



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ  
UNIVERSITY OF CRETE

# **Data - driven Innovation: Location and Big Data Analytics for Knowledge Extraction in Tourism Destinations**

A DISSERTATION PRESENTED  
BY  
KONSTANTINOS VASSAKIS  
TO  
DEPARTMENT OF ECONOMICS  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY  
IN THE SUBJECT OF  
ECONOMICS

---

University of Crete  
Rethymno, Crete  
December 2019

© 2019 - KONSTANTINOS VASSAKIS  
ALL RIGHTS RESERVED

Thesis Advisor: Petrakis Emmanuel

Professor, Department of Economics, University of Crete, [petrakis@uoc.gr](mailto:petrakis@uoc.gr)

Advisory Committee Member: Kopanakis Ioannis

Professor, Department of Management Science & Technology, Hellenic Mediterranean University, [kopanakis@hmu.gr](mailto:kopanakis@hmu.gr)

Advisory Committee Member: Georgios Tsiotas

Assistant Professor, Department of Economics, University of Crete, [tsiotas@uoc.gr](mailto:tsiotas@uoc.gr)

# **Data-driven Innovation: Location and Big Data Analytics for Knowledge Extraction in Tourism Destinations**

## **ABSTRACT**

This Ph.D. thesis is written and submitted to the Department of Economics of the University of Crete, Greece, as a partial fulfillment of my obligations as a PhD Candidate. It consists of three separate chapters that study the applications of location and big data analytics for knowledge extraction in tourism destinations.

The first chapter of the thesis entitled, "Big data Analytics: Applications, prospects and challenges", studies the role and impact of big data analytics in innovation and value creation. The tremendous increase of data through the Internet of Things (connected devices) has led to a "big-data era", where big data analytics are applicable in every sector and economy globally. The growing expansion of available data is commonly accepted, while valuable knowledge arising from the information comes from big data analysis processes. The prospects of big data analytics are important and the benefits for data-driven organizations are significant determinants for competitiveness and innovation performance. However, there are considerable obstacles to adopting a data-driven approach and get valuable knowledge through big data.

The second chapter of the thesis entitled, "Knowledge Extraction through Location & Big Data Analytics: the case of Crete", explores the knowledge extraction using location and big data analytics techniques. Nowadays, tourists generate massive volumes of data (big data) during their visit to an urban destination. However, there is little knowledge of their spatial activity and perceptions. Enterprises and organizations in hospitality and tourism are able to exploit actual behavioural data - perceptions derived from big data generated in real-time from online data sources in contrast to traditional customer surveys based on questionnaires. An innovative approach is demonstrated using the case study of Crete by integrating big data techniques, location intelligence and social media transforming tourist experiences into valuable assets (new knowledge extraction) for quicker and more efficient decision making. More specifically, the approach introduces the combination of textual and photo analytics with data derived from media sharing and textual social networks, introduces social big data analytics such as social engagement, sentiment analysis, topic/label detection combined with spatio-temporal features to provide more insights about tourist destinations. Research findings demonstrate how this novel approach of location and big data analytics, in contrast to traditional tourist surveys and conventional spatio-temporal data, can provide new and valuable knowledge. Implications arising from the study are significant assets for tourism SMEs, DMOs and other tourism stakeholders in the search of innovative marketing strategies for demonstrating the added value of destination,

strengthening destination branding and gaining a competitive advantage against other rival tourist destinations.

The third chapter of the thesis entitled, "Big Data Analytics for Tourism Destinations: A comparative analysis through Location-Based Social Networks", investigates the user-generated data in Location-based Social Networks (LBSNs) that can be a great resource of knowledge for understanding people's behaviour details and movement flows in tourism destinations. Nowadays, local authorities and tourism enterprises are using conventional methods like surveys and opinion polls for collecting data and strategic decision making. Despite the benefits of these approaches, they present significant disadvantages such as the sample size is small and they are time - consuming. Focusing on tourism and location-based social media networks, this chapter reveals a novel approach to leverage massive unstructured data for knowledge extraction. In contrast to the conventional spatio-temporal data, big social media data offer dynamically to innovation and value creation through improving the strategic decision-making process of tourism destination stakeholders. The approach integrates location and big data analytics techniques and it is implemented based upon geotagged user-generated data shared on the two largest islands in the Mediterranean Sea, the island of Crete (Greece), and the island of Cyprus that are popular summertime tourist destinations. The comparison between two tourist destinations with common characteristics provides additional insights into the potential of each destination and areas of improvement. Practical implications are arising through the efficient spatio-temporal and demographic analysis of tourist movement in both tourism destinations for improving the strategic decision making of stakeholders like local authorities and tourism SMEs leading to innovation and value creation. In addition, DMOs can leverage the new knowledge for developing innovative marketing strategies, strengthening destination branding and gaining a competitive advantage against rival tourism destinations.

# Table of contents

<b>1. Big data Analytics: Applications, prospects and challenges</b>	<b>21</b>
Introduction	22
Defining Big Data	22
Big Data Analytics	25
Big Data Analytics Applications	28
Big Data Analytics Prospects	30
Big data analytics challenges and barriers	33
Conclusions	34
<b>2. Knowledge Extraction through Location &amp; Big Data Analytics: the case of Crete</b>	<b>36</b>
Introduction	36
Literature/ Study background - framework	37
Knowledge extraction through Big Data Analytics	40
Data-driven Innovation for SMEs	41
SMEs, Location Intelligence & Tourism	43
Methodology	45
Data collection and analysis - a case study of Crete	48
Results & Discussion	50
Discussion - Conclusion	66
<b>3. Big Data Analytics for Tourism Destinations: A comparative analysis through Location-Based Social Networks</b>	<b>68</b>
Introduction	<b>68</b>
Literature / study background - framework	69
From Social Media to Big Data Analytics	71
Location and Big data analytics in Tourism	73
Methodology	77
Data collection and analysis - a case study for cities of Crete and Cyprus	80
Results and discussion	82
Conclusion	91
<b>References</b>	<b>93</b>

## List of Tables and Figures

Table 1: Conventional and Big Data analytics	26
Table 2: Big Data Analytics exploitation examples	29
Table 3: Terminology for using and analyzing data	38
Table 4: Summary of dataset (total users, posts, photos and engagement)	50
Figure 1: Challenges in Data Lifecycle	23
Figure 2: Big data trend	25
Figure 3: Process of leveraging big data	27
Figure 4: The evolution of Analytics eras	31
Figure 5: Data to knowledge process	37
Figure 6: The Analytics Process Model	39
Figure 7: The data architecture of the analysis	47
Figure 8: Study's novel approach	49
Figure 9: Type of content per month in Crete	50
Figure 10: Frequency of total posts per day in Crete	51
Figure 11: Total social engagement in Crete	51
Figure 12: Social engagement per type of content in Crete	52
Figure 13: Textual posts per day of the week and hour of day for Heraklion	52
Figure 14: Textual posts per day of the week and hour of day for Chania	53
Figure 15: Text Sentiment and entities analysis for Heraklion	54
Figure 16: Text Sentiment and entities analysis for Chania	54
Figure 17: Language and Gender of Textual Posts for Heraklion	55
Figure 18: Language and Gender of Textual Posts for Chania	55
Figure 19: Top words and hashtags of Textual Posts for Heraklion	56
Figure 20: Top words and hashtags of Textual Posts for Chania	56
Figure 21: Photo posts per day of the week and hour of day for Heraklion	58
Figure 22: Photo posts per day of the week and hour of day for Chania	58
Figure 23: Top words and hashtags of Textual Posts for Heraklion	59
Figure 24: Top words and hashtags of Textual Posts for Chania	59
Figure 25: Gender and Image Entities of Photos in Heraklion	60
Figure 26: Gender and Image Entities of Photos in Chania	60
Figure 27: Heatmap of user-generated content location in Heraklion	61
Figure 28: Heatmap of user-generated content location in Chania	62
Figure 29: Map of Heraklion - User-generated content distribution per content type	63
Figure 30: Map of Chania - User-generated content distribution per content type	63
Figure 31: Map of Heraklion - Location of the most positive textual posts	64
Figure 32: Map of Chania - Location of the most positive textual posts	64
Figure 33: Map of Heraklion - Location and type of the most engaged posts	65
Figure 34: Map of Chania - Location and type of the most engaged posts	65
Figure 35: The data flow of the analysis	79
Figure 36: The innovative approach for knowledge extraction through LBSNs	81
Figure 37: Summary of user-generated data per city and island	82
Figure 38: Type of content per month and island	83
Figure 39: Frequency of total posts per day and island	84
Figure 40: User-generated content per day of the week, island and type of content	84

Figure 41: User-generated content per hour of the day, island and type of content	<b>85</b>
Figure 42: Top words and hashtags per type for each island	<b>86</b>
Figure 43: Distribution of users' gender and type of content per island	<b>86</b>
Figure 44: Top languages per island	<b>87</b>
Figure 45: Top photo entities per island	<b>88</b>
Figure 46: Text sentiment and gender analysis per island	<b>89</b>
Figure 47: Maps of Crete - Cyprus / Location and type of the most positive and negative posts	<b>90</b>
Figure 48: Total social engagement per type of content and island	<b>90</b>
Figure 49: Map of Crete - Location and type of top engagement posts / Distribution of posts per type	<b>91</b>
Figure 50: Map of Cyprus - Location and type of top engagement posts / Distribution of posts per type	<b>91</b>



DEDICATED TO  
MY PARENTS & SISTER

# Acknowledgments

Undertaking this Ph.D. has been a truly life-changing experience for me and it would not have been possible to do without the support and guidance that I received from many people.

First and foremost, I would like to express my sincere gratitude to my Ph.D. Advisors Professors Emmanuel Petrakis and Ioannis Kopanakis for the continuous support of my Ph.D. study and related research, for their patience, motivation, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis. I could not have imagined having better advisors and mentors for my Ph.D. study.

Besides my advisors, I would like to thank Dr. Georgios Tsiotas, for his insightful comments and encouragement. I would also like to thank my colleague at the e-Business Intelligence Lab of Hellenic Mediterranean University, John Makridis, for his support and for giving me the courage to complete tasks before deadlines and for stimulating discussions.

Last but not least, I would like to express my deepest gratitude to my parents, sister, friends, and Despina Koutantou for their valuable support. They gave me enough moral support, encouragement, and motivation to accomplish my personal goals.

## SHORT BIO

Konstantinos Vassakis was born in 1989 in Athens, Greece. He holds a Bachelor's degree in Economics & Regional Development from Panteion University (2010), an MSc degree with a specialization in Business Economics and Finance with Distinction from the University of Portsmouth, UK (2012) and a Ph.D. in Economics from University of Crete (2019). He is a Scientific Associate at the e-Business Intelligence Lab of Hellenic Mediterranean University examining data-driven capabilities of SMEs. His applied research with publications in scientific journals and international conferences focuses on data management, innovation and knowledge management of enterprises.

Konstantinos Vassakis is Managing Director of H2B HUB powered by Heraklion Chamber of Commerce aiming to strengthen the Heraklion startup business ecosystem by providing the right conditions (e.g. inspiration, knowledge, collaboration capabilities & resources) for the creation of new innovative business activities. He is Co-Founder of Bizrupt, a non-for-profit organization that contributes to the building of a strong innovation ecosystem in Crete through the encouragement and promotion of innovation and sustainable entrepreneurship. He participates regularly as a speaker, mentor, and judge in competitions and events to support business ideas.

# Εκτεταμένη Περίληψη στα Ελληνικά

## ΚΕΦΑΛΑΙΟ 1: Big Data Analytics - Εφαρμογές, προοπτικές και προκλήσεις

Σήμερα, η ταχεία ανάπτυξη των τεχνολογιών που βασίζονται στο διαδίκτυο και η τεράστια αύξηση των συνδεδεμένων συσκευών μέσω του Διαδικτύου των πραγμάτων (Internet of Things - IoT), έχουν οδηγήσει σε μία ραγδαία αύξηση του όγκου δεδομένων μετονομαζοντας την εποχή που διανύουμε σε “εποχή των μεγάλων δεδομένων” (“Big Data era”). Τα μεγάλα δεδομένα αποτελούν μέρος της “οικονομίας της γνώσης” και αντιπροσωπεύουν ένα σημαντικό δυναμικό και ραγδαία εξελισσόμενο πεδίο για τις επιχειρήσεις και τους οργανισμούς, καθώς παρέχει τη δυνατότητα να αποκτήσουν πολύτιμες γνώσεις σχετικά με τις απόψεις, τις προτιμήσεις, τις ανάγκες και τις απαιτήσεις των πελατών τους.

Η ανάπτυξη του διαδικτύου με την έναρξη της Web 2.0 εποχής έδωσε στις εταιρίες ευκολότερη και πιο οικονομική πρόσβαση σε τεράστιο όγκο πληροφοριών, ενώ οι ευκαιρίες για συλλογή δεδομένων από εξωτερικές πηγές έχουν αυξηθεί ακόμη περισσότερο με την εμφάνιση του Web 3.0. Γι’ αυτό το λόγο, οι επιχειρήσεις και οι οργανισμοί ανεξαρτήτως κλάδου άρχισαν να εστιάζουν στην εκμετάλλευση των δεδομένων για την απόκτηση συγκριτικού πλεονεκτήματος. Στις μέρες μας, η εποχή των μεγάλων δεδομένων (Big Data) έχει φτάσει σε όλες σχεδόν τις επιχειρήσεις, καθώς έχουν αντιληφθεί ότι οι αποφάσεις που βασίζονται στα δεδομένα τείνουν να είναι καλύτερες και πιο ακριβείς. Παρόλα αυτά, το γεγονός ότι πολλές επιχειρήσεις σε διάφορους κλάδους εφαρμόζουν επιχειρηματικές αναλύσεις, συμπεριλαμβανομένων και αναλύσεων μεγάλων δεδομένων, δεν σημαίνει ότι όλες ωφελούνται από αυτές λαμβάνοντας χρήσιμη γνώση και πραγματική αξία από τα διαθέσιμα δεδομένα.

Η χρήση τεχνικών και εργαλείων ανάλυσης δεν αρκεί από μόνη της για να μεταμορφωθεί μια εταιρία σε εταιρία που βασίζεται στα δεδομένα (data-driven enterprise). Οι εταιρίες χρειάζεται να διαθέτουν ανθρώπους με συστηματική σκέψη για τη λήψη καλύτερων αποφάσεων βασισμένων σε δεδομένα. Η επιτυχία στο σημερινό, προσανατολισμένο στα δεδομένα, επιχειρηματικό περιβάλλον περιλαμβάνει την ικανότητα να σκέφτεται κανείς data-analytically. Καθώς ο όγκος των δεδομένων αυξάνεται διαρκώς, η γνώση και η ανάλυση δεν μπορούν να διαχωριστούν μεταξύ τους. Τόσο οι ακαδημαϊκοί όσο και οι επαγγελματίες οφείλουν να έχουν αναλυτικές ικανότητες και να κατανοούν τις επιχειρηματικές διαδικασίες.

Οι επαγγελματίες που δεν διαθέτουν στοιχειώδη κατανόηση του τρόπου σκέψης στην ανάλυση δεδομένων, δεν γνωρίζουν ουσιαστικά πως λειτουργεί μια επιχείρηση. Αν μπορούν να καταλάβουν τη

διαδικασία και στάδια εξέλιξης της, είναι ευκολότερο για αυτούς να βρουν τις κατάλληλες λύσεις στις αδυναμίες του κάθε σταδίου. Αλλά για να μπορούν οι οργανισμοί να είναι προσανατολισμένοι στα δεδομένα πρέπει να αντιμετωπίσουν διοικητικές και τεχνικές προκλήσεις.

Τα μεγάλα δεδομένα δεν έχουν να κάνουν μόνο με τον όγκο αλλά και με την ποικιλία και την ταχύτητα. Η ανάλυση τους δίνει την δυνατότητα στις επιχειρήσεις να κατανοήσουν το επιχειρηματικό τους περιβάλλον, την συμπεριφορά και τις ανάγκες των πελατών τους και τις δραστηριότητες των ανταγωνιστών τους. Χάρη στην ανάλυση των μεγάλων δεδομένων οι επιχειρήσεις μπορούν να διαμορφώσουν τα προϊόντα και τις ενέργειές τους με τέτοιο τρόπο ώστε να καλύπτουν τις ανάγκες των πελατών τους και να καινοτομούν έναντι των ανταγωνιστών κάνοντας καλύτερες προβλέψεις και λαμβάνοντας σωστότερες αποφάσεις που βασίζονται σε αποδεικτικά στοιχεία και όχι σε διαίσθηση. Οι οργανισμοί που καταφέρνουν να διαχειριστούν τις προκλήσεις και να υιοθετήσουν μια κουλτούρα που βασίζεται στα δεδομένα, μπορούν να περιμένουν καλές προοπτικές. Υπάρχουν ισχυρές αποδείξεις ότι η απόδοση μιας εταιρίας μπορεί να βελτιωθεί μέσα από αποφάσεις βασισμένες σε δεδομένα, εργαλεία και τεχνικές ανάλυσης των μεγάλων δεδομένων. Καθώς όλο και περισσότερες εταιρείες αποκτούν τις απαραίτητες δεξιότητες για την χρήση των μεγάλων δεδομένων και ασχολούνται με τις σύγχρονες τεχνολογίες, που συνεχώς εξελίσσονται, είναι πιθανόν σύντομα να μπορέσουν να ξεχωρίσουν από τον ανταγωνισμό και να αποκτήσουν αποφασιστικό συγκριτικό πλεονέκτημα.

Τα μεγάλα δεδομένα βοηθούν τις επιχειρήσεις και τους οργανισμούς να δημιουργούν αξία και να αποκτούν πρακτική γνώση ώστε να λαμβάνουν καλύτερες αποφάσεις και να φέρουν επανάσταση στις στρατηγικές και τα επιχειρηματικά τους μοντέλα (Waller & Fawcett, 2013). Η διαθεσιμότητα των μεγάλων δεδομένων που σχετίζονται με καταναλωτές δίνει στις επιχειρήσεις μοναδικές ευκαιρίες για να κατανοήσουν τις τάσεις της αγοράς, την συμπεριφορά και τις ενέργειες του καταναλωτή (Erevelles et al., 2016; Waller & Fawcett, 2013). Παρόλο που το να βασίζονται οι επιχειρήσεις στα δεδομένα δεν είναι κάτι νέο, προσφατά άρχισαν να αξιοποιούν και άλλες πηγές δεδομένων όπως κοινωνικά δίκτυα, smartphones ή σένσορες, και νέες τεχνολογίες ανάλυσης για να εκμεταλλευτούν τα δεδομένα αυτά (Hartmann et al., 2016). Τα μεγάλα δεδομένα δεν αφορούν μόνο την συλλογή και την αποθήκευση των δεδομένων, ο στόχος τους είναι η ανάλυση των δεδομένων για την λήψη γνώσης και αξίας (Bello-Orgaz et al., 2016). Στην σημερινή εποχή, οι επιχειρήσεις που δεν καταφέρνουν να προσαρμοστούν στις διαδικασίες που βασίζονται στα δεδομένα αντιμετωπίζουν το κίνδυνο να χάσουν σημαντικό ανταγωνιστικό πλεονέκτημα, ενώ η μη επαρκής χρήση των δεδομένων μπορεί να δημιουργήσει θέματα βιωσιμότητας (Brownlow et al., 2015).

Τα κοινωνικά δίκτυα έχουν μεταμορφώσει τις σχέσεις μεταξύ επιχειρήσεων, εργαζομένων, πελατών και ενδιαφερόμενων. Εξέχοντα παραδείγματα κοινωνικών δικτύων περιλαμβάνουν το Facebook (με πάνω από 2,2 δισεκατομμύρια ενεργούς χρήστες μηνιαίως), το Twitter (με πάνω από 330 εκατομμύρια

ενεργούς χρήστες μηνιαίως), το YouTube (με πάνω από 1,8 δισεκατομμύρια ενεργούς χρήστες μηνιαίως) και το Yelp, πλατφόρμα για κριτικές προϊόντων και υπηρεσιών από τους ίδιους τους καταναλωτές, όπως και blogs, forums, microblogs, πλατφόρμες διαμοιρασμού εικόνων και video (για παράδειγμα το Instagram με 800 εκατομμύρια ενεργούς χρήστες μηνιαίως). Αυτοί οι αριθμοί δείχνουν ξεκάθαρα ότι τα κοινωνικά δίκτυα παράγουν τεράστιο όγκο δεδομένων και αυτή η πληροφορία μπορεί να προσφέρει τεράστιες δυνατότητες ανάπτυξης στις επιχειρήσεις. Λόγω του αριθμού των ενεργών χρηστών και του περιεχομένου που δημιουργείται καθημερινά στις πλατφόρμες αυτές, οι οργανισμοί παρακινούνται να αναγνωρίσουν τάσεις, πιθανούς κινδύνους και ευκαιρίες, κάτι που είναι ιδιαίτερα υποκειμενικό για κάθε οργανισμό. Εκτός από την παρακολούθηση των σύγχρονων τάσεων, είναι απαραίτητο να γνωρίζουμε και τους ίδιους τους δημιουργούς του περιεχομένου, ιδιαίτερα αυτούς με την μεγαλύτερη επιρροή. Τόσο οι επιχειρήσεις όσο και οι μη-κερδοσκοπικοί οργανισμοί συλλέγουν δεδομένα που δημιουργούν οι χρήστες με στόχο να εξασφαλίσουν μια θέση στον κόσμο των πολυμέσων. Οι πελάτες μπορούν να αποκτήσουν διαφόρων ειδών δεδομένα ανταλλάσσοντας πληροφορίες μεταξύ τους μέσω των κοινωνικών δικτύων. Αυτές οι πληροφορίες παίζουν σημαντικό ρόλο στις αγοραστικές αποφάσεις. Επομένως, μια σημαντική πτυχή των κοινωνικών δικτύων είναι το χτίσιμο σχέσεων με τους καταναλωτές, τους προμηθευτές, τους διαχειριστές δημοσίων σχέσεων και τους δυνητικούς συνεργάτες.

Η ανάλυση των μεγάλων δεδομένων από τα κοινωνικά δίκτυα μπορεί να εφαρμοστεί για την ανακάλυψη νέας γνώσης που μπορεί να χρησιμοποιηθεί για καλύτερη λήψη αποφάσεων. Η ανάλυση των δεδομένων από τα κοινωνικά δίκτυα παρέχει μοναδική ευκαιρία για ανάκτηση τεράστιου αριθμού απόψεων των πελατών και εμπειριών που συμβάλλουν στην βελτίωση των παραδοσιακών στρατηγικών μέσα από την δημιουργία περισσότερο στοχευμένων ενεργειών μαρκετινγκ (Bello-Organ et al., 2016) και στρατηγικής. Η ανάλυση περιεχομένου και πληροφοριών που διαμοιράζεται στα κοινωνικά δίκτυα έχει φανεί χρήσιμη σε διάφορους κλάδους όπως μαρκετινγκ, δημόσια υγεία, δημόσια ασφάλεια (Gerber 2014), τουρισμός και πολιτική (Carvalho et al., 2017). Υπάρχουν διαφορετικές τεχνικές οι οποίες χρησιμοποιούνται για την ανακάλυψη κρυμμένων μοτίβων και γνώσεων στα δεδομένα από τα κοινωνικά δίκτυα. Μερικές από αυτές τις τεχνικές είναι: Ανάλυση κειμένου και συναισθήματος, topic modeling, natural language processing (NLP) και data mining. Η οπτικοποίηση των δεδομένων και η οπτική ανάλυση παίζουν σημαντικό ρόλο στην αποτελεσματική διαχείριση και κατανόηση των μεγάλων δεδομένων (Ribarsky et al., 2014).

Στην εποχή της ψηφιακού μετασχηματικού και καινοτομίας, οι μικρομεσαίες επιχειρήσεις (small and medium enterprises - SMEs) έχουν την ικανότητα να ανατρέψουν τα παραδοσιακά επιχειρηματικά μοντέλα χάρη στην ευελιξία και την καινοτομία τους. Ειδικότερα για τις μικρομεσαίες επιχειρήσεις που έχουν κατά κύριο λόγο περιορισμένους πόρους και δεν έχουν την πολυτέλεια να κάνουν λάθος διαχείριση, η γνώση που λαμβάνουν μπορεί να συμβάλει στην συγκέντρωση στα πιο σημαντικά

ζητήματα. Για αυτο το λόγο, η ανάλυση συναισθήματος μπορεί να θεωρηθεί μια προσοδοφόρα τεχνική που προέρχεται από τα μεγάλα δεδομένα.

Η διασύνδεση των κοινωνικών δικτύων είναι ένα εκτεταμένο εργαλείο μαρκετινγκ για να απευθυνθεί κανείς στο βασικό κοινό του. Η δυνατότητα να απαντάει κανείς γρήγορα, να αλληλεπιδρά με το κοινό του τακτικά και για θέματα που δεν σχετίζονται απαραίτητα με το προϊόν, μπορεί να ενδυναμώσει τα όρια μεταξύ των επιχειρήσεων και των πελατών τους, ενώ είναι ένας οικονομικό άμεσο κανάλι επικοινωνίας ιδιαίτερα για τις SMEs (Nobre & Silva, 2014). Οι επιχειρήσεις που δεν αφιερώνουν αρκετούς από τους πόρους τους στην διαδικασία οργάνωσης και δεν ερχονται σε επαφή με τους πελάτες τους μπορεί να αποδειχθούν περισσότερο ευάλωτες και είναι πιο πιθανό να χάσουν την ανταγωνιστικότητά τους με το πέρασμα του χρόνου. Οποιαδήποτε τουριστική επιχείρηση χρειάζεται να αξιοποιήσει τις οργανωσιακές και μαρκετίστικες στρατηγικές, τακτικές και εργαλεία για να πετύχει και να διατηρήσει ένα μόνιμο ανταγωνιστικό πλεονέκτημα (Mariani et al., 2018). Με την ένωση πολλών διαφορετικών πηγών μεγάλων δεδομένων, π χ ετερογενείς πηγές δεδομένων με τη μορφή δομημένων και μη δομημένων δεδομένων, όπως για παράδειγμα η ανατροφοδότηση των πελατών, δεδομένα κρατήσεων και δεδομένα ιστορικού αναζήτησης από την πλευρά των πελατών και των προμηθευτών μπορούν να αποκαλύψουν νέα γνώση που οι παραδοσιακές προσεγγίσεις δεν μπορούσαν να εντοπίσουν (Mariani et al., 2018). Όλες αυτές οι περίπλοκες διαδικασίες μπορούν να αξιολογηθούν και να εκτιμηθούν με την ανάλυση των μεγάλων δεδομένων τα οποία μπορούν να βοηθήσουν στην αύξηση της μακροζωίας των επιχειρήσεων.

## ΚΕΦΑΛΑΙΟ 2: Εξόρυξη γνώσης με την αξιοποίηση Location & Big Data Analytics: η περίπτωση της Κρήτης

Σήμερα, οι τουρίστες παράγουν τεράστιο όγκο δεδομένων όταν επισκέπτονται αστικούς προορισμούς. Ωστόσο, η γνώση σχετικά με την δραστηριότητα και τις αντιλήψεις τους για τον προορισμό αυτό είναι περιορισμένη. Οι επιχειρήσεις και οι οργανισμοί στους κλάδους της φιλοξενίας και του τουρισμού μπορούν να αξιοποιήσουν πραγματικά δεδομένα συμπεριφοράς - αντιλήψεις που προκύπτουν από μεγάλα δεδομένα που δημιουργούνται σε πραγματικό χρόνο από διαδικτυακές πηγές δεδομένων σε αντίθεση με τα παραδοσιακές έρευνες που βασίζονται σε ερωτηματολόγια (Heerschap et al., 2014). Ωστόσο, υπάρχει ερευνητικό κενό σχετικά με την εκμετάλλευση των μεγάλων δεδομένων στην φιλοξενία και τον τουριστικό κλάδο. Επομένως, απαιτείται η υλοποίηση καινοτόμων πρακτικών βασισμένων στην ανάλυση μεγάλων δεδομένων οι οποίες θα αναδεικνύουν την χρησιμότητα της εφαρμογής των μεγάλων δεδομένων σε αυτούς τους κλάδους (Mariani et al., 2018).

Η χρήση δεδομένων που βασίζονται στα κοινωνικά δίκτυα είναι εξαιρετικά επωφελής για πολλά ενδιαφερόμενα μέρη (stakeholders) συμπεριλαμβανομένων των πολιτών και των κυβερνήσεων. Τα δίκτυα των κοινωνικών μέσων ενημέρωσης (social media networks) μπορούν να αποτελέσουν κυρίαρχο πόρο για πολύτιμη γνώση σχετικά με τους τουριστικούς προορισμούς μέσω της συλλογής δεδομένων από τα κοινωνικά δίκτυα βάσει τοποθεσίας (Location-Based Social Networks). Ένα σημαντικό πρόβλημα είναι η έλλειψη γνώσης σχετικά με την άποψη των επισκεπτών για έναν προορισμό, καθώς η συμπεριφορά, οι ανάγκες και οι προτιμήσεις των επισκεπτών δεν είναι ορατές.

Η ευκαιρία για έκφραση των αποψεων του πελάτη δημοσίως μέσω των κοινωνικών δικτύων, χωρίς τοπικούς και χρονικούς περιορισμούς, έχει αλλάξει σε μεγάλο βαθμό τον τρόπο που οι επιχειρήσεις αντιμετωπίζουν τους πελάτες τους. Καθώς το τουριστικό προϊόν βασίζεται στην κάλυψη αναγκών και ικανοποίηση επιθυμιών, ο κλάδος του τουρισμού και της φιλοξενίας, περισσότερο από οποιονδήποτε άλλο κλάδο, πρέπει να παραμένει μια εμπειρία απόλαυσης για τον πελάτη. Οι κακές κριτικές μπορούν να βλάψουν σημαντικά την επιχείρηση. Η δημοσίευση φωτογραφιών ως αποδειξη για την λήψη κακών υπηρεσιών έχουν αρνητική επίπτωση στο πως οι νέοι δυνητικοί πελάτες θα σκεφτούν μια κράτηση ή αγορά. Η δύναμη έχει μεταφερθεί από την επιχείρηση στον πελάτη όσον αφορά τη διαφάνεια με τη μη ικανοποίηση και η δυσφορία να είναι εμφανής σε οποιοδήποτε συνδέεται στο διαδίκτυο και τα σκάνδαλα μπορούν να γίνουν viral, πράγμα που έχει ως αποτέλεσμα μεγάλο βαθμό αρνητικής δημοσιότητας και απώλεια κύρους μιας επιχείρησης.

Στις μέρες μας, είναι σημαντικό να δίνεται προσοχή στις λεπτομέρειες, καθώς η φήμη βασίζεται κυρίως στην επικοινωνία των κριτικών από ένα μεγάλο δείγμα ανθρώπων, όπου εμφανίζεται με μεγάλη λεπτομέρεια η εμπειρία σχεδόν όλων των πελατών. Οι παραδοσιακές έρευνες με ερωτηματολόγια ή

παρατηρώντας τις επισκέψεις σε τοποθεσίες, επιχειρήσεις κ.α. είναι πολύ φτωχές πηγές πληροφορίας σε σχέση με το διαδικτυακό αποτύπωμα. Η δραστηριότητα του ατόμου στα κοινωνικά δίκτυα, η οποία σχετίζεται με την αυτοέκφραση και βασίζεται αποκλειστικά στην επιθυμία του ατόμου να παρουσιάσει τον εαυτό του, είναι καταλύτης για τα μεγάλα δεδομένα (Salas-Olmedo et al., 2018). Η διαθεσιμότητα του περιεχομένου που παράγεται από τους χρήστες στο διαδίκτυο και οι νέες τεχνολογίες παρέχουν στους ερευνητές μια νέα οπτική με την οποία μπορούμε να προσεγγίσουμε την αντιλήψεις και πιθανόν και το επίπεδο ικανοποίησης των ταξιδιωτών (Alaei et al., 2017). Οι επιχειρήσεις και οι επαγγελματίες του τουρισμού μπορούν να προσαρμόσουν και να βελτιώσουν συγκεκριμένα αρνητικά σημεία που ενοχλούν την πλειονότητα των πελατών τους ενώ επίσης μπορούν να γνωρίζουν ποια μέρη της στρατηγικής τους και του επιχειρηματικού τους μοντέλου πήγαν καλά. Μπορούν επίσης να μάθουν προς ποια κατεύθυνση να καινοτομήσουν και να επεκτείνουν τις δραστηριότητες τους, τις εγκαταστάσεις τους και τις υπηρεσίες τους.

Η βελτίωση της λήψης αποφάσεων μέσω πληροφοριών για την τουριστική εμπειρία που λαμβάνονται σε πραγματικό χρόνο, μπορεί να οδηγήσει στην παρακίνηση των τουριστών ανακαλύπτοντας ανάγκες και προτιμήσεις, ανάπτυξη νέων επιχειρηματικών μοντέλων, προϊόντων και υπηρεσιών και βελτίωση του επιχειρηματικού οικοσυστήματος ευνοώντας την συνεργασία και την εμπιστοσύνη μεταξύ των μελών της κοινότητας. Στο πλαίσιο αυτό, το κεφάλαιο συνεισφέρει στις εφαρμογές ανάλυσης των μεγάλων δεδομένων στον κλάδο του τουρισμού και της φιλοξενίας προτείνοντας μια καινοτόμο προσέγγιση για την αξιοποίηση τεράστιου όγκου μη δομημένων δεδομένων από κοινωνικά δίκτυα τοποθεσίας και παρέχουν πολύτιμη γνώση για τα χωρικά, χρονικά και δημογραφικά χαρακτηριστικά των τουριστών. Ενώ έχουν παρουσιαστεί άλλοι μέθοδοι ανάλυσης δεδομένων παρατήρησης, αυτή η μελέτη παρουσιάζει μια καινοτομική προσέγγιση που συνδυάζει τεχνικές μεγάλων δεδομένων όπως text mining (ανάλυση συναισθήματος, εντοπισμός θέματος κτλ), πληροφορίες τοποθεσίας και κοινωνικά δίκτυα μεταμορφώνοντας τις εμπειρίες των τουριστών σε πολύτιμα αγαθά (απόκτηση νέας γνώσης) για γρηγορότερη και πιο αποτελεσματική λήψη αποφάσεων από τις τουριστικές SMEs και τους άμεσα ενδιαφερόμενους. Πιο συγκεκριμένα, η προσέγγιση αυτή εισάγει την ανάλυση κειμένων και εικόνων με λήψη δεδομένων από κοινωνικά δίκτυα κοινής χρήσης πολυμέσων και κειμένων, την ανάλυση των κοινωνικών δεδομένων (social big data) όπως κοινωνική διάδραση, ανάλυση συναισθήματος, εντοπισμός θεματολογίας σε συνδυασμό με τοπικά και χρονικά χαρακτηριστικά για να δώσει περισσότερη γνώση σχετικά με τους τουριστικούς προορισμούς.

Η περίπτωση της Κρήτης χρησιμοποιήθηκε για να παρουσιαστεί πως η προτεινόμενη προσέγγιση μπορεί να παρέχει χρήσιμες πρακτικές εφαρμογές στη διαχείριση τοποθεσίας, την βελτίωση και την προώθηση της εμπειρίας των τουριστών. Χρησιμοποιήθηκε ένα σετ δεδομένων από υλικό που παρήγαγαν οι χρήστες (δημοσιεύσεις κειμένων και φωτογραφιών) το οποίο κοινοποιήθηκε σε



κοινωνικά δίκτυα εντοπισμού θέσης για τις δυο μεγαλύτερες πόλεις της Κρήτης, το Ηράκλειο και τα Χανιά.

Η γνώση σχετικά με το ποιος είναι ο δημοφιλέστερος προορισμός, τι ελκύει περισσότερο τους τουρίστες και γιατί, ποιες είναι οι δημοφιλέστερες μέρες/ώρες για δημοσίευση κάθε τύπου περιεχομένου, ποιά μέρη προτιμούν να επισκέπτονται οι τουρίστες και γιατί, ποια είναι η πιο ελκυστική τοποθεσία για συγκεκριμένη μέρα/ώρα και ποια είναι τα χαρακτηριστικά των τουριστών αυτών, βοηθάει τους τουριστικούς οργανισμούς, τις τουριστές μικρομεσαίες επιχειρήσεις (SMEs) και τους υπόλοιπα ενδιαφερόμενα μέρη όπως δήμοι, περιφέρειες κτλ. να σχεδιάζουν, να υλοποιούν και να αναπτύσσουν προϊόντα και υπηρεσίες που βασίζονται στην γνώση που προέρχεται από διαδραστικές οπτικοποιήσεις δεδομένων δίνοντας τη δυνατότητα για βαθύτερη ανάλυση χωρίς να απαιτείται κάποια εξειδίκευση. Επιπλέον, τα αποτελέσματα της έρευνας παρέχουν γνώση σχετικά με την ανάπτυξη των κατάλληλων στρατηγικών για την προώθηση των πόλεων και ολοκληρου του νησιού της Κρήτης ως τουριστικού προορισμού μέσα από την διαδικτυακή προώθηση.

Τα αποτελέσματα της έρευνας καταδεικνύουν πως η νέα αυτή προσέγγιση σχετικά με την τοποθεσία και την ανάλυση των μεγάλων δεδομένων σε αντίθεση με τις παραδοσιακές έρευνες και τα συμβατικά χωρικά/χρονικά δεδομένα μπορεί να δώσει νέα και πολύτιμη γνώση. Τα συμπεράσματα της έρευνας είναι πολύτιμα για τις τουριστικές SMEs, DMOs και άλλους ενδιαφερόμενους του τουρισμού που αναζητούν καινοτόμες στρατηγικές marketing για να παρουσιάσουν την προστιθέμενη αξία ενός προορισμού και να ενδυναμώσουν την επωνυμία του αποκτώντας συγκριτικό πλεονέκτημα έναντι σε άλλους τουριστικούς προορισμούς.

### ΚΕΦΑΛΑΙΟ 3: Μεγάλη ανάλυση δεδομένων για προορισμούς τουρισμού: Μια συγκριτική ανάλυση μέσω Location - Based Social Networks

Τα δεδομένα που παράγονται από τους χρήστες στα κοινωνικά δίκτυα τοποθεσίας, Location-based Social Networks (LBSNs), μπορούν να αποτελέσουν σημαντική πηγή γνώσης για την κατανόηση των λεπτομερειών της ανθρώπινης συμπεριφοράς και την κίνηση στους τουριστικούς προορισμούς. Στις μέρες μας, οι τοπικές αρχές και οι τουριστικές επιχειρήσεις χρησιμοποιούν συμβατικές μεθόδους όπως ερωτηματολόγια και δημοσκοπήσεις για να συλλέξουν δεδομένα και να λάβουν στρατηγικές αποφάσεις. Παρά τα πλεονεκτήματά τους, οι προσεγγίσεις αυτές παρουσιάζουν σημαντικά μειονεκτήματα όπως ότι είναι αρκετά χρονοβόρες και έχουν τη δυνατότητα να συλλογής μικρού δείγματος απαντήσεων. Η βιβλιογραφία που παρουσιάζει πρακτικές μελέτες που συνδυάζουν τεχνικές ανάλυσης μεγάλων δεδομένων και κοινωνικά δίκτυα είναι σπάνια, ενώ υπάρχει έλλειψη γνώσης για την αξιοποίηση των μεγάλων δεδομένων στον κλάδο της φιλοξενίας και του τουρισμού. Επομένως, είναι αναγκαία η υλοποίηση καινοτόμων εφαρμογών ανάλυσης των μεγάλων δεδομένων τα οποία αξιοποιούν τα κοινωνικά δίκτυα θέσης για την απόκτηση νέα γνώσης σχετικά με τη συμπεριφορά και τις αντιλήψεις γύρω από τους τουριστικούς προορισμούς.

Καθώς η τάση της ψηφιοποίησης συνεχίζεται, είναι πιθανό ότι έξυπνοι τουριστικοί προορισμοί (smart tourist destinations) θα αναδειχθούν. Αυτοί θα βασίζονται στη “διαδραση ενός προορισμού με την κοινότητα των ενδιαφερομένων μερών, των κατοίκων και των τουριστών, η οποία θα βασίζεται σε δυναμικές πλατφόρμες, ροές επικοινωνίες με βάση τη γνώση και βελτιωμένα συστήματα στήριξης αποφάσεων” (Vecchio et al., 2018a). Με την πρόβλεψη ότι η πρόσβαση στα δεδομένα θα γίνει ακόμα πιο εύκολη και η πληροφορία ακόμη πιο ακριβής, θα αυξηθεί σημαντικά την ακρίβεια των προβλέψεων. Ωστόσο, στον τουριστικό κλάδο, η δυνατότητα των μεγάλων δεδομένων δεν έχει αξιοποιηθεί σε μεγάλο βαθμό. Ο κλάδος των τουρισμού και της φιλοξενίας δυσκολεύονται με καινοτόμες προσεγγίσεις οι οποίες θα προσδίδουν αξία αξιοποιώντας τα μεγάλα δεδομένα (Mariani et al., 2018).

Η έρευνα του τουρισμού μπορεί να εξελιχθεί σε μια νέα εποχή όπου οι προσεγγίσεις που βασίζονται στη θεωρία και οι πρακτικές που βασίζονται στα δεδομένα θα υποστηρίζουν η μια την άλλη για την κατανόηση ή την εξήγηση φαινομένων καθώς επίσης και τον εντοπισμό νέων διαστάσεων στις θεωρίες (Alaei et al., 2017). Στην ψηφιακή εποχή, νέες ευκαιρίες για διάδραση μπορούν να διαφανούν για τις επιχειρήσεις, τους οργανισμούς και τους τουρίστες (πελάτες) μέσω των κοινωνικών δικτύων. Καθώς όλο και μεγαλύτερος όγκος περιεχομένου δημιουργείται καθημερινά από τους ίδιους τους χρήστες, αποκτάμε όλο και περισσότερη γνώση σχετικά με τη συμπεριφορά των καταναλωτών. Ετσι, απολύτως προσαρμοσμένες στρατηγικές μπορούν να αναπτυχθούν για να καλύψουν πλήρως τις ανάγκες των πελατών βελτιώνοντας έτσι την αποτελεσματικότητα των ενεργειών μάρκετινγκ, ενώ ξεπερνιούνται εμποδία και δημιουργείται καινοτομία προς πολλά υποσχόμενες κατευθύνσεις.

Προηγούμενες μελέτες έχουν προσπαθήσει να συλλάβουν και να αναλύσετε τις τουριστικές δραστηριότητες και προτιμήσεις για την βελτίωση των στρατηγικών αποφάσεων και του προγραμματισμού στους τουριστικούς προορισμούς. Οι παραδοσιακές προσεγγίσεις όπως τα ταξιδιωτικά ημερολόγια παρουσιάζουν περιορισμούς ενώ τα κοινωνικά δίκτυα μπορούν να συλλάβουν σε μεγάλο βαθμό την συμπεριφορά των τουριστών στον χώρο και το χρόνο. Έτσι, στο παρόν κεφάλαιο γίνεται προσπάθεια να εισάγουμε μια καινοτόμα μεθοδολογία για την εξαγωγή επιπρόσθετης και αξιολογής γνώσης από το περιεχόμενο που οι χρήστες δημιουργούν και κοινοποιούν σε δημοφιλείς πλατφόρμες κοινωνικών δικτύων μέσω των κινητών τους τηλεφώνων στους υπό εξέταση τουριστικούς προορισμούς. Ο συνδυασμός των πληροφοριών τοποθεσίας (location intelligence) με την ανάλυση των μεγάλων δεδομένων παρέχει περισσότερη γνώση επιτρέποντας στις επιχειρήσεις και στα ενδιαφερόμενα μέρη να πραγματοποιούν καλύτερες προβλέψεις σχετικά με τις συμπεριφορές των ταξιδιωτών και να χρησιμοποιούν τους διαθέσιμους πόρους αναλόγως. Συγκεντρώνοντας και αναλύοντας μεγάλα δεδομένα βάσει τοποθεσίας (LBSN big data), τα ενδιαφερόμενα μέρη μπορούν να αποκτήσουν πολύτιμη γνώση για τα μοτίβα, τις προτιμήσεις και τις ανάγκες των τουριστών και των ταξιδιωτών. Τα αποτελέσματα της ανάλυσης για τα δυο νησιά μας δίνουν γνώση για τις τάσεις των τουριστών σε μέρη ανάλογα με τις ημερομηνίες και του είδους του περιεχομένου (φωτογραφίες ή κείμενο), το συναίσθημα των τουριστών (το πιο θετικό για το παραδοσιακό φαγητό και την ομορφιά των τοπίων στην Κρήτη και το πιο θετικό για την ομορφιά των τοπίων και τις δραστηριότητες στην Κύπρο) σε συγκεκριμένες περιοχές, την διάδραση του χρήστη αναλογα με το είδος του περιεχομένου σε συγκεκριμένους προορισμούς και τον εντοπισμό της θεματολογίας του παραγόμενου περιεχομένου.

Η γνώση για το ποιος είναι ο δημοφιλέστερος προορισμός, τι ελκύει τους τουρίστες περισσότερο και γιατί, ποιες είναι οι πιο δημοφιλείς ώρες και οι προτιμώμενες μέρες-ώρες για δημοσίευση περιεχομένου, ποια μέρη προτιμούν να επισκέπτονται οι τουρίστες και γιατί, ποια είναι η πιο ελκυστική τοποθεσία και ποια είναι τα χαρακτηριστικά των τουριστών, δίνει τη δυνατότητα στους τουριστικούς οργανισμούς, τις επιχειρήσεις και τους διοργανωτές να σχεδιάσουν, εφαρμόσουν και αναπτύξουν προϊόντα και υπηρεσίες με βάση τη γνώση χρησιμοποιώντας διαδραστικές οπτικοποιήσεις των δεδομένων οι οποίες βοηθούν στην βαθύτερη ανάλυση χωρίς να απαιτείται εξειδίκευση. Επιπλέον, σύμφωνα με την έρευνα γύρω από τις έξυπνες πόλεις, η βελτίωση της διαδικασίας λήψης στρατηγικών αποφάσεων από τις τοπικές αρχές κατανοώντας τις δραστηριότητες των ανθρώπων, τις προτιμήσεις, τις εμπειρίες και το πώς μετακινούνται είναι ιδιαίτερα σημαντικά για τα ενδιαφερόμενα μέρη μιας έξυπνης πόλης. Η σύγκριση μεταξύ των δυο τουριστικών προορισμών που παρουσιάζουν παρόμοια χαρακτηριστικά μπορεί επίσης να δώσει γνώση για τις πρακτικές που εφαρμόζονται επιτυχώς στο ένα και θα μπορούσαν να υιοθετηθούν και από το άλλο.

Με επίκεντρο τον τουρισμό και τα κοινωνικά δίκτυα τοποθεσίας, το κεφάλαιο στοχεύει στην ανακάλυψη μιας νέας προσέγγισης για την αξιοποίηση τεράστιου όγκου μη δομημένων δεδομένων με

σκοπό την εξόρυξη γνώσης. Σε αντίθεση με τα παραδοσιακά δεδομένα θέσης/χρονου, τα μεγάλα δεδομένα που προέρχονται από τα κοινωνικά δίκτυα προσφέρουν με δυναμικό τρόπο στην καινοτομία και την δημιουργία αξίας μέσω της βελτίωσης λήψης στρατηγικών αποφάσεων από την πολύτιμη γνώση που παράγεται σχετικά με την συμπεριφορά των ταξιδιωτών, τις εντυπώσεις τους και τις προτιμήσεις τους σχετικά με τους τουριστικούς προορισμούς. Για την επίτευξη του στόχου αυτού χρησιμοποιήθηκε μια καινοτόμα προσέγγιση με την οποία γίνεται συγκέντρωση τεράστιου όγκου μη δομημένων δεδομένων που προέρχονται από δημοφιλή κοινωνικά δίκτυα τοποθεσίας (Twitter, Foursquare, Flickr και Instagram). Πιο συγκεκριμένα, η προσέγγιση αυτή εισάγει τον συνδυασμό ανάλυσης κειμένου και φωτογραφιών που συλλέχθηκαν από δίκτυα κοινής χρήσης πολυμέσων και εισάγει αναλύσεις μεγάλων δεδομένων προερχόμενων από κοινωνικά δίκτυα, όπως social engagement, sentiment analysis, topic/label detection σε συνδυασμό με χαρακτηριστικά τόπου/χρόνου για να παρέχει περισσότερη γνώση σχετικά με τους τουριστικούς προορισμούς.

Η προσέγγιση συνδυάζει τεχνικές ανάλυσης μεγάλων δεδομένων και τοποθεσιών και χρησιμοποιεί δεδομένα γεωγραφικού εντοπισμού (geotagged) τα οποία δημοσίευσαν οι ίδιοι οι χρήστες στα δύο μεγαλύτερα νησιά της Μεσογείου, την Κρήτη (Ελλάδα) και το νησί της Κύπρου. Η σύγκριση μεταξύ δύο προορισμών με κοινά χαρακτηριστικά μπορεί να δώσει περισσότερη γνώση σχετικά με την δυνατότητα βελτίωσης της κάθε περιοχής. Πρακτικά συμπεράσματα προκύπτουν από την αποτελεσματική ανάλυση τοπικών, χρονικών και δημογραφικών δεδομένων γύρω από την κίνηση των τουριστών στους δυο τουριστικούς προορισμούς για την βελτίωση των στρατηγικών αποφάσεων των τοπικών αρχών και των τουριστικών SMEs οδηγώντας σε καινοτομία και δημιουργία αξίας. Επιπλέον, οι DMOs μπορούν να αξιοποιήσουν την νέα γνώση για τη δημιουργία καινοτόμων στρατηγικών μαρκετινγκ με στόχο την ενδυνάμωση του branding ενός προορισμού αποκτώντας έτσι ανταγωνιστικό πλεονέκτημα έναντι των ανταγωνιστικών προορισμών.

Η ανάλυση αυτή παρέχει ορισμένα συμπεράσματα για τους ενδιαφερόμενους του τουρισμού και συνεισφέρει στον ακαδημαϊκό χώρο και στις τουριστικές και επιχειρήσεις και φιλοξενίας. Ωστόσο, πρέπει να αναγνωρίσουμε ότι υπάρχουν ορισμένοι περιορισμοί στην μελέτη αυτή. Καταρχάς, χρησιμοποιείται μικρός αριθμός κοινωνικών δικτύων. Παρόλο που η έρευνα βασίζεται σε πολύ μεγάλο σύνολο δεδομένων, αφορά μόνο το διάστημα τριών μηνών και τις πόλεις των δύο νησιών (δεν περιλαμβάνονται αγροτικές ή παράκτιες περιοχές). Επιπλέον, στην έρευνα χρησιμοποιήθηκαν δεδομένα εντοπισμού θέσης (location tag) που παρήγαγαν οι ίδιοι οι χρήστες, ωστόσο υπάρχουν αρκετοί τουρίστες που δεν βγάζουν φωτογραφίες, δεν γραφουν κριτικές και δεν μοιράζονται τις εμπειρίες τους online. Επομένως προκειται για μια μελέτη που επικεντρώνεται στους τουρίστες που δημιουργούν περιεχόμενο σε κοινωνικά δίκτυα τοποθεσίας. Επιπλέον, στα κοινωνικά δίκτυα υπάρχουν κενά σε δημογραφικά χαρακτηριστικά όπως η ηλικία, το επίπεδο μόρφωσης και το εισόδημα (Li et al. 2018). Έτσι, έχουν εντοπιστεί πεδία για επιπλέον ανάλυση. Επιπλέον μελέτη μπορεί να περιλαμβάνει

την ανάλυση περισσότερων και μεγαλύτερης ποικιλίας δεδομένων όπως για παράδειγμα δεδομένα που παράγονται από σένσορες (sensor generated data) για τον εντοπισμό νέων μοτίβων. Η μελέτη παρέχει περιγραφικά αποτελέσματα μέσω της ανάλυσης των μεγάλων δεδομένων π.χ. sentiment analysis, text mining και location intelligence, ενώ μελλοντικές έρευνες θα μπορούσαν να χρησιμοποιήσουν άλλες τεχνικές.

# 1

## **Big data Analytics: Applications, prospects and challenges**

In the era of the fourth industrial revolution (Industry 4.0), big data has major impact on businesses, since the revolution of networks, platforms, people and digital technology have changed the determinants of firms' innovation and competitiveness. An ongoing huge hype for big data has been gained from academics and professionals, since big data analytics lead to valuable knowledge and promotion of innovative activity of enterprises and organizations, transforming economies in local, national and international level. In that context, data science is defined as the collection of fundamental principles that promote information and knowledge gained from data. The techniques and applications that are used help to analyze critical data to support organizations in understanding their environment and in making better decisions on time.

Nowadays, the tremendous increase of data through the Internet of Things (continuous increase of connected devices, sensors and smartphones) has contributed to the rise of a “data-driven” era, where big data analytics are used in every sector (agriculture, health, energy and infrastructure, economics and insurance, sports, food and transportation) and every world economy. The growing expansion of available data is a recognized trend worldwide, while valuable knowledge arising from the information comes from data analysis processes.

In that context, the bulk of organizations are collecting, storing and analysing data for strategic business decisions leading to valuable knowledge. The ability to manage, analyse and act on data (“data-driven decision systems”) is very important to organizations and is characterized as a significant asset. The prospects of big data analytics are important and the benefits for data-driven organizations are significant determinants for competitiveness and innovation performance. However, there are considerable obstacles to adopting a data-driven approach and get valuable knowledge through big data.

## Introduction

Data is characterized as the *“lifeblood of decision-making and the raw material for accountability. Without high-quality data providing the right information on the right things at the right time, designing, monitoring and evaluating effective policies becomes almost impossible”* (United Nations, 2014). In that context, ongoing attention to data and data-driven approaches from academics and professionals exists, since the knowledge arising from data analysis processes leads to the promotion of innovative activity, transforming organizations, enterprises and national economies.

Nowadays, in the 4th Industrial revolution era, organizations and governments focus on the development of capabilities that provide knowledge extracted from large and complex data sets, commonly known as “big data”. Big data is a buzzword in the last years in the business and economics fields, since it plays an essential role in economic activity and has strengthened its role in creating economic value by enabling new ways to spur innovation and productivity growth. Hence, the ability of management, analysis and acting is significant under the context of knowledge-based capital (KBC) that is associated with digital information, innovative capacity and economic aspects (OECD, 2015).

In that era, many enterprises independent of size, from start-ups to large organizations, attempt to obtain data-driven culture struggling for competitive advantage against rivals. Enterprises aim to leverage data generated within organizations through their operations to gain valuable insights for better, faster and more accurate decisions in crucial business issues.

The advent of Web 2.0 allows users interacting with each other on social media platforms, enabled companies getting access to big amounts of data easier and cheaper. In addition, the appearance of Web 3.0 provides considerably increased opportunities for external data collection. Mobile devices (smartphones and tablets) facilitate companies to measure even more precisely, since those devices, both Internet and mobile-enabled, have the capability to promote e.g. highly mobile, location-aware and person-centered processes and transactions. This capability will continue offering unique research challenges and opportunities through the years (Chen et al., 2012).

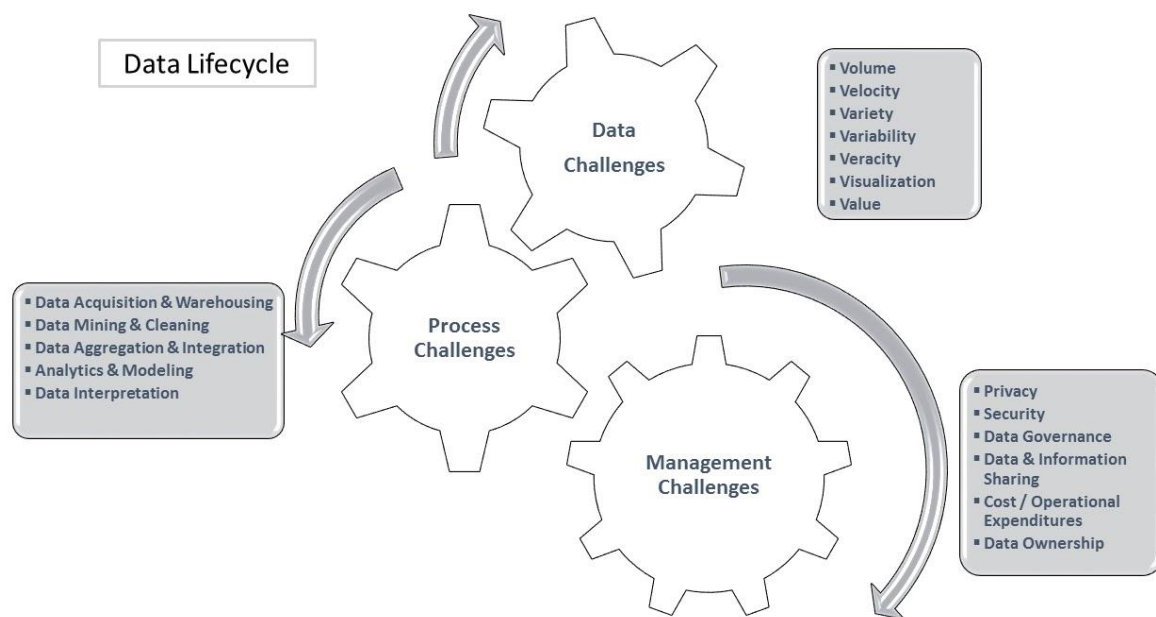
Digital enterprises like Google, Amazon and Facebook highlight the significance of big data, indicating the various ways that they can be used from supply chain to customer satisfaction highlighting the benefits of enterprises. Many enterprises started to benefit from those opportunities offered by the immense development of big data technologies. Today, enterprises in every industry sector and not limited to ICT sector, are focused on data exploitation to gain a competitive advantage, while managerial decisions rely on data-based analytics and less on the leader’s experience (Provost & Fawcett, 2013). Nonetheless, exploitation of big data needs people with skills and expertise who will be able to capture value from data insights providing significant knowledge to managers and decision-makers.

## Defining Big Data

The tremendous generation of data, expected to reach 180 zettabytes in 2025, gives data a leading role in change and growth of the 21st-century shaping a new “digital universe” with the transformation of markets and businesses (Economist, 2017). Digital information from complex and heterogeneous data coming from anywhere and at any time, introducing a new era, the era of “Big Data” (Sivarajah et al., 2017).

Big data refers to large datasets that are not able to be captured, stored, managed and analysed by typical software tools (Manyika et al., 2011). These data sets that are huge -not only in size- but also in heterogeneity and complexity (structured, semi-structured and unstructured data) include operational, transactional, sales, marketing and other data. In addition, big data includes data that comes in several formats including text, sound, video, image and more. This unstructured data is growing faster than structured and have captured 90% of all the data (Gantz & Reinsel, 2011). Therefore, new forms of processing capabilities are required for getting data insights that lead to better decision making.

On the data life cycle the challenges can be divided into three categories: data, process and management challenges (figure 1) (Sivarajah et al., 2017). Data challenges refer to the characteristics of big data including volume, velocity, variety and veracity. Process challenges are related to the techniques needed for big data acquisition, integration, transformation and analysis in order to gain insights from the big data. The data management challenges include challenges regarding data security, privacy, governance and cost/operational expenditures.



**Figure 1:** Challenges in Data Lifecycle

Big data can be characterized by the seven Vs: volume, variety, veracity, velocity, variability, visualization and value.

**Volume** refers to the large size of the datasets. It is a fact that Internet of Things (IoT) through the development and increase of connected smartphones, sensors and other devices, in combination with the rapidly developing Information and Communication Technologies (ICTs) including Artificial Intelligence (AI) have contributed to the tremendous generation of data (counting records, transactions, tables, files etc.). The speed of data is surpassing Moore's law and the volume of data generation introduced new measures for data storage i.e. exabytes, zettabytes and yottabytes.



**Variety** represents the increasing diversity of data generation sources and data formats. Web 3.0 leads to growth of web and social media networks leading to the generation of different types of data. From messages, updates, photos and videos that are posted on social media networks like Facebook or Twitter, SMS, GPS signals from smartphones, customers transactions in banking, e-business and retail, voice data in call centers etc. Many of the crucial sources of big data are comparatively novel, including mobile devices that supply huge streams of data that are connected with human behavior through their activities and locations; or web sources supplying data through comprising logs, click-streams and social media actions. Additionally, big data also differs in data types that are generated, thus big data consists of structured data (tables, records), unstructured data (text and voice), semi-structured data (XML, RSS feeds) and other data that is difficult to classify like data deriving from audio, video and other appliances.

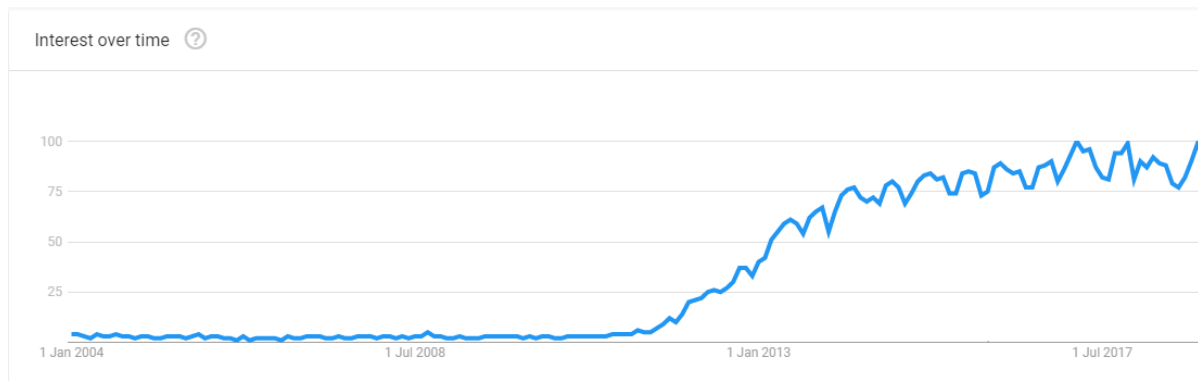
**Variability** is often confused with variety, but variability is related with rapid change of meaning. For instance, words in a text can have a different meaning according to context of a text, thus for an accurate sentiment analysis, algorithms need to find out the meaning (sentiment) of a word, taking into account the whole context.

**Velocity.** Big data is characterized by the high speed of data generation. Data generated by connected devices and web arrives in enterprises in real-time. This speed is extremely significant for enterprises in taking various actions that enable them to be more agile and gaining competitive advantage against competitors. Despite the fact that some enterprises have already exploited big data (click-streams data) to offer their customers purchase recommendations, nowadays enterprises through big data analytics have the ability to analyze and understand data taking actions in real-time.

**Veracity** of data refers to data reliability and accuracy. The data collection has data that are not clean and accurate, thus data veracity refers to the data uncertainty and the level of reliability correlated with some type of data.

**Visualization.** Data visualization is the science of visual representation of data and information. It presents quantitative and qualitative information in some schematic form, indicating patterns, trends, anomalies, constancy, variation, in ways that cannot be presented in other forms like text and tables (Friendly, 2008).

The leverage of big data can provide valuable knowledge and thus the value offered by the data analysis process can benefit enterprises, organizations, communities and consumers. Enterprises that overcome challenges and exploit big data efficiently have more precise information and are able to create new knowledge by which they can improve their strategy and business operations regarding well-defined targets like productivity, financial performance and market value (Power, 2008), while big data plays a major role in digital transformation of enterprises introducing innovations. Therefore, an increasing interest in exploitation of big data among enterprises and organizations exists.



**Figure 2: Big data trend<sup>1</sup>**

The economic benefits of big data in UK private and public sector businesses will increase from £25.1 billion in 2011 to £216 billion in 2017 (Cebr, 2012). Big data can provide more value in enterprises in various ways and is able to enhance productivity and competitiveness of enterprises. Big data is referred to the continuous growth of data and technologies that are necessary for collection, storage, management and analysis of data.

The way of thinking about businesses has changed with big data, since it changes major elements of organizations and not only management (figure 2). Big data can be a key resource for enterprises obtaining new knowledge, added value and fostering new products, processes and markets, thus data is characterized as an asset from enterprises' executives indicating the significance of data-driven approach within enterprises (Microsoft News Centre Europe, 2016). Enterprises gathered data for ages, however, nowadays more and more enterprises are analyzing the data instead of just keeping them. Hence, data-driven enterprises perform better in financial and operational terms, 5% more productive and 6% more profitable than no data-driven, gaining significant competitive precedence against their competitors (McAfee & Brynjolfsson, 2012).

## Big Data Analytics

The analysis of large datasets in enterprises, the term of big data analytics is associated with data science, business intelligence and business analytics. Data science is defined as a collection of fundamental principles that promotes taking information and knowledge from data (Provost & Fawcett, 2013).

Over the last years, data-driven approaches like Business Intelligence (BI) and Business Analytics are characterized indispensable to operating enterprises. BI is defined as the methodologies, systems and applications for collecting, preparing and analyzing data to provide information helping decision makers. In other words, BI systems are data-driven decision making systems (Burstein & Holsapple, 2008), while Business Analytics are the techniques, technologies, systems and applications that are used to analyze critical business data in order to help them to understand their business environment and take business decisions on time. The power of Business Analytics is to streamline vast amounts of data to enhance its value, while BI mainly concentrates on historical data in graphs and data table reports as a way to provide answers to queries without streamlining data and enhancing its value.

<sup>1</sup> <https://g.co/trends/CuxGn>

Business Analytics was commenced to outline the principal analytical element in BI in the late 2000's. Afterwards, the terms of big data and big data analytics have been utilised to describe analytical techniques for data- sets that are so large and complex, that need advanced data storage, management, analysis and visualization technologies. In that rapidly growing environment, the velocity of data makes the conversion of data into valuable knowledge quickly a necessity. The differences between conventional analytics and fast analytics with Big data are in analytics characteristics (type, objective and method), data characteristics (type, age/flow, volume) and primary objective (table 1) (Larson & Chang, 2016; Davenport, 2014).

	<b>Conventional Analytics</b>	<b>Big Data analytics</b>
<b>Analytics Type</b>	Descriptive, Predictive	Predictive, Prescriptive
<b>Analysis Methods</b>	Hypothesis-based	Machine learning
<b>Primary objective</b>	Internal Decision Support & Performance Management	Business Processes driver and Data-driven Products
<b>Data Type</b>	Structured & Defined (Formatted in rows & columns)	Unstructured & Undefined (Unstructured formats)
<b>Data Age/Flow</b>	> 24h Static pool of data	< Min Constant flow of data
<b>Data Volume</b>	Tens of terabytes or less	100 terabytes to petabytes

**Table 1:** Conventional and Big Data analytics

The development of the Internet and later on the connectivity coming from the web has contributed in the increase of the volume and speed of data. Since the early 2000's, Internet and Web technologies have been offering unique data collection and analysis for enterprises. Web 1.0 systems enabled enterprises to establish a web presence and offer their products/services online interacting with their customers. Web 2.0 systems, including the introduction of social media networks like Facebook, provide enterprises with more data about enterprises, products and customers. The ongoing increase of mobile devices against the number of computers introduced a new era of business analytics, including the analysis of user-generated content by social media channels. Mobile devices have the capability to promote e.g. highly mobile, location-aware and person-centered processes and transactions. Therefore, Data-driven decision making is based on data coming from all the sources of enterprises, while predictions and machine learning are based on traditional data and new innovative sources like IoT and AI.

Data analysis is the process of inspecting, cleaning, transforming and modelling data gaining useful information for suggestions and support in decision-making. It has multiple facets and approaches, encompassing diverse techniques under a variety of names, in different business, science and social science schemes, while "Big Data Analytics" refers to advanced analytic techniques, considering large and various types of datasets to examine and extract knowledge from big data, constituting a sub-

process in gaining insights from big data process. Using advanced technologies, Big Data Analytics (BDA) includes data management, open-source programming like Hadoop, statistical analysis like sentiment and time-series analysis, visualization tools that help structure and connect data to uncover hidden patterns, undiscovered correlations and other actionable insights.

The process of BDA is a resource for strategic decisions leading to significant improvements in operations performance, new revenue streams and competitiveness against rivals. In that context, the process of getting insights from big data can be divided into two phases: data management and data analysis (figure 3). Data management is related to the processes and technologies for data generation, storage, mining and preparation for analysis, while data analysis refers to the methods and techniques for analysis and interpretation of the insights coming from big data (Gandomi & Haider, 2015).



**Figure 3:** Process of leveraging big data

Analytics can be divided into four categories, ranging from descriptive and diagnostic analytics to the more advanced predictive and prescriptive analytics.

**Descriptive analytics**, based on historical and current data, is a significant source of insights about what happened in the past and the correlations between various determinants identifying patterns using statistical measures like mean, range and standard deviation.

Descriptive analytics using techniques like online analytical processing (OLAP) exploits knowledge from past experience to provide answers in what's happening in the organizations. Common examples of descriptive analytics include data visualization, dashboards, reports, charts and graphs presenting key metrics of enterprises including sales, orders, customers, financial performance etc.

**Diagnostic analytics** based also in historical data provide insights about the root-cause of some outcomes of the past. Thus, organizations can take better decisions avoiding errors and negative results of the past.

**Predictive analytics** is about forecasting and providing an estimation for the probability of a future result, defining opportunities or risks in the future. Using various techniques including data mining,

data modelling and machine learning, the implementation of predictive analytics is significant for any organization's segment. One of the most known applications of that type of analytics is the prediction of customer behavior, determining operations, marketing and preventing risk. Using historical and other available data, predictive analytics are able to uncover patterns and identify relationships in data that can be used for forecasting (Gandomi & Haider, 2015). Predictive analytics in the digital era is a significant weapon for organizations in the competitive race. Therefore, organizations exploiting predictive analytics can identify future trends and patterns, presenting innovative products/services and innovations in their business models.

**Prescriptive analytics** provide a forecasting of the impact of future actions before they are taken, answering "what might happen" as outcome of the organization's actions. Therefore, the decision-making is improved taking under consideration the prediction of future outcomes. Prescriptive analytics using high level modelling tools is able to contribute remarkably to the performance and efficiency of organizations, through smarter and faster decisions with lower cost and risk and identifying optimal solutions for resource allocation (IBM, 2017)

The advanced predictive and prescriptive analytics can play a crucial role in efficient strategic decision making dealing with significant problems of organizations like design and development of products/services, supply chain formation etc (Demirkan & Delen, 2013).

### **Big Data Analytics Applications**

Nowadays, as the growing generation of available data is a recognized trend across enterprises, countries and market segments, the majority of enterprises regardless industry is collecting, storing and analyzing data in order to capture value. Digital economy through the tremendous use of internet and digital services has transformed almost all the industry sectors, including agriculture and manufacturing, to more service-centered (Lodefalk, 2013). There are many and different sectors, like e-commerce, politics, science & technology, health, government services etc. where big data analytics are applied. Data-driven companies from various industries clarify the power of big data, making more accurate predictions leading on better decisions.

The large streams of data generated everyday need better infrastructures in order to be captured, stored and analyzed. A market with a wide supply of new products and tools designed to cover all the needs of big data has been created and it is developing rapidly (Davenport et al., 2012). There is a wide variety of analytic tools that can be used to perform BDA, among others on the basis of SQL queries, statistical analysis, data mining, fast clustering, natural language processing, text analytics, data visualization and artificial intelligence (AI). These techniques and tools provide easily and rapidly exploitation of big data.

The knowledge derived from exploitation of big data provides enterprises added value through new ways of productivity, growth, innovation and consumer surplus (Manyika et al., 2011), thus big data becomes a major determinant of competitiveness and enterprises are in need of data analysis capacity to exploit the full potential of data.

Enterprises that learn to capitalize big data utilizing real-time information coming from various sources like sensors, connected devices etc. can understand in more detail their environment and define new trends, create new and innovative products/services, respond quickly to changes and optimize their

marketing actions. The leverage of big data is able to contribute to the efficient resources' allocation and supervision, waste reduction, facilitation of new insights and higher level of transparency in different sections of enterprises from production to sales.

Therefore, BDA applications in almost every business sector exist. Applications also in politics and e-government, science and technology, security and safety, smart health and well-being exist (Chen et al., 2012). In addition, there are plenty and various types of big data applications among enterprises and industry sectors. BDA can be employed in e-commerce and marketing applications like online advertising and cross-selling, while it helps enterprises to analyze customer behavior in shaping 360-degree customer profile for implementation of targeted and optimized marketing actions to impact customer acquisition and satisfaction. It offers better understanding of customers' behavior and preferences and thus improve customer service.

Some examples of the ways BDA are exploited showing the significance of analytics in various themes (table 2) (Baesens, 2014):

<b>Marketing</b>	Market basket analysis	Recommendation systems	Customer Intelligence	Retention modeling	Customer Churn prediction
<b>Processes</b>	Supply chain analytics	Demand and supply forecasting	Business Processes analytics	HR analytics	
<b>Government</b>	Fraud detection	Terrorism Detection	Tax avoidance	Cost reduction	Social security
<b>Risk Management</b>	Credit risk modeling	Market risk modeling	Fraud detection		
<b>Web and Social media</b>	Web analytics	Social media analytics	Multivariate testing		

**Table 2:** Big Data Analytics exploitation examples

Enterprises and organizations collect large amounts of security-relevant data such as software application events, network events, people's action events. The generation of data coming from these actions are increasing rapidly per day as organizations enable logging in more sources, running more software programs, have more working employees and move to cloud solutions. Unfortunately, the volume and variety of security data quickly become overwhelming and existing analytical techniques cannot work efficiently and trustworthily. BDA applications become part of security management and monitoring, since it contributes to cleaning, preparation and analysis of various complex and heterogeneous datasets efficiently (Cárdenas et al., 2013). One of the most common uses of BDA is fraud detection, thus financial institutions, governments and phone companies use big data technologies to eliminate risk and enhance their efficacy.

In addition, BDA is widely applied in supply chain and logistics operations playing a significant role in developing supply chain strategies and supply chain operations management. BDA can support decision making through the understanding of changes in marketing conditions, identification of supply chain risks and exploiting supply chain capabilities to model innovative supply chain strategies, thereby improving the flexibility and profitability of supply chain. BDA contributes also in decision making at operational level, since it measures and analyses supply chain performance taking into account demand

planning, supplies, production, inventory and logistics. It thus improves the efficiency of operations, measures supply chain performance, reduces process alterability and contributes to the implementation of the best supply chain strategies at operational level (Wang et al., 2016).

Talking about digital and data-driven enterprises, the firsts coming in mind are Google, Amazon, Apple and Facebook. Amazon that was born digital, exploited big data achieving to disrupt traditional book market and became the leader in digital shopping. Another example of a famous born-digital firm is Google that harness data from engine search to digital marketing in order to provide and personalize search to its users, while Google and Facebook collect data providing opportunities for personalized and customized marketing.

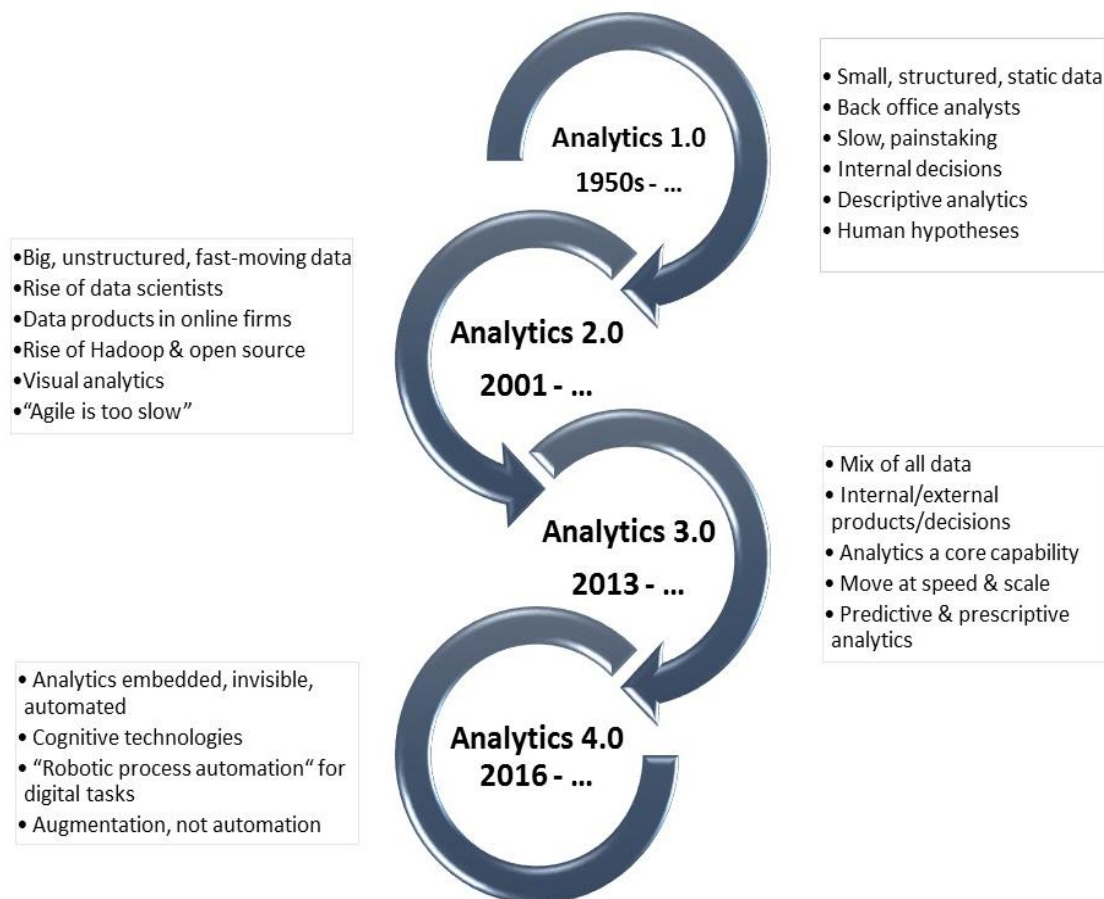
Nevertheless, traditional non-technological enterprises are also attempting to gain data-driven benefits. General Electric (GE) has developed a cloud-based platform for Industrial Internet application named “Predix” that provides real-time insights for engineers to schedule maintenance checks, improves machine efficiency and reduces downtime. GE this way provided new service value propositions in the conservative market of the oil and gas industry, while it faces its most pressing challenges: improving assets and operations productivity and eliminating the cost of tacit knowledge from an aging workforce (Winig, 2016).

Walmart and other major retailers using BDA in the entire business process, from supply-chain management to marketing, gained benefits from data. Applications of BDA are everywhere and not only in digital sectors, but also in no web-based sectors including manufacturing, agriculture, health care, energy, traveling and others. In healthcare sectors various applications of BDA exist, from quality of treatment services and cost efficiency of hospitals to improvement and predictions of patient health condition. In traveling and retail, BDA applications are able to provide customer intelligence through web and social media analytics, thus enterprises can offer personalized products/services. Additionally, in energy management the majority of the enterprises use data analytics to track and control devices achieving a more efficient energy management without services deviation.

## **Big Data Analytics Prospects**

Analytics in decision making procedure is not something new, since business analytics appeared as early as in the mid-1950s - Analytics 1.0 era - with the advent of tools that were able to generate and capture larger amounts of data in enterprises data warehouses and discover patterns more quickly than human minds with business intelligence tools. In that first era, managers gained a data-based comprehension going beyond intuition in decision making. Until mid-2000s, the rapid growth of data generation and the arrival of big data have signalled a new era - Analytics 2.0 -, where enterprises have the opportunity to leverage that data with new more powerful tools. The need of new innovative technologies appeared and enterprises moved quickly to acquire the necessary capabilities and knowledge for gaining insights from big data, with the major difference between eras being in skills required for data analysis (Davenport 2013). In the next era, analytics is an integral part of enterprises supporting decision making and enterprises move to creation of analytics-based products/services. Moving ahead, the next era - Analytics 3.0 or “data economy” -, is characterized by the tremendous increase of data generation coming from the growth of Internet of Things (IoT) with 8.4 billion connected devices in 2017 globally and 20.4 billion by 2020 (Gartner, 2017)

The most recent era -Analytics 4.0-, includes cognitive technologies including machine learning, where actions and decision making are shifted to augmentation with dynamic machine automation. The main characteristics of all these eras are appeared in figure 4 (Davenport, 2016).



**Figure 4:** The evolution of Analytics eras

In the current era of analytics, the emerging new technologies will increase the generation of data, thus enterprises and organizations have to face technical challenges in order to have access to more and better data. The worldwide revenues of big data and business analytics (BDA) will be more than \$203 billion in 2020 and banking, manufacturing, government and professional services will be the top industries in BDA investments according to International Data Corporation (IDC) (Forrest, 2016).

Therefore, enterprises should focus on capturing value from data using analytical techniques and tools. BDA can help enterprises to examine trends and discover new ones for gaining competitive advantage, introducing new and improved products. Among others, data visualization and process simulation, text and voice analytics, social media analysis, predictive and prescriptive techniques can provide valuable knowledge to enterprises, while they are able to make insights more transparent and impact any enterprise's section.

Data science and big data technologies - techniques promote data-driven decision making and thus contribute in better enterprise's performance, since the ultimate goal of data science is the improvement of decision making. Therefore, whether organizations couldn't capture value from applying data-driven decision making as their strategy, they have failed (Provost & Fawcett, 2013). There is evidence that



data-driven decision making contributes significantly and positively to enterprise's performance in terms of productivity and profitability (Brynjolfsson et al., 2011) . Data-driven approach can provide great opportunities for gaining competitive advantage, as measuring and managing more precisely business analytics can enable organizations to make better predictions and smarter decisions also to target more-effective interventions (McAfee & Brynjolfsson, 2012).

Moving to a whole new era in data analytics, organizations and enterprises are exploring new innovative strategies and techniques to remain competitive in their market. Using BDA help them to introduce new and/or improved products/services, manage more efficiently their supply chains and processes, eliminate risk through fraud detection and security improvement and exploit customer intelligence.

Applications of BDA can provide several advantages in organizations and enterprises that have an efficient data-driven approach. Big data analysis is able to provide in-depth knowledge about the different departments of an organization and thus using big data analytics for prediction making will contribute to increased performance and higher returns on investments with lower cost and risk, while more transparency is achieved.

Some of the prospects of big data analytics are:

- ❑ Gaining insights from big data analytics of all the departments of an organization to develop a comprehensive business strategy, or the entire organization. This strategy will be able to contribute to a higher level of productivity and efficiency, within the departments, but also in the whole organization with cost reduction and elimination of processes.
- ❑ Organizations can exploit more artificial intelligence (AI) technologies that are able to reinvent organizations in various ways. However, organizations should develop automations and structured analytics, before they move on the adoption of advanced AI. The integration of structured and unstructured data analytics with AI systems makes it possible to examine, explain and predict customer preferences and behavior (Harrison & O'Neill, 2017)

Data-driven innovation (DDI) relying on the knowledge-based capital, refers to innovations arising from data-driven decision processes (OECD, 2015) that lead to the discovery of new and disruptive business models, the enhancement of customer intelligence (Ryan, 2016) and the introduction of new/improved products or services. The potential of data-driven innovation big data in UK private and public sector businesses will lead to £24,1 billion contribution to the UK economy during 2012-2017 (Cebr, 2012).

Real- time analytics is a big trend enterprise need to pay attention to in the near future. Despite the challenges and issues that are addressed, it is proven that analytics-driven management has significant implications for enterprises, whether they are looking for growth, efficiency or competitive differentiation. Therefore, Big data analytics have seemingly unlimited potential to help an enterprise to grow and reveal its data potential.

The rapid growth of the demand for data analytics in combination with the lack of talent lead on collaborations and initiatives between academia and industry in order to bridge the talent gap. In that context, many universities are preparing and starting academic courses related with data science. In addition, companies realizing the potential of big data, provide training to their employees. Recently Airbnb started its own internal university called "Data University" to democratize data science and help to drive data-informed decision making.

There are different expectations from enterprises regarding big data analytics. Organizational leaders want to exploit analytics to be smarter and innovative like never before, while senior executives want to use data-driven decision making for their efficient operations (LaValle et al., 2010). Managers using a data-driven decision system (DSS), have access to historical and new data supporting them to gain insights for organization processes and resources' performance. DDS are significant not only for global organizations but also for small and medium organizations that can exploit them to their benefit (Power, 2008).

## **Big data analytics challenges and barriers**

The major challenges in adopting big data analytics from enterprises are more managerial and cultural than associated with data and technology, while the main barriers are the lack of comprehension of how to utilize big data analytics to enhance the business and the lack of management spectrum from competing priorities (LaValle et al., 2010). Studies among different industry sectors indicate that organizations use less than half of their structured data in decision making process, while less than 1% of their unstructured data is analyzed or exploited, 70% of employees have access to data they should not and 80% of analysts' time is to discover and prepare data (Dallemlue & Davenport, 2017).

**Leadership.** According to management challenges, enterprises that achieve to be successful in the data-driven era have leadership teams that determine aims, modulate achievements and ask the right questions to be answered by data insights. Despite its technological approach, the power of big data cannot be exploited without vision or human insight. Therefore, leaders of enterprises with vision and ability of revealing the future trends and opportunities, will have the ability to act innovative, motivate their teams work efficiently to achieve their targets.

**Talent management.** Enterprises in order to leverage data through big data analytics need human capital with high level of technical skills to use and exploit these systems in order to achieve exploitable knowledge for end users, mainly C-suite. People's specific skills include statistics, big data mining, master visualization tools, business-oriented mindset and machine learning. These are required to get valuable insights from big data contributing in decision making procedure (McAfee & Brynjolfsson, 2012). However, these people (data scientists, data analysts etc.) are extremely difficult to be found and thus demand for them is high. There is a challenge in finding data scientists with skills both in analytics and in domain knowledge. In general, there are existing fewer data scientists than needed (Waller & Fawcett, 2013).

**Decision making procedure.** In efficient enterprises, decision makers and knowledge derived from data exploitation are in the same place. Nonetheless, it is difficult for decision makers to handle huge amounts of data. Therefore, there is need of decision-makers having problem-solving skills and the ability to provide answers to problems with the right data or cooperation of different people in problem solving through leveraging big data (McAfee & Brynjolfsson, 2012).

**Decision making Quality.** The quality of decision making adopting a data-driven approach is a significant factor for taking advantage of the possibilities that big data analytics are offering. In that context, ensuring decision making quality is correlated with factors like data quality of big data sources, big data analytics capabilities, staff and decision-maker quality (Janssen et al., 2017). The accuracy of

big data sources is significant in providing high value in decision making eliminating wrong actions, while big data analytics capabilities are related with the utilization of the right techniques and tools from specialists with knowledge of big data analytics.

**Data-driven culture.** Another significant challenge for adopting data-driven approach is enterprise culture. The basis in obtaining data-driven culture is the capabilities to quickly condense, analyze and distribute crucial business information to decision makers. That basis is extremely significant for enhancement of business performance, while development and improvement of that capabilities empower enterprises leading to improvements in all business segments and higher returns on investments. In that context, enterprises have to adopt data-driven decision making in all issues and stop acting solely on hunches and instinct. Therefore, management must fully understand the significance of getting insights from data exploitation. In addition, for a data-driven enterprise, people who are involved in the process of data-driven decision making need to meet some requirements. Managers should be able to manage efficient data-analytics teams and projects, while marketers should be able to understand metrics and analytics in order to manage efficiently marketing activities.

**New Technology utilization.** Many enterprises conceiving the power of data, have developed technology skills in business intelligence and/or data warehousing, but the technologies of big data analytics are different and new. Therefore, enterprises have to utilize techniques and technologies that are available in order to capture value from big data. As these technologies are evolving rapidly, IT departments should be able to develop their capacity and be up-to date to that ongoing innovation. For instance, problems will emerge when database software does not support big data analytics options.

**Data privacy.** The collection of data is considered to be deeply suspicious by many people. For them, big data is an invasion of their privacy. Marketers are struggling with consumers' perception of data, as 71% of them believe that brands with access to their personal data are using it unethically, while the 58% of them have not used any digital service due to privacy concerns that lead to decision-making about the applications they download, the email addresses they share and the social media sites to use in order to connect to other websites (Chahal, 2016). Therefore, enterprises need to use safeguards in order to ensure that data is not used to violate the customers' personal privacy (Manyika et al., 2011). In that direction, data policies including privacy, security, intellectual property and liability issues, should be addressed in order to exploit big data value.

## Conclusions

The growth of the Internet with the beginning of Web 2.0 era enabled companies getting access to big amounts of data easier and cheaper, while the opportunities for external data collection have even increased with the appearance of the Web 3.0. Enterprises and organizations from all sectors began to focus on data exploitation for gaining competitive advantage.

Nowadays, the big data era has quietly settled down on almost every company, because they realized that data-driven decisions tend to be better and more accurate decisions. However, that many companies in several industries are applying business analytics including big data analytics, it doesn't mean that they all take benefit from it by getting valuable insights and real business value from the available data.

Becoming a data-driven company is more than using analytical techniques and tools. The companies need to hire people equipped with systematic thinking to promote the success in data-driven decision

making. Success in the data-oriented business environment today includes being able to think data-analytically. Since the amount of data is continuously growing, domain knowledge and analysis can't be considered as separate areas. Both academic and applied professionals of the companies are expected to have the analytical skills to understand business processes.

Employees, who don't have a basic understanding of data-analytic thinking, do not really know how will be easier for them to find suitable solutions for the weaknesses of the concerning process step. But to be able to perform data-driven, organizations have to face some challenges, both managerial and technical.

Big data is not just about data volume, but also about variety and velocity. Big data analytics have the ability to help enterprises understand their business environments, their customers' behavior and needs and their competitors' activities. Thanks to big data analytics enterprises are able to form their products and actions in order to fulfil customers' needs and innovate against rivals through better predictions and smarter decisions on the basis of evidence instead of intuition. Organizations that achieve to manage the challenges and adopt a data-driven culture, they can expect good prospects. There is strong evidence that business performance can be improved via data-driven decision making, big data technologies analytical tools and techniques on big data. As more companies learn the essential skills of using big data and how to engage with current technologies, which are continuously developing, may soon stand out from their competitors and have a decisive competitive advantage.

# **Knowledge Extraction through Location & Big Data Analytics: the case of Crete**

## **Introduction**

Nowadays, tourists generate massive volumes of data (big data) during their visit to an urban destination. However, there is little knowledge of their spatial activity and perceptions. Enterprises and organizations in hospitality and tourism are able to exploit actual behavioral data - perceptions derived from big data generated in real-time from online data sources in contrast to traditional customer surveys based on questionnaires (Heerschap et al., 2014). Nonetheless, there is a knowledge gap in the leverage of big data in hospitality and tourism, thus the implementation of innovative big data analytics applications in these domains is needed (Mariani et al., 2018).

In this context, this chapter contributes to Big Data analytics applications in the tourism and hospitality industry by proposing a novel approach to leverage massive unstructured location-based social media big data and provide valuable knowledge about spatial, temporal and demographic characteristics of tourists. While other observational data analysis methods for social media data have been presented, this study presents an innovative approach integrating big data techniques, location intelligence and social media transforming tourist experiences into valuable assets (new knowledge extraction) for quicker and more efficient decision making by tourism SMEs and stakeholders. More specifically, the approach introduces the combination of textual and photo analytics with data derived from media sharing and textual social networks, introduces social big data analytics such as social engagement, sentiment analysis, topic/label detection combined with spatio-temporal features to provide more insights about tourist destinations.

To demonstrate how the proposed approach can provide useful practical implications for applications in location management, tourist experience improvement, and promotion, the case study of Crete is used. A dataset composed of user-generated text and photos shared in location-based social media networks for the most popular cities of Crete island (Greece), Heraklion and Chania is used.

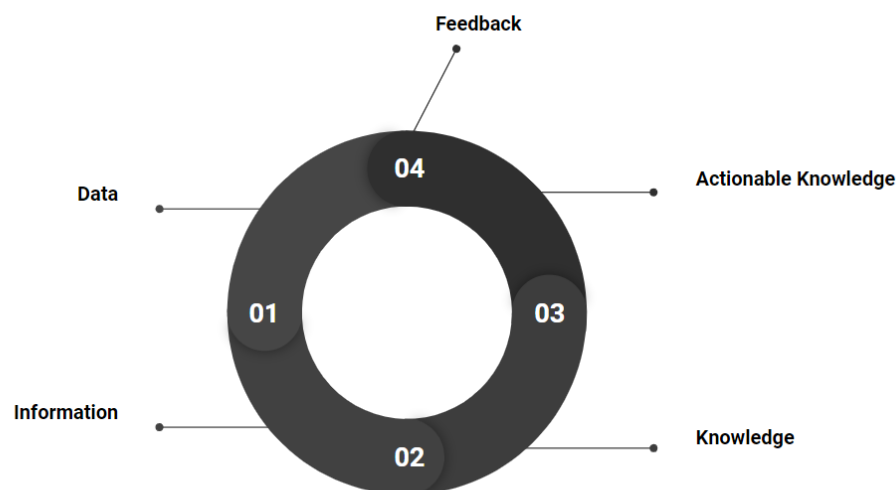
Research findings demonstrate how this novel approach of location and big data analytics in contrast to traditional tourist surveys and conventional spatio-temporal data can provide new and valuable knowledge. Implications arising from the study are significant assets for tourism SMEs, DMOs and other tourism stakeholders in the search of innovative marketing strategies for demonstrating the added value of destination and strengthening destination branding gaining a competitive advantage against other rival tourist destinations.

## Literature/ Study background - framework

Nowadays, the need for enterprises and organizations regardless of industry and size (from micro to large) to innovate is essential for their survival and evolution. Due to digital transformation, innovation is characterized as “knowledge-based” and is coming to the fore. Enterprises, especially small and medium enterprises (SMEs), can exploit knowledge fostering innovation in order to gain better performance and a competitive advantage in the globalized market. To achieve that, it is fundamental to gather and analyze the appropriate data for knowledge extraction that is needed for making actionable decisions.

Data is considered as the raw material of the 21st century and its conversion into information is significant for the quality of decision making (Elgendy & Elragal, 2016). In this digitized era, data is considered a catalyst for innovation, productivity and competitiveness. Big data, the process that includes data gathering, management and analysis to provide knowledge revealing hidden patterns, can lead to a value in many ways such as cost reduction, adaptive learning, and artificial intelligence, efficient processes, improved business decision making, new processes, and innovative activities.

Therefore, the potential of data-driven decision making is widely recognized. The data-driven process to actionable knowledge presented below (figure 5), where data is a process of collection and interpretation, whilst information provokes knowledge to decision making for action, actionable knowledge (Bumblauskas et al., 2017).



**Figure 5:** Data to knowledge process (Bumblauskas et al. 2017)

When data is transformed into actionable knowledge, it becomes valuable for organizations and enterprises in the strategic and operational decision processes. Data-driven decision making declares *“the practice of basing decisions on the analysis of data rather than purely on intuition”* (Provost & Fawcett, 2013; Brynjolfsson et al., 2011).

The history of using and analyzing data for enterprises has been developed over the years (table 3) (Davenport, 2014; Power, 2007).

Term	Time frame	Specific meaning
Decision support	1970-1985	Use of data analysis to support decision making
Executive support	1980-1990	Focus on data analysis for decisions by senior executives
Online Analytical Processing (OLAP)	1990-2000	Software for analyzing multidimensional data tables
Business Intelligence	1989-2005	Tools to support data-driven decisions, with emphasis on reporting
Analytics	2005-2010	Focus on statistical and mathematical analysis for decisions
Big Data	2010-present	Focus on very large, unstructured, fast-moving data

**Table 3:** Terminology for using and analyzing data

In recent years, the rapidly growing power of decision models has concentrated the attention of C-suite, since the increasingly advanced analytics algorithms and techniques such as clustering, neural networks with a huge amount of data opened up new ways for the improvement of decision making and business performance (Rosenzweig, 2014). Advanced data analytical techniques contribute to decision making by way of analyzing and inspecting data for hidden meaning. Due to the vast volumes of data that are available mastering them and turning them into valuable information is the challenge (Poletto et al., 2015). The data management and exploitation for quick insights can provide a better understanding of the markets, hence a competitive advantage. Data visualizations and advanced analytics can help to quickly identify patterns, trends, and opportunities that would not be efficiently identifiable by using legacy methods. This is where automation, mathematics, and art come together to maximize the value of computer processing power.

In this era, business innovation<sup>2</sup> goes far beyond R&D, while the value creation can come from different types of innovation (OECD, 2005). Big data is able to change competition by the transformation of processes, development of business ecosystems and innovation facilitation (Manyika et al., 2011). There are several modern technologies such as Big data and IoT that contribute to the evolution of data related to business transformation. Hence, a business innovation derives from the exploitation of data and analytics is determined “data-driven innovation (DDI)” (OECD, 2015).

Data-driven innovation deals with the development of new or significantly improved products/services, processes or methods. Therefore, big data leverage for decision-making is of major concern. Crucial and significant business decisions can be made based on torrents of both internal and external data (Rao & Holt 2005). There are enterprises and organizations that already benefit from leveraging data, but the bulk of them focus on internal processes and costs optimization (Manyika et al., 2011).

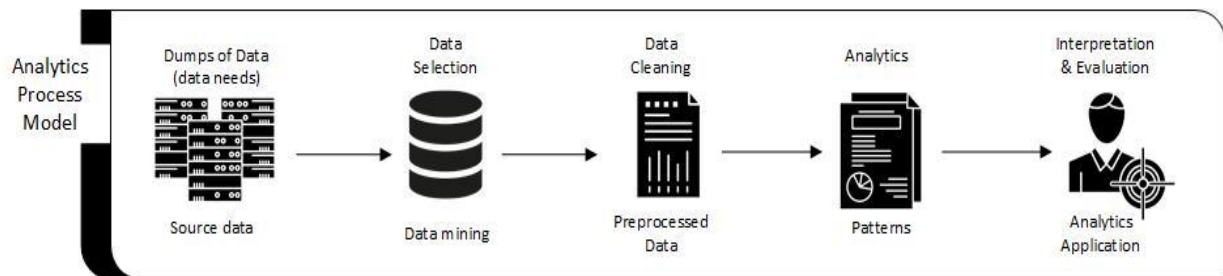
DDI can take various forms: data-driven products (new products/services developed based on data), data-intensive products (products with major ingredient data), data-driven processes (optimization of

---

<sup>2</sup> “A business innovation is a new or improved product or business process (or a combination thereof) that differs significantly from the firm's previous products or business processes and that has been introduced on the market or brought into use by the firm” (OECD & Eurostat 2018).

enterprise processes using data) and data-driven R&D processes (enhancement of R&D using data). (Zolnowski et al., 2014) examine the impact of data-driven innovation on enterprises and their business models focusing on international enterprises that implement successful data-driven innovations. They conclude that enterprises can improve their value propositions and processes in their business models fostering data-driven innovations. More specifically, patterns of cooperative value innovation and customer-centric value innovation enable a transformation from product to service-oriented offerings, while patterns of cooperative productivity improvement and company-centric productivity improvement enable an internal and external processes optimization and consequently better productivity.

The improvement in an enterprise's performance is caused by data-driven decision-making that in fact means better insights into the decision-making process in an enterprise (Sharma et al., 2014). The analytics process model (figure 6) starts with data gathering from several sources and followed by data cleaning and transformation, data analytics (patterns and insights) and concludes to interpretation and evaluation from decision makers. Data analytics is significant for an enterprise's strategy design and execution. As enterprises are becoming more analytical and embed data-driven decision making in their "DNA", they will realize significant benefits and value that are related with money (reduction of cost), time (faster and better decisions) (Davenport, 2014) and rate of return (Liebowitz, 2013).



**Figure 6:** The Analytics Process Model (Baesens, 2014)



## **Knowledge extraction through Big Data Analytics**

The benefits of data analytics to enterprises of all sizes are well documented and there is evidence about organizations' ability to improve their decision making, identify areas of cost reduction and profits gaining (Davenport, 2014; Brynjolfsson et al., 2011). Every enterprise is able to be data-driven for better decision making, however there are some challenges to be implemented by SMEs, thus, in general, it is easier to be implemented by large companies. For SMEs and start-ups, the principal benefit is the potential of rapid growth. Nowadays, the majority of SMEs managers' decision making is based on their meaning and feeling/intuition with historical results instead of being an evidence-based (data-driven) process. The tools for this "decision-making" process are usually software like Microsoft Excel. Although these tools are useful and simple, but the present limited options and prospects (Hiziroglu & Cebeci, 2013).

The adoption of big data analytics is getting more painful for smaller enterprises due to several factors. One of the most significant factors to become data-driven is IT structure enabling big data analytics. The structure consists of hardware, data processing, data analytics, data integration, data visualization, and reporting. SMEs are able to overcome these obstacles using cloud architectures for data storage and open source big data analytics projects (Coleman et al., 2016). Some of the benefits arising from using cloud-computing from SMEs are time and cost reduction, speed, elasticity, agility, flexibility and scalability (Hiziroglu & Cebeci, 2013). Recently, the most marketing tools are software as a service (SaaS) tools with the technology offered in the cloud enabling access to all enterprises regardless of size or financial resources. In that context, cloud is an affordable solution for SMEs, since it can solve the hardware and software issues. Nonetheless, the majority of SMEs do not exploit that solution, because of the lack of knowledge and understanding of cloud and security (Schaeffer & Olson, 2014).

Tech giants like Amazon, Google, and Facebook invest heavily in artificial intelligence (AI) and Machine Learning (ML) through investments internally and acquisition of start-ups. However, not only large enterprises can exploit AI. SMEs can also use AI to be transformed into data-driven organizations. The most SMEs cannot afford to hire a data scientist or a data science team in order to harvest the power of data to unlock its value, but there are self-service AI tools and applications (third parties) that can contribute in that direction. These applications can provide great solutions in significant parts of the organization such as customer relations - customer support, marketing, accounting, etc. In customer support, there is an increase of virtual customer assistants (VCA) such as chatbots that according to Gartner customer service operations will use from 2% in 2017 to 25% by 2020 (Gartner, 2018a). CRM platforms like Intercom, Salesforce and HubSpot are great choices for monitoring and improving enterprise's relationship with current and potential customers through different communication channels (social media, emails, telephone, etc.). Financial resources of SMEs are limited, thus spending a large amount of money in marketing efforts is hard. Therefore, SMEs attempt to find ways to optimize their marketing campaigns. For that purpose, there are AI tools that can contribute to better marketing results coming from implementing more effective marketing strategies such as email marketing and content marketing.

Small businesses need to be agile and innovative in order to grow, remain competitive and be successful. Therefore, quick and informed decisions are required. Thus, they need to maximize their limited resources and not make decisions based on intuition and small data. Many SMEs are overwhelmed by where to start, worried about cost and effort, while they are discouraged by stories of analytics failures. Since the leverage of big data is able for any enterprise regardless size, smaller enterprises have

advantages against larger ones since they are more flexible and without data siloed in various departments of the enterprise (Glass & Callahan, 2014).

In the past, the costs of data management (storage, process etc) and analysis were major barriers to big data exploitation, especially for SMEs. Nowadays, SMEs and start-ups are able to use a bulk of free or inexpensive resources to deal with big data. SMEs use open-source software for data management such as Hadoop, Hive and for data analysis like R, while they can use cloud computing resources like Microsoft Azure and Amazon Redshift with extremely low cost compared with the past (Davenport, 2014). In addition, infrastructure as a service (IaaS) enables allows the offering of computing resources, storage and networking capabilities that are owned and hosted by a service provider and offered to customers on-demand. The worldwide IaaS market grew 29.5% in 2017 to a total of \$23.5 billion, up from \$18.2 billion in 2016 with Amazon being in the first position of vendors in the IaaS market in 2017, followed by Microsoft, Alibaba, Google, and IBM (Gartner, 2018b).

### **Data-driven Innovation for SMEs**

In a rapidly changing “knowledge-based” global environment, enterprises competitive advantage depends on their capacity to access information, create valuable knowledge and be innovative. Innovation performance of enterprises especially for SMEs is characterized as a determinant of their survival, sustainability, and growth. However, innovation is not an easy process, but it’s complex because of the extremely competitive market conditions, continuous changes in technology and customer needs. In that context, leveraging “big data” for decision making can lead to innovation (data-driven innovation) and thus better performance.

Innovation and technologies including social media, data analytics, and business management software are significant factors for SME growth (Oxford Economics, 2013) whilst big data is, also, considered to be a major determinant in achieving that. The ability to analyze and predict customer and market demands and preferences is a great asset for SMEs and if it is applied properly, can contribute to making enterprises present increased productivity, flexibility and responsiveness, covering customer needs through better decision making (Sen et al., 2016).

The information influencing SMEs is progressively substantial; they are gathered continuously from different sources. For SMEs, imperatives of Big data can be found in internal and external environment, and organizational performance. The internal environment of a company refers to the generation of data from different departments or functions such as marketing and human resources departments. In the external environment, a tremendous amount of data that is generated to various parts of the business environment, while market competition has no geographic limits.

In that context, SMEs have to integrate external data in decision making to ensure their survival and competitiveness. According to Drucker: *“Business has only two functions: innovation and marketing”*. In the recent digitalized era, enterprises can adopt technology marketing stack - strategy bending software platforms such as customer relationship management systems, analytics tools, and data management systems, grasping streams of data that can contribute to more efficient marketing and creation of new products/services fulfilling customer needs and preferences. (Glass & Callahan, 2014). Adding new sources of data for marketing purposes including customers’ social media, mobile and geolocation data provide enhanced insights (Davenport, 2014). Therefore, big data can compromise

these various data types, help to the understanding of the environment and thus offer products and services that satisfy the various customer preferences (Mbassegue et al., 2016).

Social media is considered a resource for innovation and the development of new products, since it is a “vehicle” for the generation of customers insights, knowledge assessment, co-creation of ideas and concepts with users and also supports new product launches (Roberts et al., 2016). Therefore, knowledge discovery from social big data is receiving more and more attention from industry and academia, since big data analytics can uncover hidden opportunities, identify trends and patterns, problem areas and successes, that people may not be able to ascertain. There is evidence that huge valuable knowledge can be gained through social media networks for organizations’ strategic decisions leading to competitive advantage. An enterprise’s involvement in social media offers values for enhanced brand value (Gensler et al., 2013), sales growth (Kumar et al., 2013), customer loyalty and engagement (Zhang et al., 2017), innovation and new product development (Palacios-Marqués et al., 2015), knowledge sharing (Munar & Jacobsen, 2014), customer relationship management (CRM) (Rosman & Stuhura, 2013) and electronic word of mouth (eWoM) (Ladhari & Michaud, 2015). Enterprises can leverage data from WOM in order to improve the quality of services and products (Rhee et al., 2016). Using social media for the promotion for events and for communication actions can create trust and engagement within a huge number of stakeholders and customers (Vecchio et al. 2018a). Social media intelligence is able to allow a business to *“understand the behaviours driving the creation of online opinions with psychological and sociological perspective, evaluate the implications of behaviours on how we interpret social media and integrate them into global business strategy”* (Moe & Schweidel, 2014).

Adetunji & Carr (2016) using public Twitter data for an SME as a case study concluded that exploitation of social media data provides organizational knowledge that is related to the strategic competitive advantage of SMEs. The obstacles to leveraging big data from SMEs have been reduced significantly via the development of big data analytics frameworks. (Dittert et al., 2018) conducted a case study using a middle-sided oil-trader from southern Germany identifying that SMEs have to exploit big data analytics to ensure long-term competitiveness. SMEs present advantages against large enterprises including the manipulation and implementation of big data. The decision-making process is shorter due to the flat hierarchies. SMEs have the ability to implement a big data project more quickly and are more flexible in making the necessary adjustments such as project budget. Hence, SMEs take quicker and determined decisions gaining competitive advantages against rivals and large enterprises (Matthias et al., 2017). Nonetheless, SMEs do not leverage this opportunity in the level larger enterprises do, while due to the financial crisis, the bulk of SMEs faces significant obstacles to grow and be competitive.

## **SMEs, Location Intelligence & Tourism**

Another significant factor for SME success and its strategic competitiveness is the location (Galbraith et al., 2008), since SMEs of urban areas are more beneficial than SMEs of rural areas (Thulo, 2014). Many small companies in European Union face issues with the internationalization of their business, with only 25% of European SMEs export and even less export outside of the European Union.

Decision-making associated with the location is an everyday process. Namely, many individuals, social groups, legal entities, and companies make such decisions daily. Individuals relate them mostly to choosing a route to work, a place to socialize with friends, or going on a trip. But, for many years, especially under the influence of globalization, location-based aspects of different businesses are becoming more and more important.

It is one of the key factors of enterprise success, and its importance is growing in combination with technology, as well as various software and decision-making platforms. According to (Forbes, 2015), more than 80% of data collected, managed and used by organizations around the world has a location component. Thus, decisions concerning location play a great role in the outcomes and prosperity of a business. The envision of location data helps enterprises see relationships between customers, but also their operative data and geography (Moreno, 2015).

In this context, location data is so important for enterprises that can make vital decisions based on location information or using Location Intelligence (LI). Golfarelli et al. (2013) define LI as a set of tools and techniques which integrate a geographical dimension into business intelligence platforms, aimed at enhancing their capability of better monitoring and interpreting business events. On the other hand, Pitney Bowes, determines Location Intelligence as a capability which helps organizations make better decisions by integrating data quality, geo-demographics, and analytics with traditional mapping and geographic information systems.

LI is even more vital when it comes to SMEs since this technology helps with gaining competitiveness, especially on foreign markets. Exports also help with the increase of SMEs performance and reinforcement of sustainable growth. LI is a highly needed tool for gaining market competitiveness, whether it is in financial, retail, real estate, legal services or some other sector. Due to this, today, SMEs are showing indications of increased growth rates in market value when it comes to the greater adoption of LI. Research even shows that SMEs will comprise about 30% of market value and application of Location Intelligence, and will also have a growth rate advantage in contrast to large enterprises by the end of 2021. Moreover, SMEs benefit greatly from the application of Location Intelligence and Location Analytics in asset tracking and the optimization of scheduling, social media adaptation and so on (Prem, 2017). The potential of location analytics is high as its *“market size is expected to grow from USD 8.20 Billion in 2016 to USD 16.34 Billion by 2021, at a Compound Annual Growth Rate (CAGR) of 17.6% from 2016 to 2021”* (Markets & Markets, 2019).

Nonetheless, many enterprises still have not adopted the technology, because of the struggle with data quality and accuracy or obtaining data in real time and the battle with the derivation of disposable data from existing means. Additionally, many companies do not have the proper technology and/or employees to correctly execute analysis of the location-based data they collect. Therefore, the bulk of enterprises have a large amount of data from several sources. However, the strategy, enrichment, visualization, analysis or action does not consist of any analysis of that data.

LI provides enterprises several benefits such as Return on Investment (ROI) that is associated with the improvement of the value of existing Geographic Information System, Business Intelligence, and Data Warehouse, increased value and data quality that is achieved by Location Discovery and Location Analytics, increased quality of decisions - better decisions are made in a shorter period of time because of comprehensive and spatial vision on key business facts and data, and optimization of business process quality that is reached by the integration of location-based information or routing, mapping and geocoding.

Due to maps being remarkably befitting for the fundamental portrayal of extensive quantities of data, the conjunction of Location Intelligence and Big Data provides a bigger amount of spatial data and data analysis. Secondly, Location Intelligence contributes to the advanced and easier analysis of patterns, trends, and different potentials, as well as Mobile Business Intelligence (maps are utilized in the data analysis connected to the screen size of cell phones or tablet PCs). Thirdly, by applying Location Intelligence, SMEs can use social media to produce and analyze location-based data to offer their customers better service.

Due to smartphones being the most often used tool for searches, LI is becoming even more important for business success, since 71% of smartphone users favour ads customized to their location, it is more evident why 50% of brands are using LI to target their consumers, while the 50% of location-based searches performed via smartphones give an outcome in shop visits (Google ,2014). The ads also result in more customer purchases than other searches. This means that business strategies containing the use of LI are growing in importance. Strategies that emphasize location-based activities, such as marketing, allow small enterprises to gain a competitive advantage in the market.

Tourism, like any other economic sector, is greatly influenced by digitization and internet progress. Digitization and technology development have influenced significantly almost every economic sector, including tourism with the majority of actions from travellers and tourists being through social media. These activities generate an enormous hype of data, big data, containing valuable insights that are usually remained unexploitable. The majority of tourism processes and transactions (from trip planning, bookings to tourist feedback) are digital. The bulk of tourists and travellers use the web and social media for travel planning and acquiring trustworthy information for their travel destination (Yoo et al., 2016). Therefore, an enormous amount of data around customers is generated at tourism destinations identifying preferences and needs. However, these knowledge sources containing valuable insights remain unexploited (Höpken et al., 2013). Location-based social networks including Facebook, Foursquare, Instagram, Twitter, and Flickr have been characterized as major tools for communicating, spreading ideas, information and knowledge among their users, while each has its own peculiarities and features, but all support the addition of a location in a post.

Tourism is a service-based sector that depends on visitor satisfaction; thus, visitor's emotions and feedback are of utmost importance. Social media plays a significant role as an information source for potential visitors/tourists to collect accurate information for destination and services choice, and for enterprises/organizations that are able to exploit the same information for progressing their marketing strategies (Floris & Campagna, 2014). In addition, using location-based network data is extremely beneficial for many parties including citizens and governments and it can be life-critical. Web 2.0 applications and tools transforming the role and behaviour of travellers, make tourism stakeholders reorganize their operations and business models based on these data (Christou, 2016).

Nowadays, more and more businesses and organizations in the tourism industry use social media analytics, as they focus on gathering and analyzing social big data for creating value in marketing by developing tools such as recommendation systems. This strategy is vital for every business and organization, since it allows finding out what people are thinking about their products and services, their brand or their competitors, gaining significant information about how people behave and how they make their choices, specifying new market trends and opportunities for competitiveness sustainability, creating contextualized offerings to their customers, predicting their demands and making life-critical decisions in real-time.

## Methodology

The traditional analytical tools explore historical data, while the advanced analytics techniques are able to uncover hidden patterns and predict the future actions enabling organizations to take the necessary decisions for changes and actions to avoid crises or adopt the success. Data analytics techniques have been established in organizations and turned into one of the fastest-growing segments of Business Intelligence<sup>3</sup>.

Advanced data analytics techniques like text mining, sentiment analysis, forecasting, cluster analysis etc. can address more complex issues through knowledge extraction at deeper level. Advanced analytics utilizing multiple methods that enable the analysis of structured or unstructured big data for acquiring actionable results (Kaisler et al., 2014) and discovering new opportunities for development (Garcia et al., 2017).

Analysis of customer's behavior can provide valuable insights that are able to change products/services that are offered by an enterprise (Glass & Callahan, 2014). An important quantity of unstructured content - data generated and collected through social media networks is text. There are many text mining/ information extraction techniques including Natural Language Processing (NLP) and Machine Learning (ML) that are able to offer text analysis. The objective of these techniques is to extract entities and relationships from text providing the interpretation of new valuable knowledge (Bello-Orgaz et al., 2016). Sentiment analysis or the so-called opinion mining is a data analytics technique analyzing people's opinions, attitudes, emotions, judgments, and reviews towards objects, issues or attributes such as products, services, organizations, individuals, events and topics (Liu, 2012).

Images and videos are covering a great share in the components of social media data, mainly because of the swift rising sharing sites like YouTube and Instagram. Millions of images are uploaded weekly, and their mining can provide different precious of insights beyond what can be extracted by a text alone. *"Basic image analysis consists of the statistical analysis of tag data, demographic data and download frequency (e.g. keyword analysis for comments, most active followers, top locations). More advanced image analysis uses image processing methods, image recognition, and image tags. Similarly, the video analysis includes quantitative metrics, like amounts of users, response rate and location, analyzing voice to determining the emotional state of the user, and last but not least, a behavioral model for spoken words to determine the personality type of the user."* (Park et al., 2013).

Location is the core of the development of many cutting-edge technologies such as AI, data analytics and robotics. Due to the rapid progress of connectivity leading on an ever-connected ecosystem with

---

<sup>3</sup> Gartner (2014) Gartner Says Advanced Analytics Is a Top Business Priority. Available at: [www.gartner.com/newsroom/id/2881218](http://www.gartner.com/newsroom/id/2881218)

half of the web traffic coming from mobile devices<sup>4</sup>, location is a powerful tool for enterprises to improve customer experience, gain revenues and increase operational efficiencies. Therefore, the location-based services market size is expected to reach \$1.89 billion by 2022 (Chakravarty, 2018).

Location Intelligence was developed from GIS and incorporates the advantages of GIS into BI and uses many factors and data to recognize risks or opportunities, identify patterns, decrease expenses or increase profit and so on. Those factors can be economic, demographic, geographic and many more. In other words, the collected data is transformed into information and is then used to make companies more efficient. However, it is very demanding to determine Location Intelligence with just one definition due to many similar terms (Location Analytics, Geospatial Intelligence, GEO Business Intelligence, etc.) which appear in the literature and have almost the same explanations and examples.

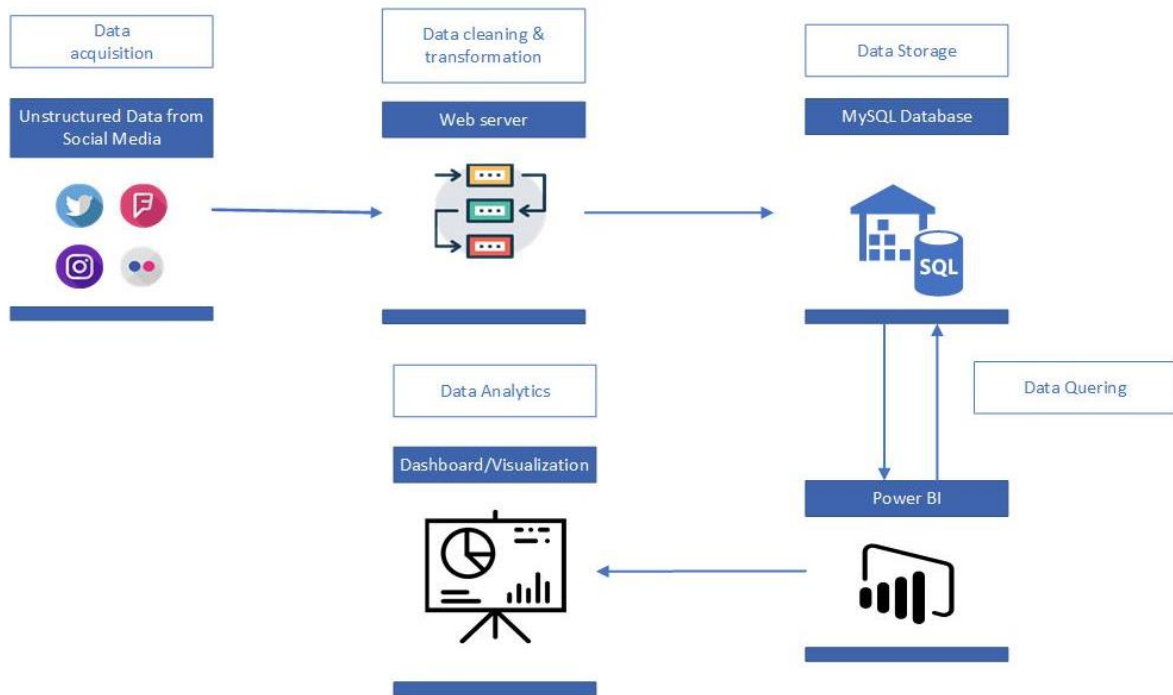
Survey-based approaches have been shown to suffer from several weaknesses and various bias such as answering particular questions covering prearranged aspects of the destination (Alaei et al., 2017). On the contrary to traditional survey-based method, social media and the availability of online user-generated content can provide cost-effective information (Hausmann et al. 2018), overcoming limitations such as sample size, time, location, nonresponse bias and self-reported errors (Mayer-Schönberger & Cukier, 2013; Veal, 2017).

The majority of previous studies in data mining from social media are using data mined from a single platform, such as Flickr (Mao, 2015; Miah et al., 2017; Donaire et al., 2014, Twitter (Adetunji & Carr, 2016; Brandt et al., 2017; Chua et al., 2016) and Instagram (Mittal et al., 2017). In this study, a combination of social networks for knowledge extraction from user-generated content with geotagged data is used. Hence, a crawler to scrape Instagram's explore page using Location-Ids as the parameters (Dhiratara et al., 2016) was built. In addition, a content analysis was conducted to define characteristics of photos shared on Instagram and Flickr (Donaire et al., 2014; Hausmann et al., 2018; Martínez et al. 2014).

In this analysis, the data flow and processes for all phases of big data analytics techniques are presented below. Twitter, Foursquare, Flickr and Instagram, four of the most popular online user-generated content social networks, are used as data sources for the analysis. These data sources provide as unstructured data/content that is generated by users and can be derived into two categories: media sharing (photos or videos) and text posts (opinion posts and reviews).

---

<sup>4</sup> Statista (2018) [www.statista.com/statistics/241462/global-mobile-phone-website-traffic-share](https://www.statista.com/statistics/241462/global-mobile-phone-website-traffic-share)



**Figure 7:** The data architecture of the analysis

The data acquisition consists of the communication of the web application programming interface (API) of each social media network periodically to acquire new geo-tagged user-generated content (data). The data cleaning and transformation of acquired data is the process for storing the necessary data and transforming them from unstructured into a structured format, to be stored in our SQL database and data warehouse.

The next phase is the data analysis for extracting useful insights from the social media content generated from users in the examining destinations. To achieve that big data analytics techniques are used. More specifically, using Google Natural Language API, several machine learning models are exploited for sentiment and entities/labels analysis on users' textual post (text mining/analytics).

Text mining/analytics provides significant information about what is being discussed/ mentioned inside the text and the feeling/sentiment about that. In addition, spatial and statistical analysis techniques are used in describing and visualizing the geographical distribution, uncovering patterns and points of interest (Floris & Campagna, 2014), while textual analytics techniques are implemented to discover knowledge that is embedded in a huge volume of social media text posts/comments shared in social networks of Twitter and Foursquare.

Another innovative addition that is added in this analysis is the image analysis using Google machine learning algorithm (Google Cloud Vision API<sup>5</sup>) - enabling to detect a broad set of labels/categories within the acquired images from the media sharing networks (Instagram and Flickr). After the acquisition, transformation, storage, and analysis of data, the next process in the data flow is data querying and data visualization. In order to provide valuable and actionable insights, it's important to

<sup>5</sup> <https://cloud.google.com/vision/docs/detecting-labels>



present the data properly in interactive, clear and easily understandable for stakeholders and decision-makers, visuals. In addition, data querying in data warehouse optimizing the time of the execution of highly complex queries over a large amount of data. In the next section, the results of the analysis are presented in interactive data visualizations using Microsoft Power BI and ESRI ArcGIS.

## **Data collection and analysis - a case study of Crete**

The dataset consists of collecting, processing and analyzing Big Data from location-based social media networks using the case study of Crete island. Crete is the largest island in Greece and the fifth largest in the Mediterranean area. It's one of the most popular tourist destinations worldwide<sup>6</sup> and its economy relies mostly on the tourism industry. In 2018, Crete as a Greek tourism destination had 5.228.000 tourist arrivals at islands; airports and that consists of the 15% of the arrivals in Greece by air (INSETE, 2019).

In contrast to the limited studies that are undertaken for Crete (Andriotis et al., 2007; Andriotis, 2011; Hosni et al., 2018; Zouganeli et al., 2012) based on traditional methodologies e.g. surveys, this approach enables to overcome the limitations arising such as limited information captured, sample size etc. Knowledge extraction from location and big data analytics is for a wider range of location covered and without limitations in seasonality, time, money, sample size and errors. The proposed approach (figure 8) contributes to the lack of innovative big data applications for tourism and to the existing literature with a novel approach exploiting big data analytics techniques and location intelligence.

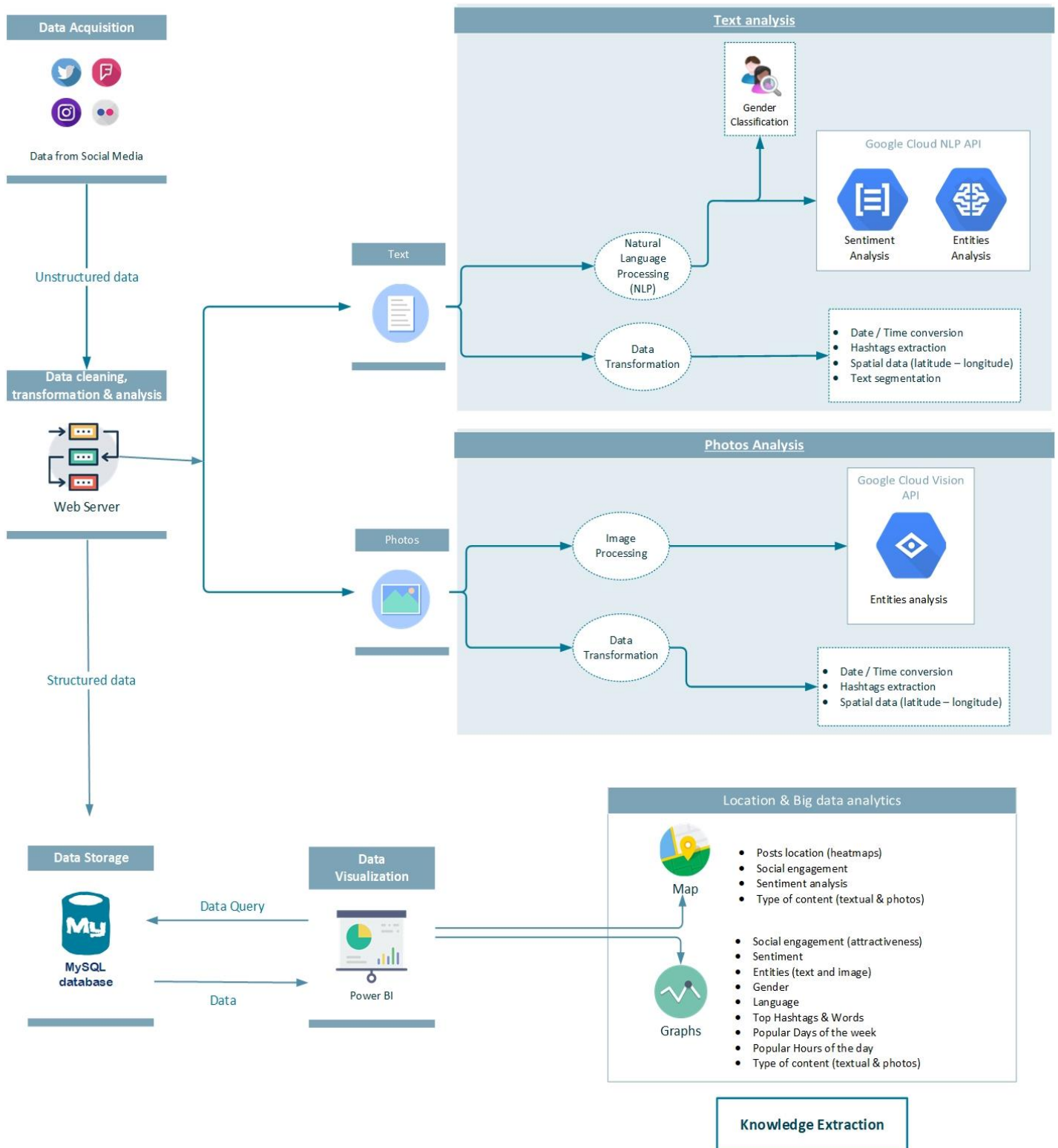
In this context, data acquisition for the case study of the tourist destination of Crete includes the largest and most visited tourist cities of Crete: Heraklion (including Knossos area), Chania, Rethymno, Agios Nikolaos (including Elounda area) and Ierapetra for the examining period of June 2018 to September 2018 (3 months of high tourism season). Nonetheless, for the purpose of the research, the presentation of the results will be on the data analysis and visualization of the two largest and most tourist-popular cities of Crete, Heraklion and Chania.

In the next phase, the user-generated unstructured data shared in location-based social networks is cleaned, transformed and analyzed, in dependence on the data type content (textual and photos). Therefore, for each type of content different big data analytics techniques are used for knowledge extraction. For text analysis, natural language processing (NLP) is used for gender classification, sentiment and entities analysis and for photos analysis, digital image processing is used. For both types of data, data transformation enables the utilization of spatio-temporal data that provides significant insights (knowledge). Spatio-temporal data and data extracted from the big data analysis of the big data for the two cities creates a pool of structured data that is stored. This stored data through queries can be visualized efficiently providing valuable insights.

More specifically, the interactive visualization of the analyzed big data for the examining tourist destination cities is in graphics (insights about tourist temporal characteristics, perspectives and demographics) and in maps (insights about tourist spatio - temporal patterns). This is the first study that spatial analytics (location intelligence) is combined with sentiment analysis, type of user-generated content and social engagement (big data social media data analytics), whilst a combination of different content type location-based social media networks are used.

---

<sup>6</sup> Tripadvisor 2019 - "Best Destinations in the World – Travelers Choice" [www.tripadvisor.com/TravelersChoice-Destinations-cTop-g1](https://www.tripadvisor.com/TravelersChoice-Destinations-cTop-g1)



**Figure 8:** Study's novel approach

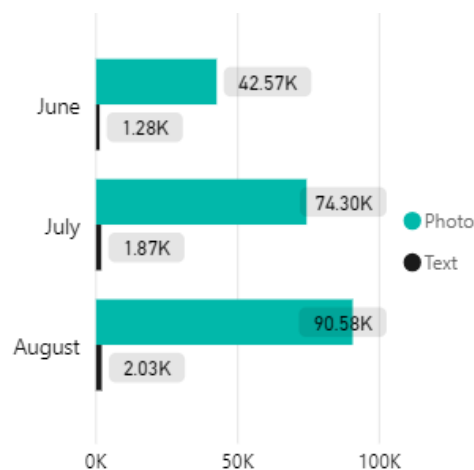
## Results & Discussion

Enterprises and destinations investigate ways to leverage big data to gain the maximum benefits in order to innovate and get competitive advantage against rivals. In that context, tourism enterprises and organizations can rely on data analytics, since it provides knowledge about customers (travellers/visitors) needs and preferences in order to provide them with better and more personalized products and services. Big data can become a key driver for offering personalized travel experiences satisfying the needs of every traveller.

From the data acquisition phase, significant number of text posts and photos for each city was collected and analyzed for knowledge discovery through the location-based social networks. It is observed that there is a huge difference between the number of textual posts (only text) and photo posts revealing the significance of including both types of posts in the analysis. The majority of travellers prefer to use media-sharing social networks to share their experiences through photos indicating the significance of using these social channels for attracting tourists and the creation of media content like photos and videos for more efficient promotion of the destinations (table 4).

City	Total Users	Total Text Posts	Total Photos	Total Engagement
Chania	16477	2703	59383	6150476
Rethymno	14788	1170	64044	5986406
Agios Nikolaos	12235	470	41755	7307735
Heraklion	11270	796	38263	4814525
Yerápetra	912	42	4129	1531457
<b>Total</b>	<b>55682</b>	<b>5181</b>	<b>207574</b>	<b>25790599</b>

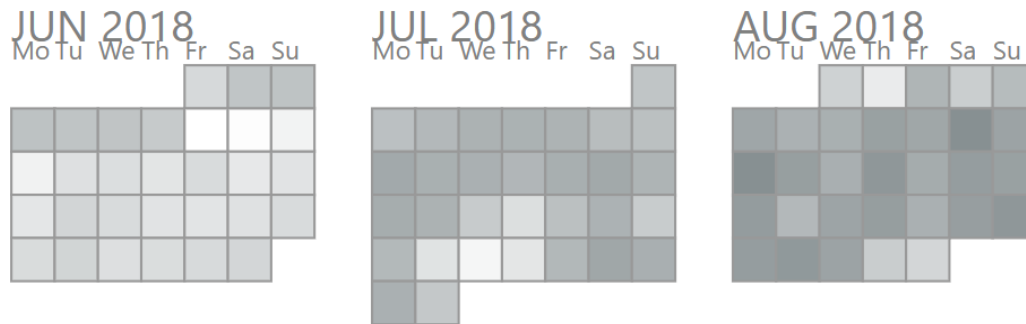
**Table 4:** Summary of dataset (total users, posts, photos and engagement)



**Figure 9:** Type of content per month in Crete

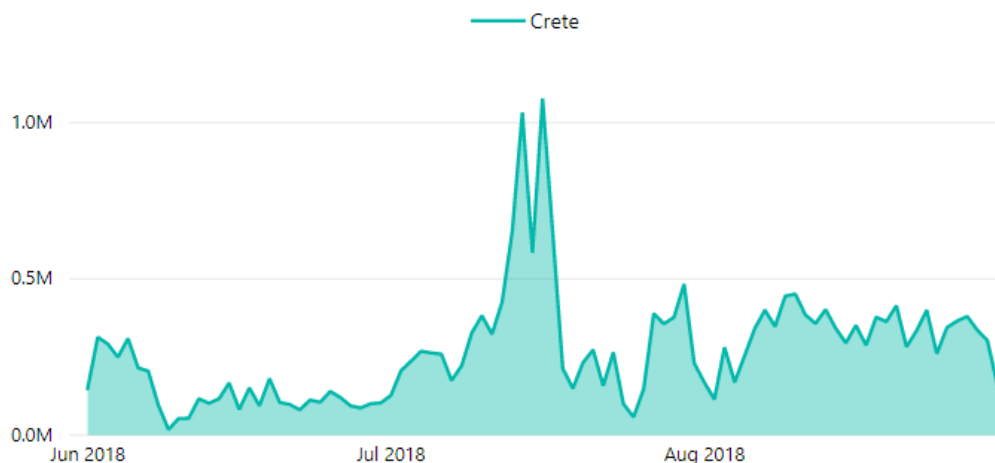
The examined period in this research consists of three months: June, July and August 2018 that are the most popular months for tourists during the year. The analysis of the acquired user-generated content provides information about the most crowded month for all the types of content. Hence, for both types of content (textual posts and photos), August is the most popular month (figure 9).

More specifically, a deeper analysis provides information about the most crowded days during the summer (figure 10), with the most crowded days presented with darker colour. This information can be very insightful for tourism enterprises/organizations in designing and developing their products/services for specific areas. For example, a travel agency can offer promotional offers for attracting more audience during the less crowded periods and thus increase their earnings, while they can shape their pricing policy based on data (data-driven pricing strategy).

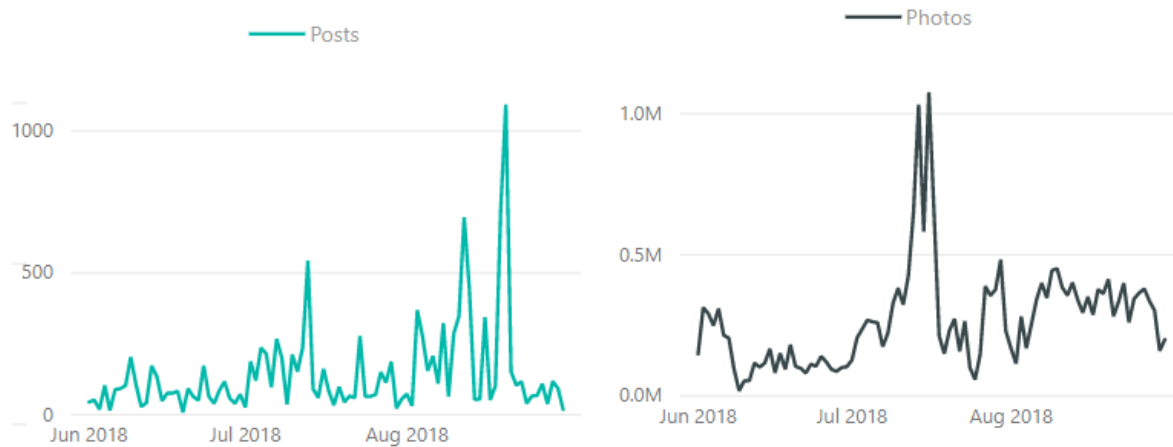


**Figure 10:** Frequency of total posts per day in Crete

One of the most significant metrics for social media analytics is total social engagement showing the summary of intentionally interactions of users on posts (likes, comments and shares) shared in social networks for various topics. In tourism, it is an indicator of the attractiveness of tourism experiences associated with a location. Social engagement is correlated with the e-word of mouth (e-WoM) or influencer marketing that are both able to enhance brand and tourism attractiveness through shared messages that induce a large number of people in the digital world (digital communities). Therefore, the analysis of social engagement can provide valuable insights about events, happenings, perspectives of users/visitors and influencers in a destination.



**Figure 11:** Total social engagement in Crete



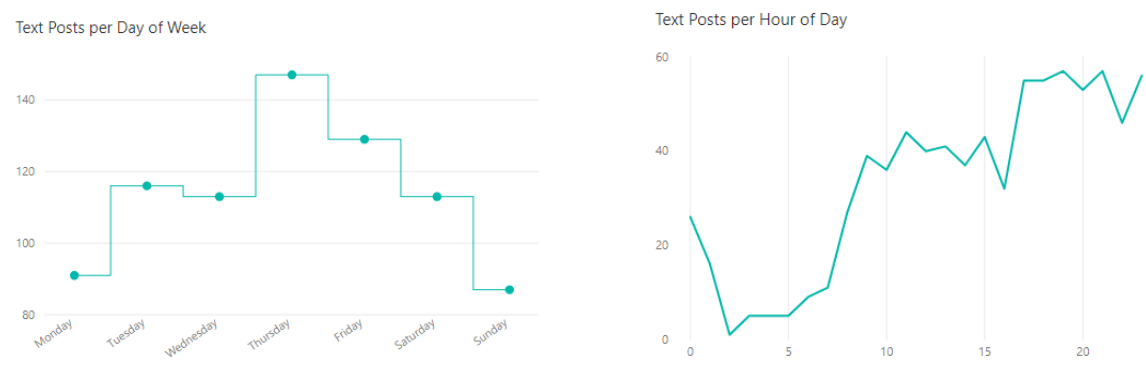
**Figure 12:** Social engagement per type of content in Crete

The total social engagement of Crete's cities presented above shows the number of interactions from social media users for the geo-tagged user-generated content that was generated in the examining period and locations. The analysis is divided into the social engagement of textual posts and photos indicating the huge difference between photos and textual posts in social engagement. Photos reach significantly larger number of engagements in contrast to textual posts, while the peaks in total social engagement are for textual posts and photos coming from social influencers, verifying the significance of influencers' marketing to enhance a destination in the digital world.

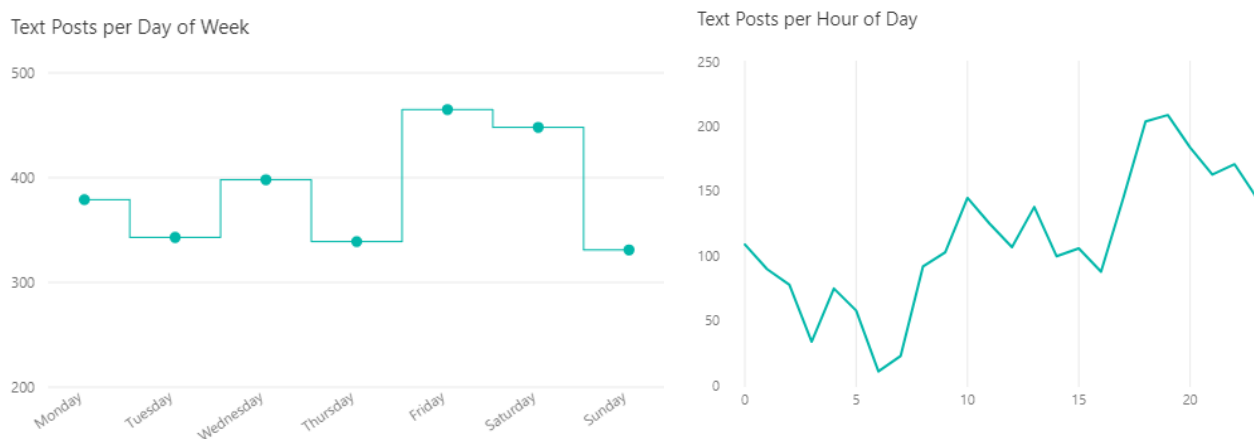
The data visualizations of the results for this research are separated into three categories: i) text analytics, ii) photo analytics and location analytics (location intelligence). In the convenience of the presentation of the results arising from the analysis, the data visualizations and analysis focuses on the two largest and most tourism-visited cities in destination of Crete, Heraklion and Chania.

### *Text analytics*

The distribution of the acquired user-generated textual data in the days of the week and hours of the day provides useful information about the date/time of sharing content in the digital world (Carvalho et al. 2017).



**Figure 13:** Textual posts per day of the week and hour of day for Heraklion



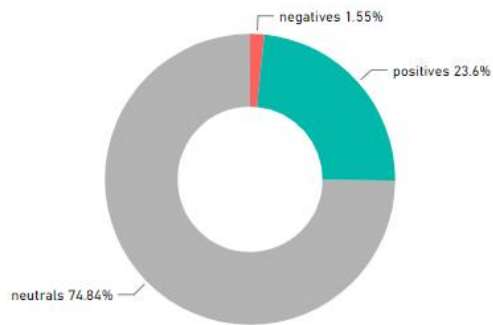
**Figure 14:** Textual posts per day of the week and hour of day for Chania

It is observed that during the examining period (figures 13 and 14) the most active days for textual posts generated from visitors in Heraklion city is Thursday and Friday and the most crowded hours of the day is from 17:00 to 21:00 and 23:00. In Chania, on the other side, the most popular days for generation of textual posts are Friday and Saturday, while the most active hours are afternoon hours, 18:00 and 19:00. It is significant for tourism enterprises or organizations that would like to maximize the visibility of their marketing efforts for offers, products/services to avoid sharing textual material during the inactive days and hours indicating that people watch and interact with text content on Friday and weekend during the afternoon and night hours. For instance, a restaurant located in Heraklion using these insights can create and share content during Thursday and Friday in the most popular hours, thus the possibility of attracting and engaging potential customers is higher.

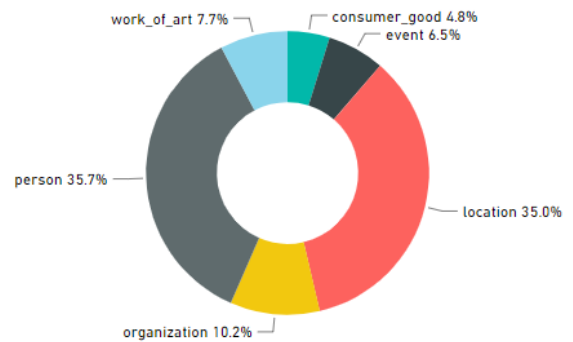
Sentiment analysis is the application of computational technologies to obtain and procure significant information from a large number of users' data. It involves the process of the extraction and analysis of opinions expressed in digital form and classified into neutral, positive and negative. In an online world where opinions are expressed and shared publicly within social media platforms, sentiment analysis is fertile ground where conversations can be utilised for the purpose of quality improvement, as such sentiment analysis is an important tool implemented by those organisations which realise the importance of such process and analysis (Abirami & Askarunisa, 2017). Sentiment analysis has been implemented strongly for business activities, such as predicting stock market movements, determining market trends, analysing product defects and managing crises. Another, and possibly more portentous, application is for auditors and accounting firms to utilise sentiment analysis as part of their continuing effort to anticipate theft, fraud, or embezzlement of company resources. Therefore, this approach has become a new standard for deterrence and prevention (Redhu, 2018; Pandey et al., 2016).

Regarding the tourism industry, visitors/travellers' online opinions/reviews play a huge role, since they induce decision making of potential travellers. However, the interpretation of sentiment analysis has "sampling bias", since an unhappy customer is more eager to express his negative opinion than a satisfied customer/visitor (Fan & Gordon, 2014) and that indicates the value of a positive opinion/feedback from a customer/visitor. Sentiment analysis can play a significant role in helping enterprises and organizations in crisis management. A negative feedback or opinion about a product/service or a destination if it is not managed has the potential to harm their image or brand significantly if these opinions go viral.

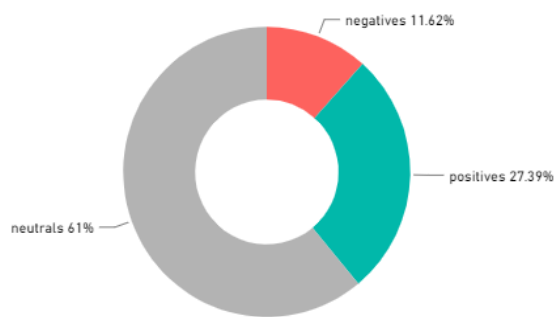
Text Sentiment Analysis



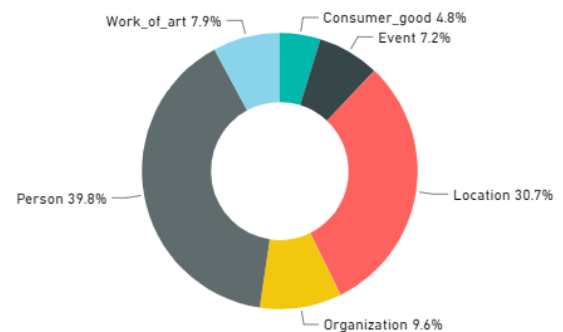
Text Entities Analysis

**Figure 15:** Text Sentiment and entities analysis for Heraklion

Text Sentiment Analysis



Text Entities Analysis

**Figure 16:** Text Sentiment and entities analysis for Chania

The sentiment analysis extracts users' sentiment from their shared textual posts. It's observed that the negative textual posts from visitors in Heraklion are in low levels, only 1.55% of the total textual posts, while the positive are 23.6%. In contrast to Heraklion sentiment analysis, Chania's analysis showed that negative posts are 11.62% of the total textual posts, while the positive posts are 27.39%. Sentiment analysis in tourism can reveal the visitors' impressions about the place and is able to provide valuable insights to stakeholders who can improve their offerings, enhance their branding and prevent issues that can be very harmful. From the collected data, it is observed that the most positive reviews/opinions for both cities Heraklion and Chania referred to the beauty of places and tastiness of local food, whilst the most negative are about customer services. Tourism Organizations/enterprises gathering insights about travellers on-site can discover what they appreciate and not, thus they can improve existing products/services or create new that will be data-driven designed.

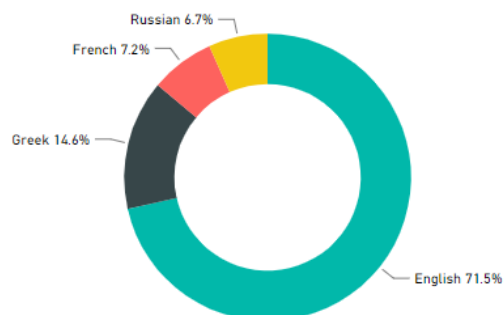
Another significant application that can provide valuable information is topic modelling that examines vast amounts of gained text in order to identify/detect the main topics or themes of that text. The revealed topics can be put into practice to provide homogeneous labels to examine the text accumulation even more. It is all performed by numbers of progressive statistical methods. Subjects uncovered by topic modelling can likewise be utilized as a source for other functions, for example, finding client

preferences, identifying rising points in discussions or online networking postings, or concluding parts of content gathering. Current progress in this field likewise enables mentioned methods to be utilized with social media sites, which gradually puts topic modelling usefulness to previously unreachable levels.

In that context, the text entities analysis provides information about the subjects that are discussed by the visitors of the city. Observation of that insights can help authorities and stakeholders to monitor the topics of discussion - interest in their areas and make strategic decisions taking that into account. For the city of Heraklion (figure 15), it's observed that the majority of textual posts are about places and people, while a small percentage is about activities/events (6.5%) and products (4.8%). For the city of Chania (figure 16), the results are in similar levels with the bulk of the posts related to places and people and a small proportion of the posts related to activities/events (7.2%) and products (4.8%). Taking these information into account, authorities and stakeholders can make better strategic decisions in order to improve the impressions of visitors in specific topics, for instance to activities that have to do with tradition or location attractiveness. The low levels of textual posts in activities/events indicate that there is a potential of providing more activities or experiences and traditional products that can get positive visitors' impressions as it is observed by sentiment analysis.

Textual analysis can provide significant information about visitors' demographics. After examining visitors' impressions and subjects of discussion, textual analysis provides the origin (language) and the gender of visitors for the two examined cities.

Language of Text posts

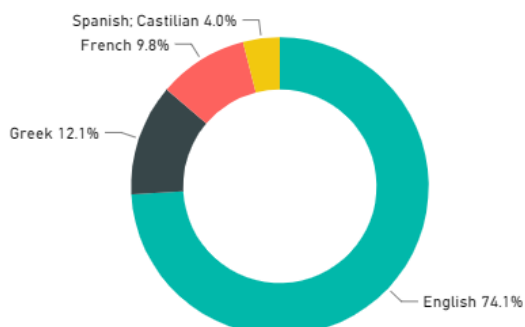


Gender for Textual Posts

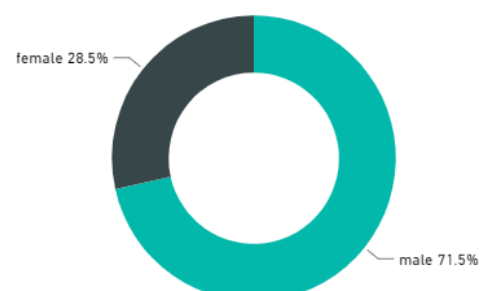


**Figure 17:** Language and Gender of Textual Posts for Heraklion

Language of Text posts



Gender for Textual Posts

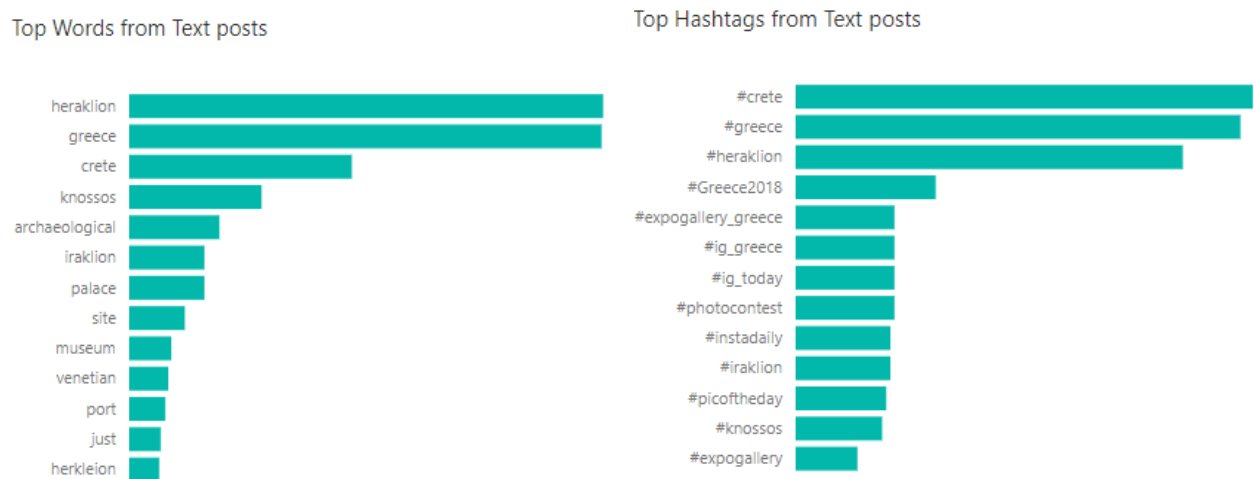


**Figure 18:** Language and Gender of Textual Posts for Chania

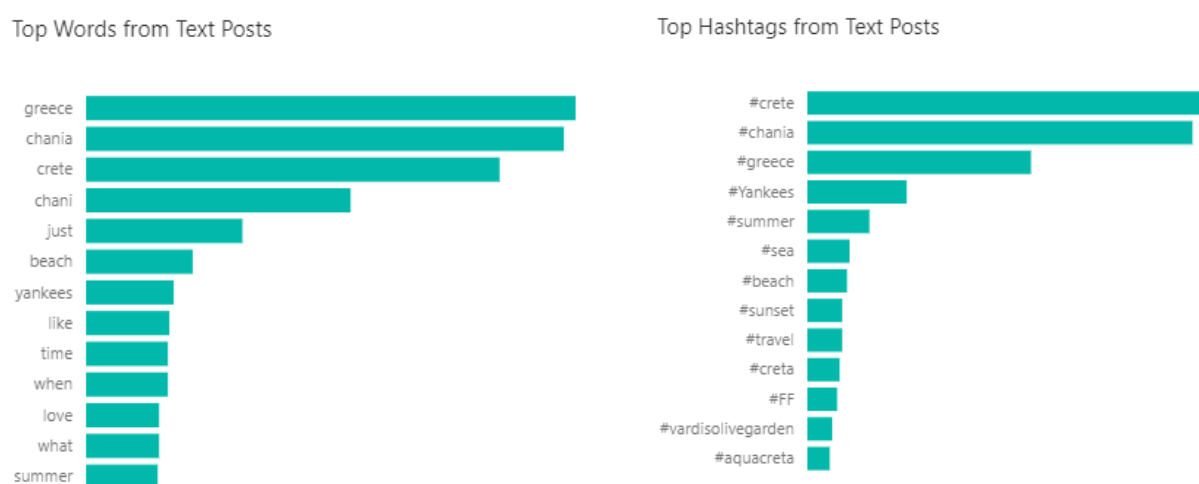


Through the analysis of textual posts, it's observed that in the city of Heraklion (figure 17) the most popular language in textual posts is English (71.5%), followed by Greek, French and Russian. The percentages of male and female visitors generating textual content are similar with male generating 54.3% of the total textual posts. In the city of Chania (figure 18), the most popular language is also English, followed by Greek, French and Spanish (Castilian), while the majority of people posting textual posts are male (71.5%). Extraction of demographic information can be valuable for strategic marketing decisions and personalization that can play a significant role in their competitiveness against rivals.

Words and hashtags are significant ingredients of the way people and enterprises interact and communicate in the digital world. The top hashtags and keywords provide information about the content that is generated from users. Therefore, they are used in tourist academic research for extracting patterns and value creation opportunities through big data (Költringer & Dickinger, 2015; Floris & Campagna, 2014; Vecchio et al., 2018b).



**Figure 19:** Top words and hashtags of Textual Posts for Heraklion



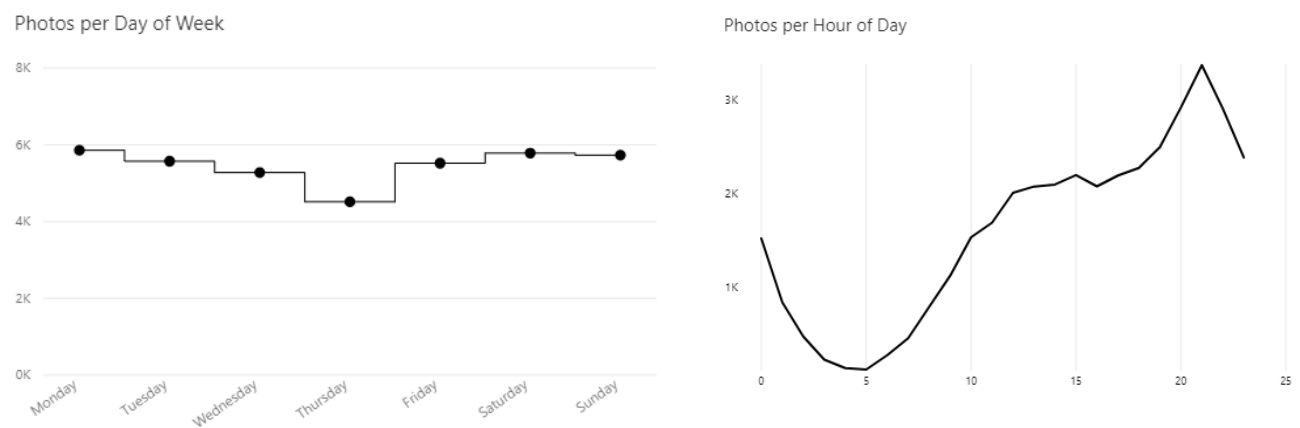
**Figure 20:** Top words and hashtags of Textual Posts for Chania

Digital word-of-mouth marketing can be powerful for enterprises and organization enabling the focus on not only collecting likes and followers, but also actively connecting and interacting with customers. In addition, hashtags can be used through the creation of share-friendly content like contests, polls or quizzes encouraging customers to engage more with a brand in social media. This is a great way to impulse impressions of a brand or destination, improve its content searchability and encourage discussions about a brand/destination. Create sharing incentives like offering customers a discount or bonuses if they make a referral about a product/service to social media, boost impressions of a brand.

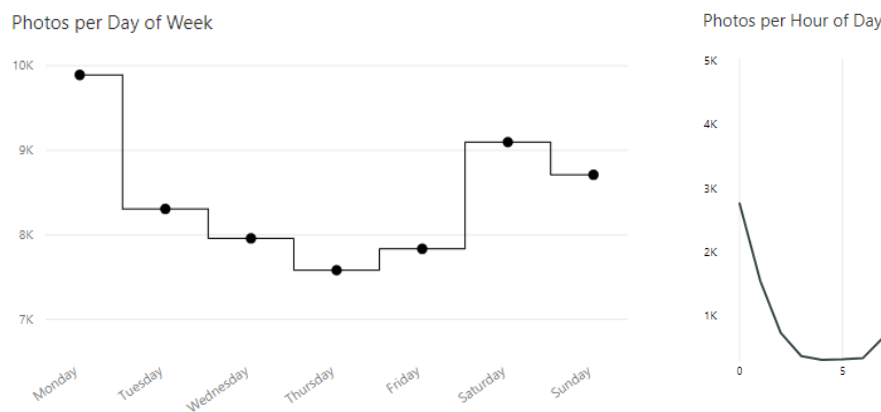
From the generated textual content in Heraklion (figure 19), it's observed that the top hashtags and words are about the island, city and country (#crete, #greece, #heraklion, crete, greece) and specific places in the area such as Knossos palace, Port and Archaeological museum. On the other hand, in Chania (figure 20), user-generated content top hashtags and words are about the island, city and country (#crete, #greece, #chania, crete, greece), specific places in the area such as a water park and travel experiences with local tradition like olive oil. In addition, the popular words and hashtags in Chania identifying that visitors enjoy spending their time at the beach, while they want to watch baseball games. Tourism SMEs such as restaurants, tourist agents and stakeholders can leverage them and discover the preferences of tourists in their area and thus they can create efficient marketing campaigns using these keywords and hashtags to promote and attract successfully potential customers of their products/services to a huge amount of users.

## Photo analytics

Images and videos are covering a great share in the components of social media data, mainly because of the swift rising of sharing sites like YouTube and Instagram. Millions of images are uploaded every day, and their mining can provide different precious insights beyond what can be extracted by a text alone. Basic image analysis consists of the statistical analysis of tag data, demographic data and download frequency. More advanced image analysis uses image processing methods, image recognition and image tags. (Park et al., 2013). In this research, two popular photo sharing social networks are used for uncovering patterns and trends of visitors are used, Instagram and Flickr. The analysis of user-generated photos in the cities of Heraklion and Chania in Crete island presents insights about tourist preferences and perspectives.



**Figure 21:** Photo posts per day of the week and hour of day for Heraklion

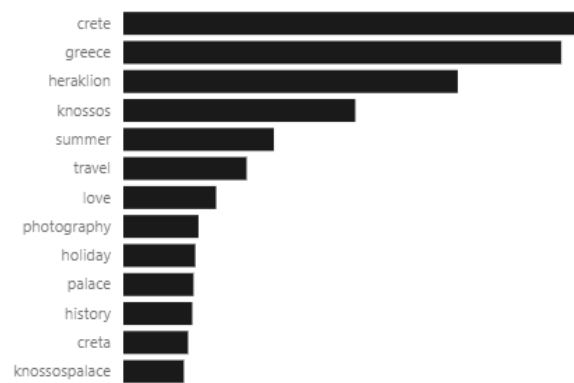


**Figure 22:** Photo posts per day of the week and hour of day for Chania

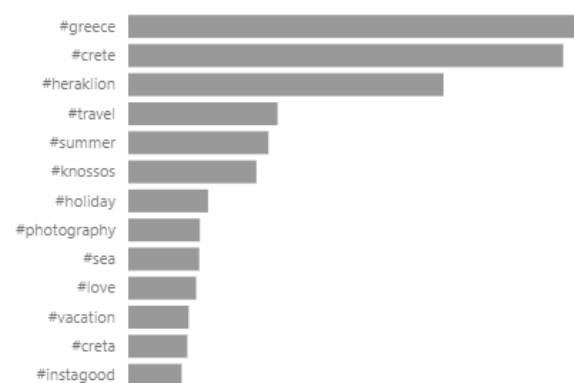
In the city of Heraklion (figure 21), it can be observed that Thursday in contrast to textual posts is the less active day for photo posts, while Monday and Saturday are the most popular days for sharing photo content in Heraklion, while the most active hour of the day for photo posting are in the afternoon and night, from 19:00 until midnight. In the city of Chania (figure 22), Monday is the most active day for photo sharing and the most popular hours are as from 19:00 to 23:00. These insights can be used by tourism SMEs and stakeholders for creating content that can be shared in the media-sharing social networks during the most popular days and hours of each destination enhancing the efficiency of their

marketing efforts. For instance, a bike-rental enterprise can create media content about its product/services and share it during the most popular days/hours engaging and attracting travellers for their offerings.

Top Words from Photo posts

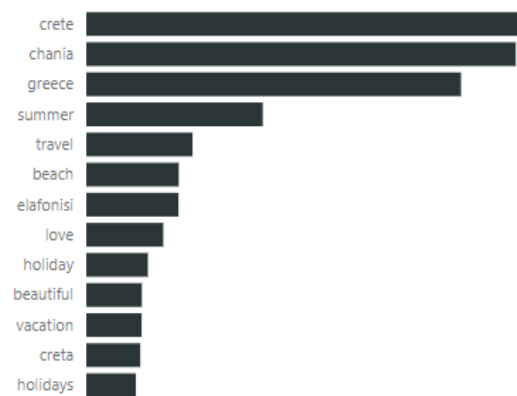


Top Hashtags from Photo posts

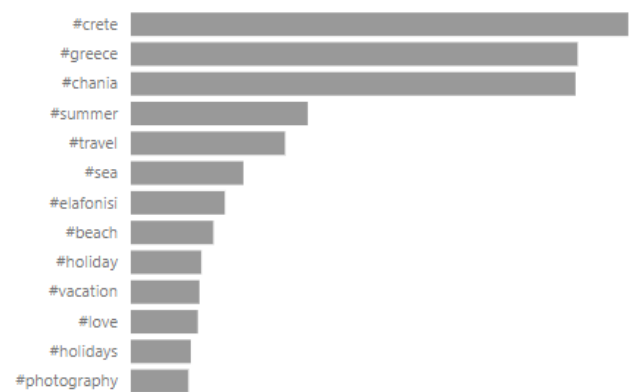


**Figure 23:** Top words and hashtags of Textual Posts for Heraklion

Top Words from Photo posts



Top Hashtags from Photo posts



**Figure 24:** Top words and hashtags of Textual Posts for Chania

Hashtags and keywords are significant for photo-sharing social networks like Instagram and Flickr that are used in this research. Users use these in order to explore and find content, thus information about the top words and hashtags used for a destination/location can be actionable knowledge for successful content sharing for customers/visitors attraction. From user-generated content in both cities of Crete island, it is observed that visitors share textual posts using hashtags and words about the island, site and country, specific places such as Knossos palace and related with the content of the photos i.e. #sea and their feeling #love, #instagood. Therefore, top words and hashtags for Heraklion (figure 23) and Chania (figure 24) identify the keywords that can be used for more efficient marketing activities.

Gender for Photos



Top Image Entities



**Figure 25:** Gender and Image Entities of Photos in Heraklion

Gender for Photos



Top Image Entities



**Figure 26:** Gender and Image Entities of Photos in Chania

Using a machine learning model, the gender of users sharing photo is identified. In that context, it can be observed that more female than male visitors of Heraklion (figure 25) and Chania (figure 26) share photos, indicating that women are more active in sharing photo content during their visit in a tourism destination. Contrary to textual posts before, it is observed that females prefer to share photos in social media in the examining tourism destinations.

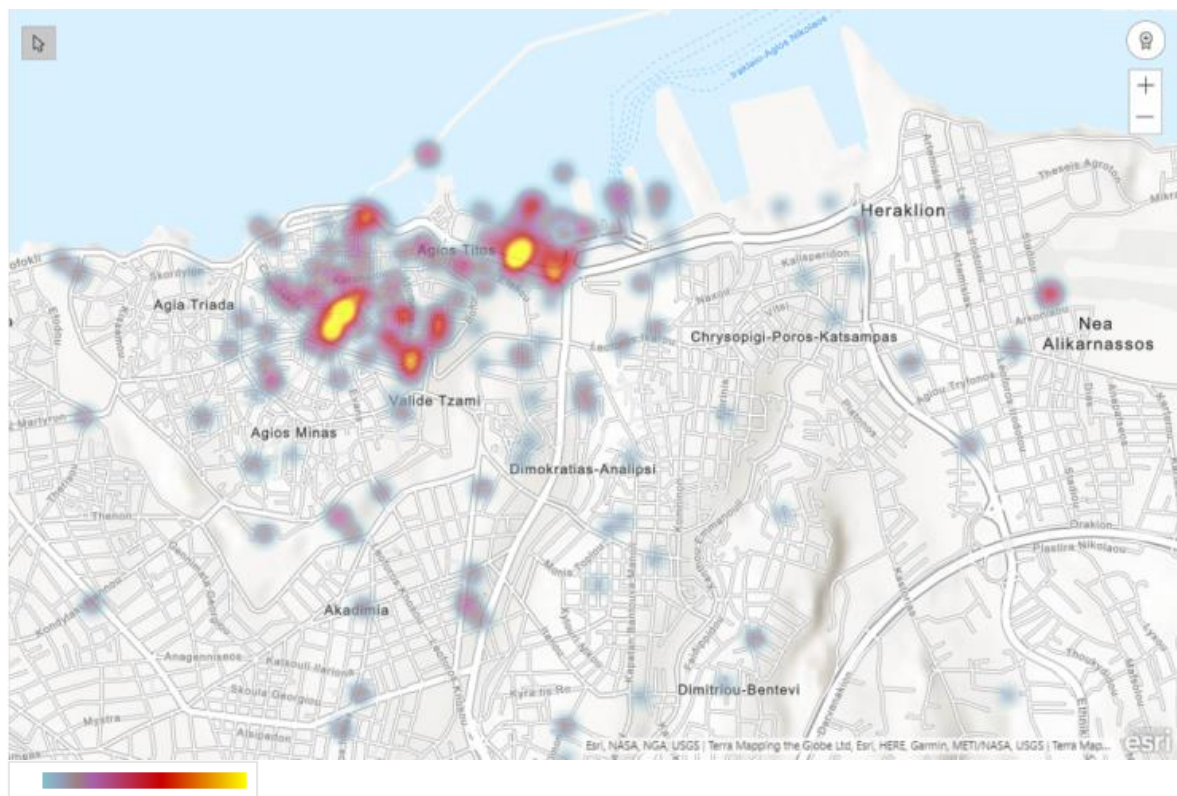
In the approach introduced in this research, image label detection is used for analyzing the images that are shared from visitors in Heraklion during the examination period. In that context, the content of photos shared is presented using word cloud that is an extremely popular non-chart visualisation (Chi et al., 2015) in use all over the web (Chang et al., 2017). The word cloud represents a simple list of words in varying sizes that correspond to word usage within the domain in question. The visualisation is typically dynamic and interactive – a click takes the user directly to a search of the keyword. The most popular labels detected in visitors' photos are related with the weather, place and people indicating that the majority of photos contain faces and people, mostly female, enjoying their holidays near the sea. Additionally, there is the detection of people characteristics indicating that people enjoy taking photographs with their presence in the view they see. An organization or enterprise can exploit these insights to run photo contest inviting users to share photos under certain hashtags or themes of the area

and offer them awards/offers for their participation. In this way, they will promote their brand and they will have the opportunity to attract them by offering them more products/services.

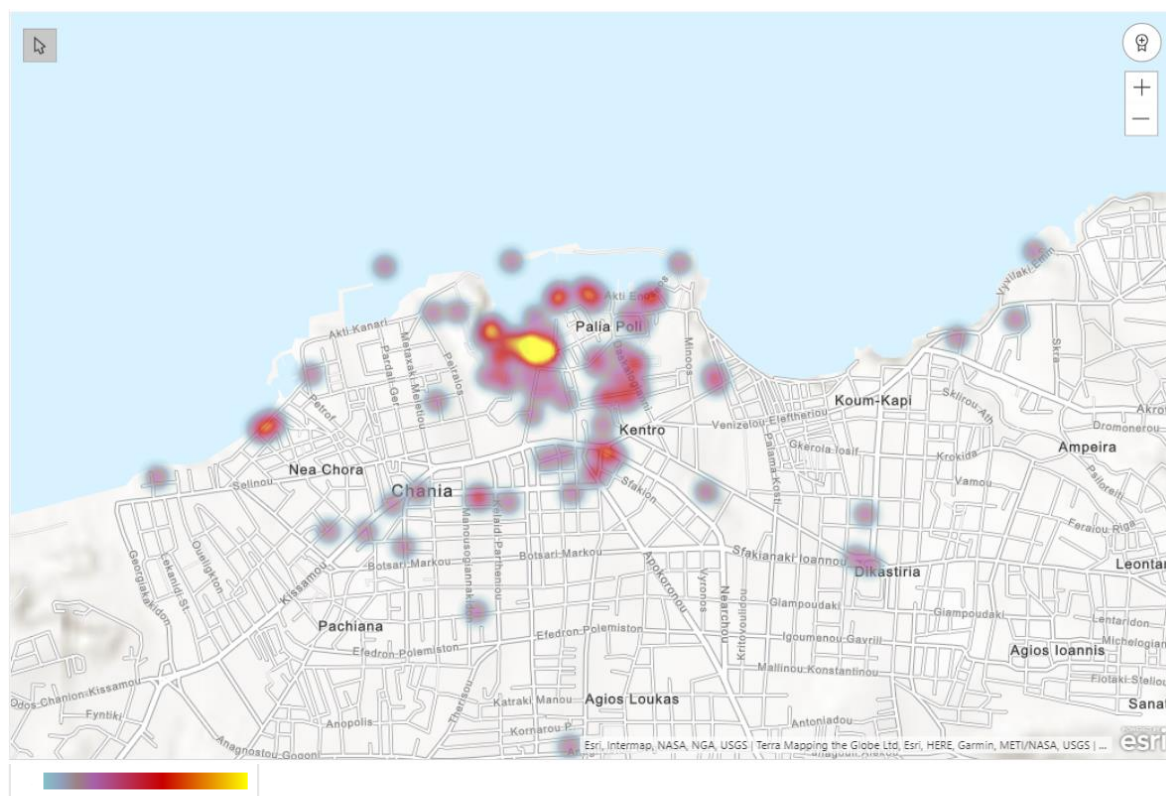
### *Location Analytics*

Location data is fuel for innovation, especially in services that are particularly location-aware such as Uber and Waze, where exact location and routing are significant to the value they provide to users. Location data and analytics results in helping businesses make smarter, more actionable decisions (Forbes, 2017). Map visualizations provide benefits through the representation of data to an easy to understand form, since *“cartographic visualisation can be claimed to solve many of the fundamental problems identified in studying spatial social distributions”* (Dorling, 2012).

Location data are able to enrich business data, since they include geocoding (latitude and longitude) and thus geo-enrichment can be achieved through incorporating valid data from government and other sources with new data by location. Thus, it can provide significant information about the characteristics of a property location. Location data can contribute significantly in visualizing business data in more efficient ways typically with maps that allow enterprises to identify patterns and trends in their data (Carvalho et al., 2017; Miah et al., 2017).



**Figure 27:** Heatmap of user-generated content location in Heraklion



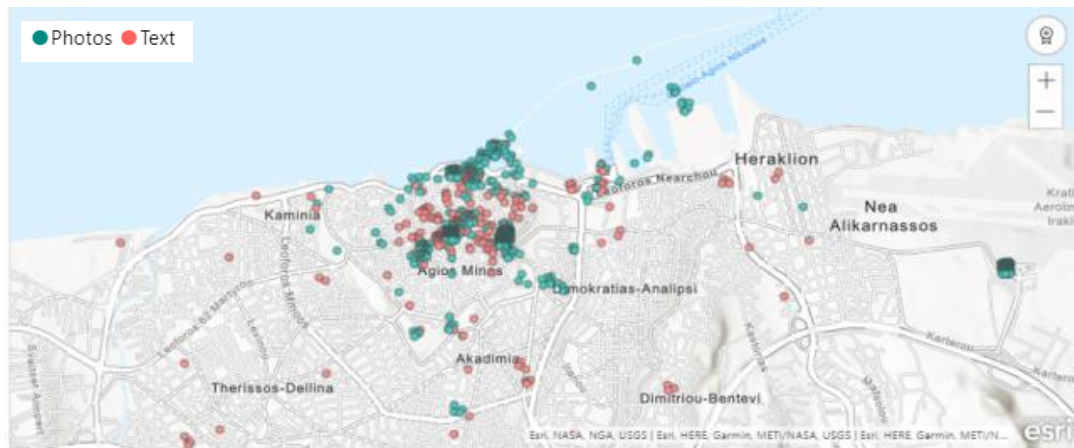
**Figure 28:** Heatmap of user-generated content location in Chania

Through heatmap, the most popular places for a given time period in a location can be observed (Abbasi et al., 2015; Liu et al., 2016), while time range can provide also the “time” (days and hours) these places are popular. These insights can be leveraged by tourism authorities and SMEs for activities focused on these specific locations. In posts heatmaps, all the acquired user-generated data (textual posts and photos) shared in the city of Heraklion (figure 27) and Chania (figure 28) are displayed. The places with the most posts are displayed with the yellow colour. The most popular places in both Heraklion and Chania are around the old town of the two cities and their ports. The stakeholders can examine the popularity during the time period they want, thus they can extract conclusions about tourist behaviour in their interest area. For instance, tourism agencies or managers can recommend or provide tours in the attractions or points of interest during off-peak times to avoid overcrowded venues that may reduce tourist satisfaction leading to bad reviews and comments.

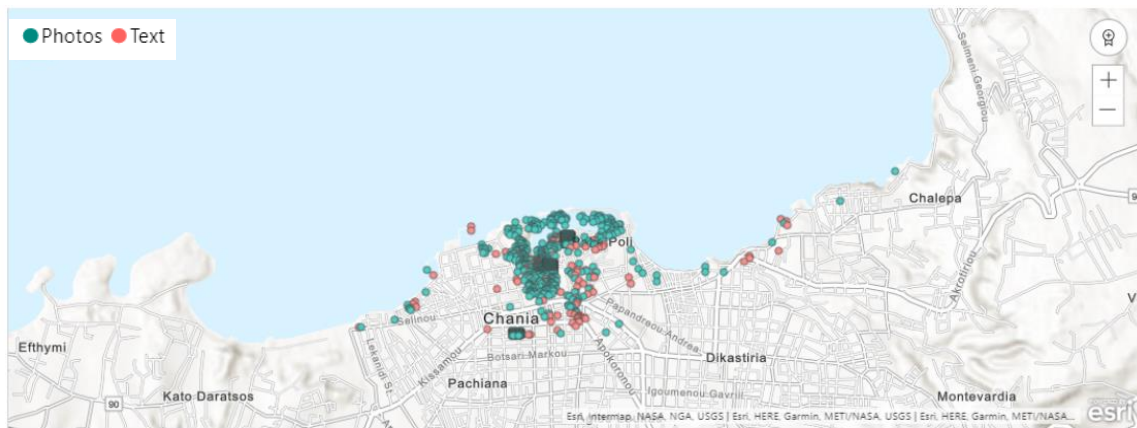
Another significant factor to be examined is the type of content shared from tourists via their mobile in social media during their visit in the city. Therefore, the type of their content, textual or photo post, is presented in the following visualizations. It is obvious that the majority of content shared is photos for both cities, with users’ content in Heraklion consisting of photos by 98.1% and 98.9% in Chania, while the frequency of posts during the days of the examining period is showing the most active days. SMEs and stakeholders leveraging these insights can observe tourist behaviour and shape their strategy and marketing in innovative and more efficient ways. Another significant insight arising from the visualization presenting spatio-temporal data derived by type of shared content is the GPS tag (latitude, longitude) of each type of content indicating the specific places in the examining places that attract the interest of tourists for taking photographs instead of share a textual message. Visual content such as photos are significant in designing material for conventional (brochures, posters etc.) and digital (website, social media channels etc) marketing. Identifying the favoured locations pictures unveil



travellers' interests and preferences, thus the locations can provide the most appropriate visual content for making their ads more relevant targeting the right audience.



**Figure 29:** Map of Heraklion - User-generated content distribution per content type



**Figure 30:** Map of Chania - User-generated content distribution per content type

While other studies examine tourist reviews and sentiments using keywords or hashtags shared on social media networks provide insights about their preferences and satisfaction (Chang et al., 2017; Költringer & Dickinger, 2015), this research introduces the integration of location intelligence with textual sentiment analysis to knowledge extraction in tourism destinations. More specifically, using the sentiment analysis and location meta-data in textual posts for Heraklion and Chania, the most positive-rated spots of each city are presented.

The insights arising from the analysis and visualisation are useful in observing spots that tourists/visitors enjoy more. For Heraklion (figure 29), it is observed that the most favourite spots are in the old town and a deeper analysis can show that tourists like walking through the city's alleys/roads and visiting town's monuments. On the other side, in Chania (figure 30), it is observed that tourists enjoy walking across the old city and old harbour. Tourism SMEs and stakeholders leveraging these insights can find where tourists enjoy more and examine why, thus they can shape more personalized products/services. The observation of the most favoured spots in a city, can provide significant insights about potential new locations for SMEs expansion and/or marketing actions.





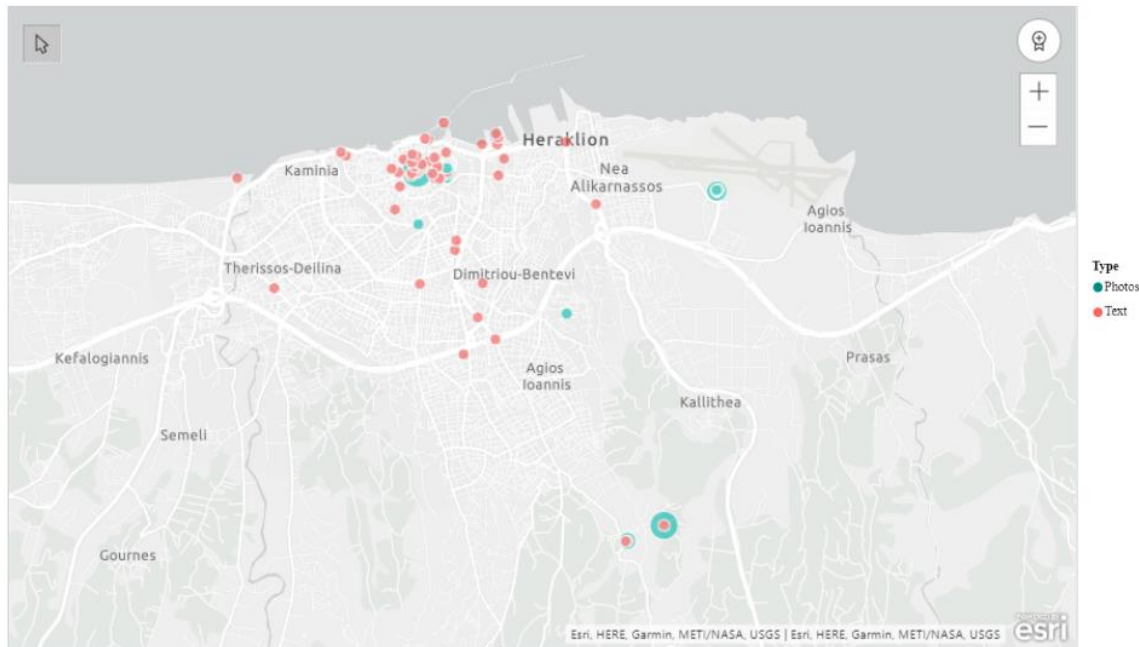
**Figure 31:** Map of Heraklion - Location of the most positive textual posts



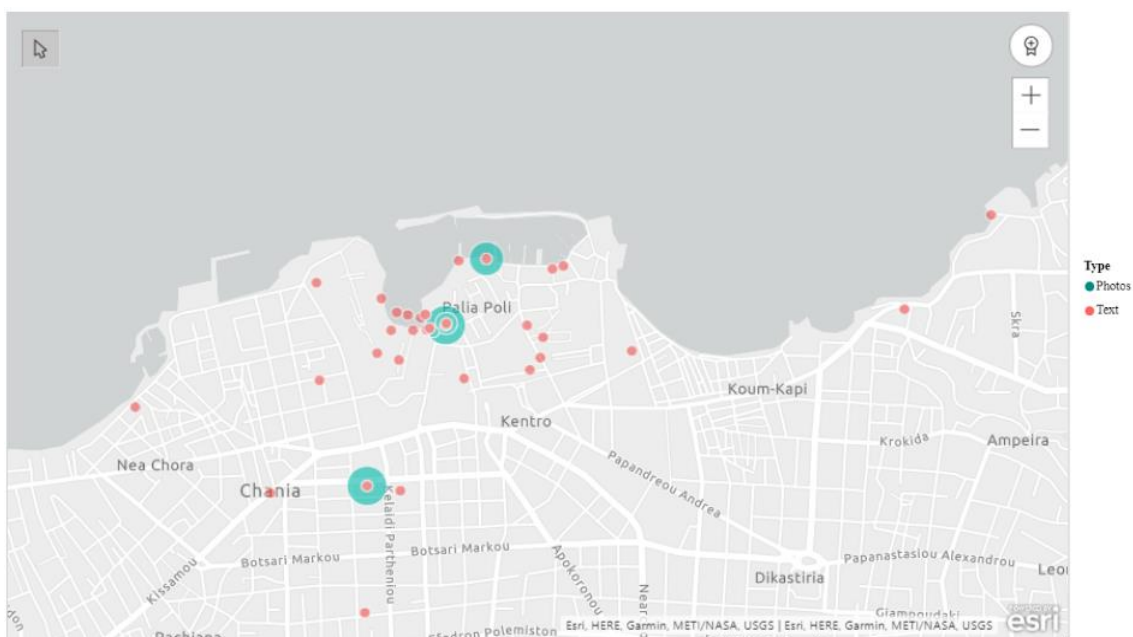
**Figure 32:** Map of Chania - Location of the most positive textual posts

The sentiment of the user-generated content (textual posts) provides insights about users' impressions and adding location of specific post can also provide information about “where” and “why” a user had a positive or a negative experience. Through the analysis of the acquired data using big data analytics including ML and AI, the enrichment of insights about a location can be achieved. More specifically, stakeholders can observe specific locations and the highest engagement content that is generated by users.

Social engagement is a metric indicating the attractiveness of a location and a tourism experience, since it summarises the social interactions of users and it is used as electronic word of mouth. Therefore, the analysis of social engagement is significant in order to empower enterprises and stakeholders with valuable insights that can transform their products/services into personalized experiences lead on more satisfied customers.



**Figure 33:** Map of Heraklion - Location and type of the most engaged posts



**Figure 34:** Map of Chania - Location and type of the most engaged posts

The analysis of the current research combines big data and location intelligence and offers insights about the exact location, the “topic/label” of discussion or photo, the social engagement and the channel

for any specific textual post or photo in the defined area and date ranges. Hence, tourist SMEs, stakeholders and local authorities are able to exploit in real-time the users behavioural perspectives and based on that (data-driven) form their strategy and actions.

## **Discussion - Conclusion**

In the age of digital innovation and transformation, small enterprises have the ability to disrupt traditional business models due to their agility and innovation. SMEs can grasp innovation and revenues due to their flexibility and their size that allows them to minimize internal corporate gaps and have more connected employees. Hence, they are able to comply with new situations and new technologies gaining competitive advantage (Newman, 2016).

Big data enables enterprises and organizations to leverage value and acquire actionable knowledge so that they can make better decisions and revolutionize their strategies and business models (Waller & Fawcett, 2013). The availability of big data related with consumers provides enterprises unique opportunities to pursue understanding of market trends, consumer behavior and activities (Erevelles et al., 2016; Waller & Fawcett, 2013). Although enterprises relying on data is not a new concept, but recently they began to leverage other data sources such as social media, smartphones or sensors, and new analytical technologies to exploit this data (Hartmann et al., 2016). Big data is not just about data collection or storage, its objective is to analyse data to get insights and leverage their value (Bello-Organ et al., 2016). Nowadays, enterprises that fail to adjust with data-driven processes face the risk of losing a critical competitive advantage, while a not efficient data utilization can lead to survival issues (Brownlow et al., 2015).

Social media networks have transformed the relationships between enterprises, employees, customers and stakeholders. Prominent examples of SM include Facebook (with over 2,2 billion monthly active users), Twitter (with over 330 million monthly active users), YouTube (with over 1,8 billion monthly active users) and Yelp; a platform for reviews of products or services provided by consumers, together with many other blogs, forums, microblogs, photo and video sharing platforms (i.e. Instagram with a whopping 800 million monthly active users). These numbers are clearly indicating that SM applications generate enormous amounts of data – this information potentially offers enormous possibilities in business growth. Because of the amount of active users on these platforms and overall content produced daily, organisations have the motivation to recognise which trends are currently evolving to identify potential risks and opportunities, which are highly subjective for every organisation. Other than monitoring current trends, it is also needed to be aware of the creators of the content themselves, specifically the most influencing ones. Both businesses and non-profit organisations collect the data produced by users with desire to secure a spot in the multimedia world. Customers are able to gain various types of shared data information from each other, by using social media. This information plays a significant role in one's purchasing decisions. Therefore, an important aspect of SM is developing relationship with customers, suppliers, PR and potential partners.

Analyzing user-generated content shared on social media can be valuable in various fields like marketing, public health, public safety (Gerber, 2014), tourism and politics (Carvalho et al., 2017). There are different analytics techniques like sentiment analysis, topic modelling, natural language processing (NLP) etc that are used to uncover hidden patterns and insights in social media data. Data visualization and visual analytics are significant for effective management and support of understanding for big data (Ribarsky et al., 2014). The evolution of decision-making exploiting information generated

by tourists in real-time, inducement of tourists' experience by discovering needs and preferences, emergence of new business models and products/services.

Social media big data analysis is able to discover new knowledge that can be used for better decision making of individuals and enterprises. Big data social media analytics provide an exclusive opportunity to retrieve a huge amount of customers' opinions and experiences contributing to the improvement of traditional strategies through the generation of more targeted marketing campaigns (Bello-Orgaz et al., 2016) and strategy.

Especially for SMEs that have mostly limited resources and cannot afford mismanagement, the insights gained can contribute to focus on the most important issues. Therefore, sentiment analysis can be considered as a profitable technique arised of Big Data. The interactivity of social networks is an extensive marketing tool for reaching out to a client base. Being able to respond quickly and hook the audience regularly by engaging with them, not necessarily on product-related topics, can strengthen the boundaries between enterprises and their customers, while it is an inexpensive direct communication channel especially for SMEs (Nobre & Silva, 2014). Enterprises that do not spend enough resources on the planning process and get engaged with their customers over time, can be seen as more vulnerable and are likely to lose their competitiveness over time. Any enterprise in tourism industry has to exploit its managerial and marketing strategies, tactics and tools to gain and maintain a sustained competitive advantage (Mariani et al., 2018). The integration of various data sources with structured and unstructured data has the ability to provide insights that cannot be revealed by data traditional analysis approaches (Alaei et al., 2017). All those complex processes can be evaluated and assessed by Big Data analytics that can help to increase the longevity of enterprises.

In this context, the purpose of this research is to reveal a novel approach to leverage massive unstructured data providing valuable knowledge (insights) for tourism SMEs and stakeholders to attain innovation and improve their value creation process through leveraging big data generated from location-based social networks. The findings of the analysis using a case study of the two destinations on the island of Crete indicating that big data user-generated content in location-based social networks can provide insightful patterns about tourists' on-site behaviour in tourism destination. Understanding tourist preferences and perspectives provides significant implications for tourism enterprises and stakeholders to efficient strategic planning and decision making.

Taking insights about which is the most popular location, what attracts tourists more and why, what are the most popular and preferred days-hours to share each type of content, where and why do tourists like to visit, which is the more attractive location and what are the characteristics of tourists, make tourism organizations, SMEs and destination managers capable of design, implement and develop knowledge-based products and services using interactive data visualizations enabling the deeper analysis without any expertise required. Additionally, the results of the study provide knowledge about developing the appropriate strategies of promotion of the cities or the whole island, Crete, as a tourist destination through online promotion.

Limitations of the study should be also identified. Firstly, only user-generated content from users located in urban cities was used, thus future research can also include rural, agricultural or coastal areas. In addition, not all tourists visiting the examining cities take photos or write reviews and share them online, thus it is a study focused on the tourists that generate content in the location-based social networks used in the study.

# **Big Data Analytics for Tourism Destinations: A comparative analysis through Location-Based Social Networks**

## **Introduction**

User-generated data in Location-based Social Networks (LBSNs) can be a great resource of knowledge for understanding people's behaviour details and movement flows in tourism destinations. Nowadays, local authorities and tourism enterprises are using conventional methods like surveys and opinion polls for collecting data and strategic decision making. Despite the benefits of these approaches, they present significant disadvantages such as time consuming and sample size.

The practical research combining big data analytics techniques and social media is scarce, while there is a lack of knowledge in leveraging big data in hospitality and tourism. Therefore, the implementation of innovative big data analytics applications leveraging location-based big social data provides new knowledge about behavioral data and perceptions in tourism destinations, is requested.

Focusing on tourism and location-based social media networks, this chapter aims to reveal a novel approach to leverage massive unstructured data for knowledge extraction. In contrast to the conventional spatio-temporal data, big social media data offer dynamically to innovation and value creation through improving strategic decision-making process of tourism destination stakeholders. Valuable knowledge is extracted about travellers' behaviour, impressions and preferences for tourist destinations. To achieve that a novel approach through harnessing massive unstructured data derived from popular location-based social networks e.g. Twitter and Instagram is developed. More specifically, the approach introduces the combination of textual and photo analytics with data derived from media sharing and textual social networks, introduces social big data analytics such as social engagement, sentiment analysis, topic/label detection combined with spatio-temporal features to provide more insights about tourist destinations.

The approach integrates location and big data analytics techniques and it is implemented based upon geotagged user-generated data shared on the two largest islands in the Mediterranean Sea, the island of Crete (Greece), and the island of Cyprus that are popular summertime tourist destinations. The comparison between two tourist destinations with common characteristics provide additional insights about the potential of each destination and areas of improvement.

Practical implications are arising through the efficient spatio-temporal and demographic analysis of tourist movement in both tourism destinations for improving strategic decision making of stakeholders like local authorities and tourism SMEs leading to innovation and value creation. In addition, DMOs can leverage the new knowledge for developing innovative marketing strategies strengthening destination branding gaining competitive advantage against rival tourism destinations.

## Literature / study background - framework

Nowadays, “Big data” is concentrating prodigious attention globally that is driven by the tremendous increase of mobile devices, social media networks and “Internet of Things” concepts. Some scholars have considered “big data” as large datasets that are not able to be captured, stored and managed with commonly used software tools (Manyika et al., 2011), however due to its great prospects, there are several definitions of that concept (Wamba et al., 2015). According to (IDC, 2012), Big Data demonstrates three major characteristics: *“the data itself, the analytics of the data, and the presentation of the results of the analytics”*, while (De Mauro et al., 2016) proposed a definition of Big Data that is compatible with the use of terms Big Data Technology & Methods: *“Big Data is the Information asset characterised by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value”*.

Some scholars and practitioners use the “Vs” to define big data and its challenges. In that context, big data challenges involve more than just managing large volumes of data, thus the “3Vs” of Big Data were proposed: volume, variety and velocity (McAfee & Brynjolfsson, 2012) and later another “V” for “value” the 4th was added in order to highlight the significance and benefits coming from leveraging big data (IDC, 2012). The 5Vs was presented with “veracity” by IBM<sup>7</sup>.

- **Volume:** it's without a doubt the major feature of big data. Large number of records or enormous volume of data using huge storage (Wamba et al., 2015)
- **Variety:** the range of data types and sources. Big data is generated from several sources and in different formats. With the availability of the storage capabilities, the volume of structured and unstructured data and its various sources will continue to burst
- **Velocity:** The speed of data variations. Frequency of data generation and/or frequency of data delivery.
- **Veracity:** the uncertainty due to data inconsistency, incompleteness, and/or model approximations, requires the analysis of large data to achieve a reliable prediction. Organizations/Enterprises have to verify the quality, governance and compliance issues of big data
- **Value:** the value coming from the insights and benefits by leveraging generated big data

Big Data creates value through its contribution to optimal price setting, optimizing supply chain, minimizing errors and improving customer satisfaction (Günther et al., 2017). The potential and benefits of big data is significant, however there are many technical challenges that should be addressed for emerging its full potential and value. The huge amount (volume) of data is the primary challenge that is recognized broadly, but there are others.

According to Moore's law, the number of transistors on a computer chip will double almost every two years. The completion of this law is that the ability to store and process data come easier, faster and cheaper. This progress coming from the latest technological advances leads to what is called big data. In other words, big data is the net of digital devices such as computers and mobile phones that create torrents of data which can be analyzed by organizations to gain valuable knowledge. There is evidence

---

<sup>7</sup> IBM - The Four V's of Big Data [www.ibmbigdatahub.com/infographic/four-vs-big-data](http://www.ibmbigdatahub.com/infographic/four-vs-big-data)

that big data has impacted any sector and has the ability to transform the business world (even traditional industries) and the world as a whole (Glass & Callahan, 2014).

By 2020, the amount of digital data generated will be more than 40 zettabytes or 5200GB (5,2 Terabyte) of data for every person on earth 2014, while the revenues of big data and business analytics market will reach \$260 billion in 2022 (IDC, 2018). According to predictions, the global data-sphere will grow to 163 ZB by 2025 with more than 25% of this data will be in real-time and more than 95% of this will be generated by Internet of Things (IoT) devices (Reinsel et al., 2017)

Data is the world's most valuable resource and the most important trend in the 21st century, since there are expectations that it will transform every human endeavour. Data has enormous benefits, but also risks. Data science is the new accelerated paradigm of discovery, but it is in its infancy and has yet to emerge facing fundamental challenges. Nowadays, the fast development of Internet-based technologies and the tremendous increase of connected devices through Internet of Things (IoT), lead to an exponential growth in the volume of datasets and thus this era has been renamed as “Big Data era”. Big Data is part of the “knowledge economy” and represents an emerging investigative field for businesses and organizations, since it provides the ability to gain valuable insights about customers views, preferences, needs, demands, attitudes etc. (Vecchio et al., 2018a).

The development of data collection, data analysis and data science are comparable to the different stages of internet development in businesses and organizations. During Web 1.0 companies were able to build a presence online, introducing their products and services to a wider variety of customers. Big Data 1.0 era was focused on establishing the means for data collection in enormous scale. Building or purchasing data warehouses, implementing the rudimental analysis skills and operational changes in companies was the first step of working with Big Data for many companies (Provost & Fawcett, 2013). Web 2.0 (social web) interactions between customers with businesses. The online communication was not one-sided anymore, since current and potential customers were able not only to get information about products and services, but also to communicate their issues, preferences and experiences via websites, social media and online surveys.

Information can play a crucial role in the performance of decision makers, especially in terms of the quality of their decisions. There are huge torrents of data that are available for enterprises and organizations to be analyzed. Therefore, it is significant decision makers to have the ability to gain valuable knowledge from big data. Using design science methodology, (Elgendy & Elragal 2016) developed a framework -“Big-Data, Analytics and Decisions” framework- examining how big data analytics can be integrated into the decision making process. Applying that framework in a retail experiment, it is found that from big data analytics implementation valuable knowledge can be extracted and leveraged to support conclusive decisions and enhance decision making process.

Big data analytics is widely acknowledged to play a significant role in enhancing business performance (Aker et al., 2016; David Kiron & Ferguson, 2014; Côte-Real et al., 2017) and capturing business value (Wamba et al., 2015). There is evidence in literature confirming the contribution of big data analytics in: price optimization and profit maximization (Davenport & Harris 2007), sales - revenue, profitability and market share (Manyika et al., 2011; Glass & Callahan, 2014; Brynjolfsson et al., 2011), return on investment (ROI) (Barton & Court, 2012) (McAfee & Brynjolfsson, 2012)), reducing cost, improvement in products/services and decision improvement (Davenport, 2014; Waller & Fawcett, 2013; Cheng et al., 2016).

## From Social Media to Big Data Analytics

*The number of people using social media networks are increasingly growing with estimations about social media users worldwide to be around 2.77 billion in 2019, up from 2.46 billion in 2017*<sup>8</sup>.

Social media (SM) is *“a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content”* (Kaplan & Haenlein, 2010). In broader sense, SM refers to *“a conversational, distributed mode of content generation, dissemination, and communication among communities”* (Zeng et al., 2010).

SM have risen, specifically in the last decade, with geometric expansion trend and currently have 3.196 billion active users with 42% penetration worldwide<sup>9</sup>. Every second, 973.000 users login on Facebook, over 1 million swipes on Tinder, and over 174.000 users scroll on Instagram<sup>10</sup>. The usage of SM creates rather large amount of information every day, such as user personal data, shared files and attitudes aimed to businesses – either positive or negative. SM have also revolutionised the means customers get in touch with these businesses and their services and play an important role in perceptions and buying decisions of customers. Therefore, nowadays, social media networks have transformed the relationships between enterprises, employees, customers and stakeholders.

SM can be defined as a *“tool which allows for the users to place and perceive information of different types and forms of publishing; thus it can be used as a tool for marketing purposes, which requires constant updating”*. SM can be an engine for obtaining the key asset – the attention of the public, this is the fundamental aspect of achieving company’s financial objectives. *“SM are continuously transforming – this applies to both the content and the technological basis and creates a demand for companies to be continuously up-to-date with the latest trends in content users”*. (Couldry, 2009; Hearn et al., 2009)

In current modern world society, the significance of social media networks is without doubt. Web 2.0 technologies and applications provided by social media networks enable the generation and distribution of information, opinions and experiences leading to an explosion of online user-generated content. Popular social media networks such as Facebook and e-commerce enterprises such as Amazon enable users to share their experiences and opinions about products and/or services through word-of-mouth (WOM) or customer reviews. Therefore, it’s extremely significant enterprises to understand the ways WOM can induce customer preferences and product sales (Zhang et al., 2012), since WOM causes unprecedented impact on business strategies and consumer behavior (Zhu & Zhang, 2010).

There are many open big data sources with SM to be one of the most representative. SM networks allow the analysis of the public's opinions about the image of an enterprise, its products/services, and the monitoring of its competition. The analysis can be very simplified, but for that, the proper tools needed. SM can be an asset for SMEs, since they allow the interaction between consumers and businesses creating value for both sides. They create proximity to customers and maintain a special relationship with them, while they also help to understand more precisely the customers’ behavior.

---

<sup>8</sup> Statista (2018) *Number of social media users worldwide from 2010 to 2021 (in billions)*

<sup>9</sup> <https://digitalreport.wearesocial.com>

<sup>10</sup> [www.socialmediatoday.com/news/this-is-what-happens-in-an-internet-minute-infographic/524426](http://www.socialmediatoday.com/news/this-is-what-happens-in-an-internet-minute-infographic/524426)



Enterprises use SM in their Customer Relationship Management (CRM), strategy, and organization activities, to market products, promote brands, communicate with their customers, and thereby develop their activities. They allow information sharing about customer expectations of business relationships, spreading information, and creating connectivity. This is called co-creation with customers via social media, since customers can collaborate with companies in the design of new products by offering new ideas, so-called innovation, they can also provide improvements to existing products. This co-creation is necessary for enterprises to achieve effective customer engagement using social media. This evolution has become a major strategic issue for enterprises as they attempt to satisfy the needs and requirements of their customers and not only at the level of product design or usage, but also at the quality, price and customer service, that are major goals for most enterprises. These data will also contribute to the product development, innovation and discovery of new opportunities.

The use of social media generates huge amounts of data on a daily basis, including consumer opinions, experiences and sentiments towards brands, products and services that is potentially of great value to enterprises (Kurniawati & Shanks, 2013), while big data can transform business models and improve the competitiveness of enterprises greatly (Bi & Cochran, 2014). In that context, social media analytics (SMA) can offer enterprises the opportunity to derive value by capturing, understanding and visualizing information on customer intelligence, changing customer tastes and preferences, campaigns performance, crisis responsiveness and influential users (Fan & Gordon, 2014). Nonetheless, social media analytics (SMA) faces some challenges related to the nature of social media big data, their collection and analysis techniques. Their process and management is not easy, since big social data is collected/stored in two forms: the structured data (or meta-data) including user profile characteristics, spatial, temporal and thematic data; the unstructured data comprises user-generated content that ranges from textual content (microblogs) to rich content includes audio and visual material (Stieglitz et al., 2014). The aggressive growth of social media has constituted to issues in data analysis techniques such as machine learning and data mining. Therefore, big data methodologies and perceptions are getting more and more popular (Bello-Orgaz et al., 2016).

SMA induces significantly every industry sector. (Abirami & Askarunisa, 2017) investigate the impact of social media on the healthcare industry. Using user generated content from social media (online reviews), they rank hospitals based on aggregated sentiment score. (Cheng & Edwards, 2015) using a visual analytic approach examine tourism-related social media posts about China providing consumers' insights for more efficient tourism marketing strategy. Travel and tourism related topics are found to be among the most popular topics in online social media networks (Miguéns et al., 2008), since travellers use the internet to gain information, share experiences, comment/review services, settle relationships with people from different destinations or buy travel products/services (Floris & Campagna, 2014).

## Location and Big data analytics in Tourism

Big data analytics (BDA) is “*a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high velocity capture, discovery and/or analysis.*” (IDC, 2012) or in other words, BDA is considered as a holistic approach that is able to manage, process and analyze the 5Vs of big data (volume, variety, velocity, veracity and value) in order to provide actionable knowledge for conveying sustained value, measuring performance and induct competitive advantages (Wamba et al., 2015). Analytics is the science of using data to build models that lead to decisions that add value to individuals or organizations. In other words, analytics is the discovery, interpretation, and communication of meaningful patterns in data. Organizations can exploit data analytics for several purposes such as the prediction and improvement of business performance, decision making, marketing optimization etc. Analytics includes social media analytics, social listening, comparative analysis, marketing analytics etc.

The tourism industry is a great application for social media analytics, since travellers’ data (opinions, reviews, ratings etc.) contain rich information for data analysis. Social media can provide real-time insights about tourists’ visiting patterns (Dhiratara et al., 2016). Tourism is a complicated industry in which travellers’ data, information and knowledge are vital determinants for destinations’ competitiveness and innovation (Vecchio et al., 2018a). Big data should be harnessed for data-driven strategic decisions and enhanced destination competitiveness. Tourism destination management is constantly transforming due to changing customer preferences/needs and technological developments. Big Data should be leveraged improving decision making process and enhancing destination competitiveness, but the field is not yet well developed (Miah et al., 2017). Understanding tourists’ perspectives in location and products/services offered to boost their satisfaction can reinforce tourism destination management (Floris & Campagna, 2014).

Vecchio et al. (2018a) using Apulia (Southern Italian region) as case study examined the value creation arising from big data associated with unusual experiences in a “*Smart Tourism Destination*”. They analyzed big data generated by social media through business analytics tools (Keyhole and Buzztrack) for supporting better planning and execution of social media strategies. Following (Secundo et al., 2017) conceptual framework, the big data social media analysis associated with the trip’s experiences in a destination and shared within a huge community of people live in the locations is able to export many value dimensions with specific techniques.

Floris & Campagna (2014) attempt to examine tourist preferences on tourism destination and services by extracting 880.000 reviews for Sardinia in five different languages coming from tourism social media networks of Booking and Tripadvisor. Applying techniques of spatial and statistical analysis to gain insights from tourists’ impressions on success factors that can be used in decision-making and tourist planning, they explore spatial dynamics of visitors’ impressions and their correlations with other variables. Their workflow of their empirical study consists of a three-step process: i) data collection and geocoding, ii) analysis of regional preferences dynamics to detect clusters and hot/cold-spots and iii) analysis of local preferences dynamics to get deeper insights on tourist options, regional territorial characteristics and tourism services quality in the examined destinations. It is found that the success of a tourist destination depends on the quality of the tourist offer but also on infrastructure, services and the territorial characteristics of the places.

The analysis of photos and images uploaded and shared on social media networks can contribute to tourism studies. (Donaire et al., 2014) analyzed Flickr pictures of a Pyrenees destination (Boí Valley). They conduct a content analysis of shared images on Flickr finding that 90,3% of the photos taken are outdoors and the rest 9,7% are indoors. In addition, through cluster analysis photographers were categorized into four groups depending on what they capture. More specifically, the groups are global tourist photographers (overall view of space), scenic (panorama mostly of natural spaces), detail hunters (spittled view of cultural areas) and monument lovers (churches). They propose for future research to follow this approach to analyze photos for other types of destinations such as urban or sun and beach destinations for clustering.

Hausmann et al. (2018) examine if social media can be used as an alternative way to traditional surveys to understand tourists' preferences for experiences in protected areas. Using data obtained from a traditional survey that was conducted in the Kruger National Park (South Africa) compared with observed preferences arising from over 13600 geotagged images shared on Instagram and Flickr by visitors of the park during the same period. They concluded that social media content can be a cost-efficient option to explore and monitor preferences for biodiversity and human activities in protected areas. (Marine-Roig & Clavé, 2015) highlight the significance of tourism user-generated content using a dataset of more than 100000 travel reviews in English by visitors of Barcelona for ten years. Using these data applying business intelligence (BI) for destination management can contribute significantly to develop marketing strategies, improve branding and provide policy implications among tourism and marketing organizations.

Many interactions between customers and enterprises are hosted in social media networks, whilst the significant advance of social media has turned online market into a conversation generating torrents of data (Chen et al., 2014). SMA contributes to the analysis of structured and unstructured data gathered from various social media channels (Vashisht & Gupta, 2015). The most popular social media platforms for analysis are social networks like Facebook and LinkedIn, microblogs like Twitter and Tumblr, media sharing like Instagram and Flickr, social news like Digg and Reddit and review sites like Foursquare and TripAdvisor. Their analysis can be separated into content-based applications with the text and its language are the most significant factors for identifying users' emotions, preferences etc. and structure-based applications with users' interests and relationships clustered into communities (Chen et al., 2014).

Nowadays, the majority of tourism processes and transactions (from trip planning, bookings to tourist feedback) are digitized, while the majority of tourists using the web and social media for travel planning and acquiring trustworthy information for their travel destination (Yoo et al., 2016). Therefore, an enormous amount of data generated at tourism destinations, identifying preferences and needs of tourists. Real-time analysis of SM data is a major driver for value creation in many industries (Vashisht & Gupta, 2015; Vecchio et al., 2018a). For instance, Park et al. (2016) examined perceptions of Asian restaurants on Twitter. Using text mining, word frequency analysis and sentiment analysis on 86,015 tweets over four months, they found that the sentiment scores of Chinese restaurants were significantly lower than others. The most positive tweets were about food quality, while negative tweets suggested problems about the service quality or food culture.

Several empirical studies have examined the user-generated content generated of social media in tourism destinations. (Chang et al., 2017) extract and visualize ratings and reviews for Hilton hotels in Tripadvisor using sentiment analysis and natural-language processing. They found the types of travellers that provided the lowest and highest ratings, the months with the lowest and highest rates and

the travellers' emotions according to the most frequently used negative or positive words. (Floris & Campagna, 2014) developed a methodology integrating TripAdvisor and Booking data to extract meaningful knowledge for tourism planning and decision-making. They investigated tourist preferences, such as the most popular destinations, reasons that people chose those destinations and what they appreciate/ignore. They found that the success of a tourism destination depends on the quality of the tourism industry and the territorial setting of the destinations. (Marchiori & Cantoni, 2015) examined the impact of the prior experience of a destination and the change in users' perceptions following exposure to user-generated content. Using a dataset of 2505 American Internet users, they found that people who are more educated or have previously visited a destination are less likely to change their opinions after being exposed to online social media content. Marine-Roig & Clavé (2015) highlighted the significance of Big Data analytics for smart destinations, examining the online image of Barcelona with a dataset of 100.000 online reviews written in English by tourists. They claimed that their analysis provided significant guidelines for the stakeholders, leading to better strategy, marketing and branding for the tourism destination.

The huge adoption of mobile devices has led to an explosion of data including user information and geolocation that are extremely valuable for marketing departments forming marketing strategies. However, the lack of technical skills does not allow the full exploitation of this data, while the data regulations create the need of filtering data that are accessible for leverage. Hence, a huge increase in use of LBSNs is observed, since users are able to express their opinions, report events, share reviews and sentiment, generating more and more user-generated content, while they are connected with others (Rathore et al., 2017). Recently, an explosive growth of LBSNs is observed, with the market leader Facebook currently having 2.06 billion monthly active users, while Instagram follows with 700 million active users and Twitter with more than 328 million users in 2017. This growth over the past years has resulted in the definition of the term Social Big Data that represents an enormous volume of data derived from social networks and their hundreds of millions of monthly active users worldwide<sup>11</sup>.

Location is dynamic and in combination with the rapidly increasing amount of real-time data coming from technology advances such 5G proliferation, enterprises will be able to get more insights by analyzing in the cloud. Location data plays a significant role in the successful exploitation of SM networks, thus *69% of company decision makers have implemented solutions to improve their location data, and 57% of them have implemented solutions for location analytics* (Belissent, 2018). Normally, SM networks analyse data by using keywords (often hashtags) but this does not give them full insight. By applying this method, they see what their customers or target groups are doing only if those groups use the keywords. So, it is possible for the network to not be aware of some keyword and hence the network will not include it into the analysis or its marketing. However, when location-based information is provided, the network has more insight into the behaviour of its customers. The networks can then analyse data not only by keywords but also by social media channels, datetime, target groups etc. For example, it is possible to choose one area of interest (one location) and see all the social media posts customers have uploaded from that location, not only the ones with keywords. This can show the network which customers are posting what and when, some keywords they are using, and the keywords the enterprise was not aware of. Then, the network can customize its offer, approach, targeting, marketing and more. But also, they can gain advantage in the market competition by recognizing locations where customers spend time.

---

<sup>11</sup> Statista (2017) "Most famous social network sites worldwide as of September 2017" [online]

Enterprises frequently apply computerized systems in location decision-making processes. One of the systems enterprises use is GIS or Geographic Information System. Pucha-Cofrep et al. (2018) define GIS as a set of tools made up of hardware, software, data and users. As stated in Fundamentals of GIS, the system allows businesses to represent, save, manage and examine digital information that helps them to create different maps. They also consider GIS as a computer model of geographic reality and its role is fulfilling identified information demands. However, GIS cannot serve as a decision-making system on its own. It is based on a geographical phenomenon and not on business matters. It helps businesses in acquiring data and creating maps, but it does not solve business problems or challenges as such. Nevertheless, when GIS, its data and maps are combined with BI or Business Intelligence, enterprises get important information. The information is mostly connected to the location or number of consumers with the highest income, the risk of natural disasters occurring where the company is located, the distance between the company and the important transport links etc.

To make key decisions based on location information, companies use Location Intelligence (LI). Location Intelligence developed from Geographic Information System and incorporates certain features of it into Business Intelligence and uses many factors and data to recognize risks or opportunities, identify patterns, decrease expenses or increase profit and so on. The information then makes the company more effective and helps in making better strategic decisions. Furthermore, the whole market based on location aspects is estimated to be worth 16.34 billion dollars by 2021 and by the same year it is estimated that 22.5 billion devices will be connected to the internet, each with a unique location. But, due to Big Data, many companies are battling to maintain all the data they acquire. It is believed that less than 1% of collected data is currently being used and the need to utilize location data will only become greater in the future. In other words, the future enterprise success depends on its potential and capacity to examine and apply location-based information. There are many enterprises that monitor social media for consumers generated content (posts, photos, etc.), while adding location data from consumers enables enterprises to react in real-time and get value in the long run. In order to exploit the full potential of data location, enterprises should carry out the following: monitoring business locations for shared social media content, identifying topics and sentiment of conversations, discovering time patterns and ensuring brand risk management through managing customers feedback, positive or negative (Harvard Business Review, 2018).

Location-based content can contribute to monitor user satisfaction in a dynamic way (Rybarczyk et al., 2018), since the evolution of location-acquisition technologies including GPS, Wi-Fi and 4G allow users share media content (texts, photos and videos) along with their position (location-tagged media content), transforming social networks to Geosocial or Location-Based Social Networks (LBSNs) (Rathore et al., 2017). “The use of locationally-referenced information as a key input in business decision-making” refers to Location Intelligence (LI) or Locational Data Analysis (Shekhar et al., 2017) adding a spatial perspective to data analysis that creates critical context to the decision-making process with the incorporation of powerful data correlation and visualization methods (Milton, 2011).

LBSN does not just add location as a new feature to an existing social structure so that people in the social network share information with a position tag, but also creates new social structures of individuals linked by interdependence derived from their physical world locations (Zheng, 2011). In a LBSN, a user's location is represented as a place e.g. street, shop, park, point of interest or building (Chorley et al., 2015) that is tagged in media content, like a post with text, picture or video providing information, to other users of the network informing them about the location of that post.

Data-driven maps and location-based applications created using LI reveal spatial relationships and correlations with other types of business data that otherwise may not have been visible. Combining this data with other types of geographic data such as population, traffic and weather provide opportunities for analyzing various spatially referenced phenomena and gain knowledge for businesses, organizations and public authorities. The analysis of data in both structured and unstructured forms, as well as the evolution in data storage, data processing, and data mining technologies lead businesses and local authorities on data-driven decision making and generation of non-obvious knowledge in real-time (Ravi et al., 2018; Vecchio et al., 2018a; Hashem et al., 2015). Enterprises and organizations can use location data to understand the significant impact of “where” in their operations. LI as a part of the whole data analysis process allows businesses or public authorities to better understand external characteristics and how these affect their activities and to gain a comprehensive overview of a phenomenon by integrating location and time dimensions with internal data (Milton, 2011).

Some prior studies have examined the role of location-based social networks in tourism. (Lee et al., 2011) used a dataset of geo-tagged Twitter messages from Japan to monitor geographical areas, finding that crowd activity can reveal expected and unexpected events. (Brandt et al., 2017) using 600,000 Twitter messages in San Francisco, examined the potential value of the spatial and semantic analysis of social media messages for smart tourism ecosystems. They found that social media analytics can reveal spatial patterns within the city related to presence, environmental and topical engagement, and these patterns contribute to value creation for smart urban tourism. (Zhou et al., 2015) proposed a method to find tourist hotspots through public Flickr images, demonstrating with the United States.

Miah et al. (2017) analyzed geotagged photos from Melbourne (Australia) shared in Flickr, demonstrate a method of analysing big data generated across social media networks for bringing more insights to enhance strategic decision making. Through this approach, destination management organisations and stakeholders can analyze and predict tourist preferences at specific places. (Psyllidis et al., 2015) demonstrated a web-based platform called SocialGlass that leverages data from several data sources including sensors, social media streams (Twitter, Instagram, Foursquare), open data from municipalities and resources from knowledge repositories. Their system using data science, semantic integration and crowdsourcing techniques can provide the mapping of demographics, people activities patterns, visitors’ preferences about specific areas in a city and destination popularity. SocialGlass offers urban analytics, interactive data visualization for data exploration, and comparison enhancing urban planning and decision making through the better understanding of cities’ dynamics.

## **Methodology**

Nowadays, big data and analytics is widely used, while data-driven approaches concentrate more and more attention from academia and almost every industry. Data visualization is the representation and presentation of data that exploits our visual perception abilities in order to amplify cognition, reveal the actual value and allowing new patterns observation.

Human visual system is extremely capable of identifying patterns, trends and relationships through processing visual information in contrast to textual or numerical form. Therefore, information that is presented visually can be extracted and analyzed by humans faster and more efficiently. Visual representation of information in form of maps, charts, images, graphs etc enables the understanding complex phenomena or detecting patterns within the large amount of data (Wexler et al., 2017).

Data visualisation exploits the high bandwidth that human brain processes visual signals. People reading text remember 10% of the information for three days, while visualizations are more likely to be recalled at 65%. Therefore, data visualization is better for transferring information than text blocks. Creating great visualisations is '*a carefully curated blend of art and science*'. Hence, a great visualisation should evoke opinion, and possibly even emotion (Ryan, 2016).

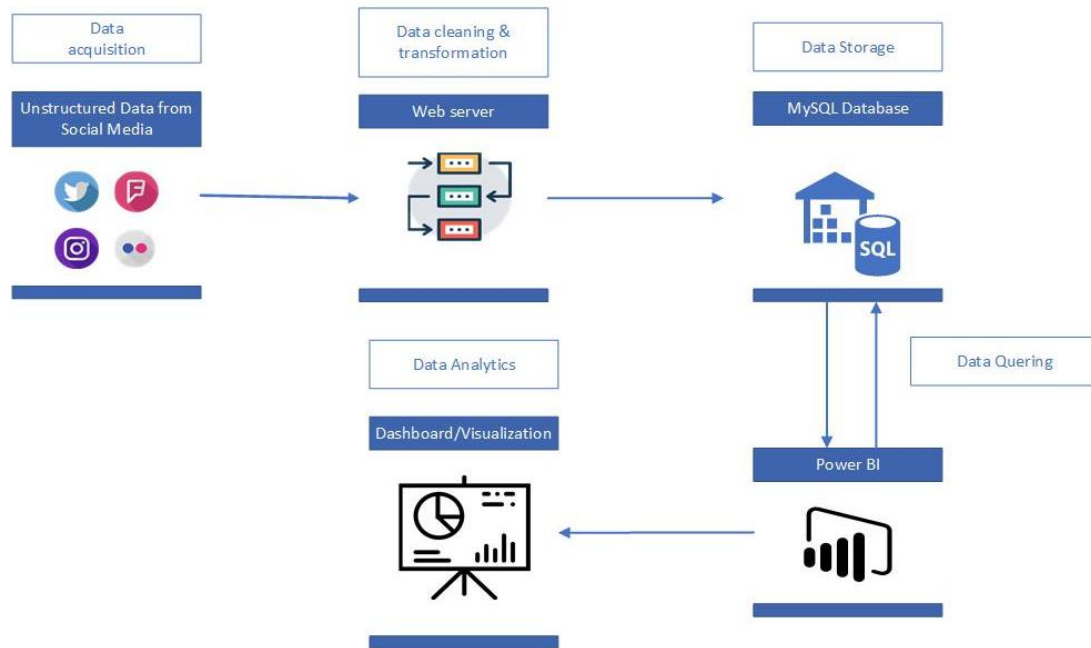
The substantial growth of data, Big Data, created the need for enterprises to adopt a data-driven culture, where data visualization is crucial, while the evolution of data every day leads on a need for new types of data visualization. Static charts provide significant insights, but they are not able to present all the knowledge of the underlying layers of data sets. Interactive visualisations increase the ability for exploration and navigation through the data providing the '*fourth visual dimension*' (Ferster & Shneiderman, 2012). Therefore, understanding, discovery capability and communication of the insights is significantly improved (Zudilova-Seinstra et al., 2008; Packham et al. 2005).

In management, visualisations assist decision makers to understand data and obtain useful information easily and quickly (Morabito, 2015). Therefore, visual analytics approaches are getting more and more popular but are not frequently used in the tourism literature. (Cheng & Edwards, 2015) demonstrate that a visual analytic approach that is primarily relied on manual processing of text, sentiment and image, can provide insights such as tourists' attitudes about the impact of travel policy news on consumers in China. (Carvalho et al., 2017) present a platform and framework named MISNIS (Intelligent Mining of Public Social Networks' Influence in Society) dealing with the issues of collecting, storing, managing, mining and visualizing Twitter data allowing any user to easily get a given subject from a very large dataset of tweets, related contents and indicators such as user impact (influence) or sentiment analysis.

Exploitation of user-generated data shared in location-based social networks can be extremely beneficial for many parties including local authorities, organizations and citizens and it can be life-critical. Web 2.0 applications and tools transforming the role and behavior of travellers, make tourism stakeholders reorganize their operations and business models based on these data (Christou, 2016). However, harvesting real-time spatial data from social media networks such as Twitter, Instagram and Flickr is challenging, since all the users of these social media networks generate an enormous volume of data (Christou, 2016). Therefore, a data analytics system is required to process offline data efficiently within a time limit and provide real-time data analysis for various social networks, including Twitter (Rehman et al., 2013; Carvalho et al., 2017; Komorowski et al., 2018; Liu et al., 2018), Flickr (Liu et al., 2018; Li et al., 2018; Donaire et al., 2014), Instagram (Schmidbauer et al., 2018; Mittal et al., 2017) and Foursquare (Mittal et al., 2017; Liu et al., 2016; Wang et al., 2015).

For the data analysis process, the system using public user-generated content from the LBSNs consists of the main sub-processes: a) data acquisition, b) data cleansing and storage, c) data querying and filtering and d) data visualization (Mahmud et al., 2016; Floris & Campagna, 2014) (figure 35). The selected LBSNs, that are used in this study are Instagram and Flickr (media sharing) and Foursquare and Twitter (textual posts and reviews). For the data acquisition process, the system communicates with the LBSNs repeatedly through their RESTful Web Services (RESTful APIs) to collect all the new generated data in an unstructured form. After acquisition, data cleansing and transformation is needed to select the required data and transform them into a structured form to be stored in the SQL database and transferred to a data warehouse in the cloud. Querying a table in a data warehouse is fast since it is optimized for analytic access patterns and processes highly complex queries overall data (Bouadi et al.,

2017). Afterward, data from each executed query is filtered to be visualized (Ferreira et al., 2013; Mahmud et al., 2016).



**Figure 35:** The data flow of the analysis

Valuable insights about people's preferences can arise from their behavioural data (Glass & Callahan, 2014). The user-generated content of social media networks is mostly in unstructured form i.e. photos, videos, text. Therefore, big data analysis techniques like text mining are used for insights/knowledge extraction through them. Therefore, user-generated content with geo-tagged tag shared in the popular social networks of Instagram, Flickr, Foursquare and Twitter is collected, processed and analyzed.

The data acquisition consists of the communication of the web application programming interface (API) of the social media networks used periodically to acquire new geo-tagged user-generated content (data). For acquiring data from Instagram, a crawler to scrape Instagram's explore page using Location-Ids as the parameters was required (Dhiratara et al., 2016). The data cleaning and transformation of acquired data is the process of storing the necessary data and transforming them from unstructured into structured format, to be stored in our SQL database and data warehouse.

The selected social networks provide unstructured data (photos and texts) with location tag generated by users. For the analysis of these data and the extraction of valuable knowledge big data techniques are used. In more detail, for text mining, Google Cloud Natural Language API and machine learning models are used for sentiment and entities/labels analysis. For image/photo analysis introduced in this study, Google machine learning model (Google Cloud Vision API) is used for the detection of objects/labels within the acquired photos collected by Instagram and Flickr. For data visualizations Microsoft Power BI and ESRI ArcGIS are used achieving the presentation of data analysis results in an interactive and easily understandable way.



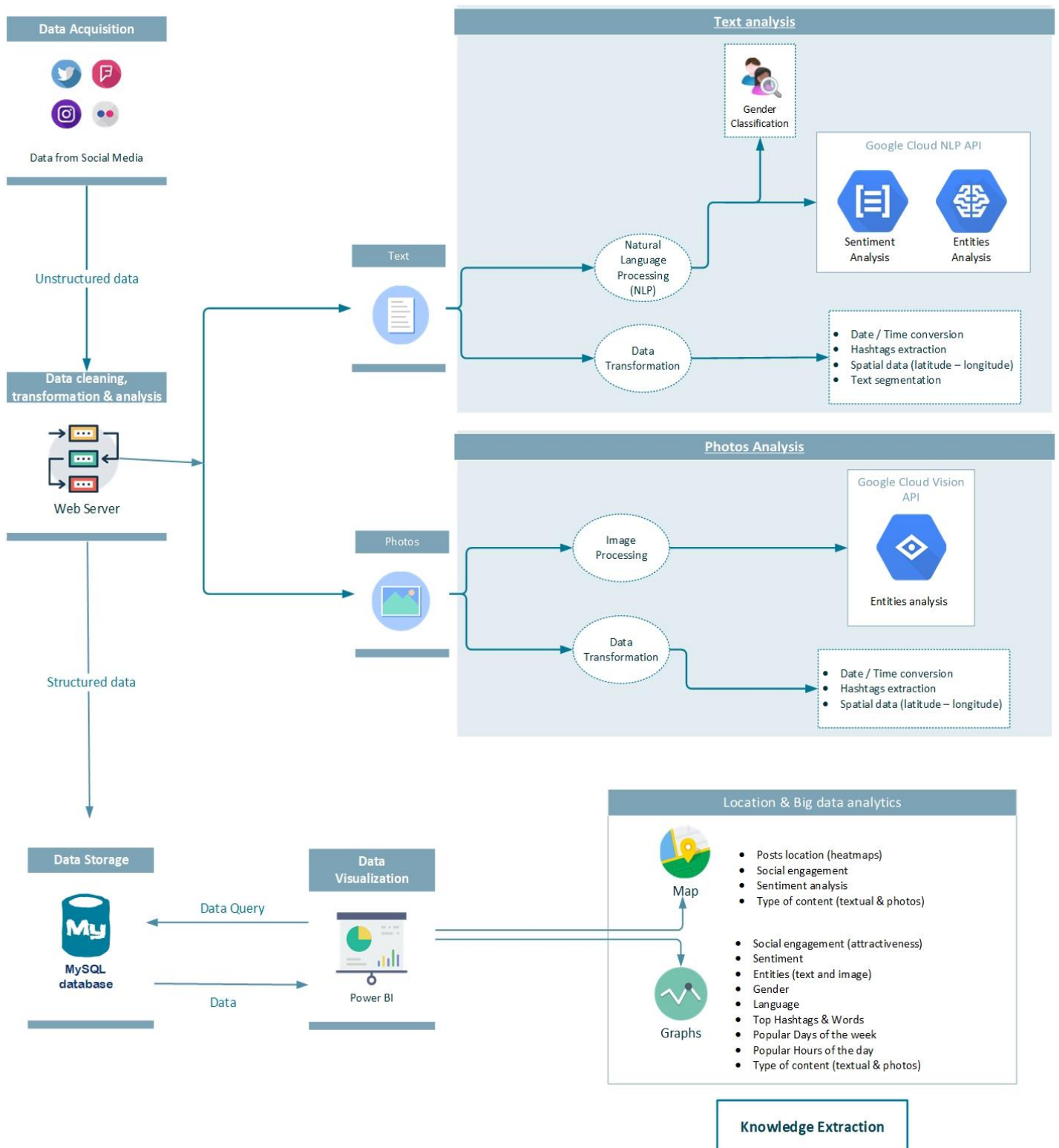
## **Data collection and analysis - a case study for cities of Crete and Cyprus**

The most popular conventional approaches rely on tourist participation and manual recordings that have limitations like high cost, limited answers and being time-consuming. Contrary to these, online social media networks can be used as a cost-efficient data source for capturing tourist behaviour and flows (Vu et al., 2018) prevailing limitations of survey and opinion polls (Veal, 2017).

The islands of Crete and Cyprus were used as a case study and the examined time period was between June 2018 and September 2018. The existing tourism research studies for the examining regions that are based on traditional approaches e.g. surveys, are for Cyprus (Clerides & Pashourtidou, 2007; Boukas & Ziakas, 2013; Farmaki, 2014; Farmaki, 2016) and for Crete (Andriotis, 2011; Hosni et al., 2018; Zouganeli et al. 2012). On the contrary, this study introduces an innovative approach for knowledge extraction arising from unstructured social big data combining location and big data analytics techniques. The examined regions are two of the largest islands in the Mediterranean Sea, the island of Crete (Greece) and the island of Cyprus. Crete is the fifth largest island and Cyprus the third largest island in Mediterranean basin. Both islands economies rely on tourism industry and on both islands the majority of tourism enterprises are SMEs.

In this context, data acquisition for the case study includes the tourists' preferences and perspectives arising from the user-generated data shared in the most tourist - popular cities of the two islands (Floris & Campagna, 2014) for the examining period of June 2018 to September 2018 (highest tourism season). After this phase, the user-generated unstructured data shared in location-based social networks is cleaned, transformed and analyzed. For the analysis of textual content, big data analytics techniques are used (gender classification, sentiment and entities analysis), while digital image processing is used for the analysis of media content (photos). For both types of content (data), data cleaning and transformation of unstructured to structured data enabling the extraction of knowledge from LBSN big data through data visualization. Interactive data visuals are able to present in an efficient and clear way unhidden patterns and insights for tourists' perceptions, spatial-temporal behavior and demographics in the examined tourism destinations. Data visualization in visual graphs and maps enabling the knowledge extraction from LBSN big data of tourists in these destinations.

More specifically, the interactive visualization of the analyzed big data for the examining tourist destination cities is in graphics (insights about tourist temporal characteristics, perspectives and demographics) and in maps (insights about tourist spatio - temporal patterns). This is the first study that spatial analytics (location intelligence) is combined with sentiment analysis, type of user-generated content and social engagement (big data social media data analytics), whilst a combination of different content type location-based social media networks are used.



**Figure 36:** The innovative approach for knowledge extraction through LBSNs

## Results and discussion

It gets to the fundamental purpose of a visualization to convey a point (Whitney, 2012) for information discovery (Ware, 2013) (Schmitz et al. 2015) or/and to aid in decision-making (Elouni et al., 2017); (Moore, 2017; Elouni et al., 2017). Tourist destination image is significant in tourists' decision-making behavior and influencing tourist destination competitiveness (Camprubi et al., 2012). User generated content shared in social media networks include text, geolocation, images or videos. The information that can be extracted from these can fuel decision making through innovation and value creation (Prodromou, 2014; Moore, 2017).

Translation of enormous amounts of data into a comprehensible form for managers is a major research challenge (Chen et al., 2012), with data visualizations be the right tool dealing with that (Huang et al., 2017). The data acquired from the LBSNs used in this study for urban areas of Crete and Cyprus during the examining period of three months (June to September 2018), as summarized in Figure 37.

The acquired data for this study are user-generated data shared in the most popular urban areas of Cyprus and Crete. More specifically, as it is presented below, the examining cities of Crete island are Chania, Rethymno, Agios Nikolaos, Heraklion and Yerapetra, while the examining cities of Cyprus island are Ayia Napa, Larnaca, Pafos, Protaras and Limassol.

City	Total Users	Total Text Posts	Total Photos	Total Engagement
Ayia Napa	17551	1072	78210	10342193
Larnaca	7965	608	32429	3478253
Pafos	6895	784	22240	1967212
Protaras	4456	396	18416	2019856
Limassol	3223	598	27590	3959445
<b>Total</b>	<b>40090</b>	<b>3458</b>	<b>178885</b>	<b>21766959</b>

City	Total Users	Total Text Posts	Total Photos	Total Engagement
Chania	16477	2703	59383	6150476
Rethymno	14788	1170	64044	5986406
Agios Nikolaos	12235	470	41755	7307735
Heraklion	11270	796	38263	4814525
Yerapetra	912	42	4129	1531457
<b>Total</b>	<b>55682</b>	<b>5181</b>	<b>207574</b>	<b>25790599</b>

Figure 37: Summary of user-generated data per city and island

A summary of the total shared content for the two islands per month and type (figure 38) indicating the most popular months of each island for each type of content (textual post or photo), while a comparison between two islands can be achieved. In a nutshell, the most popular month of Crete for both types of content are August, whilst for both types of content July is the most popular month for Cyprus. Between two islands, it is observed that Crete is more popular during June and August, while Cyprus is more popular than Crete in July.

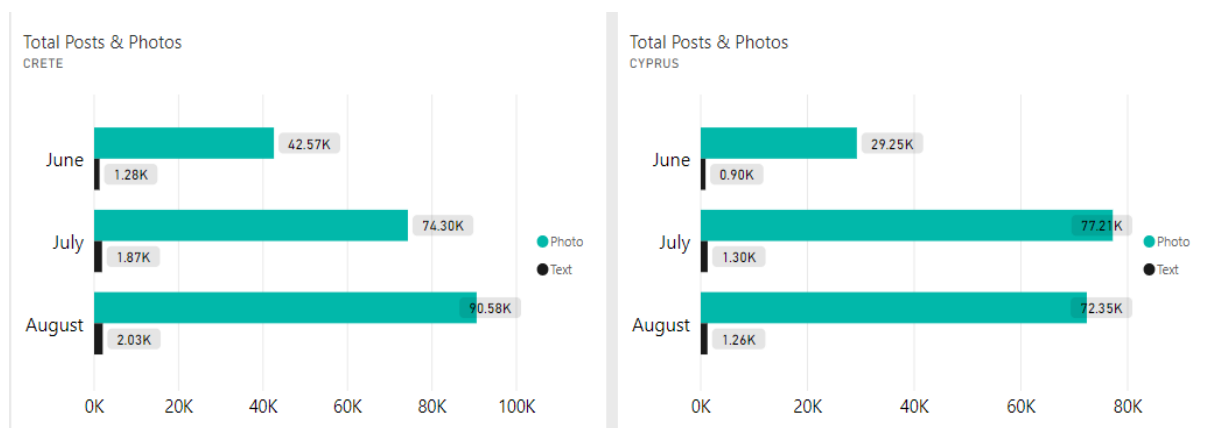
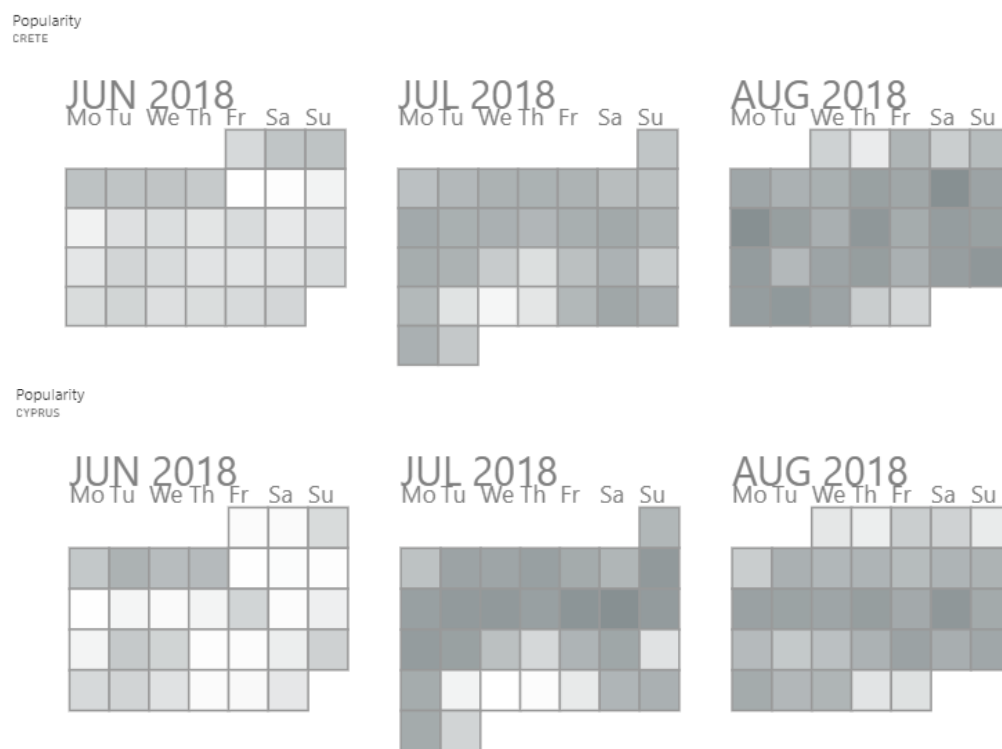


Figure 38: Type of content per month and island

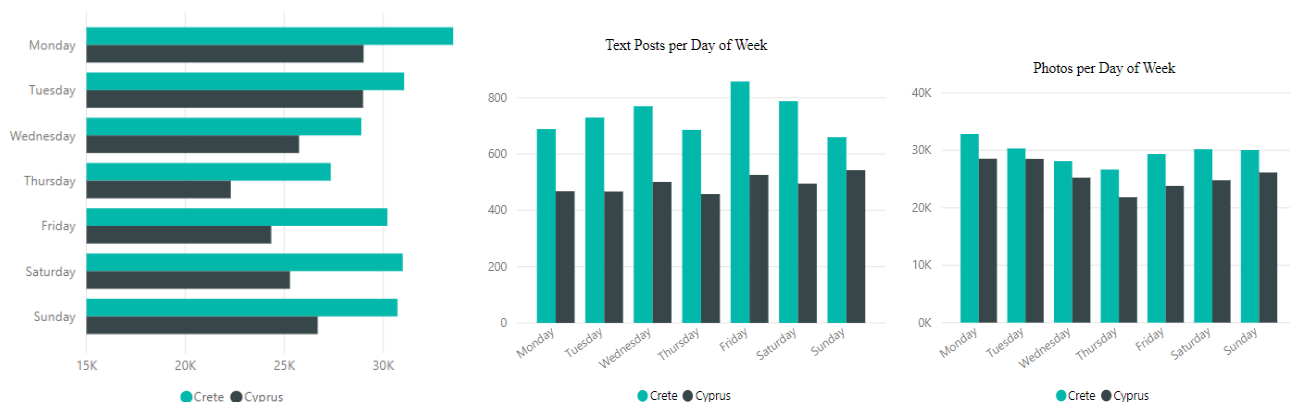
## Temporal Big Data Analytics

The analysis of user-generated data can provide significant information and uncover patterns about their temporal activity in tourist destinations like Cyprus and Crete. At first, the most popular days of the examining period for each island (figure 39) with the most popular days be more intense than others. Thus, Crete's most popular days are mostly during August, while Cyprus most popular days are in July and in one week of August.

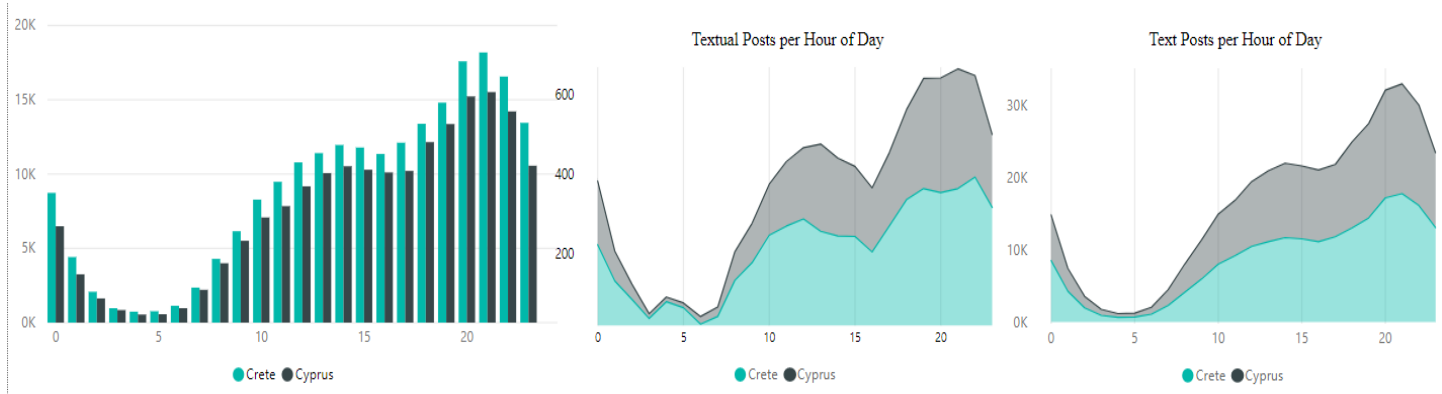
These insights can be taken into account from tourism organizations / enterprises such as travel agencies propose better itineraries satisfying more tourist needs, restaurants for efficient scheduling and promotion - giving incentives like discounts for selling their products/services in the less popular periods.



**Figure 39:** Frequency of total posts per day and island



**Figure 40:** User-generated content per day of the week, island and type of content



**Figure 41:** User-generated content per hour of the day, island and type of content

Tracking the location of tourists can be a key indicator of tourist density (Salas-Olmedo et al. 2018) on the days and times tourists visit specific places. Using BDA tourist hot spots can be identified as well as the time frame in terms of seasons preferred by tourists. From the user generated posts in location-based social networks for the examining tourist destination, temporal insights can be observed. The most popular days of the week and hours of the day for each type of content (textual and photo) can be defined and a comparison between two islands can be achieved.

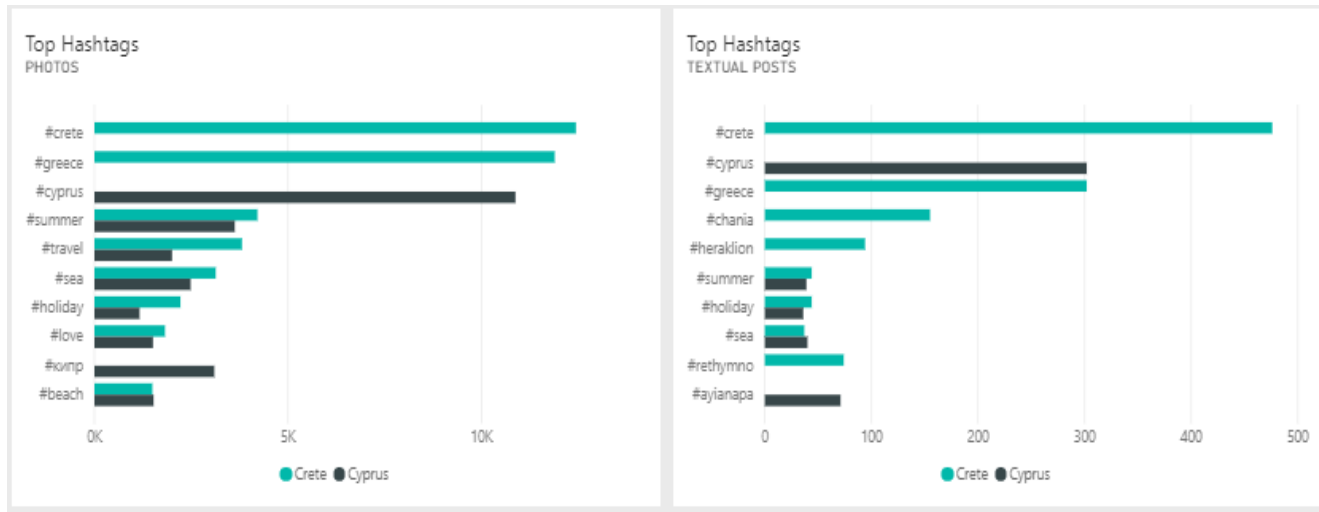
At first, the most popular days for the total generated content for both Crete and Cyprus (figure 40) are Monday and Tuesday, whilst the least popular day of the week is Thursday for both islands. Per type of content, Crete and Cyprus most popular day of the week for photo sharing are Monday and Tuesday, while for textual posts Crete's most popular days are Friday and Saturday and for Cyprus Friday and Sunday.

Additionally, the analysis provides insights about the popular hours of the day for both types of content and for each tourist destination (figure 41) thus it is observed that the most popular hours for both islands are afternoon hours after 5pm., whilst the less popular hours are after midnight to the morning hours. Studying the spatial and temporal features of tourist based on geotagged data leads on the extraction of significant knowledge/information about travellers' patterns. Based on this knowledge, tourism enterprises and stakeholders such as travel agents can be prepared to improve existing products /services or design and launching new data-driven travel packages with SMEs adjust their resources wisely (Silva et al., 2018)

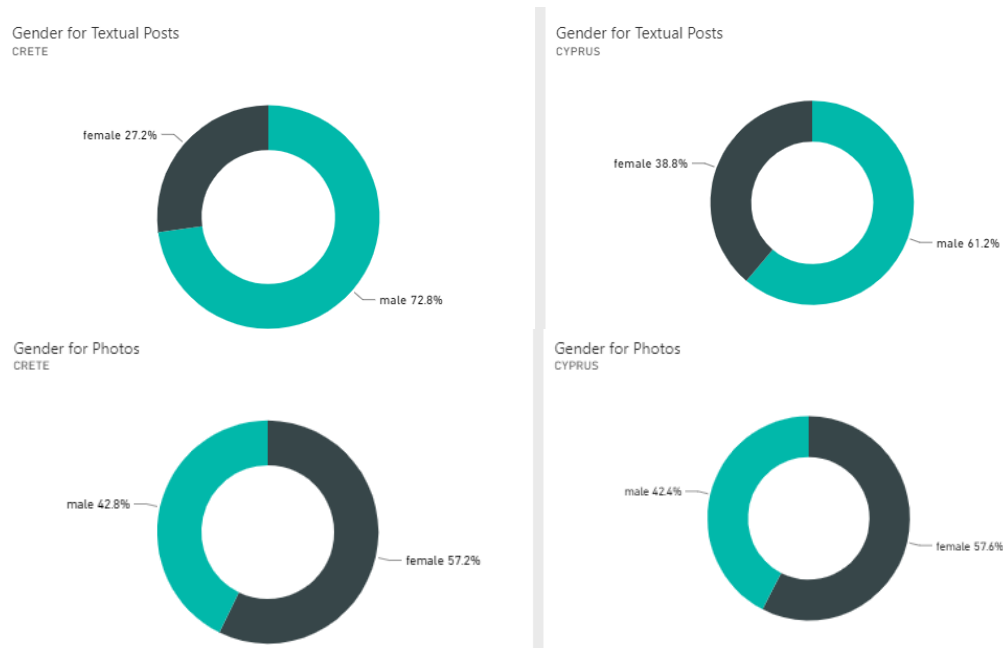
The growing use of social media has developed words and hashtags of users as valuable assets that can be utilized for attracting and engaging the right audience. Posts that contain at least one hashtag present better engagement, since nowadays most people use hashtags in social media as a resource when searching for products/services or information about specific topics/locations. Therefore, social media has become very popular allowing customers to gain the information they are looking for.

Enterprises and organizations are able to reach a broader audience using relevant hashtags that are coming from detailed research and experiment to find the right words to reach the customers. In addition, finding the trending (most popular) hashtags provide them the topic that customers are interested in. The creation of hashtag-driven content is an efficient way to brand building and engagement through promotional material about products/services and targeting the right audience. Brands engage users and collect ideas with campaigns and contests that users generate content using specific hashtags for entering a promotion or contest.

It is observed (figure 42) that the most popular hashtags for users shared photos in both islands are related with the island and country (#cyprus, #crete, #greece) and their status in these eg. #holidays, #sea #beach indicating that the majority of travellers enjoy their summer holidays on the beach. In the top hashtags from textual posts by users, it is observed that users share again hashtags related to the islands and countries, but also, they share hashtags about the cities they are eg. #chania, #heraklion, #ayianapa. These insights about the hashtags shared in each type of social network (photo-sharing or text-sharing) can be exploited for more efficient marketing campaigns, since destination managers or marketers can use them to approach a wider audience with more efficient content.



**Figure 42:** Top words and hashtags per type for each island

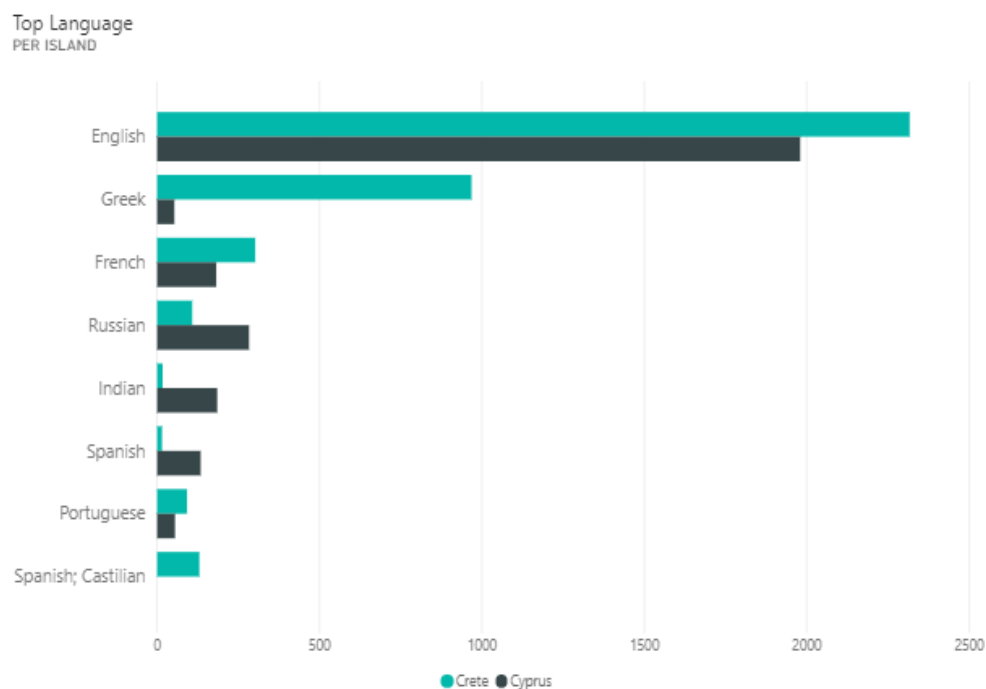


**Figure 43:** Distribution by gender of users generating content per type and island

At each tourist destination, it can be observed that there are variations in the gender sharing content in LBSNs. More specifically, in both islands (figure 43) it is observed that men mostly share textual content, while women share more photos. These demographic features in combination with other information can be beneficial i.e. for tourist destination managers and marketers in order to design and execute a marketing strategy that is more personalized for enhancing travellers visiting experience.

In order to make sense about the nation of travellers in the examining tourist destinations, text mining methods (natural language processing) used for detecting the language of textual user-generated content. Hence, the top languages of users' generated content are defined, indicating the differences of tourist language / nation visiting the tourist destinations.

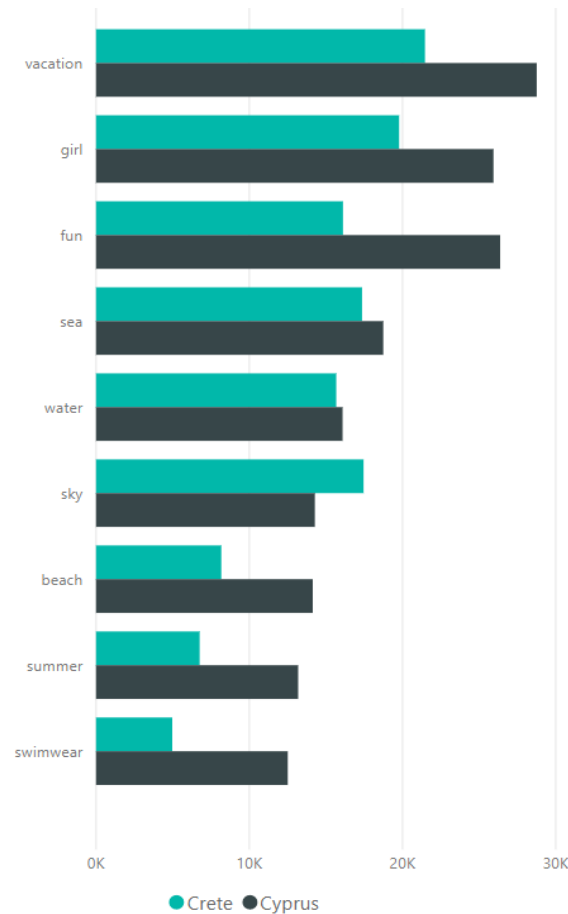
The most popular language for both islands is English, whilst the second popular language is Greek for Crete and Russian for Cyprus, indicating the origin of travellers (figure 44). The analysis of the massive user generated data enables the gathering of quantitative information about tourism destinations that are significant for the image and management of a destination.



**Figure 44:** Top languages per island

Using social media networks, tourists post an enormous number of photos. Photos with geotags that are published on photo-sharing social media networks like Instagram and Flickr provide opportunities for transforming information into knowledge. Content that is accompanied with photos offers rich information about tourist preferences and experiences, but many photos are shared without textual tags or descriptions. To discover tourist's experiences of specific places, visual (photo) processing approach, as introduced before, is used. In that context, entities detection using AI algorithms provide objects/entities that are photographed by users and shared in photo-sharing LBSNs.

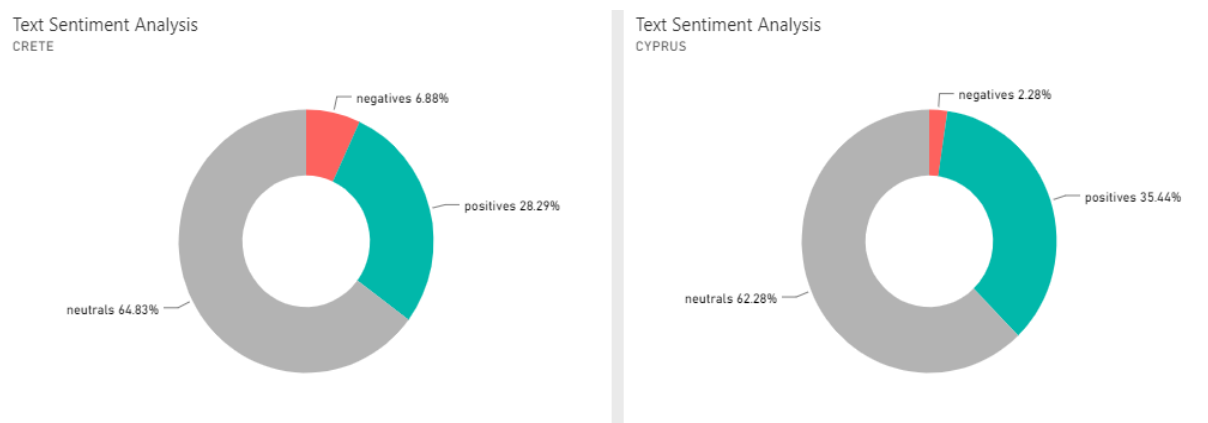
From the analysis, insights about the gender of the people that are photographed, their feelings and activities, their sentiment, location characteristics can be observed. Tourists are interested in taking and sharing photos having fun during their holidays (figure 45). The photos taking position is usually near the beach or sea/water. In addition, the most photos contain people and they are females.



**Figure 45:** Top photo entities per island



One significant contribution of this study in big data analytics applications in tourism and hospitality is the integration of location with big data analytics. Textual analysis provides significant insights about users' (tourists) sentiments (Chang et al., 2017) and the content of their discussions in a destination, whilst the negative posts can provide discovery of existing problems and directions of improvement. The knowledge extraction about the feelings of visitors for specific locations can be derived from their shared posts in the location-based social networks with sentiment analysis (Park et al., 2016). Therefore, the integration of these information with spatio-temporal data is able to uncover patterns and provide more insights in geographical dimension of the destinations examined (Floris & Campagna, 2014).



**Figure 46:** Text sentiment and gender analysis per island

For both islands it is observed that positive sentiment dominates compared to negative during the period examined (Figure 46). The most textual posts related with the cities of Crete island are positive and about the tastefulness and quality of traditional food and the beauty of some places, while for Cyprus the most positive textual posts are about the beauty of places in the cities and about events and activities eg. dancing, diving, walking.

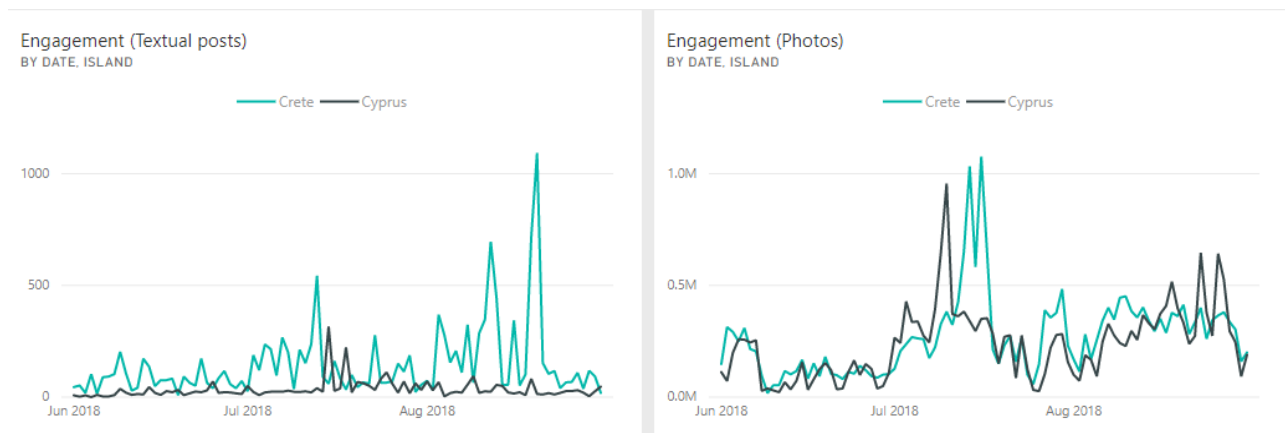
These insights are important for understanding the perceptions and preferences of travellers in specific places. For instance, local authorities and hotels can exploit the knowledge that tourists appreciate more traditional food, thus they can organize events and activities that can attract more visitors adding significant value to their tourist destination.

Digital transformation has introduced geo-marketing that enables leveraging geo-location data in various aspects of marketing overcoming challenges and providing the ability offering better customer experience. The analysis of engagement can be used for the evaluation of the attractiveness of tourism experiences associated with an event and can be a significant factor for more personalized offers focused on customer satisfaction (Vecchio et al., 2018a; Harrigan et al., 2017).



**Figure 47:** Maps of Crete - Cyprus / Location and type of the most positive and negative posts

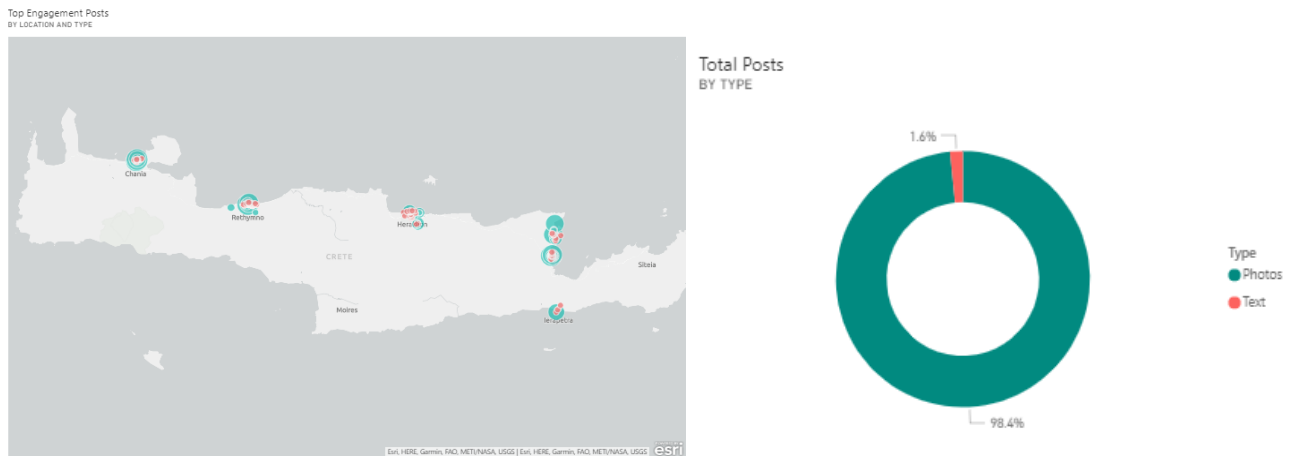
Customer or user engagement has become another key method in social media analytics, since social media users interact more and more with brands, products and organizations through several social media networks. Therefore, brand awareness and enhancement in the digital word by leveraging the power of social media play a crucial role in marketing.



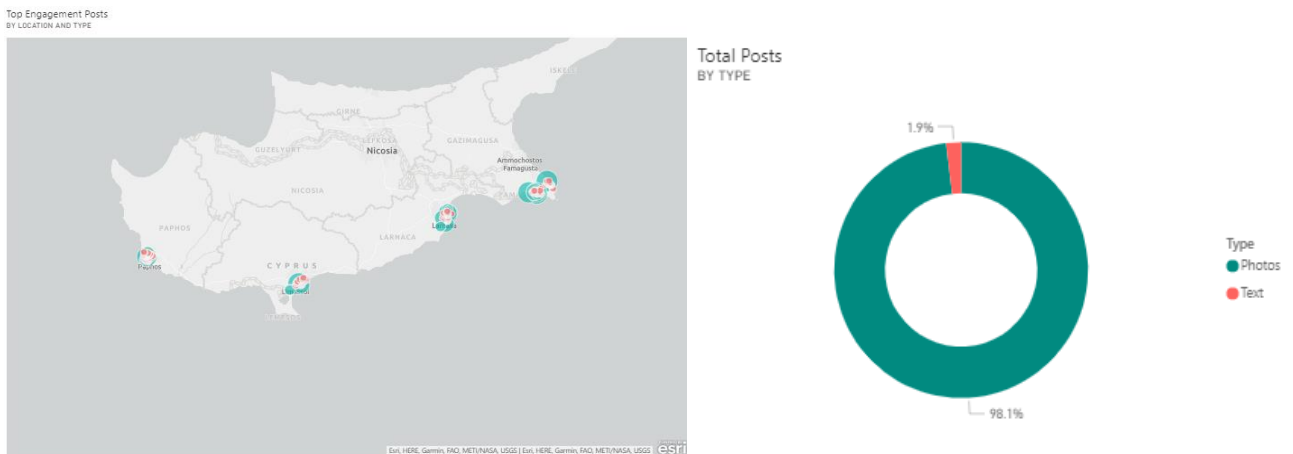
**Figure 48:** Total social engagement per type of content and island

In the same direction, social media engagement (likes, shares, and comments) in tourist destinations depicts their attractiveness not just from the travellers but from potential tourists/travellers that are impressed from the content that is shared. This is considered as online word of mouth marketing (e-word of mouth) and can enhance significantly a destination branding and attractiveness. To gain more attractive and efficient online presence, destinations or tourism enterprises/organizations can exploit influencers to share content to their online communities (Uzunoğlu & Kip, 2014).

From the social engagement arising from the user-generated data from the location-based networks, the destination of Crete has higher social engagement than Cyprus during the examination period (figure 48). During July, some peaks are observed for Crete and Cyprus, while the deeper analysis in the type of content and its social engagement is shown that is arising from users media (photo) sharing.



**Figure 49:** Map of Crete - Location and type of top engagement posts / Distribution of posts per type



**Figure 50:** Map of Cyprus - Location and type of top engagement posts / Distribution of posts per type

Using spatial analytics the locations of the examining cities with higher engagement, and hence more attraction for location-based social networks, can be observed (figure 49 and 50). The places can be used as destination branding material, while the type of user-generated content can provide information about the content the travellers share, thus in both islands the majority of travellers prefer to use photo-sharing social networks for sharing their experiences. These insights can be useful for instance in tourism enterprises like travel agencies and DMOs can connect their marketing campaigns with photos from these places, enhancing their efficiency and conversion rates.

## Conclusion

Due to digitalization and the constant shift of human life into the virtual space enhanced by technological progress such as smartphones, proliferation of high-speed mobile networks (Smith et al., 2012), humans create more data day by day. In this context, current time period is characterized as the era of data and information explosion (Breur, 2016; Demirkan & Delen, 2013). The knowledge derived from Big Data analytics can have an influence on any industry and tourism is considered as one of the industries that has the ability to benefit greatly from Big Data analytics due to the voluntary virtual actions of tourists, such as uploading photographs, paying electronically or interacting on social media.

As the digitalization trend continues, it is likely that smart tourist destinations will emerge. They will rely on “the interaction of a destination with the community of stakeholders, residents and tourists, based on dynamic platforms, knowledge intensive communication flows and enhanced decision support systems” (Vecchio et al., 2018b). It can be assumed, that the access to data will become even easier and the information more precise, which will strongly increase the accuracy of forecasts. Nonetheless, at the tourism industry, the potential of Big Data is not excessively utilized. Tourism/hospitality industry struggles with innovative approaches that have the ability to gain value leveraging big data (Mariani et al., 2018).

The opportunity to express publicly opinions with no spatiotemporal limitations through social media networks, changed the way companies deal with customers in a tremendous way. As the touristic product is defined by fulfilling needs and satisfying desires, the hospitality and tourism sector, more than other economic areas, have to maintain a pleasurable customer experience. Bad reviews and critiques can harm an enterprise significantly. The upload of a photo as evidence of receiving a bad service can have a negative impact on how new potential customers will consider a booking/buy. The power shifted from the enterprise to the customers by means of transparency. Dissatisfaction and discomfort is visible for everyone connected to the web and scandals can go viral, that results in huge negative publicity and loss of reputation. Nowadays, it's important to pay attention to details, since reputation mostly relies on word-of-mouth communication. The strongest benefit of getting reviews from a very large sample is getting detailed information about almost every customer experience.

The traditional surveys conducted with questionnaires or observing hotel or sights admissions are very poor sources of information in comparison to the virtual footprints. The self-expressional character of social media activities that rely purely on an individual's desire to present him or herself is the catalysator of Big Data (Salas-Olmedo et al., 2018). Online user-generated content and new technologies availability allow researchers approach travellers' perceptions and level of satisfaction (Alaei et al., 2017). Enterprises and tourism stakeholders can adapt, calibrate and enhance certain criticized parts that bother the majority of their clients, while also see which aspects of their strategy and business model worked out well. They can also learn in which direction to innovate and expand their activities, facilities and services.

Tourism research can be developed into a new area, where the combination of theoretical approaches and data-driven practices can lead on understanding or explaining phenomena and discovering new dimensions in theories (Alaei et al., 2017). Under the digitized era, new opportunities for interaction can emerge for enterprises, organizations and tourists (customers) through social media networks. As more user generated content is created every day, more insights about consumer behaviour can be gained. Therefore, tailor-made strategies can be developed to fully cover the needs of customers and

thus improve the effectiveness of their marketing actions, while overcoming issues and innovating in promising directions.

Prior studies have attempted to capture and analyze tourist activities and preferences for the improvement of strategic decision making and planning in tourist destinations. Conventional approaches such as travel diaries present limitations, while online social media can capture the spatio-temporal behaviour of tourists at large scale. Therefore, the present chapter attempts to introduce a novel methodology for the extraction of additional and valuable knowledge coming from user-generated content shared in popular location-based social networks via mobile in the examined tourist destinations. Integration of location intelligence with big data analytics provide more insights allowing tourism enterprises and stakeholders to better predict travellers' behaviour and allow efficient resources allocation. By gathering and analyzing LBSN big data, stakeholders can gain valuable knowledge about tourists' and travellers' patterns, preferences and needs. Results of analysis for both islands provide knowledge about tourists trends in places among dates and type of content (photo or textual), tourists sentiment (most positive about traditional food and beauty of places in Crete and most positive about beauty of places and activities in Cyprus) in specific locations, users engagement per type of content across specific destinations and topic/entity detection in generated content.

Taking insights about which is the most popular location or region, what attracts tourists more and why, what are the most popular and preferred days-hours to share each type of content, where and why do tourists like to visit, which is the more attractive location or region and what are the characteristics of tourists, make tourism organizations, enterprises and destination managers capable of design, implement and develop knowledge-based products and services using interactive data visualizations enabling the deeper analysis without any expertise required. In addition, under smart city research, the improvement of strategic decision process of local authorities by understanding people's activities, preferences, experiences and flows are extremely significant for smart city stakeholders. The comparison between two tourist destination that present similar characteristics also can provide insights about practices that are implemented successfully in one and can be adopted by the other.

This research provides some implications for tourism destination stakeholders and contributions for academia and tourism / hospitality industries. However, we need to acknowledge some of the limitations of the study. First, a limited number of social media networks are used for the research. Even though, the research relies on an extensive data set, but this only captures three months period and the urban cities of the two islands (not including rural or coastal areas). In the research, user-generated content with location -tag is used, but not all tourists visiting a destination take photos or write posts and share them online. There are still gaps in demographics characteristics in social media such as age, educational level and income (Li et al., 2018). Therefore, additional research streams are identified. User-generated content shared in other social media networks can be used additionally in the analysis. A further development of the research can include the analysis of more and various types of big data e.g. sensor generated data to diagnose new patterns. The study provided descriptive results through big data analytics e.g. sentiment analysis, text mining and location intelligence, future studies could use other big data analytics techniques to identify new patterns.

# References

- Abbasi, A., Rashidi, T.H., & Waller, S.T. (2015). Utilising Location Based Social Media in Travel Survey Methods. *Proceedings of the 8th ACM SIGSPATIAL International Workshop on Location-Based Social Networks - LBSN'15*. Available at: <http://dx.doi.org/10.1145/2830657.2830660>.
- Abirami, A.M., & Askarunisa, A. (2017). Sentiment analysis model to emphasize the impact of online reviews in healthcare industry. *Online Information Review*, 41(4), pp. 471-486.
- Adetunji, C., & Carr, L. (2016). Knowledge discovery from social media data: a case of public twitter data for SMEs. *eKNOW 2016, The Eighth International Conference on Information, Process, and Knowledge Management*, Venice, Italy. 24 - 28 Apr 2016. pp. 119-125
- Akter, S., Wamba, S.F., Gunasekaran, A., Dubey, R., & Childe, S.J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment?, *International Journal of Production Economics*, 182, pp. 113-131.
- Alaei, A.R., Becken, S., & Stantic, B. (2017). Sentiment Analysis in Tourism: Capitalizing on Big Data. *Journal of Travel Research*, 58(2), pp. 175-191.
- Andriotis, K. (2011). A comparative study of visitors to urban, coastal and rural areas. Evidence from the Island of Crete. *European Journal of Tourism Research*, 4(2), pp.93-108.
- Andriotis, K., Agiomirgianakis, G., & Mihiotis, A. (2007). Tourist Vacation Preferences: The Case of Mass Tourists to Crete. *Tourism Analysis*, 12(1), pp.51-63.
- Baesens, B. (2014). *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*, John Wiley & Sons.
- Barton, D., & Court, D. (2012). Making Advanced Analytics Work for You. *Harvard Business Review*. Available at: <https://hbr.org/2012/10/making-advanced-analytics-work-for-you> [Accessed: July 26, 2018].
- Belissent, J. (2018). AI Takes Location To A New Level. *Forrester*. Available at: <https://go.forrester.com/blogs/ai-takes-location-to-a-new-level/> [Accessed: September 15, 2019].
- Bello-Orgaz, G., Jung, J.J., & Camacho, D. (2016). Social big data: Recent achievements and new challenges. *An international journal on information fusion*, 28, pp.45-59.
- Bouadi, T., Cordier, M., Moreau, P., Quiniou, R., & Salmon-Monviola, J. (2017). A data warehouse to explore multidimensional simulated data from a spatially distributed agro-hydrological model to improve catchment nitrogen management. *Environmental Modelling & Software*, 97, pp. 229-242.
- Boukas, N., & Ziakas, V. (2013). Exploring perceptions for Cyprus as a sustainable golf destination: motivational and attitudinal orientations of golf tourists. *International Journal of Sport Management and Marketing*, 14(1-4), p.39.
- Brandt, T., Bendler, J., & Neumann, D. (2017). Social media analytics and value creation in urban smart tourism ecosystems. *Information & Management*, 54(6), pp. 703-713.
- Breuer, T. (2016). Statistical Power Analysis and the contemporary “crisis” in social sciences. *Journal of Marketing Analytics*, 4(2-3), pp.61-65.
- Brynolfsson, E., Hitt, L.M., & Kim, H. (2011). Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? *SSRN Electronic Journal*. Available at: <http://dx.doi.org/10.2139/ssrn.1819486>.

- Bumblauskas, D., Nold, H., Bumblauskas, P., Igou, A. (2017). Big data analytics: transforming data to action. *Business Process Management Journal*, 23(3), pp.703-720.
- Burstein, F., & Holsapple, C. (2008). *Handbook on Decision Support Systems 1: Basic Themes*, Springer Science & Business Media.
- Camprubi, R., Guia, J., & Comas, J. (2012). Analyzing Image Fragmentation in Promotional Brochures. *Journal of Hospitality & Tourism Research*, 38(2), pp.135-161.
- Cárdenas, A., Pratyusa, M., & Sreeranga, R. (2013). Big Data Analytics for Security. *IEEE Security & Privacy*, 11 (6), pp. 74-76.
- Carvalho, J.P., Rosa, H., Brogueira, & Batista, F. (2017). MISNIS: An intelligent platform for twitter topic mining. *Expert systems with applications*, 89, pp. 374-388.
- Cebr, (2012). *Data equity: Unlocking the value of big data*. Available at: [https://www.cebr.com/wp-content/uploads/2013/03/1733\\_Cebr\\_Value-of-Data-Equity\\_report.pdf](https://www.cebr.com/wp-content/uploads/2013/03/1733_Cebr_Value-of-Data-Equity_report.pdf) [Accessed November 6, 2017].
- Chahal, M. (2016). Marketers overestimate consumers' attitude to data. *Marketing Week*. Available at: [www.marketingweek.com/2016/06/23/marketers-overestimate-consumers-attitude-to-data/](http://www.marketingweek.com/2016/06/23/marketers-overestimate-consumers-attitude-to-data/) [Accessed June 18, 2017].
- Chakravarty, S. (2018). Why location is fundamental to all business processes. *Geospatial World*. Available at: [www.geospatialworld.net/article/how-location-is-fundamental-to-all-business-processes](http://www.geospatialworld.net/article/how-location-is-fundamental-to-all-business-processes) [Accessed September 15, 2019].
- Chang, Y.C., Ku, C.H., & Chen, C.H. (2017). Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. *International Journal of Information Management*, 48, pp. 263-279.
- Cheng, M., & Edwards, D. (2015). Social media in tourism: a visual analytic approach. *Current Issues in Tourism*, 18(11), pp.1080-1087.
- Cheng, S., Zhang, Q., & Qin, Q. (2016). Big data analytics with swarm intelligence. *Industrial Management & Data Systems*, 116(4), pp. 646-666.
- Chen, H., Chiang, R., & Storey, V.C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), pp.1165-1188.
- Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, 19(2), pp.171-209.
- Chi, M.T., Lin, S.S., Chen, S.Y., Lin, C.H., & Lee, T.Y. (2015). Morphable Word Clouds for Time-Varying Text Data Visualization. *IEEE transactions on visualization and computer graphics*, 21(12), pp.1415-1426.
- Chorley, M.J., Whitaker, R.M., & Allen, S.M., (2015). Personality and location-based social networks. *Computers in Human Behavior*, 46, pp. 45-56.
- Christou, E. (2016). *Social Media in Travel, Tourism and Hospitality: Theory, Practice and Cases*, Routledge.
- Chua, A. Servillo, L., Marcheggiani, E., & Moere, A. (2016). Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy. *Tourism Management*, 57, pp. 295-310.

- Clerides, S.K., & Pashourtidou, N., (2007). Tourism in Cyprus: Recent Trends and Lessons from the Tourist Satisfaction Survey. *Cyprus Economic Policy Review*, 1 (2), pp. 51-72.
- Coleman, S., Göb, R., Manco, G., Pievatolo, A., Tort-Martorell, X., & Reis, M. S. (2016). How Can SMEs Benefit from Big Data? Challenges and a Path Forward. *Quality and Reliability Engineering International*, 32(6), pp. 2151-2164.
- Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of business research*, 70, pp. 379-390.
- Couldry, N. (2009). Does “the Media” Have a Future? *European journal of disorders of communication: the journal of the College of Speech and Language Therapists, London*, 24(4), pp. 437-449.
- Dallemule, L., & Davenport, T. (2017). The 2 Types of Data Strategies Every Company Needs. *Harvard Business Review*. Available at: <https://hbr.org/2017/05/whats-your-data-strategy> [Accessed: June 18, 2017].
- Davenport, T.H. & Harris, J.G. (2007). *Competing on Analytics: The New Science of Winning*, Harvard Business Press.
- Davenport, T.H., Barth, P., & Bean, R. (2012). How “Big Data” is Different. *MIT Sloan Management Review*, 54(1), pp. 22-24.
- Davenport, T. (2013). Analytics 3.0. *Harvard Business Review*. Available at: <https://hbr.org/2013/12/analytics-30> [Accessed June 21, 2017].
- Davenport, T. (2014). *Big Data at Work: Dispelling the Myths, Uncovering the Opportunities*, Harvard Business Review Press.
- Davenport, T. (2016). *Analytics and IT New Opportunity for CIOs*, Harvard Business Review.
- David Kiron, P.K., & Ferguson, R.B. (2014). The Analytics Mandate [MIT Sloan Management Review, *MIT Sloan Management Review*. Available at: <https://sloanreview.mit.edu/projects/analytics-mandate/> [Accessed: August 1, 2018].
- De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library review. Great Britain. Forestry Commission*, 65(3), pp. 122-135.
- Demirkan, H., & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision support systems*, 55(1), pp. 412-421.
- Dhiratara, A., Yang, J., Bozzon, A., & Houben, G.J. (2016). Social Media Data Analytics for Tourism-A Preliminary Study. Available at: [www.semanticscholar.org/paper/Social-Media-Data-Analytics-for-Tourism-A-Study-Dhiratara-Yang/f91608556d523999295b6b0e93d3b76f66375743](http://www.semanticscholar.org/paper/Social-Media-Data-Analytics-for-Tourism-A-Study-Dhiratara-Yang/f91608556d523999295b6b0e93d3b76f66375743) [Accessed: July 13, 2018].
- Dittert, M., Härting, R.C., Reichstein, C., & Bayer, C. (2018). A Data Analytics Framework for Business in Small and Medium-Sized Organizations. *Intelligent Decision Technologies 2017*, pp. 169-181.
- Donaire, J.A., Camprubí, R., & Galí, N. (2014). Tourist clusters from Flickr travel photography. *Tourism Management Perspectives*, 11, pp. 26-33.
- Dorling, D. (2012). *The Visualization of Spatial Social Structure*, John Wiley & Sons, Ltd.
- Economist (2017). Data is giving rise to a new economy. *The Economist*. Available at:



- [www.economist.com/news/briefing/21721634-how-it-shaping-up-data-giving-rise-new-economy](http://www.economist.com/news/briefing/21721634-how-it-shaping-up-data-giving-rise-new-economy) [Accessed: October 6, 2017].
- Elgendy, N., & Elragal, A. (2016). Big Data Analytics in Support of the Decision Making Process. *Procedia computer science*, 100, pp. 1071-1084.
- Elouni, J., Ltifi, H., Ayed, M., & Masmoudi, M. (2017). Temporal Patterns Visualization for Knowledge Acquisition in Dynamic Decision-Making Environment. In *Intelligent Systems Design and Applications* . pp. 841–851.
- Evergreen, S., & Metzner, C. (2013). Design Principles for Data Visualization in Evaluation. *New Directions for Evaluation*, (140), pp. 5-20.
- Fan, W., & Gordon, M.D. (2014). The power of social media analytics. *Communications of the ACM*, 57(6), pp. 74-81.
- Farmaki, A. (2014). Satisfaction with the rural tourism product of Cyprus. *International Journal of Tourism Policy*, 5(4), p. 249.
- Farmaki, A. (2016). Managing Rural Tourist Experiences: Lessons from Cyprus. *The Handbook of Managing and Marketing Tourism Experiences*, pp. 281-298.
- Ferreira, N., Poco, J., Vo, H.T., Freire, J., & Silva, C.T. (2013). Visual Exploration of Big Spatio-Temporal Urban Data: A Study of New York City Taxi Trips. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), pp. 2149-2158.
- Ferster, B. & Shneiderman, B. (2012). *Interactive Visualization: Insight through Inquiry*, MIT Press.
- Floris, R., & Campagna, M. (2014). Social media data in tourism planning: Analysing tourists satisfaction in space and time. *Proceedings of 19th International Conference on Urban Planning, Regional Development and Information Society*, pp. 997-1003.
- Forbes (2015). Location Intelligence Leads to Business Insights and Creates Competitive Advantage. *Forbes*. Available at: <https://www.forbes.com/sites/forbespr/2015/03/12/location-intelligence-leads-to-business-insights-and-creates-competitive-advantage/> [Accessed: June 26, 2018].
- Forbes (2017). Forbes Insights: The Power of Place. *Forbes*. Available at: [https://www.forbes.com/forbesinsights/pitney\\_bowes\\_power\\_of\\_place/](https://www.forbes.com/forbesinsights/pitney_bowes_power_of_place/) [Accessed: September 15, 2019].
- Forrest, C. (2016). Big data, business analytics to hit \$203 billion by 2020, says IDC report. *TechRepublic*. Available at: <https://www.techrepublic.com/article/big-data-business-analytics-to-hit-203-billion-by-2020-says-idc-report/> [Accessed: February 5, 2019].
- Friendly, M. (2008). The Golden Age of Statistical Graphics. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 23(4), pp. 502-535.
- Galbraith, C.S., Rodriguez, C.L., & DeNoble, A.F. (2008). SME Competitive Strategy and Location Behavior: An Exploratory Study of High-Technology Manufacturing. *Journal of Small Business Management*, 46(2), pp.183-202.
- Gandomi, A. & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International journal of information management*, 35(2), pp. 137-144.
- Gantz, J., & Reinsel, D. (2011). *Extracting Value from Chaos*, IDC. Available at: [www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf](http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf) [Accessed: June 14, 2017].

- Garcia, M.G.P., Medeiros, M.C., & Vasconcelos, G.F.R. (2017). Real-time inflation forecasting with high-dimensional models: The case of Brazil. *International journal of forecasting*, 33(3), pp. 679-693.
- Gartner (2017). *Gartner Says 8.4 Billion Connected “Things” Will Be in Use in 2017...* Available at: <https://www.gartner.com/en/newsroom/press-releases/2017-02-07-gartner-says-8-billion-connected-things-will-be-in-use-in-2017-up-31-percent-from-2016> [Accessed: February 5, 2019].
- Gartner (2018a). *Gartner Says 25 Percent of Customer Service Operations Will Use Virtual Customer Assistants by 2020*. Gartner. Available at: <https://www.gartner.com/en/newsroom/press-releases/2018-02-19-gartner-says-25-percent-of-customer-service-operations-will-use-virtual-customer-assistants-by-2020> [Accessed: September 14, 2019].
- Gartner (2018b). *Gartner Says Worldwide IaaS Public Cloud Services Market Grew 29.5 Percent in 2017*. Available at: <https://www.gartner.com/en/newsroom/press-releases/2018-08-01-gartner-says-worldwide-iaas-public-cloud-services-market-grew-30-percent-in-2017> [Accessed: September 14, 2019].
- Gensler, S., Völckner, F., Liu-Thompkins, Y., & Wiertzd, C. (2013). Managing Brands in the Social Media Environment. *Journal of Interactive Marketing*, 27(4), pp. 242-256.
- Gerber, M.S. (2014). Predicting crime using Twitter and kernel density estimation. *Decision support systems*, 61, pp. 115-125.
- Glass, R. & Callahan, S. (2014). *The Big Data-Driven Business: How to Use Big Data to Win Customers, Beat Competitors, and Boost Profits*, John Wiley & Sons.
- Golfarelli, M., Mantovani, M., Ravaldi, F., & Rizzi, S. (2013). Lily: A Geo-Enhanced Library for Location Intelligence. In *Lecture Notes in Computer Science*. pp. 72-83.
- Google (2014). Going Local: How Advertisers Can Extend Their Relevance With Search. *Think with Google*. Available at: <https://www.thinkwithgoogle.com/advertising-channels/search/how-advertisers-can-extend-their-relevance-with-search/> [Accessed September 14, 2019].
- Günther, W.A. Mohammad, H., Mehrizi, R., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*, 26(3), pp. 191-209.
- Harrigan, P., Evers, U, Miles, M., & Daly, T. (2017). Customer engagement with tourism social media brands. *Tourism Management*, 59, pp. 597-609.
- Harrison, N., & O'Neill, D. (2017). If Your Company Isn't Good at Analytics, It's Not Ready for AI. *Harvard Business Review*. Available at: <https://hbr.org/2017/06/if-your-company-isnt-good-at-analytics-its-not-ready-for-ai> [Accessed: June 22, 2017].
- Hartmann, P., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management*, 36(10), pp. 1382-1406.
- Harvard Business Review (2018). Enhancing Customer Insights with Public Location Data. *Harvard Business Review*. Available at: <https://hbr.org/2018/06/enhancing-customer-insights-with-public-location-data> [Accessed: July 20, 2018].
- Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A., & Khan, S.U. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47, pp. 98-115.

- Hausmann, A., Toinoven, T., Slotow, R., Tenkanen, Moilanen, A., Heikinheimo, V., & Di Minin, E. (2018). Social Media Data Can Be Used to Understand Tourists' Preferences for Nature-Based Experiences in Protected Areas: Social media data in protected areas. *Conservation Letters*, 11(1).
- Hearn, G., Foth, M., & Gray, H. (2009). Applications and implementations of new media in corporate communications. *Corporate Communications: An International Journal*, 14(1), pp.49–61.
- Heerschap, N., Ortega, S., & Priem, A. (2014). Innovation of tourism statistics through the use of new big data sources. In *12th global forum on tourism statistics, Prague, CZ*. Available at: <https://www.itu.int/en/ITU-D/Statistics/Documents/events/wtis2014/002INF-E.pdf>.
- Hiziroglu, A., & Cebeci, H.İ. (2013). A Conceptual Framework of a Cloud-Based Customer Analytics Tool for Retail SMEs. *Periodicals of Engineering and Natural Sciences*, 1(2), pp. 44-50.
- Höpken, W., Fuchs, M., Höll, G., Keil, D., & Lexhagen, M. (2013). Multi-Dimensional Data Modelling for a Tourism Destination Data Warehouse. *Information and Communication Technologies in Tourism 2013*, pp. 157-169.
- Hosni, H., Kievit, H., Drakos, P., Angelakis, G., & Van Dijk, G. (2018). Measuring Tourist Attitudes and Consumer Behavior. *International Journal of Food and Beverage Manufacturing and Business Models*, 3(2), pp. 29-41.
- Huang, A., Gallegos, L., & Lerman, K. (2017). Travel analytics: Understanding how destination choice and business clusters are connected based on social media data. *Transportation Research Part C: Emerging Technologies*, 77, pp. 245-256.
- IBM (2017). How to leverage the power of prescriptive analytics to maximize ROI. *IBM Big Data & Analytics Hub*. Available at: <https://www.ibmbigdatahub.com/blog/how-leverage-power-prescriptive-analytics-maximize-roi> [Accessed: February 5, 2019].
- IDC (2018). Revenues for Big Data and Business Analytics Solutions Forecast to Reach \$260 Billion in 2022, Led by the Banking and Manufacturing Industries, According to IDC. *IDC: The premier global market intelligence company*. Available at: [www.idc.com/getdoc.jsp?containerId=prUS44215218](http://www.idc.com/getdoc.jsp?containerId=prUS44215218) [Accessed: August 27, 2018].
- IDC (2012). The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East by John Gantz and David Reinsel sponsored by EMC. *IDC*. Available at: <https://www.emc.com/leadership/digital-universe/2012iview/index.htm> [Accessed August 7, 2018].
- INSETI (2019). Eiserxomenos Tourismos. Available at: [http://www.inseti.gr/Portals/0/meletes-INSETI/01/2019/2019\\_Eiserxomenos\\_Tourismos-2019.pdf](http://www.inseti.gr/Portals/0/meletes-INSETI/01/2019/2019_Eiserxomenos_Tourismos-2019.pdf) [Accessed August 4, 2019].
- Janssen, M., van der Voort, H., & Wahyudi, A., 2017. Factors influencing big data decision-making quality. *Journal of business research*, 70, pp. 338-345.
- Kaisler, S.H., Espinosa, A., Armour, F., & Moneyet, W. (2014). Advanced Analytics -- Issues and Challenges in a Global Environment. *47th Hawaii International Conference on System Sciences*.
- Kaplan, A.M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons*, 53(1), pp. 59-68.
- Kirk, A. (2012). *Data Visualization: A Successful Design Process*, Packt Publishing Ltd.
- Knaflic, C.N. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals*, John Wiley & Sons.

- Költringer, C., & Dickinger, A. (2015). Analyzing destination branding and image from online sources: A web content mining approach. *Journal of Business Research*, 68(9), pp. 1836-1843.
- Komorowski, M., Do Huu, T., & Deligiannis, N. (2018). Twitter data analysis for studying communities of practice in the media industry. *Telematics and Informatics*, 35(1), pp. 195–212.
- Ko, S., Maciejewski, R., Jang, Y., & Ebert, D.S. (2012). MarketAnalyzer: An Interactive Visual Analytics System for Analyzing Competitive Advantage Using Point of Sale Data. *Computer graphics forum: journal of the European Association for Computer Graphics*, 31(3), pp. 1245-1254.
- Kumar, V., Bhaskaran, V., Mirchandani, R., & Shah, M. (2013). Practice Prize Winner—Creating a Measurable Social Media Marketing Strategy: Increasing the Value and ROI of Intangibles and Tangibles for Hokey Pokey. *Marketing Science*, 32(2), pp. 194-212.
- Ladhari, R., & Michaud, M. (2015). e-WOM effects on hotel booking intentions, attitudes, trust, and website perceptions. *International Journal of Hospitality Management*, 46, pp. 36-45.
- Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, 36(5), pp. 700-710.
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2010). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, 52(2), pp. 3-22.
- Lee, R., Wakamiya, S., & Sumiya, K. (2011). Discovery of unusual regional social activities using geo-tagged microblogs. *World Wide Web*, 14(4), pp. 321-349.
- Li, D., Zhou, X., & Wang, M. (2018). Analyzing and visualizing the spatial interactions between tourists and locals: A Flickr study in ten US cities. *Cities*, 74, pp. 249-258.
- Liebowitz, J. (2013). *Business Analytics: An Introduction*, CRC Press.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers.
- Liu, H., Ge, Y., Zheng, Q., Lin, R., & Li, H. (2018). Detecting global and local topics via mining twitter data. *Neurocomputing*, 273, pp. 120-132.
- Liu, Y., Wang, H., Li, G., Gao, J., Hu, H., & Li, W.S. (2016). ELAN: An Efficient Location-Aware Analytics System. *Big Data Research*, 5, pp. 16-21.
- Lodefalk, M. (2013). Servicification of manufacturing - evidence from Sweden. *International Journal of Economics and Business Research*, 6(1), p.87.
- Mahmud, S., Iqbal, R., & Doctor, F. (2016). Cloud enabled data analytics and visualization framework for health-shocks prediction. *Future Generation Computer Systems*, 65, pp. 169-181.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, McKinsey Global Institute, United States.
- Mao, T. (2015). Mining One Hundred Million Creative Commons Flickr Images Dataset to Flickr Tourist Index. *International Journal of Future Computer and Communication*, 4(2), pp. 104-107.
- Marchiori, E., & Cantoni, L. (2015). The role of prior experience in the perception of a tourism destination in user-generated content. *Journal of Destination Marketing & Management*, 4(3), pp. 194-201.

- Mariani, M., Baggio, R., Fuchs, M., & Höepken, W. (2018). Business intelligence and big data in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, 30(12), pp. 3514-3554.
- Marine-Roig, E., & Clavé, S.A. (2015). Tourism analytics with massive user-generated content: A case study of Barcelona. *Journal of Destination Marketing & Management*, 4(3), pp. 162-172.
- Markets & Markets (2019). Location Analytics Market by Solutions & Services - 2024 | Markets and Markets. Available at: [https://www.marketsandmarkets.com/Market-Reports/location-analytics-market-177193456.html?gclid=CjwKCAjwyMfZBRAXEiwAR3gMz5T\\_cJ3vxd1pFtasRqeH0BtUOneuuVC31Uc1-0hANuSOAKnvTvaZxoCd1IQAvD\\_BwE](https://www.marketsandmarkets.com/Market-Reports/location-analytics-market-177193456.html?gclid=CjwKCAjwyMfZBRAXEiwAR3gMz5T_cJ3vxd1pFtasRqeH0BtUOneuuVC31Uc1-0hANuSOAKnvTvaZxoCd1IQAvD_BwE) [Accessed: September 14, 2019].
- Martínez, M.P.L., Berrozpe, T.I., & Lasarte, M.P. (2014). Image-focused social media for a market analysis of tourism consumption. *International Journal of Technology Management*, 64(1), pp. 17-30.
- Matthias, S., Graue, C., & Britzelmaier, B. (2017). Big Data in SME - Findings of an empirical study. In *10th Annual Conference of the EuroMed Academy of Business*. pp. 1732–1750.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A Revolution that Will Transform how We Live, Work, and Think*, Houghton Mifflin Harcourt.
- Mbassegue, P., Escandon-Quintanilla, M.-L., & Gardoni, M. (2016). Knowledge Management and Big Data: Opportunities and Challenges for Small and Medium Enterprises (SME). In *IFIP Advances in Information and Communication Technology*. pp. 22-31.
- McAfee, A. & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review*, 90(10).
- Miah, S.J., Vu, H.Q., Gammack, J., & McGrath, M. (2017). A Big Data Analytics Method for Tourist Behaviour Analysis. *Information & Management*, 54(6), pp. 771-785.
- Microsoft News Centre Europe (2016). Go bigger with big data - Microsoft News Centre Europe. *Microsoft News Centre Europe*. Available at: <https://news.microsoft.com/europe/2016/04/20/go-bigger-with-big-data/> [Accessed: February 5, 2019].
- Miguéns, J., Baggio, R., & Costa, C. (2008). Social media and Tourism Destinations: TripAdvisor Case Study. In *Proceedings of the IASK International Conference on "Advances in Tourism Research."*
- Milton, S. (2011). Location Intelligence - The Future Looks Bright. *Forbes*. Available at: <https://www.forbes.com/sites/stevemilton/2011/10/11/location-intelligence-the-future-looks-bright/> [Accessed: July 9, 2019].
- Mittal, V., Kaul, A., Gupta, S.S., & Arora, A. (2017). Multivariate Features Based Instagram Post Analysis to Enrich User Experience. *Procedia computer science*, 122, pp. 138-145.
- Moe, W.W. & Schweidel, D.A., 2014. *Social Media Intelligence*, Cambridge University Press.
- Moore, J. (2017). Data Visualization in Support of Executive Decision Making. *Interdisciplinary Journal of Information, Knowledge, and Management*, 12, pp. 125-138.
- Morabito, V., 2015. *Big Data and Analytics, Strategic and Organizational Impacts*. Springer.
- Moreno, H. (2015). The Where Factor: Location Intelligence and the Competitive Edge. *Forbes*. Available at: <https://www.forbes.com/sites/forbesinsights/2015/04/22/the-where-factor-location-intelligence-and-the-competitive-edge/> [Accessed June 26, 2018].

- Munar, A.M., & Jacobsen, J.K.S. (2014). Motivations for sharing tourism experiences through social media. *Tourism Management*, 43, pp. 46-54.
- Newman, D. (2016). How Small Companies Can Innovate Like Big Enterprises. *Entrepreneur*. Available at: <https://www.entrepreneur.com/article/272561> [Accessed October 1, 2018].
- Nobre, H. & Silva, D. (2014). Social Network Marketing Strategy and SME Strategy Benefits. *Journal of Transnational Management*, 19(2), pp. 138-151.
- OECD (2005). *Defining innovation*. Available at: [www.oecd.org/site/innovationstrategy/defininginnovation.htm](http://www.oecd.org/site/innovationstrategy/defininginnovation.htm) [Accessed July 20, 2019].
- OECD (2015). *Data-Driven Innovation Big Data for Growth and Well-Being: Big Data for Growth and Well-Being*, OECD Publishing.
- OECD & Eurostat (2018). *The Measurement of Scientific, Technological and Innovation Activities Oslo Manual 2018 Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition*, OECD Publishing.
- Oxford Economics (2013). How successful SMEs are reinventing global business. Available at: <https://www.oxfordeconomics.com/my-oxford/projects/238969> [Accessed: September 16, 2019].
- Packham, I.S.J., Rafiq, M.Y., Borthwick, M.F., & Denham, S.L. (2005). Interactive visualisation for decision support and evaluation of robustness, in theory and in practice. *Advanced Engineering Informatics*, 19(4), pp. 263-280.
- Palacios-Marqués, D., Merigó, J.M., & Soto-Acosta, P. (2015). Online social networks as an enabler of innovation in organizations. *Management Decision*, 53(9), pp. 1906-1920.
- Pandey, A., Verma, R., Bariha, D., & Dewangan, V. (2016). Sentiment Analysis Through Text Mining- A Review. *International Journal Of Engineering And Computer Science*, pp. 18665-18657.
- Park, J., Fink, C., Barash, V., & Cha, M. (2013). Emoticon Style: Interpreting Differences in Emoticons Across Cultures. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*.
- Park, S.B., Jang, J., & Ok, C.M. (2016). Analyzing Twitter to explore perceptions of Asian restaurants. *Journal of Hospitality and Tourism Technology*, 7(4), pp.405-422.
- Poleto, T., de Carvalho, V.D.H. & Cabral, A. (2015). The Roles of Big Data in the Decision-Support Process: An Empirical Investigation. In *Lecture Notes in Business Information Processing*. pp. 10-21.
- Power, D.J. (2007). A Brief History of Decision Support Systems. Available at: <http://dssresources.com/history/dsshistory.html> [Accessed: July 29, 2018].
- Power, D.J. (2008). Understanding Data-Driven Decision Support Systems. *Information Systems Management*, 25(2), pp. 149-154.
- Prem, R. (2017). Growing Adoption of Location Analytics by SMEs is Projected to Drive LI & LA (Location Intelligence & Location Analytics) Market Growth - industryarc blog. *industryarc blog*. Available at: <http://industryarcblog.com/growing-adoption-location-analytics-smes-projected-drive-li-la-location-intelligence-location-analytics-market-growth/> [Accessed: June 26, 2018].
- Prodromou, T. (2014). Drawing Inference from Data Visualisations. *International Journal of Secondary Education*, 2(4), pp. 66-72.

- Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big data*, 1(1), pp.51-59.
- Psyllidis, A., Bozzon, A., Bocconi, S., Bolivar, C.T. (2015). A Platform for Urban Analytics and Semantic Data Integration in City Planning. In *Communications in Computer and Information Science*. pp. 21–36.
- Pucha-Cofrep, F., Fries, A., Cánovas-García, F., Oñate-Valdivieso, F., & González-Jaramillo. (2018). *Fundamentals of GIS: Applications with ArcGIS*, Loja: Ediloja.
- Rao, P., & Holt, D. (2005). Do green supply chains lead to competitiveness and economic performance? *International Journal of Operations & Production Management*, 25(9), pp. 898-916.
- Rathore, M.M., Paul, A., Ahmad, A., Imran, M., & Guizani, M. (2017). Big data analytics of geosocial media for planning and real-time decisions. *2017 IEEE International Conference on Communications (ICC)*, 21-25 May 2017, Paris, France.
- Ravi, K., Khandelwal, Y., Krishna, B.S., & Ravi, V. (2018). Analytics in/for cloud-an interdependence: A review. *Journal of Network and Computer Applications*, 102, pp. 17-37.
- Redhu, S. (2018). Sentiment Analysis Using Text Mining: A Review. *International Journal on Data Science and Technology*, 4(2), pp. 49-53.
- Rehman, N.U., Weiler, A., & Scholl, M.H. (2013). OLAPing social media. *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining - ASONAM '13*. Available at: <http://dx.doi.org/10.1145/2492517.2500273>.
- Reinsel, D., Gantz, J., & Rydning, J. (2017). *DataAge 2025 - The Digitization of the World | Seagate UK*, IDC. Available at: <https://www.seagate.com/gb/en/our-story/data-age-2025/> [Accessed: July 9, 2019].
- Ribarsky, W., Wang, D.X., & Dou, W., 2014. Social media analytics for competitive advantage. *Computers & graphics*, 38, pp. 328-331.
- Roberts, D.L., Piller, F.T., & Lüttgens, D., 2016. Mapping the Impact of Social Media for Innovation: The Role of Social Media in Explaining Innovation Performance in the PDMA Comparative Performance Assessment Study. *Journal of Product Innovation Management*, 33, pp. 117-135.
- Rosenzweig, P. (2014). The benefits--and limits--of decision models. *McKinsey & Company*. Available at: <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/the-benefits-and-limits-of-decision-models> [Accessed June 14, 2018].
- Rosman, R., & Stuhura, K. (2013). The Implications of Social Media on Customer Relationship Management and the Hospitality Industry. *Journal of Management Policy and Practice*, 14(3), pp. 18-26.
- Ryan, L. (2016). *The Visual Imperative: Creating a Visual Culture of Data Discovery*, Morgan Kaufmann.
- Rybarczyk, G., Banerjee, S., Starkling-Szymanski, M.D., & Shaker, R.R. (2018). Travel and us: the impact of mode share on sentiment using geo-social media and GIS. *Journal of Location Based Services*, 12(1), pp. 40-62.
- Salas-Olmedo, M.H., Moya-Gómez, B., García-Palomares, J.C., & Gutiérrez, J. (2018). Tourists' digital footprint in cities: Comparing Big Data sources. *Tourism Management*, 66, pp. 13-25.

- Schaeffer, D.M., & Olson, P.C. (2014). Big Data Options For Small And Medium Enterprises. *Review of Business Information Systems (RBIS)*, 18(1), pp. 41- 46.
- Schmidbauer, H., Rösch, A., & Stieler, F. (2018). The 2016 US presidential election and media on Instagram: Who was in the lead? *Computers in Human Behavior*, 81, pp. 148-160.
- Schmitz, P., Cooper, A., Kruger, T., Speed, K., Barkhuizen, M., Lochner, H., & Linnen, C. (2015). Space-Time Visualization for Investigative and Forensic Purposes. In *Lecture Notes in Geoinformation and Cartography*. pp. 267–281.
- Secundo, G., Del Vecchio, P., Dumay, J., and Passiante, G. (2017). Intellectual capital in the age of Big Data: establishing a research agenda. *Journal of Intellectual Capital*, 18(2), pp. 242-261.
- Sen, D., Ozturk, M., & Vayvay, O. (2016). An Overview of Big Data for Growth in SMEs. *Procedia - Social and Behavioral Sciences*, 235, pp.159–167.
- Sharma, R., Mithas, S. & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, 23(4), pp. 433-441.
- Shekhar, S., Xiong, H., & Zhou, X. (2017). *Encyclopedia of GIS / SpringerLink*, Springer.
- Silva, F.B., Herrera, M.A., Rosina, K., Barranco, R.R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management*, 68, pp. 101-115.
- Sivarajah, U., Kamal, M.M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of business research*, 70, pp. 263-286.
- Smith, M., Szongott, C., Henne, B., & Von Voigt, G. (2012). Big data privacy issues in public social media. *2012 6th IEEE International Conference on Digital Ecosystems and Technologies (DEST)*.
- Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social Media Analytics - An Interdisciplinary Approach and Its Implications for Information Systems. *Business & Information Systems Engineering*, 6(2), pp. 89-96.
- Thulo, L. (2014). Location, location, location - SME. *SME*. Available at: <http://www.smesouthafrica.co.za/Location-location-location/> [Accessed June 26, 2018].
- United Nations (2014). *A world that counts. Mobilizing the data revolution for sustainable development*, New York: United Nations.
- Uzunoglu, E., & Kip, S.M. (2014). Brand communication through digital influencers: Leveraging blogger engagement. *International Journal of Information Management*, 34(5), pp. 592-602.
- Vashisht, P., & Gupta, V. (2015). Big data analytics techniques: A survey. *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*. Available at: <http://dx.doi.org/10.1109/icgciot.2015.7380470>.
- Veal, A.J. (2017). *Research Methods for Leisure and Tourism*, Pearson UK.
- Vecchio, P.D., Mele, G., Ndou, V., & Secundo, G. (2018a). Creating value from Social Big Data: Implications for Smart Tourism Destinations. *Information processing & management*, 54(5), pp. 847-860.
- Vecchio, P.D., Mele, G., Ndou, V., & Secundo, G. (2018b). Open Innovation and Social Big Data for Sustainability: Evidence from the Tourism Industry. *Sustainability*, 10(9), p. 3215.



- Vu, H.Q., Li, G., Law, R., & Zhang, Y. (2018). Tourist Activity Analysis by Leveraging Mobile Social Media Data. *Journal of Travel Research*, 57(7), pp. 883-898.
- Waller, M.A., & Fawcett, S.E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), pp. 77-84.
- Wamba, S.F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How “big data” can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, pp. 234-246.
- Wang, G., Gunasekaran, A., Ngai, E.W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, 176, pp. 98-110.
- Wang, L., Gopal, R., Shankar, R., & Pancras, J. (2015). On the brink: Predicting business failure with mobile location-based checkins. *Decision Support Systems*, 76, pp. 3-13.
- Ware, C. (2013). *Information Visualization: Perception for Design*, Elsevier.
- Wexler, S., Shaffer, J., & Cotgreave, A. (2017). *The Big Book of Dashboards: Visualizing Your Data Using Real-World Business Scenarios*, John Wiley & Sons.
- Whitney, H. (2012). *Data Insights: New Ways to Visualize and Make Sense of Data*, Newnes.
- Winig, L. (2016). GE’s Big Bet on Data and Analytics | MIT Sloan Management Review. *MIT Sloan Management Review*. Available at: <https://sloanreview.mit.edu/case-study/ge-big-bet-on-data-and-analytics/> [Accessed February 5, 2019].
- Yoo, K.-H., Sigala, M., & Gretzel, U. (2016). Exploring TripAdvisor. *Open Tourism*, pp.239–255.
- Zeng, D., Chen, H.C., Lusch, R., & Li, S.H. (2010). Social Media Analytics and Intelligence. *IEEE intelligent systems*, 25(6), pp. 13-16.
- Zhang, M., Guo, L., Hu, M., & Liu, W. (2017). Influence of customer engagement with company social networks on stickiness: Mediating effect of customer value creation. *International journal of information management*, 37(3), pp. 229-240.
- Zhang, Z., Li, X., & Chen, Y. (2012). Deciphering word-of-mouth in social media. *ACM Transactions on Management Information Systems*, 3(1), pp. 1-23.
- Zheng, Y. (2011). Location-Based Social Networks: Users. *Computing with Spatial Trajectories*, pp. 243-276.
- Zhou, X., Xu, C., & Kimmons, B. (2015). Detecting tourism destinations using scalable geospatial analysis based on cloud computing platform. *Computers, Environment and Urban Systems*, 54, pp.144-153.
- Zhu, F. & Zhang, X. (2010). Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *Journal of marketing*, 74(2), pp. 133-148.
- Zolnowski, A., Christiansen, T., & Gudat, J. (2016) Business model transformation patterns of data-driven innovations. In 24th European Conference on Information Systems, ECIS 2016. pp. 1-16.
- Zouganeli, S., Trihas, N., Antonaki, M., & Kladou, S. (2012). Aspects of Sustainability in the Destination Branding Process: A Bottom-up Approach. *Journal of Hospitality Marketing & Management*, 21(7), pp. 739-757.

Zudilova-Seinstra, E., Adriaansen, T., & van Liere, R. (2008). *Trends in Interactive Visualization: State-of-the-Art Survey*, Springer Science & Business Media.