UNDERSTANDING TRAVELLERS' CHOICES USING DATA ANALYTICS

Kariyawasam Jalaththanthri Prabhash Sri Malinda Jayasekera

(179115T)

Degree of Master of Business Administration in Information Technology

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

March 2019

UNDERSTANDING TRAVELLERS' CHOICES USING DATA ANALYTICS

Kariyawasam Jalaththanthri Prabhash Sri Malinda Jayasekera

(179115T)

Thesis submitted in partial fulfilment of the requirement for the degree of Master of Business Administration in Information Technology

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

March 2019

DECLARATION

I declare that this is my own work and this thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

The above candidate has carried out research for the Masters thesis under my supervision.

Name of the supervisor: Dr. Amal Shehan Perera

Signature of the supervisor: Date:

Name of the co-supervisor: Prof. Sanath Jayasena

Signature of the co-supervisor: Date:

ABSTRACT

Travel and tourism industry is one of the largest and growing industries in the world that depends on choices and demands of travellers. The identification of these choices and demands will provide benefits to both service providers in the industry and travellers. The use of data analytics to achieve this has been discussed briefly over the years using different types of data. The findings of these studies were inconclusive due to limitations in the selected data types, features and analysis techniques. This research aims to overcome these limitations by identifying the factors that impact the choices of travellers, establishing a feature framework to identify those choices, finding the feasibility of using time series forecasting to predict travellers' demand and proposing the use of data analytics in travel insurance. The limitations in previous studies and the unavailability of necessary data for research have increased the importance of using data analytics in travel insurance, an industry within travel and tourism industry. This research achieves its objectives by conducting a study with data from the UK, one of the best performing outbound markets in the world. The data was analysed using data analytics techniques to find the destination and travel mode choices of travellers and two other subgroups, travellers with medical conditions and cruise travellers. The number of outbound trips and the visitors for destinations were forecasted for a year to find the feasibility of using time series forecasting to predict travellers' demands. The results of the analysis confirm that a traveller's age, group type they choose to travel under, and their health have an impact on their destination and travel mode choices, and the two choices have an impact on each other. The study finds that time series forecasting is a reliable demand forecasting technique when a large data set is available.

Keywords: Outbound Markets, Travel Demand, Travel Patterns, United Kingdom

ACKNOWLEDGEMENT

I wish to express my sincere gratitude to all those who helped me in successfully completing my research study on “Catering to United Kingdom’s Outbound Market”.

First and foremost, I wish to thank my research supervisors Dr. Amal Shehan Perera, Head of Department, Department of Computer Science and Engineering, University of Moratuwa, and Prof. Sanath Jayasena, Associate Professor, Department of Computer Science and Engineering, University of Moratuwa for their continuous support, encouragement and attention extended to me in realising the research objectives.

I wish to convey my special gratitude to Ms. Dinushka Peiris, Director, Research, Sri Lanka Tourism Development Authority for validating the findings of the research. Furthermore, I extend my thanks to Ms. Nirmala Chandrasiri for her help in proofreading the thesis and colleagues at my workplace for all their help.

My heartfelt appreciation goes to my father, mother and brothers for all the encouragement, care and unconditional support given to me throughout the good and bad times during this research.

Finally, I wish to convey my heartfelt thanks to all those who helped me in many ways for bringing this study a success.

TABLE OF CONTENTS

DECLARATION i

ABSTRACT ii

ACKNOWLEDGEMENT iii

TABLE OF CONTENTS iv

LIST OF FIGURES vii

LIST OF TABLES x

LIST OF ABBREVIATIONS xvi

LIST OF APPENDICES xvii

1. INTRODUCTION 1

1.1. Background 1

1.1.1. Travel & Tourism Industry 1

1.1.2. Outbound Markets and Growth 2

1.1.3. Demand and Customer Centric Approach in Travel and Tourism 2

1.1.4. Benefits of Using Big Data, Business Intelligence (BI) and Business Analytics (BA) in Travel, Tourism and Travel Insurance 3

1.2. Problem Statement 4

1.3. Research Objectives 4

1.4. Research Design 4

1.5. Outline 5

1.6. Summary 5

2. LITERATURE REVIEW 6

2.1.1. Business Intelligence (BI) 6

2.1.2. Business Analytics (BA) 6

2.1.3. Big Data 7

2.1.4. Big Data & BI Awareness in Travel & Tourism 7

2.2. Real-world Applications 8

2.2.1. Data 8

2.2.2. Features 12

2.2.3. Analysis Techniques for Demand Prediction 15
2.2.4. Accuracy Measures 23
2.3. Summary 25
3. RESEARCH METHODOLOGY 26
3.1. Exploratory Data Analysis 26
3.2. Methodology Framework 26
3.3. Conceptual Framework 27
3.3.1. Age Groups 28
3.4. Population and Sample 28
3.5. Data Pre-processing 29
3.6. Data Analysis 30
3.6.1. Descriptive Analytics 30
3.6.2. Predictive Analytics 32
3.7. Accuracy Measures 34
3.8. Summary 35
4. DATA ANALYSIS 36
4.1. Travellers 37
4.1.1. Group Type 41
4.1.2. Age 54
4.1.3. More than one Destination 65
4.2. Travellers with Medical Conditions 69
4.2.1. Group Type 74
4.2.2. Age Group 79
4.2.3. Cruise 84
4.3. Cruise Travellers 91
4.3.1. Group Type 97
4.3.2. Age 101
4.3.3. Medical Condition 106
4.4. Summary of Findings 111

4.5. Summary 115
5. CONCLUSIONS AND RECOMMENDATIONS 116
5.1. Conclusions 116
5.1.1. Feature Framework 116
5.1.2. Time series Forecasting 118
5.2. Recommendations 118
5.3. Limitations and Future Work 120
5.4. Summary 121
REFERENCES 122
APPENDIX A - Factors Influencing Destination and Travel Mode Choices 129
APPENDIX B - Definitions of Generations 130
APPENDIX C - Cruise Travellers & Total Visitors to Destinations 132
APPENDIX D - Accuracy of Time Series Forecasting 133
APPENDIX E - Interview with the Industry Expert 138

LIST OF FIGURES

Figure 1.1: Economic impact and employment opportunities in travel and tourism industry 1
Figure 1.2: Growth rates of best performing outbound markets in 2016 2
Figure 2.1: Knowledge destination framework architecture 8
Figure 2.2: Forecasting framework for search trend data 11
Figure 2.3: Framework for knowledge discovery 16
Figure 2.4: Framework for Traveller Behaviour Collecting and Travel Recommender Innovative System 17
Figure 2.5: Demand forecasting model 20
Figure 3.1: Methodology Framework 27
Figure 4.1: Number of UK travellers over time 37
Figure 4.2: Number of UK travellers received by the ten most visited destinations over time 40
Figure 4.3: Number of UK travellers that travelled in different group types over time 42
Figure 4.4: Number of UK travellers that travelled as couples received by the ten most visited destinations by such travellers from 2014 to 2017 47
Figure 4.5: Number of UK travellers that travelled as individuals received by the ten most visited destinations by such travellers from 2014 to 2017 49
Figure 4.6: Number of UK travellers that travelled as groups received by the ten most visited destinations by such travellers from 2014 to 2017 50
Figure 4.7: Number of UK travellers that travelled as families received by the ten most visited destinations by such travellers from 2014 to 2017 52
Figure 4.8: Number of UK travellers that travelled as single parent families received by the ten most visited destinations by such travellers from 2014 to 2017 53
Figure 4.9: Number of travellers that belonged to different age groups over time 55
Figure 4.10: Number of Generation Z travellers received by the ten most visited destinations by Generation Z travellers from 2014 to 2017 60
Figure 4.11: Number of Millennials received by the ten most visited destinations by Millennials from 2014 to 2017 61

Figure 4.12: Number of Generation X travellers received by the ten most visited destinations by Generation X travellers from 2014 to 2017 62
Figure 4.13: Number of Baby Boomers received by the ten most visited destinations by Baby Boomers from 2014 to 2017 63
Figure 4.14: Number of Silent Generation travellers received by the ten most visited destinations by Silent Generation travellers from 2014 to 2017 64
Figure 4.15: Number of UK travellers that travelled to more than one destination during a trip over time 66
Figure 4.16: Number of UK travellers from different group types that travelled to more than one destination during a trip over time 67
Figure 4.17: Number of UK travellers from different age groups that travelled to more than one destination during a trip over time 68
Figure 4.18: Number of UK travellers with medical conditions travelled over time 69
Figure 4.19: Number of travellers with medical conditions received by the ten most visited destinations by such travellers from 2014 to 2017 73
Figure 4.20: Number of travellers with medical conditions that travelled in different group types over time 75
Figure 4.21: Number of travellers with medical conditions that belonged to different age groups from 2014 to 2017 80
Figure 4.22: Age distribution of travellers with medical conditions that travelled on a cruise during their trip 84
Figure 4.23: Number of UK travellers with medical conditions that travelled on a cruise during their trips over time 85
Figure 4.24: Number of UK travellers with medical conditions that did not travel on a cruise during their trips over time 85
Figure 4.25: Number of UK travellers without any medical conditions that travelled on a cruise during their trips over time 86
Figure 4.26: Number of UK travellers without any medical conditions that did not travel on a cruise during their trips over time 86
Figure 4.27: Number of UK travellers that travelled on a cruise during their trip over time 91

Figure 4.28: Number of UK travellers that did not travel on a cruise during their trip over time 92
Figure 4.29: Number of cruise travellers received by the ten most visited destinations by cruise travellers from 2014 to 2017 96
Figure 4.30: Number of cruise travellers that travelled in different group types over time 98
Figure 4.31: Number of cruise travellers that belonged to different age groups from 2014 to 2017 103
Figure 4.32: Age distribution of cruise travellers that had medical conditions 106
Figure 4.33: Number of cruise travellers with medical conditions over time 107
Figure 5.1: Feature Framework 116

LIST OF TABLES

Table 3.1: Feature Selection 27
Table 3.2: Definition of generations used in the research 28
Table 3.3: Number of outbound trips originated from UK over the years 29
Table 4.1: Most visited destinations by travellers from 2014 to 2017 38
Table 4.2: Number of travellers that travelled in each group type as a percentage of total number of travellers from 2014 to 2017 41
Table 4.3: Most visited destinations by couples as a percentage of total number of travellers that travelled as couples received by all the destinations from 2014 to 2017 43
Table 4.4: Most visited destinations by individuals as a percentage of total number of travellers that travelled as individuals received by all the destinations from 2014 to 2017 44
Table 4.5: Most visited destinations by groups as a percentage of total number of travellers that travelled as groups received by all the destinations from 2014 to 2017 44
Table 4.6: Most visited destinations by families as a percentage of total number of travellers that travelled as families received by all the destinations from 2014 to 2017 45
Table 4.7: Most visited destinations by single parent families as a percentage of total number of travellers that travelled as single parent families received by all the destinations from 2014 to 2017 45
Table 4.8: Number of travellers belonging to each age group as a percentage of total number of travellers from 2014 to 2017 54
Table 4.9: Most visited destinations by Generation Z as a percentage of total number of Generation Z travellers received by all the destinations from 2014 to 2017 56
Table 4.10: Most visited destinations by Millennials as a percentage of total number of Millennials received by all the destinations from 2014 to 2017 56
Table 4.11: Most visited destinations by Generation X as a percentage of total number of Generation X travellers received by all the destinations from 2014 to 2017 57

Table 4.12: Most visited destinations by Baby Boomers as a percentage of total number of Baby Boomers received by all the destinations from 2014 to 2017 57
Table 4.13: Most visited destinations by Silent Generation as a percentage of total number of Silent Generation travellers received by all the destinations from 2014 to 2017 58
Table 4.14: Number of travellers in each group type that travelled to more than one destination on a trip as a percentage of total number of travellers that travelled to more than one destination on a trip from 2014 to 2017 65
Table 4.15: Number of travellers belonging to each age group that travelled to more than one destination on a trip as a percentage of total number of travellers that travelled to more than one destination on a trip from 2014 to 2017 65
Table 4.16: Most visited destinations by UK travellers with medical conditions from 2014 to 2017 70
Table 4.17: Most visited destinations by UK travellers without any medical conditions from 2014 to 2017 70
Table 4.18: Number of travellers with medical conditions that travelled in each group type as a percentage of total number of travellers with medical conditions from 2014 to 2017 74
Table 4.19: Number of travellers without any medical conditions that travelled in each group type as a percentage of total number of travellers without any medical conditions from 2014 to 2017 74
Table 4.20: Most visited destinations by travellers with medical conditions that travelled as couples as a percentage of total number of travellers with medical conditions that travelled as couples received by all the destinations from 2014 to 2017 76
Table 4.21: Most visited destinations by travellers with medical conditions that travelled as individuals as a percentage of total number of travellers with medical conditions that travelled as individuals received by all the destinations from 2014 to 2017 77
Table 4.22: Most visited destinations by travellers with medical conditions that travelled as groups as a percentage of total number of travellers with medical

conditions that travelled as groups received by all the destinations from 2014 to 2017 77
Table 4.23: Most visited destinations by travellers with medical conditions that travelled as families as a percentage of total number of travellers with medical conditions that travelled as families received by all the destinations from 2014 to 2017 78
Table 4.24: Most visited destinations by travellers with medical conditions that travelled as single parent families as a percentage of total number of travellers with medical conditions that travelled as single parent families received by all the destinations from 2014 to 2017 78
Table 4.25: Number of travellers with medical conditions from each age group as a percentage of total number of travellers with medical conditions from 2014 to 2017 79
Table 4.26: Number of travellers without any medical conditions from each age group as a percentage of total number of travellers with medical conditions from 2014 to 2017 79
Table 4.27: Most visited destinations by Generation Z travellers with medical conditions as a percentage of Generation Z travellers with medical conditions received by all the destinations from 2014 to 2017 81
Table 4.28: Most visited destinations by Millennials with medical conditions as a percentage of Millennials with medical conditions received by all the destinations from 2014 to 2017 82
Table 4.29: Most visited destinations by Generation X travellers with medical conditions as a percentage of Generation X travellers with medical conditions received by all the destinations from 2014 to 2017 82
Table 4.30: Most visited destinations by Baby Boomers with medical conditions as a percentage of Baby Boomers with medical conditions received by all the destinations from 2014 to 2017 83
Table 4.31: Most visited destinations by Silent Generation travellers with medical conditions as a percentage of Silent Generation travellers with medical conditions received by all the destinations from 2014 to 2017 83

Table 4.32: Most visited destinations by UK travellers with medical conditions that travelled on a cruise during their trips as a percentage of total number of travellers with medical conditions that travelled on a cruise during their trips received by all the destinations from 2014 to 2017 87
Table 4.33: Most visited destinations by UK travellers with medical conditions that did not travel on a cruise as a percentage of total number of travellers with medical conditions that did not travel on a cruise received by all the destinations from 2014 to 2017 87
Table 4.34: Most visited destinations by UK travellers without any medical conditions that travelled on a cruise during their trips as a percentage of total number of travellers without any medical conditions that travelled on a cruise during their trips received by all the destinations from 2014 to 2017 88
Table 4.35: Most visited destinations by UK travellers without any medical conditions that did not travel on a cruise as a percentage of total number of travellers without any medical conditions that did not travel on a cruise received by all the destinations from 2014 to 2017 89
Table 4.36: Most visited destinations by UK travellers that travelled on a cruise during their trip from 2014 to 2017 93
Table 4.37: Most visited destinations by UK travellers that did not travel on a cruise during their trips from 2014 to 2017 93
Table 4.38: Number of cruise travellers that travelled in each group type as a percentage of total number of cruise travellers from 2014 to 2017 97
Table 4.39: Number of travellers that did not travel on a cruise during their trips in each group type as a percentage of total number of travellers that did not travel on a cruise during their trips from 2014 to 2017 97
Table 4.40: Most visited destinations by travellers that travelled on a cruise during their trips as couples as a percentage of total number of travellers that travelled on a cruise during their trips as couples received by all the destinations from 2014 to 2017 99
Table 4.41: Most visited destinations by travellers that travelled on a cruise during their trips as individuals as a percentage of total number of travellers that travelled on

a cruise during their trips as individuals received by all the destinations from 2014 to 2017 99
Table 4.42: Most visited destinations by travellers that travelled on a cruise during their trips as groups as a percentage of total number of travellers that travelled on a cruise during their trips as groups received by all the destinations from 2014 to 2017 100
Table 4.43: Most visited destinations by travellers that travelled on a cruise during their trips as families as a percentage of total number of travellers that travelled on a cruise during their trips as families received by all the destinations from 2014 to 2017 100
Table 4.44: Most visited destinations by travellers that travelled on a cruise during their trips as single parent families as a percentage of total number of travellers that travelled on a cruise during their trips as single parent families received by all the destinations from 2014 to 2017 101
Table 4.45: Number of cruise travellers belonged to each generation as a percentage of total number of cruise travellers from 2014 to 2017 101
Table 4.46: Number of travellers that did not travel on a cruise during their trip belonged to each generation as a percentage of total number of traveller that did not travel on a cruise during their trip from 2014 to 2017 102
Table 4.47: Most visited destinations by Generation Z cruise travellers as a percentage of total number of Generation Z cruise travellers that visited all the destinations from 2014 to 2017 104
Table 4.48: Most visited destinations by Millennial cruise travellers as a percentage of total number of Millennial cruise travellers that visited all the destinations from 2014 to 2017 104
Table 4.49: Most visited destinations by Generation X cruise travellers as a percentage of total number of Generation X cruise travellers that visited all the destinations from 2014 to 2017 105
Table 4.50: Most visited destinations by cruise travellers who are Baby Boomers as a percentage of total number of cruise travellers who are Baby Boomers that visited all the destinations from 2014 to 2017 105

Table 4.51: Most visited destinations by Silent Generation cruise travellers as a percentage of total number of Silent Generation cruise travellers that visited all the destinations from 2014 to 2017 106
Table 4.52: Most visited destinations by cruise travellers with medical conditions as a percentage of total number of cruise travellers with medical conditions received by all the destinations from 2014 to 2017 107
Table 4.53: Most visited destinations by cruise travellers without any medical conditions as a percentage of total number of cruise travellers without any medical conditions received by all the destinations from 2014 to 2017 108
Table 4.54: Most visited destinations by travellers that did not travel on a cruise and had medical conditions as a percentage of total number of travellers that did not travel on a cruise and had medical conditions received by all the destinations from 2014 to 2017 109
Table 4.55: Most visited destinations by travellers that did not travel on a cruise and without any medical conditions as a percentage of total number of travellers that did not travel on a cruise and without any medical conditions received by all the destinations from 2014 to 2017 110

LIST OF ABBREVIATIONS

Abbreviation	Description
GDP	Gross Domestic Product
UK	United Kingdom
CLV	Customer Lifetime Value
BI	Business Intelligence
BA	Business Analytics
DSS	Decision Support System
EU	European Union
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
MAPE	Mean Absolute Percentage Error
MASE	Mean Absolute Scaled Error
PII	Personally Identifiable Information

LIST OF APPENDICES

Appendix	Description	Page
APPENDIX A	Factors Influencing Destination and Travel Mode Choices	129
APPENDIX B	Definitions of Generations	130
APPENDIX C	Cruise Travellers & Total Visitors to Destinations	132
APPENDIX D	Accuracy of Time Series Forecasting	133
APPENDIX E	Interview with the Industry Expert	138