from this information
supply chain and the data sharing – like deeper knowledge of the market, identification of fraud, stronger relationships between the companies and better operational performance – the risk of severe misuse of personal data is nowadays stronger than ever and any privacy breach can result to a harmful customer impact. Ethical questions on the responsibility and liability of the company who initially gathers and transfers users’ personal data also arise (Martin, 2015). The challenge in this case is for companies to find the perfect balance and to ensure that the fine line between using Big Data and ensuring privacy rights is not crossed. (U. Sivarajah, 2016)

Security challenges are similar to those of traditional data; Big Data are vulnerable to hackers and cyberattack, malware poses a threat to data security when at the same time security controls which can ensure the resilience of information, are not always adequate. Confidentiality, integrity and accountability of Big Data should not be taken for granted, especially when considering the nature and characteristics of Big Data (high variety, volume and velocity) and ubiquity of data sources (U. Sivarajah, 2016). At the same time security breaches pose a threat not only for customers but for the companies as well. According to the study “2017 Cost of Data Breach” (Ponemon, 2017) which is published annually by the Ponemon Institute, the average size of data breach (meaning the number of sensitive or confidential data stolen, lost or put at risk) increased by 1.8% in 2017 and the cost of this breach for the businesses and the organizations is estimated around $3.62 million (Ponemon, 2017). It’s worth mentioning that the study for 2017 included 11 countries and two regional samples: United States, United Kingdom, Germany, Australia, France, Brazil, Japan, Italy, India, Canada, South Africa, the Middle East (including the United Arab Emirates and Saudi Arabia) and ASEAN region (including Singapore, Indonesia, the Philippines and Malaysia).

Understanding the gravity of the situation, legislators are striving to respond to the new era’s challenges. On European level, a major step towards the strengthening of customers’ fundamental rights in the digital age is the General Data Protection Regulation (GDPR) (EU) 2016/679 adopted by the European Parliament and the Council in 2016. This Regulation as being part of the Digital Single Market Strategy, responds to Big Data’s privacy and security challenges, by strengthening citizens’ rights to data protection with regard to the further processing and free movement of their personal
data. From the one side, people are guaranteed free and easy access to their data and from the other side, businesses and organizations are obliged to inform their customers about data breaches that could have a negative effect on them (Frangoul, 2017). GDPR will apply as law and will harmonize the data protection legislation in all Member States of EU; this clearly constitutes an advantage for businesses since they have to conform to only one law rather than 28. The Regulation entered into force on May 2016 and will apply from 25 May 2018 (Commission, n.d.).

Data Ownership
Closely linked with privacy and security issues is also data ownership. For example, a key question that often arises is “who is the owner of the data of a Facebook account? The social media provider - in that case Facebook - or the user of the account?” It is generally perceived that both parties, the social media provider and the user own the data (U. Sivarajah, 2016).

Referring to European Union law, the status of data ownership rights is still ambiguous and unclear. According to the Digital Economy Working Paper 2017 published by the Joint Research Centre of European Commission, there is a definition of partial and limited ownership data rights in the EU Database Directive (1996), the General Data Protection Regulation (2016), as well as in some provisions in the Trade Secrets Protection Directive (2016) and in general contract law. From the one side, the Database Directive recognizes some limited property rights to data collectors, although those rights are restricted to a certain degree by the jurisprudence of European Court of Justice (ECJ). From the other side, the newly adopted GDPR recognizes certain rights to data subjects, such as the right not to be subject to data processing without an informed consent, the prohibition to use personal data for other purposes than originally intended, the right of the data subject to access its data, the right to be forgotten and the right to data portability. However GDPR doesn’t recognize to ‘data subject’ full ownership rights. Article 4(1) of the GDPR defines personal data as "any information relating to an identified or identifiable natural person ('data subject')". This definition implies that there a second category, that of non-personal data. Social links and online
reviews are included in this category and their ownership still remains undefined (Nestor Duch-Brown, 2017).

In general data ownership is a complex issue since it brings to surface deeper concerns about control and accuracy of data (U. Sivarajah, 2016) posing a significant challenge for legislators and business companies.

**Cost/Operational Expenditures**

When it comes to Big Data implementation, one of the biggest challenges that companies are confronted with, relates to the operational expenditures.

Taking the example of a tourism company which wants to leverage Big Data analytics, the options are basically two: the company can either create an internal Data Analysis Department or assign to a specialized company the running of Big Data analytics on its behalf. In the first case, the company can take advantage of open-source software that is free (e.g Hadoop) or inexpensive. Here, the cost is mostly related to the purchase of technological equipment – plus its maintenance – and to the employment of data-skilled personnel, since Big Data analytics are labor-intensive (Group, 2013). In the second case, the company can use the services and the know-how of a Cloud and Data Platform provider. Advanced analytics, Big Data and machine learning are embedded in the services provided and the cost in this case, although it may be less than that in the first option, is still not negligible.

Concluding, both options entail high operational expenditures; therefore, cost minimization poses a significant challenge since it can deter a company from Big Data implementation. (U. Sivarajah, 2016)

**Data skilled personnel**

Among the challenges that every company faces when coming to Big Data analytics is the lack of skilled personnel. Nowadays there is a shortage of data scientists, meaning professionals who possess all necessary skills and competences in order to transform the gathered data into clear insights (U. Sivarajah, 2016) (A. Lambrecht, 2015). This
shortage is directly interrelated with the great number and the high level of theoretical and practical skills that are required for practicing this profession (Figure 5).

Starting with the necessary technical skills, a data scientist should have expertise in a variety of fields such as applied mathematics, statistical programming, algorithms, Big Data processing platforms (Hadoop, Spark, Flink etc), structured (SQL) and unstructured (3-5 top NoSQL DBs) data, coding, machine and deep learning, data modeling, data mining, predictive analytics and data visualization as well as traditional software and research (Mayo, 2016). Normally data scientists are highly educated – having Master’s degree or PhDs when their first degree is usually Mathematics and Statistics, Computer Science and Engineering (Burtch, 2014).

Parallel with the technical skills, a data scientist is required to have a set of non-technical expertise. Strong interpersonal skills such as communication skills and the ability to tell compelling stories (Bean, 2016) are crucial for the successful translation of data to non-technical audience, such as the company’s managers and the marketing department. Decision-making and problem solving skills and the ability to work as part of a team are also equally important. At the same time a data scientist should possess intellectual curiosity and business acumen, meaning the ability to understand the needs of the market in which the company operates, in this case the tourism industry. Only then he/she will be able to pose the right questions and get the right answers from Big Data analytics.

Given the diverse skills that a data scientist should possess, it is evident that locating such professionals is rather difficult or they may prefer working in data-rich online businesses over travel firms. (Group, 2013)

Concluding with literature review, it became evident that Big Data can revolutionize global travel and tourism industry, by introducing innovative tools, maximizing strategic benefits and creating opportunities for fast movers (WTTC, 2014). However these benefits are closely interrelated with equally important risks and companies are confronted with the challenge to find the best possible strategy for managing Big Data analytics; in-depth research, continuous information and meticulous strategic planning could become the key is for this strategic aim.
Figure 5: Skills and Competences of Data Scientists
Source: Mayo, M., 2016
Methodology

For the purposes of this research both *secondary and primary* data were used as sources of information.

**Secondary Data**

As already demonstrated in the previous chapter, scientific journal and newspaper articles, researches and reports, published books etc. were gathered and analyzed as secondary source of information. These data were enhanced with documented business case studies from tourism companies around the world, who implemented successfully Big Data analytics into their business strategy.

**Primary Data**

The desired outcome of the primary research was to provide a complete and rounded understanding on Big Data’s role in tourism industry with particular emphasis on Greek tourism companies. For that reason a strong effort was made to gather data that depict the views of key Big Data stakeholders, such as the tourism companies and organizations, Data scholars and professional Data scientists. As primary source of information, two types of data were collected and analyzed:

i. *Quantitative* data in the form of questionnaires with close-ended questions, gathered directly from representatives of tourism companies and

ii. *Qualitative* data in the form of in depth face-to-face and phone interviews

More explicitly:

I. QUANTITATIVE DATA

- *Sample of Population*: Exhibitors of the international tourism exhibition *Philoxenia*.

  Importance of research setting: *Philoxenia* is the longest running and the largest tourism exhibition in Greece. It is organized annually in Thessaloniki and its goal is the promotion and the further internationalization of Greek tourism industry.
▪ **Gathering Method**: On-site distribution of questionnaires that were anonymous and consisted of 12 series of questions

▪ **Selection Method**: The respondents were selected according to the following two (2) criteria:
  - The exhibitor should be a company or an organization; therefore participants like embassies, unions and sector associations were excluded.
  - The exhibitor should operate in tourism industry; therefore participants like advertising companies, academic institutions, publication and press companies were excluded.

▪ **Number of distributed questionnaires**: 60

▪ **Collected responses**: 39

▪ **Encountered Problems**: Many company representatives were reluctant to participate in the survey. This was mainly due to the following facts:
  - They were not yet familiar with the term “Big Data”
  - Big data wasn’t in their scope of responsibility and the person in charge (Business Manager etc.) was often engaged in B2B Meetings or didn’t come in Philoxenia at all.
  - In general, it was observed that the term “Big Data” was often intimidating — mostly for reasons such as breach of privacy - and deterring the exhibitors from participating in the survey.

▪ **Time of conducted research**: 10/11/2017 – 12/11/2017

▪ **Venue**: TIF HELEXPO

II. **QUALITATIVE DATA**

As part of this research, two (2) in depth interviews were conducted as well:
  - one face-to-face interview with a Data scholar &
  - one phone interview with a professional Data scientist

The interviews were structured and consistent of seven (7) questions that were posed to both respondents. Their names will not be mentioned for privacy reasons. Finally, it should be mentioned that one more Big Data analytics company was approached for an interview, however with no success.
Primary Research Results

This chapter focuses on the presentation of the primary research results. It starts with the analysis of quantitative data, gathered in the form of questionnaires and proceeds with the findings of the qualitative research, meaning the two in-depth interviews.

Quantitative research – Results and Analysis

Question 1: “In which of the following sectors of tourism industry is your organization operating?”

The majority of companies who took part in the survey is operating in the accommodation sector (28%), followed by tour operators (13%) and companies that provide transportation & travel/tourism services (12%) (Figure 6). It should be mentioned however that since many participants enjoy a wide network of operations, more than one sectors were often selected by the same respondent as answers to this question.

![Figure 6: Sector of Business Operations](image-url)
Question 2: “Where is the base of your organization’s operations?”

The segmentation of participants was made according to Kallikratis administrative division. As it is depicted in Figure 7, the greater number of exhibitors originate from Central Makedonia with 39% and Attiki with 29%. It is worth noting also, that 8% of the respondents originate from Europe, a rather low percentage considering the international character of Philoxenia tourism exhibition.

**BASE OF BUSINESS OPERATIONS**

![Pie chart showing the distribution of business operations across different regions.](image)

**Figure 7: Base of Business Operations**

Question 3: “What is the number of your employees?”

Aim of this question was to indicate the size of the companies that participated in the survey. As it is depicted in Figure 8, in the first position are companies small to medium-size with 6-20 employees (24%), followed by rather big companies with 101-500 employees (14%).
Question 4: “What sources of information you use, when planning your organizational strategy and to what extent?”

Referring to the sources of Data that companies are using when planning their business strategy, *internal business data and transactions* are on the top of managers’ preferences (20%), followed by *website reviews* (16%) and *social media* (15%) (Figure 9).

![Figure 8: Number of company’s employees](image)

![Figure 9: Sources of Data employed by the company](image)
Question 5: The purpose of this question was to identify the level of Big Data’s implementation into the company’s business strategy. It is interesting to observe that the two most preferred answers are at opposite poles: whilst 37% of the companies rely mostly on managerial experience, 31% are already experienced with Big Data analytics. (Figure 10).

![Big Data's Level of Use](image)

**Figure 10:** Big Data’s level of use in Company’s business Strategy

Question 6: “To what extent do you believe that big data analytics can create a competitive advantage for a tourism firm?”

The majority of respondents (51%) strongly agrees that Big Data can become a powerful tool for achieving competitive advantage against industry’s rivals. 41% of companies’ representatives “agree”, when only 8% is simply neutral. It is worth mentioning that none of the survey participants disagreed with the given fact (Figure 11).

![Big Data as Competitive Advantage](image)

**Figure 11:** Big Data as Competitive Advantage
In the following two questions (n. 7 & 8), the respondents were asked to evaluate the key benefits and challenges that are related to Big Data use.

**Question 7:** “According to your opinion, how important are the following benefits of big data analytics for a tourism firm?”

As it is easily observed in figure 12, all given factors were identified as almost equally important. The facilitation in the design of new or better products and services was selected by the respondents as the most valuable trait of Big Data (27%), when internal decision support (26%) and the improvement of customer relationships through personalization of services (25%) came second and third in the respondents’ answers.

![Benefits of Big Data](image)

**Figure 12: Benefits of Big Data**

**Question 8:** “According to your opinion how important are the following challenges of big data analytics?”

As the most important challenge in Big Data use was singled out the lack of experts (e.g. data scientists) and of technological equipment (27%). Privacy and security issues were also classified high in the list of Big Data risks (25%) (Figure 13).
The next two questions (n. 9 & 10) are related to data-skilled personnel. Question 9: “Do you consider crucial the recruitment of personnel with strong data skills in order to gain competitive advantage in tourism industry?”

Since as major challenge of Big Data was identified the lack of experts (Figure 14), it was foreseeable that the majority of respondents (62%) would answer affirmatively to the question concerning the recruitment of data skilled personnel and its value for the company.

**Importance of data skilled personnel**

![Figure 14: Importance of Data Skilled Personnel](image-url)
Question 10: “Is your organization employing a data skilled personnel?”

Consequently, the majority of the companies that participated in the survey is already employing data skilled personnel (63%) (Figure 15).

Figure 15: Employment of data skilled personnel

Question 11: “How would you rate the importance of improving the overall big data analytics performance of your organization?”

The respondents were asked to evaluate the need for their company’s further improvement on Big Data analytics. Taking into account the nature of all previous given answers, that put significant value on Big Data, it wasn’t unexpected that 46% of participants consider this improvement as “Important” and 33% as “Very Important” for their company’s strategic performance (Figure 16).

Figure 16: Improvement in Big Data analytics Performance
Question 12: “How interested would you be in learning more about the benefits and challenges of big data analytics in tourism sector?”

The final question of the research was aiming to identify participants’ interest in acquiring a more profound knowledge of Big Data phenomenon and of the necessary steps for implementing it into their business strategy. The results were very positive: 56% of respondents were interested in learning more about Big Data, 38% very interested and only 3% showed no interest at all (Figure 17).

![Interest for further information about Big Data](chart.png)

Figure 17: Interest for further information about Big Data

Concluding with the findings of the quantitative research, it is apparent that the majority of Greek tourism companies, which participated in the survey, acknowledges Big Data as an important and decisive strategic asset. Not all participants are experienced with this innovative tool and the impediments that come along, however it is very encouraging that the greater part of them has a keen interest in learning more about this digital era’s phenomenon, its benefits and challenges.
Qualitative research – Presentation of Interviews

With the aim to provide a holistic approach of Big Data phenomenon, the research conducted also two in-depth interviews of data experts, those of a data scholar and a data scientist. The questions posed were trying to identify the respondents’ views concerning the benefits and challenges of Big Data, its role in companies’ strategy and the importance of employing data skilled personnel; at the end, the respondents were asked to provide some recommendations for better Big Data implementation on behalf of Greek tourism companies.

Starting with the most important benefits, both respondents pointed out that Big Data provides clear insights and profound knowledge in customers’ needs and wants, driving innovation and facilitating the companies in the improvement of their services, refining at the same time the whole experience of the customer. Personalization was also recognized as key attribute of Big Data. In particular, Big Data can become a powerful tool for the tailoring of a company’s products and services to customers’ individual preferences and for the establishment of a mutual beneficial relationship between company and customers at any time, in any place and from any device. Finally, both respondents referred to the improvement of internal business’s operations and decision-making process, since entrepreneurs are able to be informed fast and accurate about any problem that arises and come up with real-time decisions. According to data scientist, through Big Data analytics the operational efficiency is improved and new revenue streams are created.

On the opposite site, the respondents were asked to identify the biggest challenges of Big Data for a company. Privacy and security issues have been pinpointed as two of the major challenges of Big Data. Both respondents referred to the General Data Protection of Regulation (GDPR) adopted by European Union in 2016 with the aim to harmonize data privacy law across Europe. As they both explained, every company which wants to
leverage Big Data is striving at the moment to adjust its internal Big Data policy to EU Regulation. The deadline ends at 25th May 2018 and after that date, heavy fines are going to be opposed to any company that doesn’t comply with EU Regulation. Furthermore, both respondents agreed that Big Data is requiring a considerable big investment of money; this investment concerns not only the appropriate technological equipment (that also needs maintenance) but also the employment of experienced data skilled personnel. At this point, the scholar highlighted as a challenge the need for transition from the traditional to a data-driven company culture that is mostly characterized by experimentation, agility and a data-driven mindset. He concluded by adding as an extra key factor - especially for Greek tourism companies – the access to such amount of data that could be characterized as “Big”. As he noted, Greek tourism firms are usually small to medium size and unlike companies like Booking and Trivago, their operations don’t generate the necessary amount of data that would allow them to easily and successfully implicate in Big Data analytics.

To the following question whether the respondents consider that Big Data analytics can offer a competitive advantage to a company, both answers were affirmative. As they explained, Big Data is a state of the art innovation, the ‘new-new thing’ that can significant ameliorate the strategy of a business, providing a profound knowledge about the customer’s needs and wants and becoming the basis for growth and enhancement of productivity and creativity. According to data scientist, forward thinking entrepreneurs in all sectors, including travel industry, have already incorporated Big Data analytics into their business operations and achieved very positive results. And he concluded by presenting the examples of leading tourism companies like Booking, Trivago and AirBnB, which are basing their operations exclusively on Big Data analytics.

Proceeding with the next question, the respondents were asked about the skills and competences that a data scientist should possess. Both acknowledged that data scientist is a very demanding profession, since it requires a great number of skills and competences. Applied mathematics, traditional software, statistical programming,
algorithms, databases, machine and deep learning, data modeling, data mining, predictive analytics and data visualization are only some of the necessary knowledge that a data scientist needs to possess. Secondly, he/she should have strong interpersonal skills such as communication skills, decision-making and problem solving skills and the ability to work as part of a team. Both respondents referred to Venn Diagrams and in particular that of Steven Geringer, according to whom looking for a data scientist “it would be like searching for a “Unicorn”, the mythical beast with magical powers, who is rumored to exist but never actually seen”.

According to the respondents’ opinion, there isn’t adequate number of data scientists to cover the market needs. As the scholar added, this is the reason why many entrepreneurs are choosing - or are forced - to employ professionals who don’t possess all the necessary hard and soft skills of a data scientist, with the aim to train them while they are employed in the company. Both asserted that there are plenty of job opportunities for every young professional who wants to specialize further on Big Data analytics. The data scientist highlighted however that solely the university degree doesn’t cover the full range of necessary qualifications and that data science requires a continuous personal effort and training.

Continuing with the next question, the respondents were asked to share their opinion whether they believe that entrepreneurs – in particular in Greece – are well informed about the benefits and challenges of Big Data. Accordingly, if they believe that Greek companies, especially those operating in tourism industry, are well prepared for implementing Big Data analytics into their business operations. The opinions here were divergent. From the one side, the scholar said that on world-wide level, entrepreneurs are intrigued by Big Data benefits and are putting emphasis on the deeper understanding of how analytics can boost business’s operations. He asserted that there are several examples from foreign companies who implemented successfully Big Data analytics and improved their overall performance. In Greece however entrepreneurs’ information on Big Data is very limited and consequently its use is not so widespread. According to the
scholar’s opinion, for the majority of Greek companies, especially for those operating in tourism industry, there is still a lot to be done. He believes that this is closely related to a number of factors such as the small - medium size of Greek tourism companies that doesn’t allow the generation of Big amount of Data from internal operations, the high expenditures required and the overall Greek economic crisis that doesn’t create a fruitful ground for investment and innovation. For all the above mentioned reasons he believes that – with some exceptions of course - Big Data in Greece is still in their infancy.

On the other side, the data scientist claimed that entrepreneurs, also in Greece, are nowadays well informed about Big Data. They are intrigued about the value it can bring to their companies and they show genuine interest in learning more about the ways of implementing it into their business strategy. Concerning the level of preparation, he said that this depends from the type of business. For example big hotel chains and start-ups might be much better prepared for Big Data analytics comparing to a small family-owned hotel.

Finally, the respondents were asked to suggest ways for the easiest and most effective implementation of Big Data analytics by a Greek tourism firm. From his side, the scholar suggested that it would very beneficial for the tourism entrepreneur to cooperate with a Cloud and Data platform provider, who possesses the necessary knowledge and expertise on Big Data analytics. Thus the entrepreneur can concentrate on the core business areas, leaving the experts to deal with analytics and turn data into insight for the tourism company. He concluded by noting that contrary to an investment in technological equipment and data-skilled personnel, the money required in this case is considerably less and most often “you pay as you go”.

The data scientist, on the other side, highlighted that the key issue when we talk about Big Data implementation, is to find the correct question that needs to be answered. This means that a tourism company should first locate the problem that needs to be solved and then check whether the solution can be found through Big Data analytics. He concluded by indicating that there is an interconnection between the amount of data that the company generates and the successful implementation of Big Data analytics. He
compared the telecommunication companies – which produce plenty of data - to Greek tourism companies, which are mainly small to medium sized and they don’t generate such Big amount of data.

Concluding with the presentation of qualitative research results, it is evident that both respondents consider Big Data analytics as a powerful tool for Greek tourism companies. Of course there are challenges that companies need to overcome, but the competitive advantage they will gain is worth the try. It seems there is dichotomy of opinions on the level of information and preparation concerning Big Data implementation by Greek tourism companies; one explanation could be the different scale of tourism businesses, from small and family-owned to big travel chains and corporations. Future research could enlighten more this particular point of interest.
Discussion and Conclusion

So far it became clear that Big Data can become a very powerful tool that can alter the core strategy of a business, driving growth and innovation. Both primary and secondary research highlighted as the most valuable benefits of Big Data, the personalization of services, the delivery of new and competitive products and services, the support of internal operations and decision-making process as well as the improvement of revenue management and pricing strategy. Referring to the quantitative research, it is very encouraging and optimistic that all these factors were identified as equally beneficial by the companies’ representatives (Figure 12). Although only 31% of survey participants are already experienced with Big Data analytics, with 37% of the companies to rely mostly on managerial experience (Figure 10), Big Data is still evaluated as a determinant attribute of business strategy.

From the other side, there are unavoidably some challenges deriving from Big Data use. According to secondary research, as Management key challenges have been identified privacy and security issues, data ownership, cost and operational expenditures along with the shortage of data scientists. Primary research concluded in the same results, adding two more factors. In particular, quantitative research pinpointed as another Big Data challenge the difficulty in achieving accuracy of results that is related to the Big Data characteristics (e.g. data volume, variety, velocity, veracity, volatility, quality, discovery and dogmatism); the qualitative research indicated also the difficulty in gaining access to data produced from sources other than company’s operational activities. This challenge is based mostly on the technical aspects of capturing data.

Both primary and secondary research concluded that regardless of the challenges and risks, Big Data can provide a significant competitive advantage to a travel company and the business leaders who want to invest in technological progress and innovation, will be rewarded by gaining a significant piece of the niche market. This view is reflected also in the answers given by the representatives of Greek tourism companies, since 46% of them consider as “Important” and 33% as “Very Important” the need for their company’s further improvement on Big Data analytics. Additionally 56% of them are “Interested” and
36% “Very interested” in acquiring a more profound knowledge of Big Data phenomenon and of the necessary steps for Big Data implementation into their business strategy.

So the question that arises is “what are those steps and course of action that tourism companies could follow for the successful Big Data implementation?”

Based on the qualitative research and the opinions expressed by the data scientist and the data scholar, there are two basic steps that tourism firms need to take.

Starting with the data scientist’s recommendation, it is imperative for the entrepreneurs as first step “to find the correct question that needs to be answered”. This means that a tourism firm - before even implements Big Data analytics - should firstly locate the problem that needs to be solved and then examine whether the solution can be found by the use of Big Data tools. As the both interviewees noted, there is a strong interconnection between the amount of data that the company generates and the successful implementation of Big Data analytics. Greek tourism companies are mainly small to medium sized and they don’t always generate such Big amount of data that could be used for further analysis. An experienced professional could assist the tourism companies in identifying their key objectives and the role of Big Data in the accomplishment of their strategic goals.

The second step, based on the data scholar recommendation, would be the cooperation of the tourism firm with a Cloud and Data platform provider, meaning a company which possess the necessary expertise and know-how in Big Data analytics. This recommendation can be proved very beneficial and useful when a tourism company doesn’t have previous experience on data analysis. Hence, the owner or manager of the tourism firm can focus on the core business operations, assigning to professional data scientists the task to transform data into valuable insights. An additional benefit in this case is also the reduced expenditure. According to the data scholar “contrary to an investment in technological equipment and data-skilled personnel, the money required in this case is considerably less and most often you pay as you go”. That means, the tourism firm doesn’t not pay a fix price but the cost is flexible and vary according to the services provided.
Another determinant factor, as indicated by both primary and secondary research, is the assembling of Big Data skills. The significance of this strategic step becomes obvious in the responses of companies’ representatives, since 62% of them consider the recruitment of data skilled personnel of fundamental importance for implementing Big Data analytics (Figure 14). The data experts who participated in the interviews are holding the same opinion, stressing out however the shortage of experienced data analysts, that is related to the demanding nature of this profession. At this point, it should be accentuated that this strategic step is mostly recommended for tourism companies that possess a level of experience in data management and are aiming in creating an internal Data analysis Department (Group, 2013).

Finally one more recommendation – based on researchers and academics - that could be proved especially useful for large tourism companies such as hotel chains, airline companies etc which already implement Big Data analytics in their internal operations, would be the establishment of Data Governance. Considering the heterogeneous and unstructured nature of Big Data that often makes its further categorization and processing very demanding, one can conclude that there is a need for a more holistic approach. For that reason, IT managers highlight nowadays the need for Data Governance, meaning a formal Big Data strategy followed by the companies, which describes what data is warehoused and analyzed, ensuring thus the quality of data mined and assessed (U. Sivarajah, 2016). Especially in large tourism companies or companies that deal with complex data analysis, there is an imperative need for managers and data scientists to “decide how to decide” (Thomas, n.d.).

According to Data Governance Institute (Governance, n.d.), Data Governance is “a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods.” In other words, Data Governance includes all the necessary processes and decision making for classifying, organizing and managing data (Wise, 2016). Its principal aim is the facilitation of decision-making inside the tourism company, the minimization of business’s operational obstacles, the creation of transparent and standard processes, the implementation of a common
approach to data management, the cost reduction and the increase of operational effectiveness (Data Governance Institute, n.d.).

According to DAMA - the Data Management Association International and its Data Management Body of Knowledge (DAMA DMBOK) Data Governance is comprised by nine different disciplines such as Data Quality Management, Data Security Management and Data Development as depicted in Figure 18 (The Global Data Management Community, n.d.). All together are assuring that Big Data analytics are implemented with the most effective and valuable manner.

Figure 18: Data Governance Disciplines

Source: The Global Data Management Community-Body of Knowledge
Taken into account the new General Data Protection Regulation (GDPR) (EU) 2016/679 and the need for every company to comply to its provisions, the importance of Data Governance is nowadays more crucial than ever. Data Governance principles as integrity between its participants, transparency, auditability and accountability of decisions and operations, check-and-balances between managers and data scientists and standardization of processes (Data Governance Institute, n.d.) can assist the companies efficiently in adapting their internal rules and operations to European data protection law.

At this point however, one can wonder: “Is Big Data by itself enough for a company to gain competitive advantage in the market?” or “Is Big Data the panacea to every problem a company faces?”

The answer to these questions could be found in the Framework of Knowledge Management. Knowledge Management (KM) is a concept that was introduced around three decades ago, in 1990 (Koenig, 2012). One of the most popular definitions of KM was given in 1998 by Duhon (Gartner Group) who stated that "Knowledge management is a discipline that promotes an integrated approach to identifying, capturing, evaluating, retrieving, and sharing all of an enterprise's information assets. These assets may include databases, documents, policies, procedures, and previously un-captured expertise and experience in individual workers." Another equally popular definition is given by Davenport, according to whom “Knowledge management is the process of capturing, distributing, and effectively using knowledge.” Principal aim of Knowledge management is to improve company’s performance (Oleg Tilchin, 2013).

Organizational theorist Prof. Russel Ackoff developed the DIKW Hierarchy (Data, Information, Knowledge and Wisdom Hierarchy), as it is depicted in figure 19. In the base of the pyramid is “Data”, meaning the raw objective facts about an event. Big Data belong also in this category. On the next level of the pyramid is “Information”, meaning the data that it has been processed and interpreted. At this point analytics have transformed Big Data into insights. “Knowledge” follows on the third level of DIKW Hierarchy and according to the KM literature, it is divided two categories:
a. The explicit knowledge, meaning the type knowledge that is or can be set out in tangible form (Koenig, 2012), is formalized and codified and therefore it is easy to be identified and stored; Sources of explicit knowledge are databases, documents etc (Frost, 2017) (Oleg Tilchin, 2013).

b. The tacit knowledge, meaning the intuitive knowledge that is mostly based on people’s experience, values, beliefs, ideas and is transferred through socialization and mentoring (Frost, 2017) (Oleg Tilchin, 2013). According to academics and researchers tacit knowledge is considered as the most valuable form of knowledge and it is a prerequisite for innovation and progress.

Finally, on the top of the Pyramid is “Wisdom”, that it is achieved through explicit and tacit knowledge with insight.

Figure 19: DIKW Hierarchy - The Knowledge Pyramid

Source: CHARLES STURT UNIVERSITY
https://www.csu.edu.au/division/dit/eal/portfolios/information
As it is understood - based to the Framework of Knowledge Management - Big Data by itself isn’t enough for a company to gain competitive advantage and differentiate from its business rivals. It needs also tacit knowledge and empathy (Fontana, 2010). Both these attributes are especially important in tourism industry, which is customer-centered and where customer care as well as relationship building are essential parts of the services provided. Primary concern of tourism companies should be the experiences they provide because travelers might forget what companies offered them through technology but they will always remember how they made them feel. This is where the key to real customer loyalty lies (Megatrends Defining Travel in 2017, 2017).

Hence, the basis to competitive advantage can be found in the effective pairing of technology with human intelligence (Megatrends Defining Travel in 2017, 2017). Big Data, like every other technological innovation, is a tool that it can bring significant benefits to its users, e.g. a tourism company. However, it could be proved much more powerful and effective when it is used by a well-trained, inspired and empathic leader, who can understand the needs and wants of travelers and who strives for building an true emotional connection between the tourism company and its customers, as flesh and blood individuals (Megatrends Defining Travel in 2017, 2017) (WTTC, 2014).

Concluding with this research, it would be an omission not to refer briefly the key Big Data Trends. 2018 is expected to be the Year of Transition (Dataflod, n.d.) from the Third to the Fourth Industrial Revolution (Schwab, 2016). The First Industrial Revolution was based on water and steam power, the Second on electric power and the Third on electronics and information technology. Nowadays we are on the edge of the Fourth Industrial Revolution, the Digital Revolution, which is characterized by an amalgam of technological breakthroughs. This Revolution is evolving at an exponential pace and is disrupting all sectors of industry, tourism among them. The growth of Internet of Things and the plethora of electronic devices with unprecedented processing power and storage capacity are opening the doors to knowledge and information sharing, becoming the source of even “Bigger” Data; Artificial Intelligence (e.g. self-driving cars, virtual assistants) will continue to mature and is expected to become even smarter, when machine learning (SabreLabs, 2017), deep learning as well as Cloud platforms and prescriptive analytics will
gain on importance in the future, providing deep insights and driving economic growth. Finally, privacy and security issues are expected to be on the epicenter of interest for businesses, organizations and governments.

In this context of continuous technological evolution, businesses are by all means strongly influenced. According to William Schmarzo, CTO, Dell EMC Services Big Data, known as the "Dean of Big Data" among the key 2018 Trends will be “the continued transformation of Big Data from an IT ‘task’ to a Business mandate” (Mayo, 2017). With the unceasing introduction of technological innovations from the one side and the growing need for personalization of products and services from the other, both supply and demand side are transformed, goading companies in all industry sectors to become agile and transcend their traditional structures and operations to a more innovative market approach.

Tourism companies are part of this disrupting evolution. To the hands of an inspired and creative business leader, digital capabilities and data-based services can become a powerful tool for increasing the value of services provided, allowing at the same time the companies to gain a deeper understanding of the constantly changing environment, drive knowledge and competition.

**Limitation and Further Research**

As is the case with any research, readers should examine the presented results within a context of certain limitations. Due to the difficulties already explained in methodology chapter, the sample of primary research, both quantitative and qualitative was rather small, when the word limitation posed an important obstacle to the further elaboration of Big Data benefits and challenges. Hence, a future and more detailed research based in a larger sample and context could shed more light on the use of Big Data by tourism companies, especially those operating in Greece. To that end and taking into account the fast-paced technological developments as well as the aforementioned Big Data trends, it
would be interesting to investigate to *what degree* Greek tourism companies adapted to the changing environment and incorporated Big Data tools into their strategy, in a two-three years’ time framework from now. Future research could also examine *how GDPR*, as it is been adopted by European Union, *affected tourism companies* – both on national and European level - concerning Big Data implementation. In any case, Big Data is a technological innovation that affects the way companies do business and interact with customers, creating thus an imperative need for further investigation.
Bibliography

- IPSOS., 2015. How British Airways is utilizing Big Data and Analytics to enhance customer experience. Date 23. 05.2015 Kenya, CIO EAST AFRICA
- Baggio, R., 2016. *Big Data, Business Intelligence and Tourism: a brief analysis of the literature*, Bocconi University, Milan, Italy: s.n.
- British, A., n.d. *Customer Analytics*. [Website] Available at:


Group, A. I., 2013. *At the Big Data Crossroads: turning towards a smarter travel experience*, s.l.: s.n.


[Access date 03 01 2018].


 MOROZ, A., 2017. HOW AIRLINES CAN BENEFIT FROM BIG DATA. [Website]
Available at: https://sigma.software/about/media/how-airlines-can-benefit-big-data
[Access date 03 01 2018].


 n, I. G., 2016. Top 10 Strategic Technology Trends 2017. [Website]
Available at: https://www.gartner.com/smarterwithgartner/gartners-top-10-technology-trends-2017/
[Access date 05 12 2017].

 Panetta, K., 2016. Gartner’s Top 10 Strategic Technology Trends for 2017. [Website]
Available at: https://www.gartner.com/smarterwithgartner/gartners-top-10-technology-trends-2017/
[Access date 15 12 2017].


 Reddy, T., 2016. 5 ways to turn data into insights and revenue in 2016. [Website]
Available at: https://www.ibm.com/blogs/watson/2016/02/5-ways-to-turn-data-into-insights-and-revenue-in-2016/
[Access date 05 01 2018].

 Sabre Labs, 2017. Emerging Technology in Travel, s.l.: Available at: https://www.sabre.com/labs/emergingtech/


Available at: https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/
[Access date 07 01 2018].

Available at: https://blog.safetyculture.com/industry-trends/three-hotel-chains-using-big-data-exceed-revenue-projections
[Access date 05 12 2017].

Siodmok A., 2017. *From best practice to next practice*
Available at: https://openpolicy.blog.gov.uk/2017/07/24/from-best-practice-to-next-practice/
[Access date 01 12 2017].


Stem, n.d. *Profile - InterContinental Hotels*. [Website]
Available at: http://www.stem.com/profile-intercontinental-sf/
[Access date 01 12 2017].

Available at: https://www.kdnuggets.com/2016/10/battle-data-science-venn-diagrams.html
[Access date 7 12 2017].

Available at: https://dama.org/content/body-knowledge
[Access date 05 01 2018].


UNECE, 2013. *Classification of Types of Big Data*. [Website]
Available at: https://statswiki.unece.org/display/bigdata/Classification+of+Types+of+Big+Data

Available at: https://www.cio.com/article/3116812/business-intelligence/understanding-data-governance.html
[Access date 01 12 2017].


Appendix

Questionnaire

Descriptive statistics

Relative Frequency

Interviews
Big Data in Tourism

Big Data as a term describes the large volume of different types of data included in multiple sources such as social media, web transactions, business processes, sensors, digital pictures, GPS signals etc. The importance of Big Data relies mostly upon its further analysis. This survey examines the role of Big Data analytics in tourism industry, their benefits and challenges for the organizations, along with the skills that are necessary for exploiting Big Data.

1. In which of the following sectors is your organization operating? Tick all that apply.

   [ ] Accommodation
   [ ] Travel/Tourism Services
   [ ] Transportation
   [ ] Research/Consulting
   [ ] Tour operators
   [ ] Online Travel Agency (OTA)
   [ ] Tourism Organization/Tourism Authorities
   [ ] Information Systems
   [ ] Education
   [ ] Other:

2. Where is the base of your organization’s operations?

3. What is the number of your employees? Mark only one oval.

   [ ] 1-5
   [ ] 6-20
   [ ] 21-50
   [ ] 51-100
   [ ] 101-500
   [ ] >500
   [ ] Other: __________________________
4. What sources of information you use, when planning your organizational strategy and to what extent?
(On a scale 1-3, 1=less 3=more) Mark only one oval per row.

<table>
<thead>
<tr>
<th>Source of Information</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Media</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business transactions and internal data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Website reviews</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emails</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open data/Public Sector</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weblogs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Videos/Pictures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other: ________________________________</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Would you consider that your organization
Mark only one oval.

- Relies mainly on managerial experience
- Is aiming at exploring big data analytics in the near future
- Has started recently applying big data analytics
- Is experienced with big data analytics and is following a big data strategy

6. To what extent do you believe that big data analytics can create a competitive advantage for a tourism firm? Mark only one oval.

- Strongly agree
- Agree
- Neither agree or disagree
- Disagree
- Strongly disagree

7. According to your opinion, how important are the following benefits of big data analytics for a tourism firm? (on a scale 1-5, 1=less 5=more) Mark only one oval per row.

<table>
<thead>
<tr>
<th>Benefit</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Facilitate the design of new / better product and services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) Better decision support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) Boost the financial performance of the organization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) Build better customer relationships</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
8. According to your opinion how important are the following challenges of big data analytics? (on a scale 1-5, 1=less 5=more) Mark only one oval per row.

<table>
<thead>
<tr>
<th>Challenge</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy/security issues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of experts &amp; technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty in achieving accuracy of results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operational expenses/Overall cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

9. Do you consider crucial the recruitment of personnel with strong data skills in order to gain competitive advantage in tourism industry? Mark only one oval.

- Yes
- No
- Maybe

10. Is your organization employing a data skilled personnel? Mark only one oval.

- Yes
- No
- Other: ______________________

11. How would you rate the importance of improving the overall big data analytics performance of your organization? Mark only one oval.

<table>
<thead>
<tr>
<th>Importance</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

12. How interested would you be in learning more about the benefits and challenges of big data analytics in tourism sector? Mark only one oval.

- Very interested
- Interested
- Neutral
- Not interested
APPENDIX B

DESCRIPTIVE STATISTICS
<table>
<thead>
<tr>
<th>In which of the following sectors of tourism industry is your organization operating?</th>
<th>Where is the base of your organization’s operations?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accommodation</td>
<td>Travel/Tourism Services</td>
</tr>
<tr>
<td>Mean</td>
<td>0.23</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.02</td>
</tr>
<tr>
<td>Median</td>
<td>0.23</td>
</tr>
<tr>
<td>Mode</td>
<td>0.23</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.01</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>0.00</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-2.11</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.00</td>
</tr>
<tr>
<td>Range</td>
<td>1.00</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.00</td>
</tr>
<tr>
<td>Maximum</td>
<td>1.00</td>
</tr>
<tr>
<td>Sum</td>
<td>19.00</td>
</tr>
<tr>
<td>What sources of information you use, when planning your organizational strategy and to what extent? (on a scale 1-3, 1=very low, 3=very high)</td>
<td>Does your organization follow a Big Data strategy?</td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>What is the number of your employees?</td>
<td></td>
</tr>
<tr>
<td>Social Media</td>
<td>Business transactions and internal data</td>
</tr>
<tr>
<td>Employees</td>
<td>1-500</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.01</td>
</tr>
<tr>
<td>Median</td>
<td>0.01</td>
</tr>
<tr>
<td>Mode</td>
<td>0.01</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.01</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>0.01</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.01</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.01</td>
</tr>
<tr>
<td>Range</td>
<td>0.01</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.01</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.01</td>
</tr>
<tr>
<td>Sum</td>
<td>0.01</td>
</tr>
<tr>
<td>Count</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Facilitize the design of new/better products and services</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------------------------------------------</td>
</tr>
<tr>
<td>Standard Error</td>
<td>4.41</td>
</tr>
<tr>
<td>sunshine</td>
<td>4.00</td>
</tr>
<tr>
<td>Holiday</td>
<td>4.50</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.50</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>0.50</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.43</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.00</td>
</tr>
<tr>
<td>Mean</td>
<td>572.00</td>
</tr>
<tr>
<td>Sum</td>
<td>58.00</td>
</tr>
</tbody>
</table>
APPENDIX C

RELATIVE FREQUENCY
Appendix C: Relative Frequency

1) In which of the following sectors of tourism industry is your organization operating?

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accommodation</td>
<td>19</td>
<td>28%</td>
</tr>
<tr>
<td>Travel/Tourism Services</td>
<td>8</td>
<td>12%</td>
</tr>
<tr>
<td>Transportation</td>
<td>8</td>
<td>12%</td>
</tr>
<tr>
<td>Research/Consulting</td>
<td>7</td>
<td>10%</td>
</tr>
<tr>
<td>Tour Operators</td>
<td>9</td>
<td>13%</td>
</tr>
<tr>
<td>Online Travel Agency (OTA)</td>
<td>4</td>
<td>6%</td>
</tr>
<tr>
<td>Tourism Organization/Tourism Authorities</td>
<td>6</td>
<td>9%</td>
</tr>
<tr>
<td>Information Systems</td>
<td>3</td>
<td>4%</td>
</tr>
<tr>
<td>Education</td>
<td>3</td>
<td>4%</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>3%</td>
</tr>
<tr>
<td>SUM</td>
<td>69</td>
<td>100%</td>
</tr>
</tbody>
</table>

2) Where is the base of your organization’s operations?

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attiki</td>
<td>11</td>
<td>48%</td>
</tr>
<tr>
<td>Rodos</td>
<td>3</td>
<td>13%</td>
</tr>
<tr>
<td>Drama</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>Ilia</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>Achaia</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>Thessalia</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>Ionian Islands</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>Epirus</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>Europe</td>
<td>3</td>
<td>13%</td>
</tr>
<tr>
<td>SUM</td>
<td>23</td>
<td>100%</td>
</tr>
</tbody>
</table>
Appendix C: Relative Frequency

3) **What is the number of your employees?**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>7</td>
<td>18%</td>
</tr>
<tr>
<td>6-20</td>
<td>9</td>
<td>24%</td>
</tr>
<tr>
<td>21-50</td>
<td>6</td>
<td>16%</td>
</tr>
<tr>
<td>51-100</td>
<td>6</td>
<td>16%</td>
</tr>
<tr>
<td>101-500</td>
<td>7</td>
<td>18%</td>
</tr>
<tr>
<td>&gt;500</td>
<td>3</td>
<td>8%</td>
</tr>
<tr>
<td>SUM</td>
<td>38</td>
<td>100%</td>
</tr>
</tbody>
</table>

4) **What sources of information you use, when planning your organizational strategy and to what extent?**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Media</td>
<td>70</td>
<td>15%</td>
</tr>
<tr>
<td>Business transactions and internal data</td>
<td>91</td>
<td>20%</td>
</tr>
<tr>
<td>Website reviews</td>
<td>74</td>
<td>16%</td>
</tr>
<tr>
<td>Emails</td>
<td>56</td>
<td>12%</td>
</tr>
<tr>
<td>Sensors</td>
<td>26</td>
<td>6%</td>
</tr>
<tr>
<td>Open data/Public Sector</td>
<td>59</td>
<td>13%</td>
</tr>
<tr>
<td>Weblogs</td>
<td>34</td>
<td>7%</td>
</tr>
<tr>
<td>Videos/Pictures</td>
<td>37</td>
<td>8%</td>
</tr>
<tr>
<td>Other</td>
<td>11</td>
<td>2%</td>
</tr>
<tr>
<td>SUM</td>
<td>458</td>
<td>100%</td>
</tr>
</tbody>
</table>
5) **Would you consider that your organization:**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relies on Managerial Experience</td>
<td>17</td>
<td>37%</td>
</tr>
<tr>
<td>Aims at using BD</td>
<td>8</td>
<td>17%</td>
</tr>
<tr>
<td>Is starting to use BD</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td>Is Experienced with BD</td>
<td>14</td>
<td>30%</td>
</tr>
<tr>
<td>SUM</td>
<td>46</td>
<td>100%</td>
</tr>
</tbody>
</table>

6) **To what extent do you believe that big data analytics can create a competitive advantage for a tourism firm?**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly agree</td>
<td>20</td>
<td>51%</td>
</tr>
<tr>
<td>Agree</td>
<td>16</td>
<td>41%</td>
</tr>
<tr>
<td>Neither agree or disagree</td>
<td>3</td>
<td>8%</td>
</tr>
<tr>
<td>SUM</td>
<td>39</td>
<td>100%</td>
</tr>
</tbody>
</table>
Appendix C: Relative Frequency

7) According to your opinion, how important are the following benefits of big data analytics for a tourism firm?

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facilitate the design of new / better product and services</td>
<td>172</td>
<td>27%</td>
</tr>
<tr>
<td>Achieve better decision support</td>
<td>166</td>
<td>26%</td>
</tr>
<tr>
<td>Boost the financial performance of the organization</td>
<td>146</td>
<td>23%</td>
</tr>
<tr>
<td>Build better customer relationships</td>
<td>161</td>
<td>25%</td>
</tr>
<tr>
<td>SUM</td>
<td>645</td>
<td>100%</td>
</tr>
</tbody>
</table>

8) According to your opinion how important are the following challenges of big data analytics?

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy/security issues</td>
<td>121</td>
<td>25%</td>
</tr>
<tr>
<td>Lack of experts &amp; technology</td>
<td>128</td>
<td>27%</td>
</tr>
<tr>
<td>Difficulty in achieving accuracy of results</td>
<td>113</td>
<td>24%</td>
</tr>
<tr>
<td>Operational expenses/Overall cost</td>
<td>114</td>
<td>24%</td>
</tr>
<tr>
<td>SUM</td>
<td>476</td>
<td>100%</td>
</tr>
</tbody>
</table>
9) *Do you consider crucial the recruitment of personnel with strong data skills in order to gain competitive advantage in tourism industry?*

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>24</td>
<td>62%</td>
</tr>
<tr>
<td>No</td>
<td>2</td>
<td>5%</td>
</tr>
<tr>
<td>Maybe</td>
<td>13</td>
<td>33%</td>
</tr>
<tr>
<td>SUM</td>
<td>39</td>
<td>100%</td>
</tr>
</tbody>
</table>

10) *Is your organization employing a data skilled personnel?*

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>24</td>
<td>63%</td>
</tr>
<tr>
<td>No</td>
<td>14</td>
<td>37%</td>
</tr>
<tr>
<td>SUM</td>
<td>38</td>
<td>100%</td>
</tr>
</tbody>
</table>

11) *How would you rate the importance of improving the overall big data analytics performance of your organization?*

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very important</td>
<td>13</td>
<td>33%</td>
</tr>
<tr>
<td>Important</td>
<td>18</td>
<td>46%</td>
</tr>
<tr>
<td>Quite important</td>
<td>8</td>
<td>21%</td>
</tr>
<tr>
<td>SUM</td>
<td>39</td>
<td>100%</td>
</tr>
</tbody>
</table>
12) How interested would you be in learning more about the benefits and challenges of big data analytics in tourism sector?

<table>
<thead>
<tr>
<th>Factors</th>
<th>Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very interested</td>
<td>15</td>
<td>38%</td>
</tr>
<tr>
<td>Interested</td>
<td>22</td>
<td>56%</td>
</tr>
<tr>
<td>Not interested</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>Neutral</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>SUM</td>
<td>39</td>
<td>100%</td>
</tr>
</tbody>
</table>
Interview N.1

Respondent: Data scholar

1. *Interviewer:* According to your opinion, what are the most important benefits of Big Data?

   Respondent: There are numerous benefits that derive from Big Data. First of all it provides clear *insights* and profound knowledge in customers’ needs and wants, facilitating thus the companies in the creation of new or better products and in the improvement of their services, refining at the same time the whole experience of the customer. Another key attribute of Big Data is *personalization*. In particular, Big Data can become a powerful tool for the tailoring of a company’s products and services to customers’ individual preferences and for the establishment of a mutual beneficial relationship between company and customers at any time, in any place and from any device. Finally, *internal business’s operations* can also be improved thanks to Big Data analytics since entrepreneurs are able to be informed fast and accurate about any problem that arises and come up with real-time decisions.

2. *Interviewer:* On the opposite site, what would you consider as the biggest challenges of Big Data for a company?

   Respondent: To start with, *privacy and security issues* have always been two of the major challenges of Big Data. This is also reflected in European Union policy. In particular, in 2016 the EU adopted the General Data Protection Regulation (GDPR) with the aim to harmonize data privacy law across Europe. This practically means that every company which wants to leverage Big Data is obliged to comply with GDPR requirements; the deadline ends at 25th May 2018 and after that date heavy fines are going to be opposed to any company that doesn’t comply. Furthermore, I would say that Big Data is requiring a considerable *big investment* of money. This investment concerns not only technological equipment (that also needs maintenance) but also the employment of experienced data skilled personnel.
At the same time, I consider of major importance the need for transition from the traditional to a data-driven company culture that is mostly characterized by experimentation, agility and a data-driven mindset. Concluding, I would like to add that a key challenge - especially for Greek tourism companies – is gaining access to such amount of data that could be characterized as “Big”. Greek tourism firms are usually small to medium size and unlike companies like Booking and Trivago, their operations don’t generate the necessary amount of data that would allow them to easily implicate in Big Data analytics.

3. **Interviewer:** Do you consider that Big Data analytics can offer a competitive advantage to a company?

   **Respondent:** Yes, of course. Big Data is a state of the art innovation that can significant ameliorate the strategy of a business, providing a profound knowledge about the customer’s needs and wants. It allows the entrepreneurs to keep track of the constant changes in the market, which is a goal not easily reachable. In the hands of an inspired entrepreneur Big Data can become the tool for achieving innovation and gaining competitive advantage against company’s business rivals.

4. **Interviewer:** What skills and competences should possess a data scientist?

   **Respondent:** Data scientist is a very demanding profession, since it requires a great number of skills and competences. Applied mathematics, traditional software, statistical programming, algorithms, databases, machine and deep learning, data modeling, data mining and data visualization are only some of the necessary competences that a data scientist needs to possess. At the same time equally important are the strong communication skills, the ability to understand market needs and to think outside of the box, the story-telling and the ability to work as part of a team. At this point, I would like to refer to Venn Diagrams and in particular that of Steven Geringer, according to whom a data scientist is like a “Unicorn”, the legendary beast that has magical powers and it is rumored to exist but is never actually seen.
5. Interviewer: Do you believe that there is adequate number of data scientists to cover the market needs?
Respondent: Unfortunately no. For that reason, many entrepreneurs are choosing - or are forced - to employ professionals who don’t possess all the necessary hard and soft skills of a data scientist, with the aim to train them while they are employed in the company. On the other side, these are good news for young professionals, who are interested in data analytics. They can capitalize on this market gap and specialize further in data science, having thus stronger chances for finding an interesting and well-paid work position.

6. Interviewer: Do you believe that entrepreneurs – in particular in Greece – are well informed about the benefits and challenges of Big Data? Accordingly, are the companies, especially those operating in tourism industry, well prepared for implementing Big Data analytics into their business operations?
Respondent: On world-wide level, entrepreneurs are intrigued by Big Data benefits and are putting emphasis on the deeper understanding of how analytics can boost business’s operations. Therefore we can easily find several examples from foreign companies who implemented successfully Big Data analytics and improved their overall performance.
In Greece however entrepreneurs’ information on Big Data is very limited and consequently its use is not so widespread. For the majority of Greek companies, especially for those operating in tourism industry, I would say that there is still a lot to be done. This is closely related to the small - medium size of Greek tourism companies that doesn’t allow the generation of Big amount of Data from internal operations, the high expenditures required and the overall Greek economic crisis that doesn’t create a fruitful ground for investment and innovation. For all the above mentioned reasons it wouldn’t be an exaggeration to say that – with some exceptions of course - Big Data in Greece is still in their infancy.

7. Interviewer: Do you have any suggestions for the easiest and most effective implementation of Big Data analytics by a Greek tourism firm?
Respondent: To a tourism entrepreneur I would suggest to address to a Cloud and Data platform provider, who possesses the necessary knowledge and expertise on Big Data. Thus he can concentrate on the core business areas, leaving the experts to deal with analytics and turn data into insight for the tourism company. It is worth noting also that contrary to an investment in technological equipment and data-skilled personnel, the money required in this case is considerably less and most often the system “you pay as you go” is applied.
1. **Interviewer:** According to your opinion, what are the most important benefits of Big Data?

Respondent: Big Data yields actionable *insights* to market needs, becoming thus an important tool for driving innovation and business growth. Especially for tourism firms Big Data facilitates the delivery of a more *tailor-made experience* and promotes *personalization*, meaning the customization of products and services according to traveler’s needs and wants. Furthermore Big Data have the potential to support the *internal decision-making process* of a company and to guide the business manager towards *data-driven decisions*; through Big Data analytics the operational efficiency is improved and *new revenue streams* are created.

2. **Interviewer:** On the opposite site, what would you consider as the biggest challenges of Big Data for a tourism company?

Respondent: Big Data analytics is not an easy task. Of course there are many Big Data tools in the market, like Hadoop - whose use is often free of charge - however those tools require experienced data scientists, who can transform the data into valuable insights. Hence, one of the key challenges that a company faces is the *investment* in the appropriate technological equipment and the employment of a well-trained data-skilled personnel.

Moreover, *security and privacy issues* are a very challenging parameter of Big Data. This became particularly important with the adoption of the General Data Protection Regulation (GDPR) by European Union. Starting from May 2018, this Regulation aims to harmonize the data protection law in all Member States of EU. All companies are striving at the moment to adjust their internal Big Data policy to EU Regulation, otherwise they will be charged with heavy fines.
3. **Interviewer:** Do you consider that Big Data analytics can offer a competitive advantage to a company?

Respondent: Yes, of course. Big Data is the *new-new thing*. It can become the basis for growth and enhancement of productivity and creativity. Forward thinking entrepreneurs in all sectors, including travel industry have already incorporated Big Data analytics into their business operations and achieved very positive results. Representative examples are leading tourism companies like Booking, Trivago and AirBnB, which are basing their operations exclusively on Big Data analytics.

4. **Interviewer:** What skills and competences should possess a data scientist?

Respondent: Data science is a very demanding profession. In the first place, a data scientist should have expertise in applied mathematics and statistics, machine learning, computer science, predictive analytics, visualization and traditional research. Secondly, he/she should possess strong interpersonal skills such as communication skills, decision-making and problem solving skills and the ability to work as part of a team. As you understand however, a professional data scientist who possesses all these skills is really hard to be found. According to the Venn Diagram of Steven Geringer, “it would be like searching for a “Unicorn”, the mythical beast with magical powers, who is never actually seen”.

5. **Interviewer:** Do you believe that there is adequate number of data scientists to cover the market needs?

Respondent: No, unfortunately there is a big gab in the market, especially in Greece. Although the demand is high, is really difficult to find an experienced data scientist. This also means that there are plenty job of opportunities for every young professional who wants to specialize further on Big Data analytics. However, I would like to highlight that data science requires a *continuous personal effort and training*; solely the university degree doesn’t cover the full range of necessary qualifications.
6. **Interviewer: Do you believe that entrepreneurs – in particular in Greece – are well informed about the benefits and challenges of Big Data? Accordingly, are the companies, especially those operating in tourism industry, well prepared for implementing Big Data analytics into their business operations?**

   **Respondent:** Based on my experience, I would say that entrepreneurs, also in Greece, are nowadays well informed about Big Data. They are intrigued about the value it can bring to their companies and they show genuine interest in learning more about the ways of implementing it into their business strategy. Concerning the level of preparation, I would say that this depends from the type of business. For example big hotel chains and start-ups might be much better prepared for Big Data analytics comparing to a small family-owned hotel.

7. **Interviewer: Do you have any suggestions for the easiest and most effective implementation of Big Data analytics by a tourism firm?**

   I would say that the most important thing, when we talk about Big Data implementation, is to find the correct question that needs to be answered. This means that a tourism company should first locate the problem that needs to be solved and then check whether the solution can be found through Big Data analytics. This is closely linked with the amount of data that the company generates. Unlike companies in other sectors, e.g. telecommunications, Greek tourism companies are mainly small to medium sized and as far as I am aware, they don’t generate such Big amount of data. This can pose an obstacle in the easy and successful implementation of Big Data analytics.