

Analysis on the travel and behavioural patterns of UK Residents, to produce a web-based visualisation tool

Student Name

Supervisor: **name**

Final Year Report

COMP1682 Final Year Project

Programme Title: BSc Business Information Systems

Word count: 14,986

Abstract

This project will produce visualisation in form of a dashboard which will provide an insight into the travel and behavioural patterns of UK residents. It will contain information such as number of nights spent, expenditure, purpose of visit, etc. The information will focus on the continent/countries like Europe, south America, Asia, etc. The information gathered from the visualisation will be used to issue recommendations to improve the services provided to their UK customers who travel to their countries/continents.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors for the support and assistance with my project. Also, I appreciate the support from my family and friends throughout this year.

1 Table of Contents

Contents

1	Table of Contents	4
2	Introduction	6
2.1	<i>Background</i>	6
2.2	<i>Aims</i>	6
3	Objectives	7
4	Approach	9
4.1	<i>Justification of the suitability of a Methodology or a Framework followed.</i>	9
5	Literature Review.....	10
5.1	<i>Approach to literature searching</i>	10
5.2	<i>The success of the tourism industry</i>	10
5.3	<i>The current issues with the tourism industry</i>	10
5.4	<i>Technological impact on the Tourism Industry</i>	11
5.5	<i>The usefulness of Business Intelligence in the Tourism Industry</i>	12
5.6	<i>The use of Data Visualisation for this Industry</i>	12
5.7	<i>The importance of using various graphs and charts to display different types of data</i>	13
5.8	<i>Conclusions</i>	14
6	Further Research	14
6.1	<i>Business intelligence</i>	14
6.2	<i>Business intelligence Tools</i>	14
6.3	<i>Jakob Nelson heuristics</i>	15
7	Product Research.....	17
8	Legal, Social, Ethical and Professional Issues and Considerations	20
9	Requirements.....	21
9.1	<i>Analysis of requirements</i>	21
9.2	<i>Comparison of systems</i>	23
9.3	<i>Business requirements</i>	24
9.3.1	<i>Rich Picture</i>	24
9.4	<i>Functional requirements with MoSCoW prioritisation</i>	25
9.5	<i>Non-functional requirements</i>	27
9.6	<i>Time Boxing</i>	28
10	Design	29
10.1	<i>Use-case Diagram</i>	29
10.1.1	<i>Primary and Secondary Use-case scenario</i>	30
10.2	<i>Entity Relationship Diagrams</i>	32
10.3	<i>Low-fidelity wireframes for prototypes</i>	35
11	Implementation	38
11.1	<i>Creation of Datawarehouse using ETL process</i>	38
11.2	<i>1st Prototype – Microsoft Excel Dashboard</i>	46
11.3	<i>2nd Prototype – Tableau Dashboard</i>	50
11.4	<i>Final product – WIX Website Containing Dashboard</i>	54
12	Testing.....	58
13	Evaluation.....	58
13.1	<i>Summarised Key findings from the project</i>	58
13.2	<i>Recommended measures produced using key findings</i>	59
13.3	<i>Usability Evaluation - Cognitive Walkthrough</i>	61
13.4	<i>Project Evaluation</i>	62
13.5	<i>Personal Evaluation</i>	62
13.6	<i>Conclusion</i>	63

14	References.....	64
15	Bibliography	65
16	Appendix A – Project Proposal	67
17	Appendix B (Implementation – Datawarehouse).....	74
18	Appendix C (Implementation – Prototype 1 – Excel Dashboard)	87
19	Appendix D (Implementation – Prototype 2 –Tableau Dashboard).....	94
20	Appendix E (Implementation – Final product – Wix Website containing Dashboard)	101
21	Appendix F (Testing)	104
22	Appendix G (Interaction with Supervisor and Blog Entries)	110

2 Introduction

2.1 Background

The Tourism industry has been growing at a significant rate, due to the improvements in the technologies, customer values and a decrease in the cost of travelling. "In the annual analysis of the global economic impact of Travel & Tourism, the sector is shown to account for 10.4% of global GDP and 313 million jobs, or 9.9% of total employment, in 2017."(Manzo G, 2018) There are different kinds of businesses booming in the Tourism industry such as commercial airline, cruise, travel agency, tour organiser, local tourist information Centre, etc. All the businesses in this industry prioritise one thing in common when it comes to building a successful or profitable business which is data. There are millions of data created every second and they can be used for a variety of purposes. For example, Data collected from a commercial airline business can be used for predicting what the customers' needs/likes are and what will make them come back to use their services again.

Collecting and storing these datasets will be an easy task to accomplish. However, the biggest challenge lies in breaking these data down into useful information. The data collected cannot be understood by people who are making the business decision, and this can have a negative impact on the growth of their business. With the help of useful information, it can provide businesses to gain insight into their customer's preferences. To produce useful information such as hidden patterns or customer preferences, there are various tools/techniques which can be used but the project will focus on using business intelligence tools. Rouse describes Business intelligence as "... a technology-driven process for analysing data and presenting actionable information to help executives, managers and other corporate end users make informed business decisions." (Rouse M, 2017)

The project will focus on using business intelligence tools and datasets collected from an online source to produce a visualisation using a dashboard interface which will provide an insight into UK residents travel and behavioural patterns. The information gathered from this visualisation tool will be used to provide the countries/continents with some recommendation which may improve the services they provide to UK customers. This dashboard interface will then be available online and can be accessed by stakeholders who are concerned in the tourism industry in their countries, to gain an insight into the UK customers who travel to their respected countries. They can then use this information to tailor their services to accommodate UK customers.

Keywords: [Tourism, Business intelligence, Visualisation, Dashboard, Data warehousing, Spreadsheet, Tableau]

2.2 Aims

The aim of this project is to produce web-based visualisation in form of a dashboard which will provide an insight into the travel and behavioural patterns of UK residents. Recommendation will be produced for countries/continents using the visualisation tool.

3 Objectives

The timeframe is given in [Weeks. Days]. The productive working hours per day is 6 hours.

Research [3.1]

3.1 Conduct a Feasibility Study [0.3]

- 3.1.1 Research the current need for data analysis in the industry [0.3]

3.2 Research about the tourism industry [0.5]

- 3.2.1 Research why the industry is booming over the years [0.2]
- 3.2.2 Research the types of businesses in the industry [0.2]
- 3.2.3 Research the issues in the industry [0.1]
- 3.2.4 Research on data collected in this industry [0.1]

3.3 Research Business intelligence and visualisation [0.6]

- 3.3.1 Research into the advantage and disadvantage of Business intelligence and visualisation [0.2]
- 3.3.2 Research the types of visualisation [0.1]
- 3.3.3 Research types of business intelligence tools needed for the project (Data warehousing, spreadsheet and tableau) [0.1]
- 3.3.4 Compare similar work done using BI tools. [0.2]

3.4 Research relevant topic areas [1.0]

- 3.4.1 Research Jakob Nielsen heuristics [0.1]
- 3.4.2 Research Data mining and Big Data [0.2]
- 3.4.3 Research ETL Process [0.2]
- 3.4.4 Research tools, techniques and process needed to create ERD [0.1]
- 3.4.5 Research Legal, Social, Ethical and Professional issues with the project [0.1]

Analysis [2.4]

3.5 Design rich picture relating to the project [0.1]

3.6 Business and Product Requirements [1.4]

- 3.6.1 Produce Business requirements [0.5]
- 3.6.2 Produce Functional requirements using MoSCoW [0.3]
- 3.6.3 Produce Non-Functional Requirements [0.3]

3.7 Design Use-Case diagram [0.1]

3.8 Design data warehouse model [0.5]

- 3.8.1 Design a conceptual ER Diagram [0.1]
- 3.8.2 Design a logical ER Diagram [0.1]
- 3.8.3 Design a physical ER Diagram [0.3]

Design & Implementation [7.2]

3.9 Produce visual concepts for prototypes using Low-fidelity wireframe [0.2]

3.10 Create a Data warehouse [1.1]

- 3.10.1 Use ETL Process to create a data warehouse using MySQL [1.0]
- 3.10.2 Test the Data warehouse using SQL queries. [0.1]

3.11 Produce prototype using Excel Spreadsheet [2.1]

- 3.11.1 Transfer data from the Data warehouse to excel spreadsheet [0.1]

3.11.2 Using the data, create a visualisation using pivot tables and charts to produce useful information [1.3]

3.11.3 Combine the visualisation together to create a dashboard interface [0.2]

3.11.4 Test the dashboard interface [0.1]

3.12 Create Final Visualisation tool using Tableau [2.3]

3.12.1 Transfer dataset to Tableau [0.2]

3.12.2 Produce visualisation in form of a dashboard [1.1]

3.12.3 Test Dashboard interface [0.2]

3.12.4 Publish Visualisation tool to the web [0.5]

3.13 Create website to host dashboard [1.3]

3.13.1 Create website and Transfer dashboard to the webpage [1.0]

3.13.2 Add necessary features to the website [0.3]

Testing [1.4]

3.14 Produce Testing Documentation using various testing techniques [1.4]

3.14.1 Test MySQL Data warehouse [0.3]

3.14.2 Test Excel Spreadsheet Dashboard [0.4]

3.14.3 Test Tableau Dashboard and Website [0.4]

Evaluation and conclusion [1.5]

3.15 Evaluation and conclusion for Project [1.5]

3.15.1 Summaries the key finding from the project [0.3]

3.15.2 Produce recommendation using the information gathered from the visualisation tool [0.3]

3.15.3 Usability evaluation using cognitive walkthrough [0.1]

3.15.4 Project evaluation [0.1]

3.15.5 Personal Evaluation [0.1]

3.15.6 Write a conclusion for the project [0.2]

3.15.7 Write future development of the system [0.1]

4 Approach

The methodology which was chosen for this project is DSDM atern which is the latest version of DSDM. DSDM is the agile method which stems from RAD, and the traditional DSDM methodology contains 12 principles which can be followed to have a successful project within a shorter period and fixed resource. However, the latest DSDM atern only has eight basic principles which can be followed if you want flexibility in the project. When following this methodology, it allows the project to produce a minimal viable product by different techniques such as timeboxing, MoSCoW, etc. There is also less documentation done when using this methodology as it is the focus is to give the product needed for the business. This will save much time because you do not have to get the task approved by the higher-ups in the company.

For this project, many different techniques from DSDM will be used to ensure the project finishes on time with a useable product. Prototyping will be used to ensure that the issues are found in each prototype so that the final product will have less problem overall. MoSCoW will be used to priorities the important requirement first and will help to finish the project before the deadline.

4.1 Justification of the suitability of a Methodology or a Framework followed.

The project will be using the Agile DSDM Atern to allow flexibility when proceeding with the project. Techniques such as MoSCoW prioritisation, Timeboxing, Prototyping will be used to deliver the project on-time and only produce the requirements which are needed to run the business. As the project does not have a fixed requirement, the project will make sure to be adaptable to new changes if any arises during the project stages. Following the DSDM atern principles will benefit this project by delivering on time and not compromise on the quality of the application because it will only contain the essential components needed to answer the queries. The project will use iterative development and timeboxing together to produce the visualisation tool (dashboard) which will contain the necessary functions and features and will deliver on time. There are some disadvantages when it comes to using DSDM atern but tailoring the principles to suit this project will allow it to become successful and will encounter fewer problems.

By using this methodology, the project will be able to produce a minimum viable system in a few months which can offer some benefits earlier and can gain an idea into the possibilities of the system. The prototypes created using this methodology will make sure less problems are encountered in the next prototype/products. Every iterative, new prototype will be created which will improve the overall product at the end of the project.

5 Literature Review

5.1 Approach to literature searching

To conduct this research, it is crucial to make sure that all the information which are collected are peer-reviewed and cited by many academic people. The research will focus on the tourism industry and what issues it is facing currently by looking at important academic journals and articles. I will also investigate the definition and purpose of business intelligence and why it is needed for this industry by reading essential books which are related to this topic area. I also explored visualisation and why it is essential to use different types of graphs and charts. All this information has been gathered by using google scholar, emerald insight database and other databases which contains academic articles, journals and reports.

5.2 The success of the tourism industry

Currently, the tourism industry has a significant impact around the world by boosting the country's economy enormously and allowing new technologies to emerge also providing "income creation and generation of jobs" (Agaraj and Murati, 2009). This is further explained by (Alsos, G. A et al, 2014) "Tourism plays an important role in economic development by sustaining employment and bringing in foreign currency". The word 'Travel' and 'tourism' have two different meaning; travel means that you will travel for a more extended period and tourism is meant to be about travelling for pleasure and other purposes. However, both have one thing in common which is travelling, and they can be interlinked when talking about the tourism or travel industry. Over the years, the industry got popular with the help of new and improved innovations and various additional strategy. However, the tourists themselves have made the industry popular by spending more on travelling than usually.

As reported by the world tourism organisation, the worldwide expenditures from international tourists for 2016 were 1.362 trillion US dollars (Data.worldbank.org, 2019). This means that international tourists spent about 113.5 billion US dollar per month in the tourism industry which is a significant increase from the past decade. Because of this, the majority of the countries around the world invest their resources on tourism to ensure that they can grow and benefit from this industry sector. Furthermore, according to UNWTO, tourism also helps the countries with preservation of their culture and environment which is an essential aspect for all the countries around the world and will ensure the industry does not go down easily (E-unwto.org, 2018). However, UNWTO also found out that the purpose of visit for tourist has changed over the years from business purpose being 13% to leisure/holiday orientated which 55% (E-unwto.org, 2018). This will mean that people are willing to spend more on their holidays and it will keep the industry profitable if it is done correctly.

5.3 The current issues with the tourism industry

The tourism industry has grown into an enormous industry with many advancements in both the economy and technology. However, there are some fundamental issues which halt the industry from progressing any further.

A few years ago, the industry was providing products and services to its customers using face to face communication and TV advertisement. However, as technology has advanced, customers can plan and book their holidays without the help from most of the tourism industry. Furthermore, Pantano and Pietro support this by stating that "... the web is the main primary source of information for searching or booking suitable travel destinations." (Pantano and Pietro, 2013). This indicates that some of the tourism industry (such as travel agency, ticket office, etc.) can be affected negatively because of technology advancement and customers taking responsibility of booking holiday and making plans into their own hands. The tourism industry needs to make sure they are keeping up with the current technologies which the customers are using in their day-to-day lifestyle and ensuring they are using the critical tools needed for keeping up with competitors.

The industry also produces a significant amount of data which are very valuable. According to (Fuchs, Höpken and Lexhagen, 2014), "an organization's value is limited by the amount of knowledge within it" and the useful information gathered from the data can be crucial in finding insight into their customer's preferences and understand which action to take to make the business successful. The tourism industry needs to ensure that they are using the necessary techniques and tools to process the data into useful information as it can affect their business if not taken seriously. David supports this by stating that "Good information is essential for effective operation and decision making at all levels in businesses." (Kaye, 1995).

The main concern for the tourism industry and tourists is safety and security. According to Istvan and Krisztina, security and safety around the world have decreased due to "terrorist acts, local wars, natural disasters, epidemics and pandemics ..." (Kővári and Zimányi, 2010). Furthermore, Michael supports this point by saying "Tourists behaviour and, consequently, destinations, are deeply affected by perceptions of security and the management of safety, security and risk" (Hall, Timothy and Duval, 2012). This means that the industry must make sure that the security and safety of their services or products need to be robust and secure for the customers to gain trust in the industry. The security can be referred to both physical and online as more crimes are occurring through the web and in the physical location of the business.

5.4 Technological impact on the Tourism Industry

Technology has been a crucial component for the tourism industry and has allowed it to change the way people look at the industry. It also produced various benefits and opportunity for the industry (such as better communication, lower operating costs, better productivity, etc).

Furthermore, Dimitrios and rob supported this point by saying that "Due to the popularity of Internet applications, most tourism organisations ... have embraced Internet technologies as part of their marketing and communication strategies" (Buhalis and Law, 2008). Dr Dimitrios adds more to this by stating "As information is the life-blood of the travel industry, effective use of ITs is pivotal" (Buhalis, 2000). The industry has now become dependent on technology to be functional. Technologies have helped in finding hidden opportunities which normally cannot be found manually by employees or employers. According to Haiyan and han, "big data can be used to tailor marketing campaigns and find business model inefficiencies." (Song and Liu, 2017). Big data is one of the technologies which the industry is utilising to get a competitive edge and run their business efficiently. It allowed businesses to use their existing data and find useful information within a few hours. This means that they do not have to wait for more extended period to find out what they are doing wrong in the business and can quickly make a critical decision using the information from

using technologies like big data. However, they can also collect necessary data from their customers who leave digital footprint when using the internet or other services.

Technology advancement has allowed customers to make their own choices and made the job of the tourism industry easier. All the services provided to the customers have been faster and straightforward and because of this customer are more willing to spend on the industry without overthinking. Furthermore, Dimitrios and rob believed that “The Internet-enabled consumers to engage directly with suppliers and to challenge the role of intermediaries” (Buhalis and Law, 2008). Because of this, the industry has grown steadily, and anyone who does not keep up with the technology advancement and customer’s preference will shut down eventually.

5.5 The usefulness of Business Intelligence in the Tourism Industry

Business intelligence (BI) contains all the necessary techniques, tools and technologies needed for collecting, analysing and visualisation of raw data from businesses which then can be used for making strategic and operative decision making (Dedić and Stanier, 2016). Business intelligence allows businesses to understand the data which they have collected through customers, internal sources and external sources. The data will be transformed into useful information which the businesses can use to make a critical decision. It also allows the business to view the real-time progress by producing essential reports which can be used by the higher-ups. The information will also be more straightforward to read and understand when the data is processed using BI tools which is the main benefit as it allows the higher-ups in the business to make a decision quickly and efficiently.

The tourism industry is collecting a variety of data from customers digital footprint to receipts of purchase. Typically, these data would not be useful, but with the help of BI tools, it can be possible to find many useful insights about their customers. According to (Mariani et al., 2018), BI allows the tourism industry to understand the needs and wants of the customers and believes the businesses main approach nowadays is customer-centric. The BI tools can be used by small to large businesses for various purpose. It can help with understanding what the business has done from the past, present and future to see if there is any correlation between their actions taken and results produced. The BI tools can produce useful information (such as graphs, tables, answers to queries, etc) which can be useful for the industry. Database and visualisation are mainly used when trying to find insights and hidden patterns in data. The BI tools also include spreadsheets, OLAP, data mining, etc.

5.6 The use of Data Visualisation for this Industry

One of the techniques from business intelligence is visualisation. There are many methods when it comes to visualisation, but all of them require data to be effective. Different visualisation methods are needed for different purposes, and it is essential to make sure that the right type of visualisation is chosen to get useful insight from the data. The types are information visualisation, data visualisation and scientific visualisation. For this report and project, the focus will be on data visualisation as it is the main component in creating a dashboard for this industry. The data collected from various sources can be converted into useful information by displaying in a graphical

image, and this method is called data visualisation. According to (S and Naik, 2017), “The data are represented in a graphical format to achieve effective communication”, and this proves that it can be a useful technique for any industry that wants to understand the data which they have collected.

The main advantage of data visualisation allows managers and higher executives to understand the information without needing to understand the whole topic. The tourism industry is a fast-paced and complex field, so using visualisation will ensure the higher-ups understand what the data is indicating to them about their business. According to (Sas.com, 2018) who supports this point, stated that people who view the visualisation will be engaged and will get the message of the visualisation quicker. All the visualisation can also be combined to create a dashboard which will contain all the necessary information needed to view the progress of the overall business.

Furthermore, (Few, 2006) states that a dashboard purpose is to display information in a single-screen and allow users to efficiently monitor the information so that they can accomplish their tasks. The tourism industry will gather a variety of data and using visualisation with a dashboard interface will allow the business to see all the essential information needed to make critical decisions on one screen. It will save time and resources as the dashboard can be checked anytime, they need to decide for the business. It can also be used to find any trends in the tourism market which is very important as the industry's trends changes very offend depending on the technology advancement and consumers preferences and needs (Sas.com, 2018). The benefits gained from using data visualisation is significantly high. However, the importance of data accuracy needs to be considered when it comes to the use of data visualisation for decision making. (Cotton, 2014) States that inaccurate data “can cause real detriment to a business and its bottom line” and it can occur by people using data visualisation with inaccurate data. Furthermore, (Hassan, 2003) elaborates cotton’s points by stating that inaccurate data can lead to various problems within a business. When inaccurate data is used, it can challenge the usefulness of decisions making and increase the doubt on trustworthy and reliability of the business. The customers who are affected by the decisions will also be disappointed and lose confidence in the business.

5.7 The importance of using various graphs and charts to display different types of data

All the data which are collected from various sources are mostly in numbers which are too complicated to understand when trying to read them. That is why graphs and charts are used for translating the data into useful and digestible information which people can understand. (Slutsky, 2014) supports this point by saying “The purpose of a graph is to present data that are too numerous or complicated to be described adequately in the text and in less space.”. This means that graphs and charts to help with understanding the data more clearly and allowing necessary information to be displayed in smaller spaces. Furthermore, Slutsky elaborated about why it is also essential to choose the right graphs and charts when trying to display different types of information. One type of graph cannot display all the information accurately and can be harder to understand what it is trying to convey to the user.

5.8 Conclusions

To conclude, having done thorough research about the tourism industry, business intelligence and visualisation, I have gained a good understanding into why business intelligence will be a good choice for displaying information and producing reports and graphs. Furthermore, I understood the current issues which are occurring in the tourism industry and why business intelligence can be useful for the industry. In addition, the BI tools can also be used by both smaller to larger businesses in the tourism industry as the cost to produce useful information can vary depending on their budget and resource. Moreover, using visualisation ensures that the manager and the higher executives understand the information quickly and can make a critical decision without needing to understand all the data.

All the research done on the topic has influenced the use of visualisation for the tourism industry to understand the trends and hidden patterns for each country. The data needed for the visualisation are mostly collected from internal sources. However, for the project, I will collect the necessary data from the external source from the internet. I will also use a dashboard interface to combine all the essential visualisations needed for the tourism industry by adding different types of graphs and charts. For future development, I may research more into big data and different techniques which can be used to process and display data differently.

6 Further Research

6.1 Business intelligence

Business intelligence's main purpose is to use various technologies, processes and application to analyse data and produce valuable information which can be used by managers, executives and end users to make business decision (Rouse, 2017). Furthermore, it contains tools and techniques which can be used for collecting and translating the data which then can be used for creating reports that can show insight into the data which was collected. The advantage of business intelligence is that it can produce useful information from historical and present dataset. This enables the managers and executives to identify what problem the business is facing and view what it is doing right. Moreover, current organisations are more willing to focus their resources on BI tools as the advantages greatly outweighs the negatives. Some of the advantages are competitive edge, improved data quality, quicker and easier reports, cost efficient depending on the dataset size and better business decision (Janoschek, 2017). All this combined will enable a business to stay a step ahead of their competitors and help increase customer satisfaction. However, there are some disadvantage which can affect the business negatively if not taken seriously. The main disadvantage is incorrect use of BI tools and techniques because the cost to implement and use BI tools are much higher if the business is very large. The business will need to pay for hardware and software cost, implement the tools and techniques cost and employing people cost. If the BI tools were to be used incorrectly, it can be a huge waste of resources which can affect the business operations and limit them from moving forward (Charles, 2014).

6.2 Business intelligence Tools

The popularity of business intelligence in the business field as allowed it to gain more tools and techniques which have been updated or created to work with the current needs in the data

analytics field. There has been numerous types of software, processes and technologies which can be used for collecting, modifying and transforming data. BI tools can contain dashboards, visualisations, ETL process, OLAP, Data mining, etc (Softwareadvice, 2019). There are many different software's in the market for small to big organisations and it is important to choose the right software to ensure that it can be used properly in the company. The most common software which can be used for business intelligence is excel spreadsheet as it provides the basic essential to create reports and interactive dashboard. Small companies and individual can use this software to create a basic reports and dashboards to gain insight into the data which they have collected. However, if you want to publish the dashboard to the internet, then a software called tableau can be used which has gained a good reputation in the BI community. It is crucial to consider the types of database which will be used to create the dashboard as poor quality and unstructured data will significantly affect the outcome of the reports from the dashboard. Many industries use MySQL, SQL Server and Oracle Server as their main DBMS, and it is important to choose the appropriate DBMS when trying to produce quality data.

The common process which is used with Datawarehouse is ETL process which is used for cleaning the raw data and getting them ready for reporting stage. ETL stands for Extract, transform and load. The process starts from the raw data being extracted from different sources which can be internal, primary or secondary sources. The data will then be transformed by arranging them appropriately and converting from unnormalised data to 3NF. The final step is to load the data into a Datawarehouse and all the data from the 3NF table will need to be converted into 2NF. 2NF allows for faster and easier querying, but the trade-off is the need for more storage space which will require a lot of resources from the business or individual. After coming the ETL process, the Datawarehouse will be ready to be used for analytics. This process ensures that the Datawarehouse is updated regularly with new data without making a lot of mistake which can happen when working with different kinds of data from different sources or DBMS (oracle, 2019).

Dashboard is another component of business intelligence and it is used for “track KPIs, metrics, and other key data points relevant to a business, department, or specific process.” (Klipfolio, 2019). The dashboard ensures that complex data sets used in the data visualisation are made easier for the user to understand and allows the users to view all the performance at first sight of the dashboard. Dashboard will allow the user (such as executive or managers) to gain insight into the industry without needing to understand all the complex data and can be a useful tool when trying to make a business decision. There are many different types of dashboard (such as timeframe, audience and objective) which are used in different situation and it is important to choose the appropriate type when creating a dashboard.

Data mining can be a very valuable tool when trying to analyse a large set of data. Patterns and insight can be found in complex massive dataset which can take a longer period when trying to use other techniques to analyse the data. It can also be used for various other application such as AI application, machine learning, etc. It is important to consider this technique when dealing with a massive dataset which cannot be handle using normal methods. Big data is another technique which similar data mining is, but it focuses on analysing the dataset to get insights which can be used for business decisions and strategic moves. Big data can be useful for finding issues within the industry, predict trends and patterns, analyse behaviour of user, etc.

6.3 Jakob Nelson heuristics

Before creating a system, it is important to plan and design the user interface as it can allow us to

see what the system should look like and what it requires to make it functional. This can be done easily by following rules and guidelines which have been tested by many UI experts. One of the popular guidelines for UI design is Jakob nelson's 10 heuristics. It is not compulsory to follow them, but it can help make sure the new system's user interface has the necessary features and functions which will allow the system to have less problems and attract more users.

1) Visibility of system status

The system needs to keep the user informed about what the system is doing, and this should be done by providing the feedback within sensible time.

2) Match between system and the real world

The system needs to use language which the user understands by using different words and phrases. The information needs to flow naturally and logically.

3) User control and freedom

If the user wants to undo their mistake, there needs to be an exit option to leave the state which they don't want to be in. It should be easy and quick without a lot of dialogue.

4) Consistency and standards

The user interface should be consistent by following and using similar actions and language to ensure the users are not confused. Consistency allows the user to learn the system faster.

5) Error Prevention

Preventing error is better than error message and it will ensure the user will encounter less problem when using the user interface. The error cannot be solved, then an appropriate error message should be displayed before the user commits the action.

6) Recognition rather than recall

Reduce the number of things the user must remember when using the user interface. Make the instruction for the system clear and visible.

7) Flexibility and efficiency of use

Ensure to have different functions and features for both experienced and inexperienced users. Different users will want to use the system faster as they may be familiar with the system.

8) Aesthetic and minimalist design

Any information in the user interface should be relevant and useful. Irrelevant information can confuse the user and provide a bad experience. Minimalist design ensures the user can concentrate on the content of the system.

9) Help users recognise, diagnose and recover from errors

Error message should be easy to understand and use basic language. The system should explain or show the problem properly and provide recommended actions to take to solve the problem.

10) Help and documentation

Provide help and documentation for users for the system and can be used to find out how to use the system.

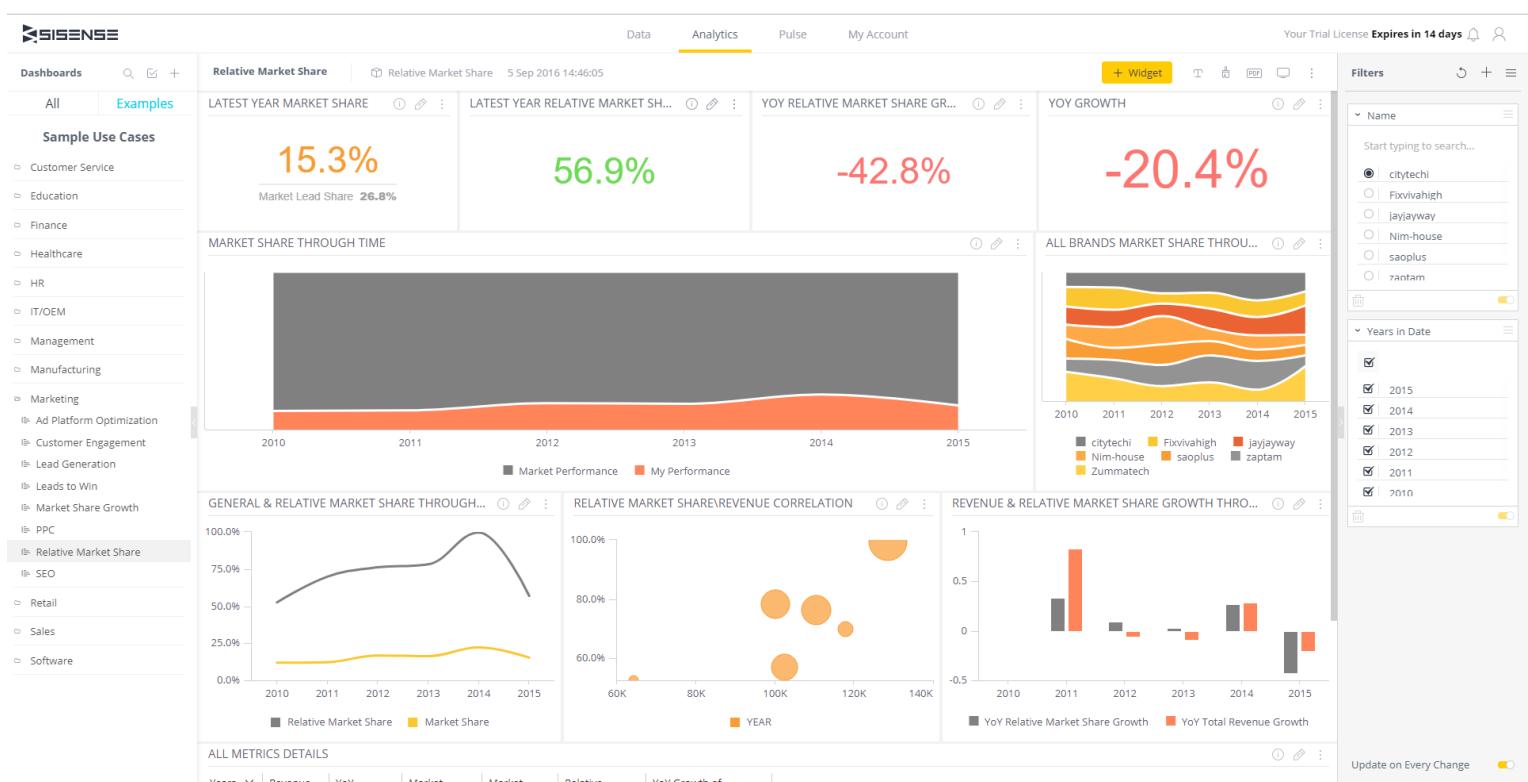
(Nielsen, 1994)

Conclusion - After conducting further research, more requirement will be added to ensure the overall project is good. The nelson heuristics will be used to find out the usability of the system created for this project. The data warehouse will also be used for this project which will contain all the data sets required for the creation of the dashboard. The ETL process will be used to create and populate the data warehouse.

7 Product Research

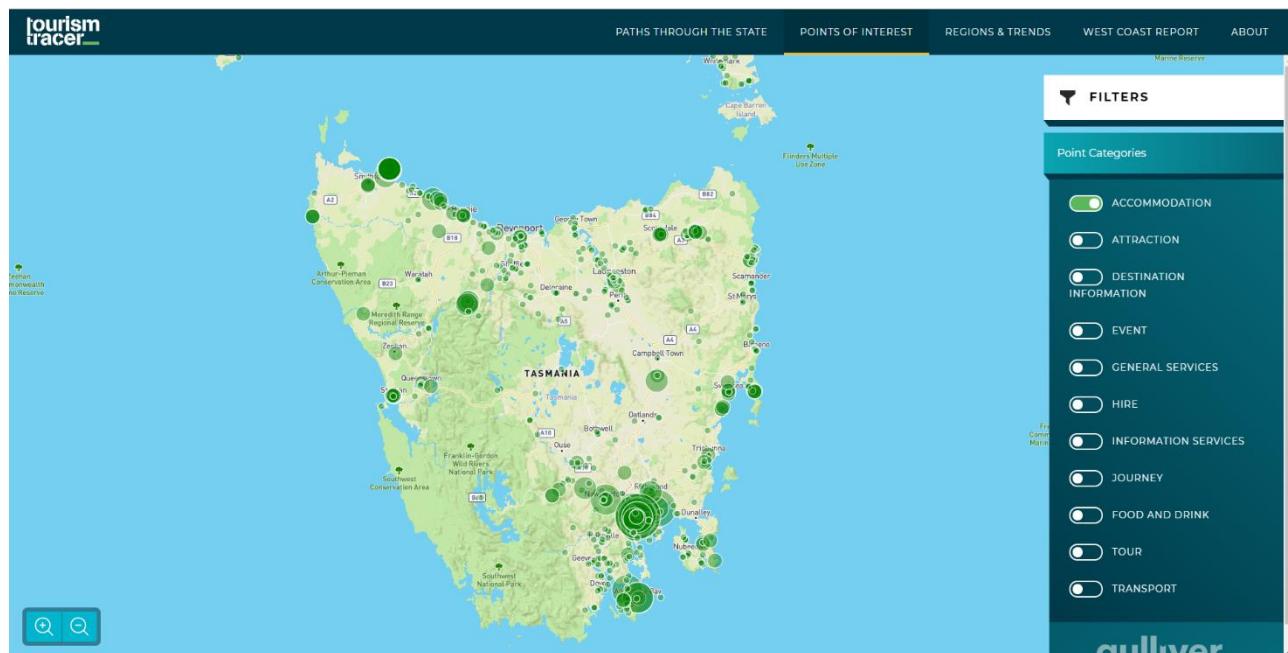
There are many different types of product which are very similar to the product from this project, and it is important to understand and learn from these products so that it will help us find new issues which we can avoid when trying to create the product. Conducting this research will also produce new ideas which can be added to the product to make it more effective. This project will investigate different types of products and compare what they have in common and what they are doing differently. With this information, I can then decide on how to go about designing and implementing them into my product.

Sisense Dashboards for Marketing: <https://www.sisense.com/dashboard-examples/marketing/>



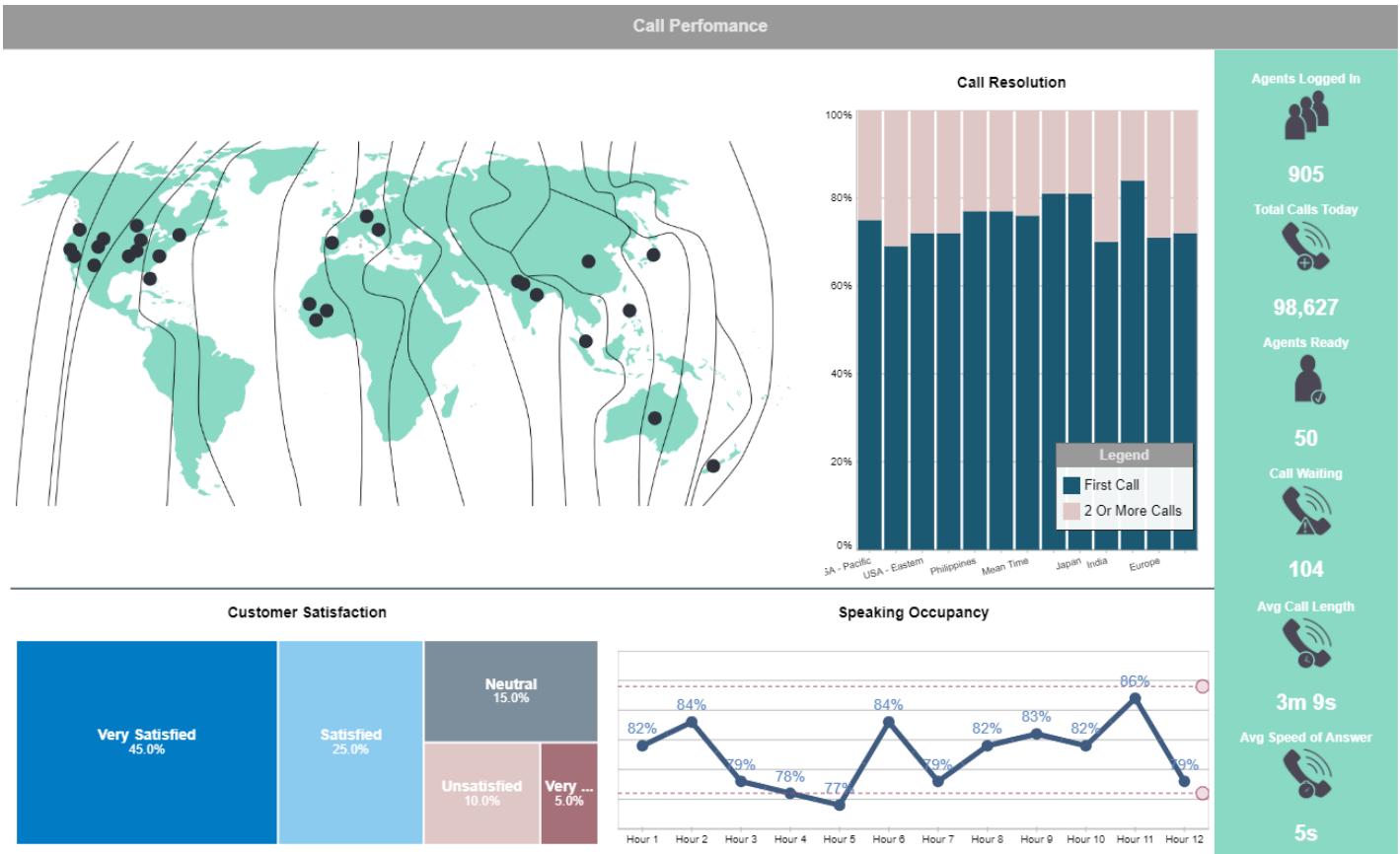
This product is from sisense, and they produce different types of the dashboard for different purposes (such as marketing, IT, education, health, etc.) which can be accessed using the internet. All the dashboard created by them is web-based meaning that it can only be accessed using the internet. This product contains all the necessary information about marketing and uses different types of graphs to display crucial information. The product ensures that different data requires different graphs and charts. The colour and font used in the dashboard allow the user to read the information clearly and does not distract them with unwanted content. However, the disadvantages are the use of buttons and checkboxes. There is a lot of them in the dashboard which makes it harder to focus and can be difficult for new users to understand how to use the web application.

Tourism Tracer – Point of interest : <https://tasmania.tourismtracer.com/points-of-interest>



This dashboard is produced by tourism tracer who gather information about different countries from around the world. The image at the top shows all the points of interest for Tasmania, and they also produce more information by using the toggle buttons in the side of the screen. This style of the dashboard is interactive and makes it easier to understand the data quicker. The user will also find it enjoyable to use because it uses a realistic map to show the whole state. The design they follow for the dashboard is minimal and simplistic which allows a different type of users to use the dashboard. The disadvantage of using this dashboard is that there is no information about the points of interest which means that people will not be able to use this dashboard to gather information about the country.

Idashboards – Call centre dashboard: <https://www.idashboards.com/dashboard-examples/call-center-dashboard/>



This dashboard only focuses on the call performance of the call centres from around the world. As you can see, the dashboard uses minimal design which ensures the user can use the dashboard with ease and prevent errors from occurring. The colours and fonts used are very similar to make sure that all the details are consistent which means that the user will be able to focus on understanding the information from the dashboard. The language used here mimics the real-world language by using simple words. The dashboard also includes important information on the right side and making sure that irrelevant does not exist which can distract and waste the user time. However, the dashboard does not provide information about how to use the dashboard which is important for first-time user of dashboard interface.

Conclusion

To conclude, each of the products which I have looked at has different methods of displaying information. Majority of the dashboard uses minimal colours and similar fonts to ensure the users have an easier time when trying to read the information.

After conducting this research and comparing different products, I was able to understand what others are doing and what to include in my final product. For my final product, I will follow the sisense style by using a variety of graphs and charts and try to use maps to display information which enjoyable to view and easier to understand. As my project is related to tourism, using a map will be essential when trying to create my own dashboard. The research has also helped with understanding the need for colours and fonts and all the products have similar colour theory to

keep it consistent. It can also affect the system's success when it is used by the end user. I will make sure to choose the appropriate colours and fonts by looking into the user groups for this product.

8 Legal, Social, Ethical and Professional Issues and Considerations

The dataset needed for this project will be collected from an online source. The project will follow the **data protection** laws and regulations by only allowing only authorised users to gain access to the data used in this project. Also, ensuring the data cannot be accessed, modified or deleted by unauthorised users.

The project will also respect the **copyright** law by following the terms and condition, to use the dataset in their website. Before using a dataset, it is professional to request permission from the owner of the dataset to get their **consent**. Having contacted the owner of the dataset which is needed for the project, he replied by saying that the dataset can be used and adapted under the terms of (Open government license - OGL) and any dataset used in the project must need to be referenced. (**See Appendix A**)

Equality will be taken into consideration when analysing issues related to gender and age. **Culture-related issues** should also be considered when providing recommendations to the countries because each country may have a different view of thinking.

Social issues can be both negative and positive, and it is mostly controlled by people. There are many issues which stem from social issues such as work, health, education, age, etc. it is vital to investigate **unemployment issues** for this project because the web application tool may change the industries perspective and can lay off their employees because of the results they found.

One of the ethical issue is **IT professionals misusing their power**. The IT professional will have control over all the data in the system and it is important to follow the law and guidelines when handling the data or information from the business. Another issues is also **misrepresenting the true** because providing the incorrect information can have a negative impact on the people who depend on the information.

Ethical and professional issues will be less relevant, has the size of the workforce for this project is smaller.

9 Requirements

Requirements help to understand what tasks are needed for completing the project and ensures that everyone who is working for the project understands what must be done to meet the client's needs from the project. It is important to begin a project by doing requirement analysis as it allows the stakeholders to voice their problems with their current system and what they want the new system to do. Understanding the customers problems is the key to finding out what the new system will require. However, there are a variety of requirements which support different parts of the project (such as, functional requirement, non-functional requirement, business requirement, and so forth). First, it is crucial to figure out all the high-level requirement for this project without doing prioritising.

9.1 Analysis of requirements

When starting a project, it is essential to think about the requirements of the proposed system and the whole project. A list of high-level requirements will be required before starting the implementation stage, and it should contain all the requirement of the new system. This will be created by regularly communicating with the stakeholders in meetings and researching similar products from competitor.

	High-level requirement	Justification
1	Produce graphs and tables using the data from Data warehouse.	The whole system is a dashboard and without graphs and tables, it will not be useful to anyone. The dashboard needs to be use the data from Data warehouse or other sources to create the graphs and tables.
2	Data warehouse (tourism related data)	The dashboard main source is data and it is essential to have a Data warehouse to store all the data which have been collected for the dashboard. The Data warehouse can also be used for backing up the data which is helpful if you lose your data in the dashboard.
3	Access dashboard using the internet	The dashboard will contain useful information about UK tourist in different countries and continents and it is important for stakeholders from around the world to have access to this information as it will allow them to improve their business. If the dashboard is not online, then no one can see it or use it for getting insights into the data.
4	Let end users customise the data with filters.	The dashboard will come with default settings to show basic information. However, the stakeholder may want to use the data different by selecting specific

		dataset to produce a different insight. Filters in graphs and tables will grant flexibility when it comes to getting insights.
5	Standalone website for dashboard	Even if the dashboard can be accessed using the internet, it is not in a private website which can be used for producing revenue by including adverts. Having a website to show the dashboard will give flexibility in terms of the types of content which can be publish and follow terms and condition which are created by ourselves. The website can have features such as comments section, login, register, etc.
6	Automatic creation of report	When the system is finished, the data which it is using will be old after a few months. However, if it automatically creates new reports from new and update dataset, then it can benefit stakeholders from around the world for a longer period.
7	Display recommendation for the countries/Continents.	All the graphs and charts in the dashboard can show various types of data, but it is important for stakeholders to understand what the data is implying. The dashboard should display recommendations for each graph/charts/table on what action to take.
8	Register and login	User can use the accounts to access the website and the dashboard. It can also be used for communicating with other users and admin. There are many features which the user and the admin can benefit from having accounts.
9	Read and Write Comments	Writing and reading comments can be useful for admin and users because the knowledge and opinion are shared with everyone and could help gain new insight which the dashboard can't find.

9.2 Comparison of systems

Having done product research, there are various systems which are similar to this project. Each system from the research have similar features and functions with a few unique features and it will be compared together to find out new requirements for the system.

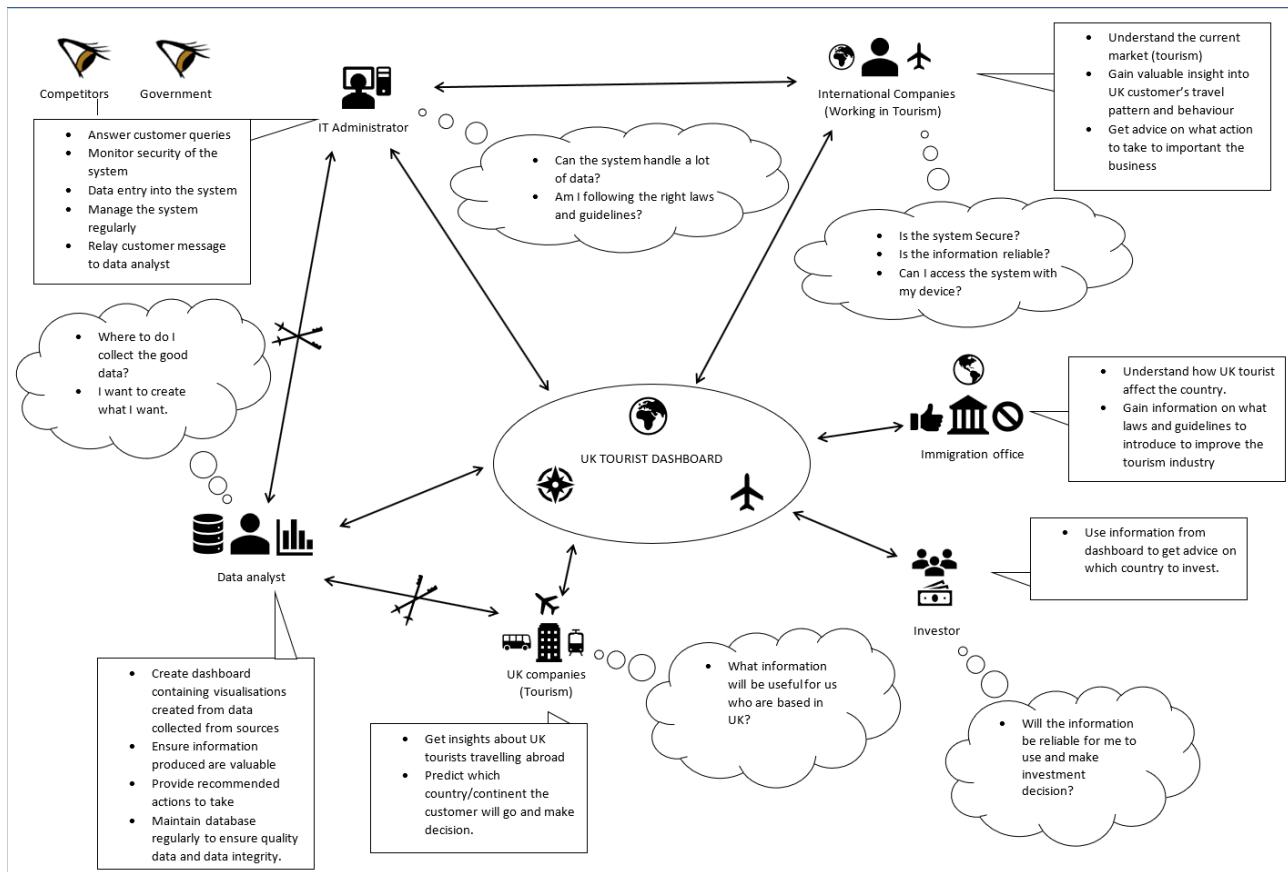
Sisense Dashboard	Tourism Tracer	Idashboards
Filter for each dataset and can add new filter	Filter for each dataset	Basic interactive map for selecting specific data.
Various graphs, charts and values (line, bar and scatter graphs)	Realistic map used for visualisation	Different graphs and charts used for representing different data types.
Basic colour scheme and design	Simplistic design and colour scheme	Bitesize information on the corner to understand the data with a glance.
Different font size, but similar font size	Same font size and style	Similar font style and colour for simplicity
Interactive graphs and charts	Interactive map with clickable content	Basic layout to keep it easy to use
Explanation for each graphs, charts and values.	Less information available about the data	Limited information about the graphs and charts
Too much information on the dashboard, so harder to concentrate	Not easy to understand the information	Too basic and does not allow flexibility when it comes to customising the data.
Not user-friendly for beginner users		

CONCLUSION

The research conducted on three different products as helped with getting new requirements which can be used when the prototype is being created. Each product provides different benefits and negatives, but it is important to choose the appropriate features and functions which will be added to the dashboard. After comparing all the product, sisense dashboard was chosen to be follow when creating the dashboard for this project. This is because it contains all the similar functions and features a dashboard will contain plus the negatives can be changed by added new ideas which can fix those issues. The prototype dashboard will contain basic colour scheme, similar fonts style and size and various graphs and charts. The map component from Idashboards will be used for the prototype dashboards because it provides easy to use map will information that are simple to access.

9.3 Business requirements

9.3.1 Rich Picture



Explanation of Rich Picture

This is a diagram which shows how the actors will react if the product was introduced. The diagram can be used to identify the issues which may arise when releasing the product to the environment. The overview of the product in the environment will aid in understanding the possible issues which can happen and by drawing this rich picture, more information has surfaced which can help with producing requirements for the project.

Each rich picture features have different purpose and meaning. The 'swords' indicates that there is a conflict between two actors and issues may arise between them in the future. IT administrator has a conflict with the data analyst because both have different opinion on how the system should be run also what content it should contain. The UK tourism companies has conflict with the data analysis because most of the data will only be useful for people from other countries and not the UK companies. The 'eye' is used to show the outside view who will have interest with the product but will not interact with the internal components of the system. The government and competitor are added in the rich picture because the government will want to make sure the system does not break any data protection law or any other law. On the other hand, the competitor will also want to see what their rival has in their product which they can use in their own system to boost sales. The thought bubbles will explain what the actors are thinking, and the square speech box will contain information about their responsibility/purpose when using the system.

The **business requirement** is necessary for understanding what a business must do to keep it running and successful. It will help to understand what the business goals are and what needs to be done for the business to be successful. However, it is not the same as a functional requirement which focuses on the functionality of the system and its process. A Business requirement does not change very often during the project as it will explain the core of the business. Here are the business requirements which must be taken into consideration when working on this project:

When undertaking the project, it is crucial to follow all the necessary LSEP guidelines and laws to ensure the project finishes successful without any complications. It is essential to obey the data protection and privacy law when collecting sensitive data from primary and secondary sources because the project mainly revolves around the use of data. Another business requirement is the importance of data collection. The project's purpose is to collect data and convert them into useful information which the industry can use and it is important to make sure that when collecting the data, the quality and accuracy of the data is looked at regularly to ensure the information produced are reliable. Finally, it is important to produce valuable recommendation for specific countries using the information gathered from the dashboard. These recommendations can be very vital for many of business in the tourism industry, so it is important to produce useful information which can give an insight into their industry in their country.

9.4 Functional requirements with MoSCoW prioritisation

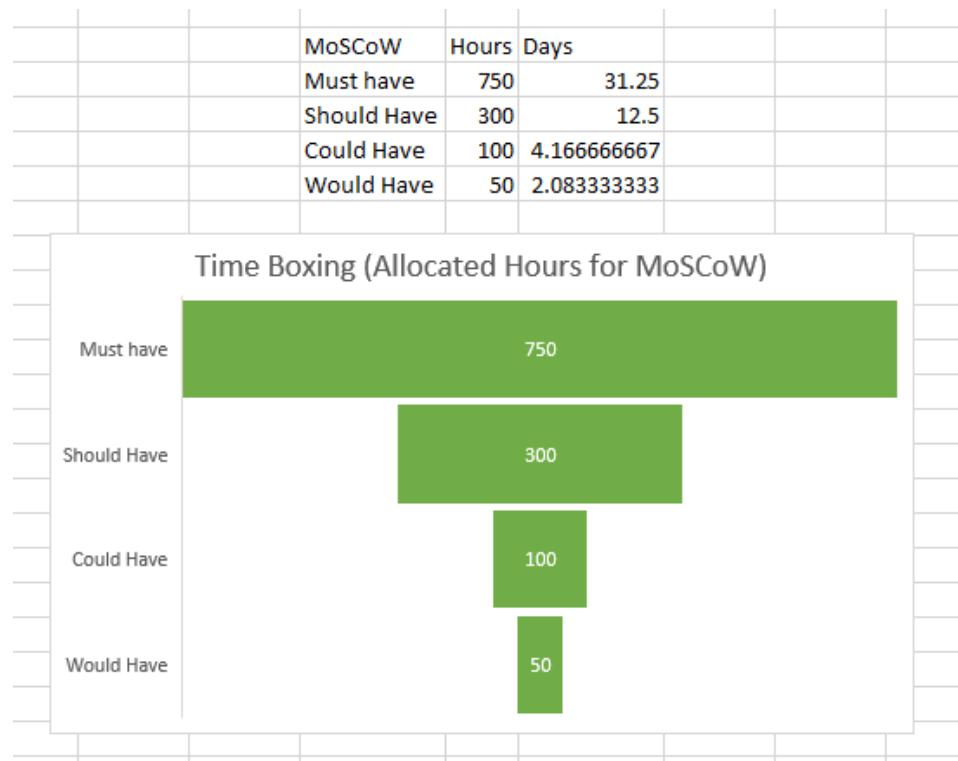
No	Functional requirements	MoSCoW	Justification
1	Produce graphs and tables using the data from Data warehouse.	Must have	This is a 'must have' because the focus of the final system is to produce graphs, charts, tables and values using the data from the Data warehouse. Without this function, the system will be useless.
2	Data warehouse (tourism related data)	Must have	The Data warehouse will contain all the data required for the creation of the dashboard. This is a 'must have' as it will be used for creating the graphs and charts, but mainly store the data collected from various sources.
3	Access dashboard using the internet	Should Have	This is a 'should have' seeing that the dashboard purpose is to give recommendation to stakeholders from around the world. Without the internet, no one can view the dashboard and it will be useless.
4	Let end users customise the data with filters.	Should have	This is a 'Should have' because the dashboard will contain visualisations which will have default information which can be useful for many end users. However, it is important to give the user the option to customise the

			data to get a different insight relating to their business.
5	Standalone website for dashboard	Could Have	The system will be published online and can be viewed by anyone from around the world using third party website. This is a 'Could have' because the current proposed system can be viewed without the need for a standalone website and the project deadline is very short. However, if there is enough time to create the website, then it will be produced as it can provide many benefits for the system.
6	Automatic creation of reports	would have	This is a 'would have' because after the dashboard is publish online, it may not be updated regularly and may become useless when the data is not updated. Having an automated process of collecting, storing, creating and updating the dashboard will ensure that future user can still benefit from the system. However, with the limited time frame and resources, it cannot be created this time but could be done in the next iterative cycle.
7	Display recommendation for the countries/Continents.	Should have	This is a 'should have' because the graphs, chart, table and values in the dashboard may not be understood by every user. Having some recommendations next to each content will make it easier for users to understand the data and can be used for making decision for their business.
8	Register and login	Should have	This is a 'should have' since the user will want to store any information which they acquired using the system. They can also use the account to communicate with the admin or data analyst or other users in the system.
9	Read and Write Comments	Could have	This is a 'could have' because it is not necessary to have this function in the system but may be useful for users to write and read comments about the dashboard. With limited time, this could not be possible to implement in the system.

9.5 Non-functional requirements

No	Non-functional requirements	Justification
1	Interactive graphs and charts.	The graphs and charts should be interactive so that users are engaged and interaction with the dashboard will allow more ways to view the information in the system.
2	Data warehouse should be backed up regularly	The Data warehouse will contain all the data which are required for the dashboard to function. Regular backup will ensure that if the data are corrupted or deleted, the backup can help to start the system again with minimal work.
3	Robust security for both the Data warehouse and the system to protect from malicious attacks and viruses.	Both the dashboard and Data warehouse can be accessed using the internet and it is important to ensure the data which are stored and used are protected from malicious attacks and viruses. Secure system will ensure the system will run for longer and earn the trust from the end users.
4	The system should be able to be accessed from different (OS) platforms.	The dashboard is intended for all stakeholders relating to the tourism industry and they will be in different parts of the world. However, not all users will use the same type of system and OS. It is important to make sure that the system can be accessed in different types of OS and browsers.
5	Data protection and misuse law must be followed when handling data.	When working with data, it is vital to follow the appropriate laws and guidelines to ensure the data is being handled correctly.
6	Maintain Data integrity by regularly using processes and software's.	After the dashboard is created and published, the data used for the system will be getting useless over a period and it is important to regularly update and maintain the Data warehouse by adding new or updated data to make the dashboard useful for the stakeholders and end users. Appropriate software and process should be followed when trying to maintain integrity of the data.
7	The system should be user-friendly and easy to understand.	The system should be easy to use as not all end users will understand how to use the system for the first time. Ease of use will ensure the end user learn how to use system and get the information they need quicker.

9.6 Time Boxing

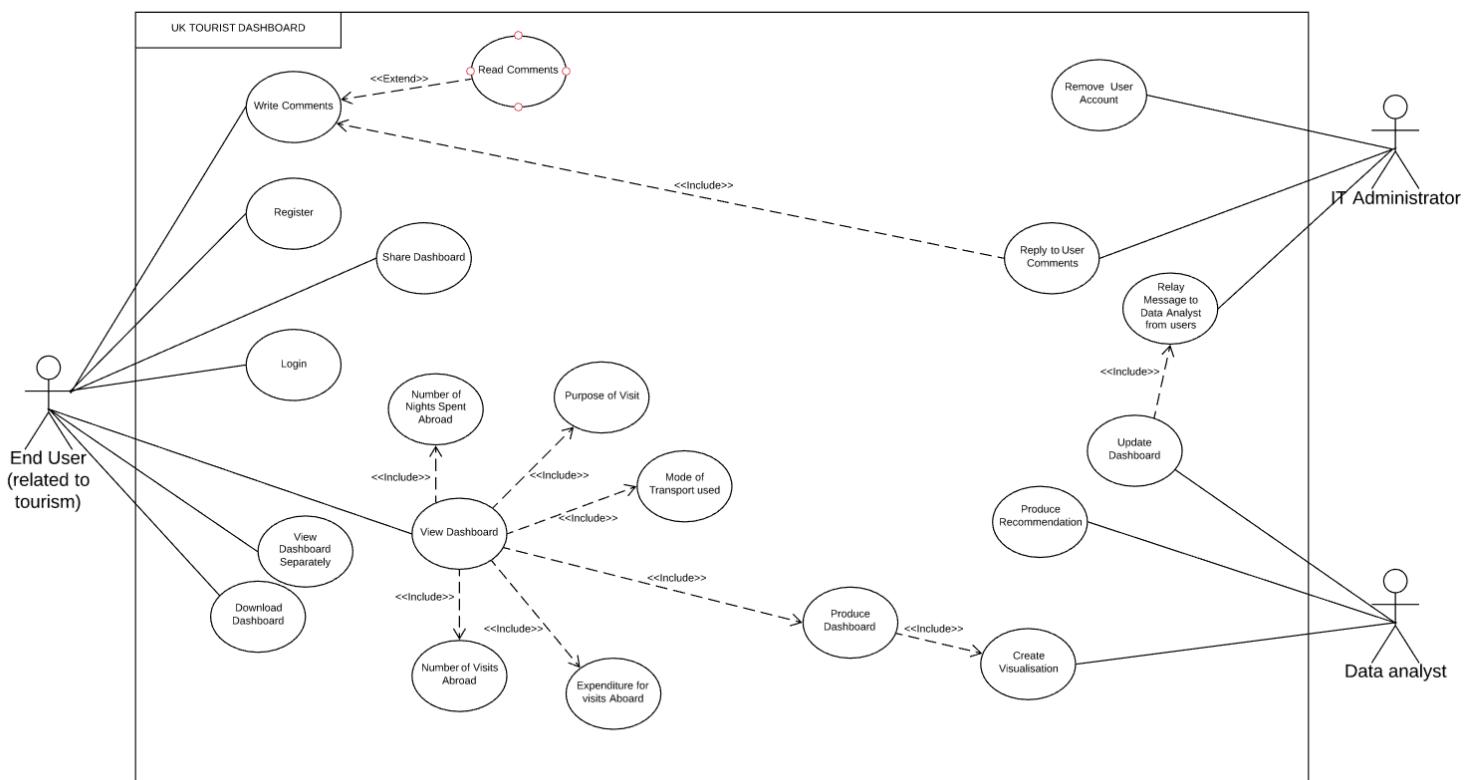


Time boxing can help understand which group of tasks should be prioritised and how much time can be put into them. Here you can see the duration for each section of MoSCoW and all the hours add up to 1200 which is equivalent to 5 months of productive working months. It is important to only use the hours listed here when continuing with the project.

10 Design

10.1 Use-case Diagram

Use case diagram is very important when you want to understand how the system will be used by the users. The user interaction with the system will help with understanding which functions and features are needed for the system. Use case diagram stems from UML and according to (Lucidchart, 2019) “a use case diagram can summarize the details of your system's users (also known as actors) and their interactions with the system”. When designing the use case diagram, it is essential to use different symbols and notations (such as, Use case, actors, association, system boundary boxes and packages).



To prepare for the use case diagram, it is crucial to identify the actors who will be interacting with the proposed system. As you can see in the diagram above, I have included actors who I believe will be using the system. Majority of the actors will be interacting with the same use cases. I have included all the use cases which are appropriate for the ‘user’ actor.

The ‘view dashboard’ has many ‘<<include>>’ relationships because each of the use cases are individual visualisation which make up the dashboard and the dashboard cannot exist without them. The administrator focuses on maintenance of the system and relays any message which the data analyst did not see from the users. The Data analyst collects the necessary data and produces visualisation which can then be combined to produce a dashboard. Using the dashboard, he can produce recommendation which the user can use to make decisions.

Use-case diagram - Assumptions

- ‘View dashboard separately’ was added because user may want to view only the dashboard to get insight quickly without looking at individual graphs and charts.
- ‘Produce recommendation’ was added in the use case diagram because after producing the dashboard, the data analyst can produce recommendation which the user can use to improve their business decision.
- ‘Relay message to data analyst from users’ was added since the user message is mainly looked by the administrator first, he/she can relay user message which are useful to the data analyst directly.

10.1.1 Primary and Secondary Use-case scenario

Primary Use-case Scenario 1

Use case:	Share dashboard
Actors:	End user
Precondition:	The end user should be in the website containing the dashboard.
Flow of events:	
Primary scenario	
1. The use case starts when the user is logged on to the system. 2. The end user goes to dashboard page. 3. The end user selects the ‘share option’ in the page. 4. The end user chooses their preferred sharing method and enters the required information. 5. The user case ends when the end user has finished sharing the dashboard and they are back to the dashboard page in the website.	
Postcondition:	When sharing is complete, a confirmation is sent to the end user’s email.

Possible Secondary Scenario which could occur

- The user did not know their username and password.
- The user did not have any account with any of sharing methods in the dashboard.
- The website could not share the dashboard.
- The dashboard does not function properly.

Primary Use-case Scenario 2

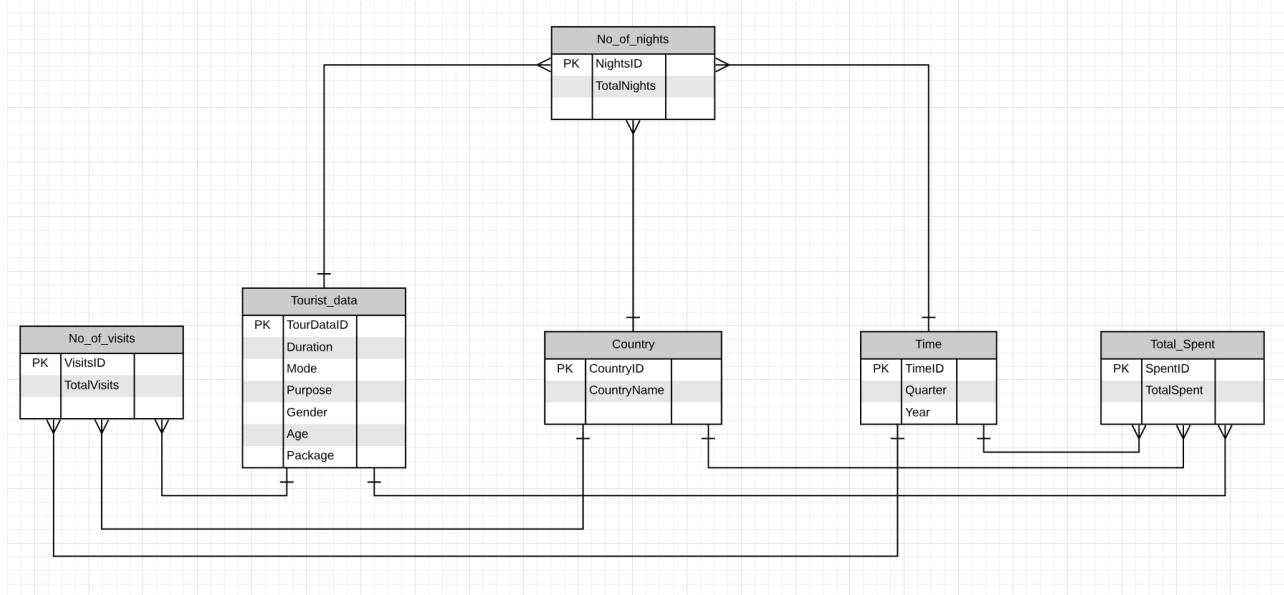
Use case:	Create visualisation
Actors:	Data analyst
Precondition:	The appropriate data should be stored in the database.
Flow of events:	
Primary scenario	
1. The use case starts when the data analyst connects the database with the system used for creating visualisation.	
2. The data analyst chooses essential attributes from the database using the system to create the visualisation	
3. The data analyst combines the data together to create graphs, chart or maps.	
4. The use case ends when the data analyst creates a set of graphs, chart or maps which are suitable for dashboard	
Postcondition:	The visualisations are combined to form a dashboard

Possible Secondary Scenario which could occur

- The data analyst does not have the necessary data to create the visualisation.
- The system does not create the visualisation which the data analyst wanted.
- The graphs, charts or maps were not valuable or could not be understood by the user.
- The attributes needed for the visualisation did not exist in the database.

10.2 Entity Relationship Diagrams

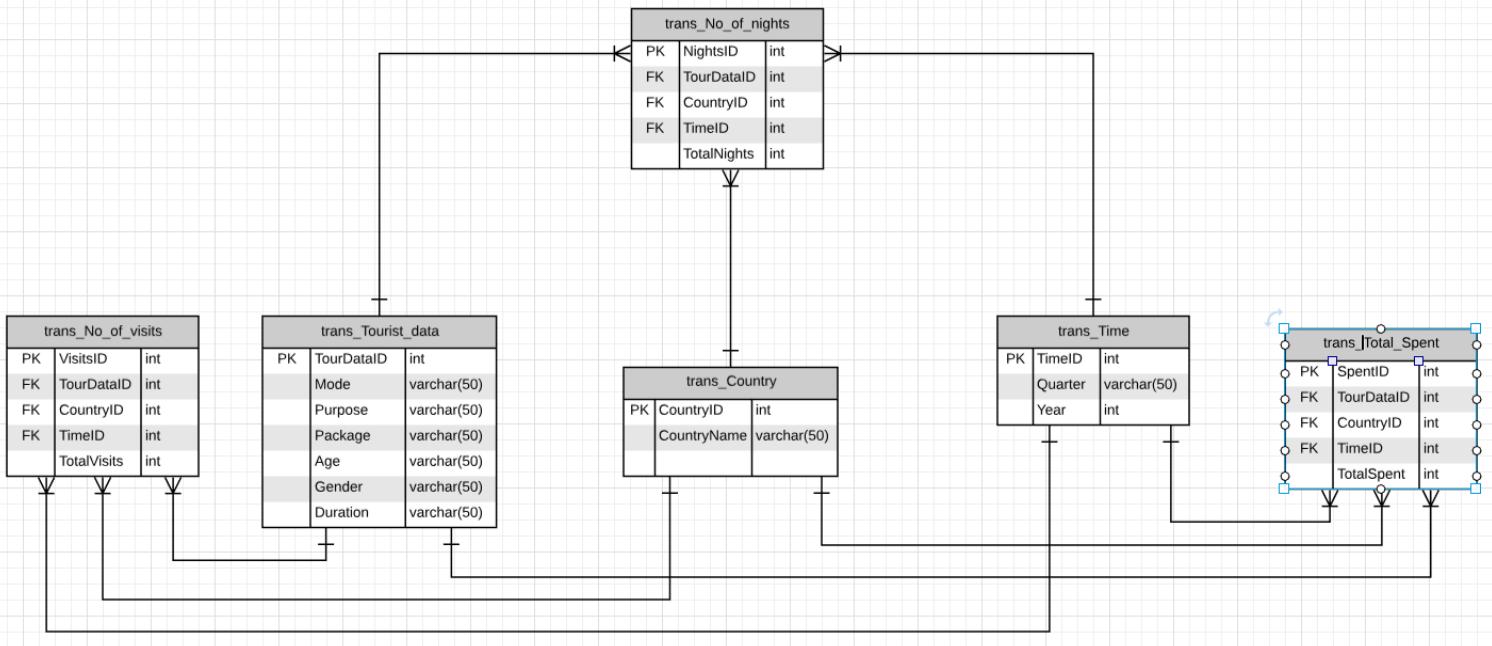
Conceptual ERD



This ERD was created using Lucidchart since it allows the user to export sql code from the diagram which can be used in the database system to create the tables without writing it themselves.

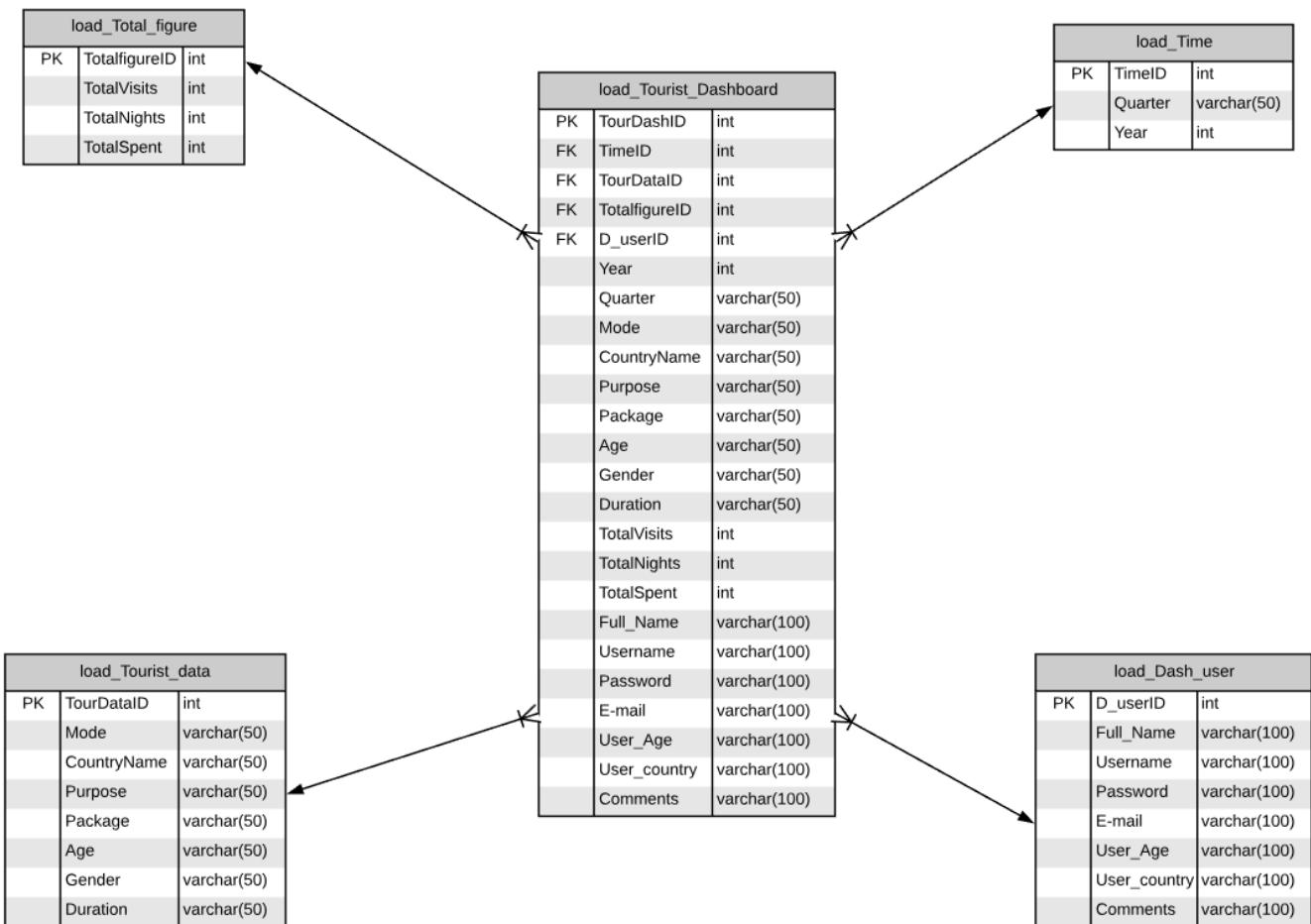
The conceptual ERD contains all the attributes and primary keys which are related to the data collected in ONS (office of national statistics). Each table will contain different attributes from the dataset, but no datatypes and foreign keys were added in the diagram because this is the early stage of ER diagram. The relationships between the tables has also been established which can be seen in the diagram above. This ERD must have more information about datatypes, values and foreign keys which can be added in the next diagram. Without this information, the database cannot be created.

Logical ERD (3NF)



After creating the conceptual ERD, the next step is to start designing the Logical ERD. The diagram was created in third normal form because the transformation phase requires the database to be in 3NF which will reduce the redundancy of the data stored in the database. The disadvantage of using 3NF is having difficulty when querying the database because the data will be complex and will be harder to write specific sql commands to get the right information. All the entities (tables) hold attributes, primary keys, foreign keys, relationships, datatypes, values and table name. 'Varchar' was used for all the attributes which do not require calculation to be done on them and 'Int' was used for all the attributes which are required to be analysed. The foreign keys have also been added to tables which uses data from other tables.

Physical ERD (2NF)



This is the physical ER diagram which is the final design needed for creating the Data warehouse. This ERD uses second normal form (2NF) and star formation was used when designing the diagram. The 2NF makes querying easier because of all the data being accessed from one table in the database. The disadvantage of using 2NF is increased redundancy which can mean that the database will use more storage and processing power than usual. All four tables in each corner will be combined to create the table called 'load_tourist_dashboard' also known as 'fact table'. All SQL queries must use the fact table to get the information they require from the database. Lucidchart can transform the diagrams into sql codes which can be used in the database to create the tables.

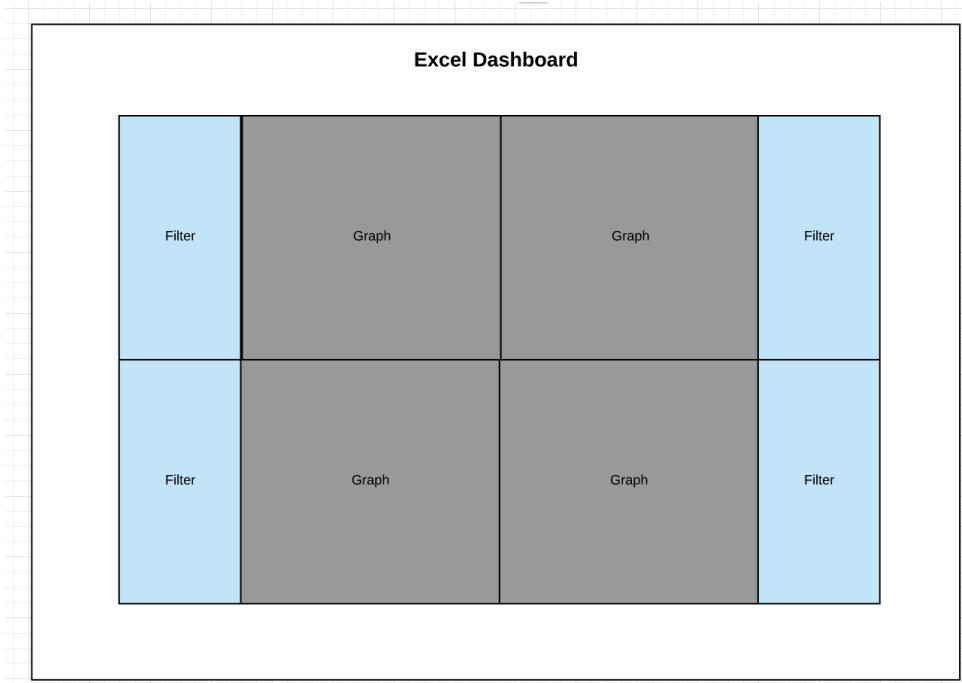
Assumptions

- The ERD does not contain 'nationality' attribute because all the data collected for the prototypes only contain UK resident data.
- 'Countryname' attribute was added because all the data collected only relate to UK resident's travel pattern

10.3 Low-fidelity wireframes for prototypes

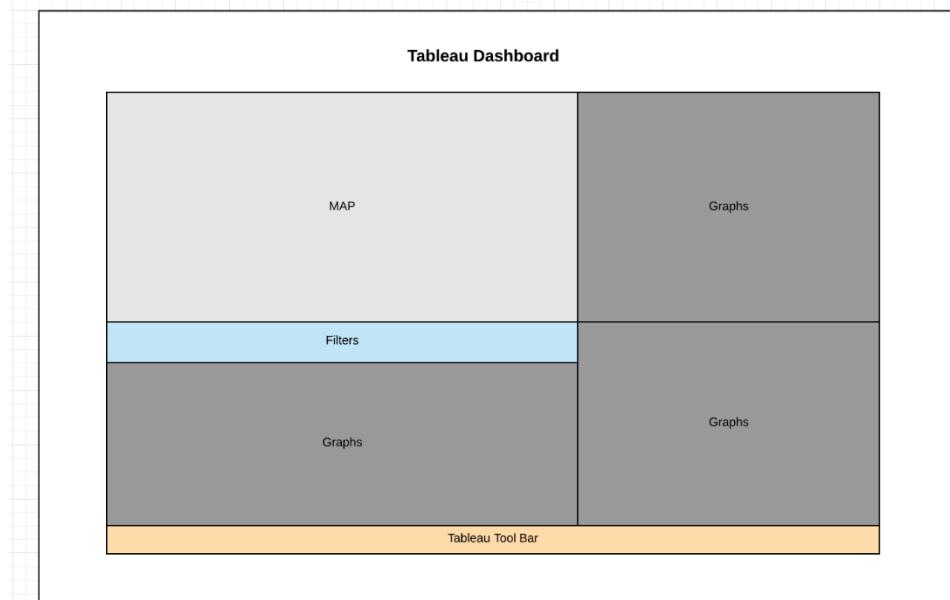
Before proceeding into the implementation stage, it is important to have an idea on how the system will look like and what features it will include. This will make it easier to create the prototypes as it will help save time when starting the prototypes and it can help reduce mistake from happening if it is fixed in the designing stage.

Low-fidelity wireframe for Microsoft Excel Dashboard



This is the wireframe design for Microsoft excel dashboard and only few graphs were added as it is only the first prototype.

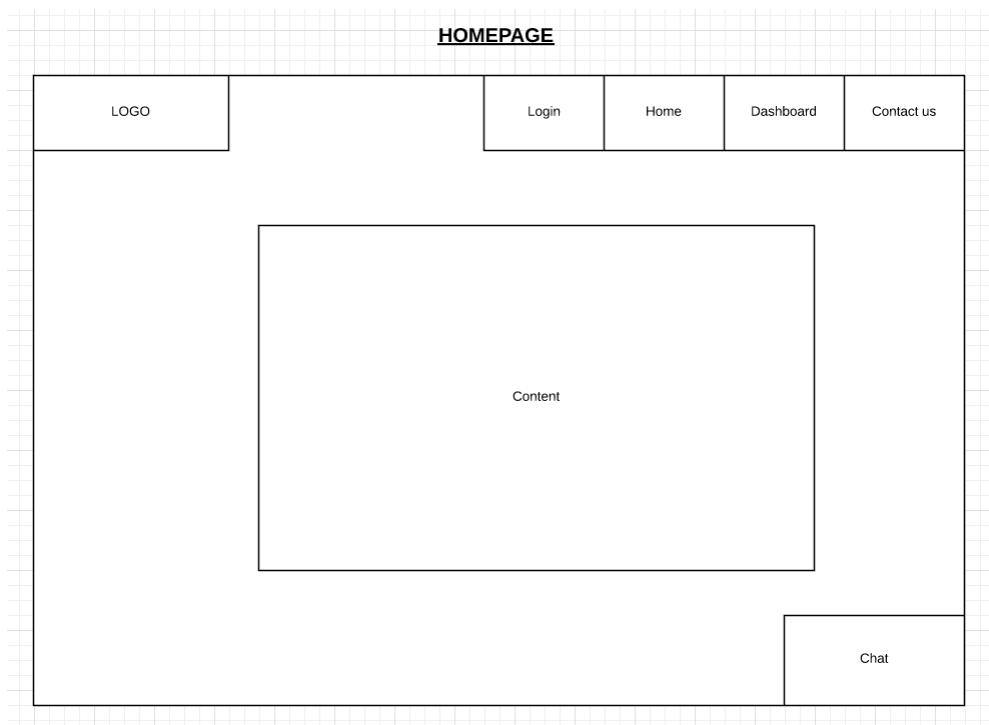
Low-fidelity wireframe for Tableau Dashboard



This wireframe contains the location of the graphs, filters, map and tableau tool bar. This wireframe will be used to create the dashboard in tableau.

Low-fidelity wireframe for website

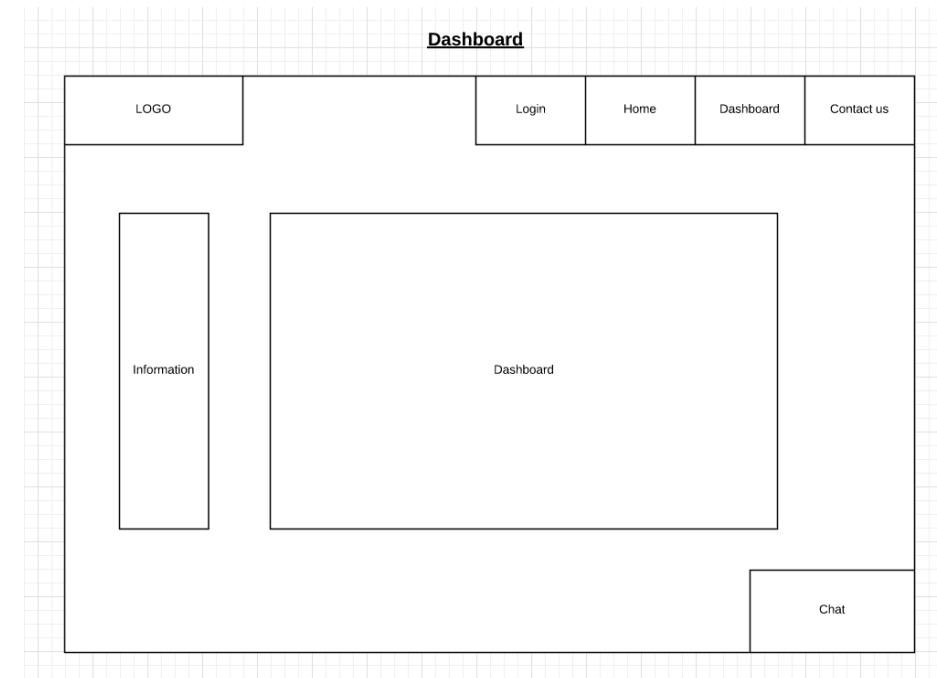
These designs will be used when creating the website for the dashboard.



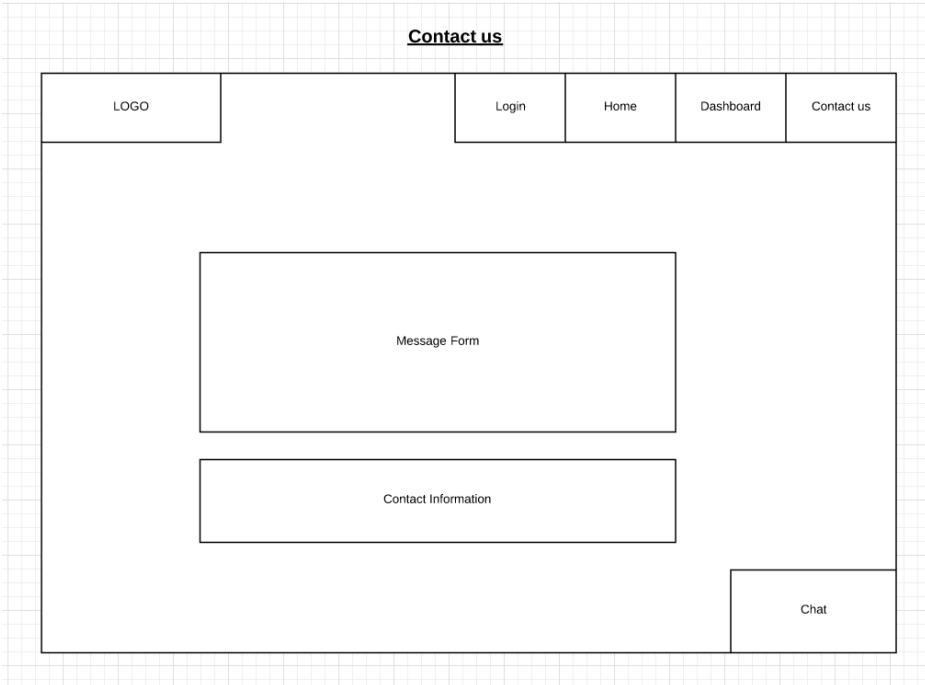
Each navigation bar will have different purpose. The login button can be used to login/logout from the website. The home button will bring user back to the homepage.

The content in the middle will contain some information about the website.

The user can also chat with the admin using the chat function below.



The dashboard will be in the centre and will be linked with the tableau dashboard. The information in the side will explain how to use the dashboard and purpose of the graph/map.



The message form can be used by the user to message any question about the dashboard to the admin. The contact information allows user to contact admin or data analyst with different methods.

All the designs above could change when producing the prototypes because of new ideas and design preference gained during the creations of the prototypes.

11 Implementation

Having done the designs for the prototype, it is time to start the implementation of the dashboard. However, the first step is to collect crucial data and get them ready for the first prototype. All the data used for the project was collected from ONS (office of national statistics). All the dataset contain data from surveying 250,000 people est*. This dataset contains travel information about the UK resident's, and it is not modified. All the datasets contain information which are related to the tourism industry and will be valuable in getting insight into the Uk tourist travel behaviour and patterns.

ETL process is used to get the dataset ready to be transferred to a Datawarehouse. There are 15 files and all of them combined contain 15 years of travel data.

[Click here to view the appendix containing all images and explanation for Datawarehouse](#)

11.1 Creation of Datawarehouse using ETL process

ETL - Extraction Phase

Table 2

Visits abroad by UK residents

	North America	Europe	of which EU	of which EU15	of which Other EU	Other Countries	Total World
Annual							
2008	4,629	54,424	50,081	45,000	5,081	9,958	69,011
2009	3,652	45,944	42,396	38,144	4,252	9,018	58,614
2010	3,653	42,565	38,925	34,784	4,140	9,344	55,562
2011	3,668	44,065	40,487	36,105	4,381	9,103	56,836
2012	3,394	44,217	40,777	36,320	4,457	8,926	56,538
2013	3,389	45,319	41,958	37,251	4,707	9,085	57,792
2014	3,677	47,025	43,834	38,784	5,051	9,380	60,082
2015	3,912	51,743	48,113	42,202	5,911	10,065	65,720
2016	4,133	56,320	52,954	45,342	7,613	10,362	70,815
2017	3,916	58,158	54,680	46,376	8,303	10,699	72,772
Quarterly							
2008	q1	1,055	10,422	9,677	8,687	990	2,886
	q2	1,224	14,894	13,808	12,473	1,335	2,430
	q3	1,351	18,656	16,888	15,115	1,773	2,460
	q4	999	10,452	9,707	8,725	982	2,181
2009	q1	732	8,218	7,645	6,822	822	2,334
	q2	959	12,785	11,831	10,687	1,144	2,217
	q3	1,146	16,292	14,864	13,411	1,453	2,339
	q4	815	8,649	8,057	7,224	833	2,128
2010	q1	639	7,174	6,599	5,884	715	2,349
	q2	939	11,450	10,530	9,413	1,117	2,222
	q3	1,130	15,827	14,262	12,759	1,503	2,589
	q4	945	8,114	7,535	6,729	806	2,184
2011	q1	665	7,388	6,829	6,072	757	2,282
	q2	1,001	12,359	11,397	10,189	1,208	2,252
	q3	1,138	15,725	14,258	12,714	1,544	2,399
	q4	865	8,593	8,002	7,129	873	2,170
2012	q1	640	7,227	6,604	5,846	759	2,393
	q2	934	12,343	11,431	10,269	1,162	2,243
	q3	1,033	15,952	14,631	12,975	1,657	2,234
	q4	787	8,695	8,110	7,230	879	2,057



	Annual	North America	Europe	EU	EU15	Other EU	Other Countries
2008		4629.25	54424.26	50080.52	44999.99	5080.53	9957.53
2009		3652.08	45944.08	42396.29	38144.31	4251.98	9018.08
2010		3653.21	42565.02	38924.68	34784.43	4140.25	9344.03
2011		3667.92	44065.11	40486.57	36105.22	4381.36	9102.87
2012		3393.74	44217.43	40776.65	36319.91	4456.75	8926.43
2013		3388.58	45319.30	41958.49	37251.44	4707.06	9084.60
2014		3676.71	47025.35	43834.18	38783.52	5050.65	9379.82
2015		3912.41	51742.69	48112.91	42202.27	5910.65	10064.81
2016		4133.14	56319.59	52954.29	45341.70	7612.58	10362.05
2017		3915.94	58157.53	54679.69	46376.50	8303.19	10698.63
Quarterly							
2008		1055.10	10422.18	9677.41	8687.17	990.24	2885.95
		1224.39	14894.04	13807.56	12473.05	1334.51	2430.14
		1350.85	18655.77	16888.27	15114.77	1773.50	2460.47
		998.92	10452.28	9707.29	8725.00	982.29	2180.97
2009		732.24	8218.48	7644.56	6822.47	822.09	2334.07
		959.14	12784.68	11831.20	10686.97	1144.23	2217.34
		1145.83	16291.57	14863.65	13410.91	1452.74	2339.05
		814.87	8649.35	8056.87	7223.95	832.92	2127.61
2010		638.94	7174.10	6598.70	5883.58	715.12	2349.38
		939.05	11449.96	10529.51	9412.88	1116.63	2222.49
		1130.23	15827.18	14261.60	12758.85	1502.75	2588.63
		944.99	8113.77	7534.87	6729.12	805.75	2183.54
2011		664.50	7387.76	6829.12	6072.45	756.67	2282.35
		1000.65	12359.42	11397.13	10189.43	1207.70	2251.76
		1137.86	15725.13	14258.40	12714.03	1544.37	2398.64
		864.90	8592.80	8001.92	7129.30	872.62	2170.11
2012		640.11	7226.54	6604.40	5845.85	758.55	2392.96
		934.13	12343.30	11430.96	10268.79	1162.17	2243.04
		1032.68	15952.33	14631.38	12974.85	1656.54	2233.83
		786.81	8695.26	8109.91	7230.43	879.48	2056.60
2013		552.69	7396.75	6776.72	5975.19	801.53	2243.81
		928.10	12638.85	11718.91	10567.15	1151.76	2360.20
		1110.88	16541.58	15350.28	13568.43	1781.85	2470.00
		796.90	8742.11	8112.58	7140.67	971.92	2010.59
2014		618.81	7649.34	7020.75	6193.86	826.89	2371.88
		941.95	13329.85	12447.95	11134.63	1313.31	2432.58

The excel file in the left is no suitable for Datawarehouse and contains some unnecessary data. After removing some of the data and cleaning up the file, it will look like the image in the right side. This is one of the examples of extraction phase of ETL. After editing and modifying the excel files, the files were saved in CSV format. This is done because CSV file can be easily imported into database. More excel files can be found in [appendix B1](#)

After all the files are saved as CSV files, it is time to create the table in MySQL. This image shows the table being created and it does not require any ERD design. This one table will contain all the data required for the project. ‘Cwe_UKtouristdata’ indicates that it is in extraction phase. ‘CREATE TABLE’ SQL command can be used but this function is easier to use to create the table.

The screenshot shows the phpMyAdmin interface for creating a new table. The table name is set to 'cwe_UKtouristdata'. The table structure consists of 14 columns:

Name	Type	Length/Values	Default	Collation	Attributes	Null	Index
UKtourdataID	INT		None				
Year	INT		None				
Quarter	VARCHAR	100	None				
Mode	VARCHAR	100	None				
Country	VARCHAR	100	None				
Purpose	VARCHAR	100	None				
Package	VARCHAR	100	None				
Age	VARCHAR	100	None				
Gender	VARCHAR	100	None				
Duration	VARCHAR	100	None				
Visits	INT		None				
Nights	INT		None				
Spent	INT		None				

After importing a few csv files, the table will look like the image below. After importing all the files into the extraction table, the total records imported was 267,243 which contains 15 years of tourism data. The extraction phase is finished, and the next phase is transformation. Importing can be found in **appendix B5**

The screenshot shows the MySQL Workbench interface. On the left, the database structure is visible with the 'information_schema' and 'mdb_kk5572a' databases selected. Under 'Tables', the 'cwe_UKtouristdata' table is highlighted. The main pane displays the table data with 17093 rows. The columns are: UKtouristdataID, Year, Quarter, Mode, CountryName, Purpose, Package, Age, Gender, Duration, Visits, Nights, and Spent. The data shows various travel details for Austria from 2004.

	UKtouristdataID	Year	Quarter	Mode	CountryName	Purpose	Package	Age	Gender	Duration	Visits	Nights	Spent
1	1	2004	Jan-Mar	Air	Austria	Holiday	Independent	0-15	Male	4-13 nights	548	3834	337412
2	2	2004	Jan-Mar	Air	Austria	Holiday	Independent	0-15	Female	4-13 nights	4858	36957	2073617
3	3	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	1-3 nights	1564	4693	242457
4	4	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	4-13 nights	5327	36686	2333298
5	5	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	14-27 nights	746	11187	37291
6	6	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Male	1-3 nights	1226	2732	278602
7	7	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Male	4-13 nights	2425	19566	1020525
8	8	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Female	1-3 nights	1948	5845	837722
9	9	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Female	4-13 nights	6108	44183	9475629
10	10	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	1-3 nights	1695	4803	435872
11	11	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	4-13 nights	4559	35147	2054804
12	12	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	14-27 nights	703	12660	351680
13	13	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Female	1-3 nights	2263	5564	370786
14	14	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Female	4-13 nights	6690	52008	4747390
15	15	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	1-3 nights	2671	6549	335918
16	16	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	4-13 nights	3956	26486	3042096
17	17	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	14-27 nights	1610	22545	1983955
18	18	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Female	1-3 nights	347	693	52937
19	19	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Female	4-13 nights	4735	30081	2547179
20	20	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Male	1-3 nights	281	561	100162
21	21	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Male	4-13 nights	1226	7740	625156
22	22	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Male	14-27 nights	1037	14523	1278041
23	23	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Female	1-3 nights	1020	2040	179524
24	24	2004	Jan-Mar	Air	Austria	Holiday	Independent	65 & over	Male	1-3 nights	737	2211	194586
25	25	2004	Jan-Mar	Air	Austria	Holiday	Independent	65 & over	Male	4-13 nights	1427	9403	869410

The screenshot shows the phpMyAdmin interface. The database 'mdb_kk5572a' is selected. The table 'cwe_UKtouristdata' is shown with its details: 267,243 rows, InnoDB type, latin1_swedish_ci collation, and 27.6 MiB size. A circled '267,243' is highlighted in the 'Rows' column.

Table	Action	Rows	Type	Collation	Size	Overhead
cwe_UKtouristdata	Browse Structure Search Insert Empty Drop	267,243	InnoDB	latin1_swedish_ci	27.6 MiB	-

End of extraction phase.

ETL – Transformation Phase

```

1 CREATE TABLE `trans_Country` (
2     `CountryID` int AUTO_INCREMENT,
3     `CountryName` varchar(50),
4     PRIMARY KEY (`CountryID`)
5 );
6
7 CREATE TABLE `trans_No_of_nights` (
8     `NightsID` int AUTO_INCREMENT,
9     `TourDataID` int,
10    `CountryID` int,
11    `TimeID` int,
12    `TotalNights` int,
13    PRIMARY KEY (`NightsID`),
14    KEY `FK` (`TourDataID`, `CountryID`, `TimeID`)
15 );
16
17 CREATE TABLE `trans_Time` (
18     `TimeID` int AUTO_INCREMENT,
19     `Quarter` varchar(50),
20     `Year` int,
21     PRIMARY KEY (`TimeID`)
22 );
23
24 CREATE TABLE `trans_Total_Spent` (
25     `SpentID` int AUTO_INCREMENT,
26     `TourdataID` int,
27     `CountryID` int,
28     `TimeID` int,
29     `Totalspent` int,
30     PRIMARY KEY (`SpentID`),
31     KEY `FK` (`TourdataID`, `CountryID`, `TimeID`)
32 );
33
34 CREATE TABLE `trans_Tourist_data` (
35     `TourdataID` int AUTO_INCREMENT,
36     `Mode` varchar(50),
37     `Purpose` varchar(50),
38     `Package` varchar(50),
39     `Age` varchar(50),
40     `Gender` varchar(50),
41     `Duration` varchar(50),
42     PRIMARY KEY (`TourdataID`)
43 );
44
45 CREATE TABLE `trans_No_of_visits` (
46     `VisitsID` int AUTO_INCREMENT,
47     `TourDataID` int,
48     `CountryID` int,
49     `TimeID` int,
50     `TotalVisits` int,
51     PRIMARY KEY (`VisitsID`),
52     KEY `FK` (`TourDataID`, `CountryID`, `TimeID`)
53 );
54
55

```

The 3NF ERD created in Lucidchart can produce sql code which can be used to create tables in phpMyAdmin. See appendix B7

Before running the sql command into phpMyAdmin, auto increment is added to each primary key in the sql code. It will automatically create the ID for each record when inserting them into the table. Running this command created all the tables for the transformation phase.

From the left-hand side of the image below, all the tables for the transformation phase have now been created. The table will now be populated with data from the extraction table using SQL codes. ‘INSERT INTO SELECT’ is being used for duplicating and transferring data from one table to another. This is being used for the ‘trans_Country’ table and the data is coming from ‘cwe_UKtouristdata’. The ‘DISTINCT’ is added to avoid repeated data from being transferred into the new table. After running the command, the table will contain all the crucial data.

The screenshot shows the MySQL Workbench interface. On the left, the database structure is displayed under the 'Tables' node for the 'mdb_kk5572a' database. The 'trans_Country' table is highlighted with a black oval. On the right, the 'Run SQL query/queries on table mdb_kk5572a.trans_Country:' window is open, containing the following SQL code:

```

1 insert into trans_Country (CountryName)
2 select DISTINCT (cwe_UKtouristdata.CountryName) from cwe_UKtouristdata

```

This code uses an 'INSERT INTO SELECT' statement to insert unique country names from the 'cwe_UKtouristdata' table into the 'trans_Country' table. Below the code, there are buttons for 'SELECT *', 'SELECT', 'INSERT', 'UPDATE', and 'DELETE'. A checkbox for 'Bind parameters' is also present.

At the bottom of the interface, a message box indicates: 'Show query box' and '211 rows inserted. Inserted row id: 211 (Query took 0.1625 seconds.)'. The SQL query shown is:

```
insert into trans_Country (CountryName) select DISTINCT (cwe_UKtouristdata.CountryName) from cwe_UKtouristdata
```

Below this, a results grid displays the data in the 'trans_Country' table. The columns are 'CountryID' and 'CountryName'. The data consists of 21 rows, each representing a country:

CountryID	CountryName
1	Austria
2	Belgium
3	Bulgaria
4	Croatia
5	Czech Republic
6	Cyprus EU
7	Denmark
8	Estonia
9	Finland
10	France
11	Germany
12	Gibraltar
13	Greece
14	Hungary
15	Iceland
16	Irish Republic
17	Italy

To the right of the table, a descriptive text block states:

This is one of the transformation tables which has been populated using the ‘INSERT INTO SELECT’ sql command. The same process was being used to populate all the other tables and can be found in [appendix B10](#). The SQL code used for all tables can be found in [appendix B11](#).

End of transformation phase.

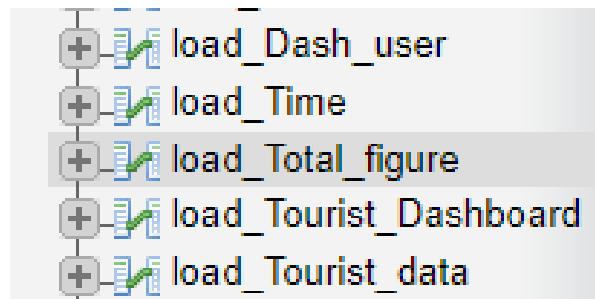
ETL - Loading Phase

This is the final step of the ETL process and 2NF ERD diagram is required to create the tables for the loading phase. 2NF creates more redundancy in the Datawarehouse, but it can help do the querying faster which makes it simpler when used with other applications. The diagram was created using Lucidchart and the sql code created using the diagram can be exported to phpMyAdmin to create the tables. Exporting can be found in [appendix B12](#)

```
Run SQL query/queries on database mdb_kk5572a

1 CREATE TABLE `load_Dash_user` (
2   `D(userID` int AUTO_INCREMENT,
3   `Full_Name` varchar(100),
4   `Username` varchar(100),
5   `Password` varchar(100),
6   `E-mail` varchar(100),
7   `Age` varchar(100),
8   `User_country` varchar(100),
9   `Comments` varchar(100),
10  PRIMARY KEY (`D(userID`)
11 );
12
13 CREATE TABLE `load_Total_figure` (
14   `TotalfigureID` int AUTO_INCREMENT,
15   `TotalVisits` int,
16   `TotalNights` int,
17   `TotalSpent` int,
18   PRIMARY KEY (`TotalfigureID`)
19 );
20
21 CREATE TABLE `load_Time` (
22   `TimeID` int AUTO_INCREMENT,
23   `Quarter` varchar(50),
24   `Year` int,
25   PRIMARY KEY (`TimeID`)
26 );
27
28 CREATE TABLE `load_Tourist_data` (
29   `TourDataID` int AUTO_INCREMENT,
30   `Mode` varchar(50),
31   `CountryName` varchar(50),
32   `Purpose` varchar(50),
33   `Package` varchar(50),
34   `Age` varchar(50),
35   `Gender` varchar(50),
36   `Duration` varchar(50),
37   PRIMARY KEY (`TourDataID`)
38 );
39
```

After pasting the sql code, auto increment is added to all primary keys to allow the table to automatically create new ID's for each new record inserted. The image below shows all the tables which are created in phpMyAdmin



The same process used for the transformation phase will be used for the loading phase. 'INSERT INTO SELECT' command is used for duplicating and transferring data from the transformation table to the new loading tables. Here you can see the sql command for 'load_time' and running the code will populate the 'time' table

```
Run SQL query/queries on table mdb_kk5572a.load_Time: ⚙

1 insert into load_Time (load_Time.Quarter, load_Time.Year)
2 select trans_Time.Quarter,trans_Time.Year from trans_Time
```

	TimeID	Quarter	Year
	1	Jan-Mar	2004
	2	Jan-Mar	2004
	3	Jan-Mar	2004
	4	Jan-Mar	2004
	5	Jan-Mar	2004
	6	Jan-Mar	2004
	7	Jan-Mar	2004
	8	Jan-Mar	2004
	9	Jan-Mar	2004
	10	Jan-Mar	2004
	11	Jan-Mar	2004
	12	Jan-Mar	2004
	13	Jan-Mar	2004
	14	Jan-Mar	2004
	15	Jan-Mar	2004
	16	Jan-Mar	2004
	17	Jan-Mar	2004
	18	Jan-Mar	2004
	19	Jan-Mar	2004

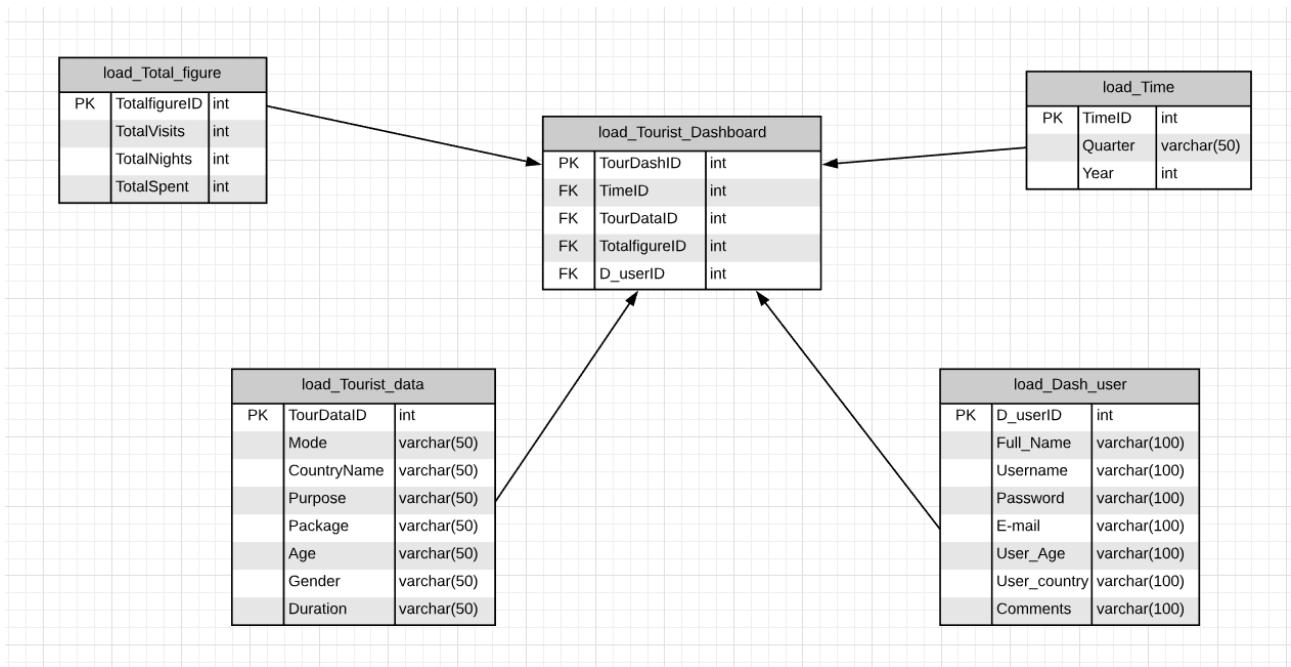
This table is called load_time and contains all the time records for the data in the data warehouse. The same process was used to populate the other loading tables and can be found in [appendix B16](#). All the SQL codes can be found in [appendix B17](#)

This is the final loading table which contains all the data required to start the prototype. This is called the fact table and all sql queries should use this table to get the information they require. This Datawarehouse is now ready to be used for the prototypes. The ETL process is now finished and the first prototype can be created.

	TourDashID	TimeID	TourDataID	TotalfigureID	D(userID)	Year	Quarter	Mode	CountryName	Purpose	Package	Age	Gender	Duration	TotalVisits	TotalNights	TotalSpent
	1	1	1	1	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	0-15	Male	4-13 nights	548	3834	337412
	2	2	2	2	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	0-15	Female	4-13 nights	4858	36957	2073617
	3	3	3	3	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	1-3 nights	1564	4693	242457
	4	4	4	4	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	4-13 nights	5327	36686	2333298
	5	5	5	5	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	14-27 nights	746	11187	37291
	6	6	6	6	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Male	1-3 nights	1226	2732	278602
	7	7	7	7	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Male	4-13 nights	2425	19566	1020525
	8	8	8	8	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Female	1-3 nights	1948	5845	837722
	9	9	9	9	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Female	4-13 nights	6108	44183	9475629
	10	10	10	10	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	1-3 nights	1695	4003	435872
	11	11	11	11	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	4-13 nights	4559	35147	2054804
	12	12	12	12	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	14-27 nights	703	12660	351680
	13	13	13	13	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Female	1-3 nights	2263	5564	370786
	14	14	14	14	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Female	4-13 nights	6690	52008	4747390
	15	15	15	15	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	1-3 nights	2671	6549	335918
	16	16	16	16	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	4-13 nights	3956	26486	3042096
	17	17	17	17	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	14-27 nights	1610	22545	1983955
	18	18	18	18	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Female	1-3 nights	347	693	52937
	19	19	19	19	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Female	4-13 nights	4735	30081	2547179
	20	20	20	20	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Male	1-3 nights	281	561	100162
	21	21	21	21	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Male	4-13 nights	1226	7740	625156
	22	22	22	22	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Male	14-27 nights	1037	14523	1278041
	23	23	23	23	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Female	1-3 nights	1020	2040	179524
	24	24	24	24	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	55 & more	Male	1-3 nights	737	2211	194586

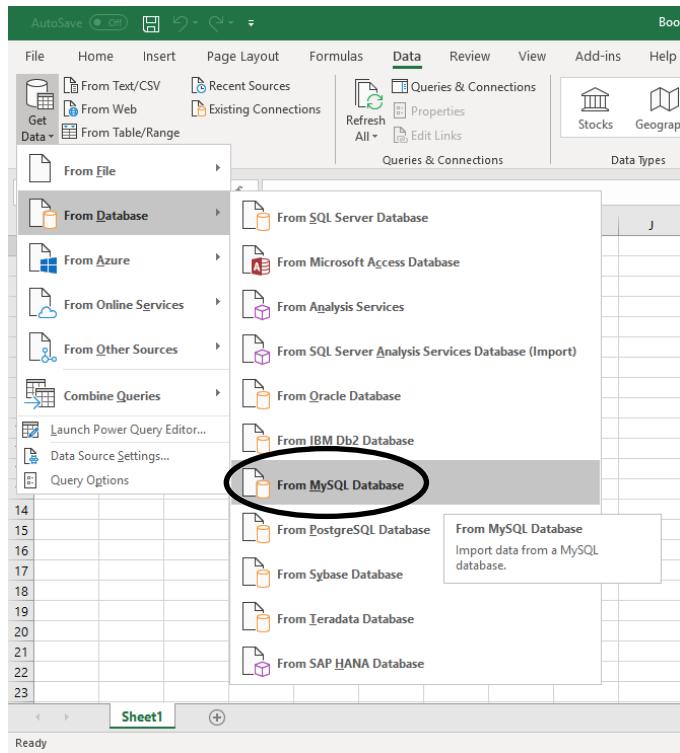
End of Loading phase

The ERD used for the Datawarehouse above is different from the traditional ERD. Normally it should not contain all the data in one table. However, with limited time and getting errors from querying using the traditional ERD, the traditional design was not used. This is original ERD for the loading phase which currently does not work. The original ERD was modified to create a working loading tables and the modified ERD can be found in the [design section](#) of this report.



11.2 1st Prototype – Microsoft Excel Dashboard

[Click here to view the appendix for Excel Dashboard](#)



Data warehouse is ready to be used and the next step is creating the first prototype in excel. First, the data from phpMyAdmin needs to be exported into excel. The appropriate database system must be chosen. Required security information must be entered to gain access to the database system and can be found in **appendix C2 and C3**.

Select the fact table from phpMyAdmin which contains all the crucial data. The table name is 'load_Tourist_Dashboard' and clicking the 'load' button will automatically import the data from the data warehouse into the excel spreadsheet.

TourDashID	TimeID	TourDataID	TotalfigureID	D(userID)	Year
1	1	1	1	null	201
2	2	2	2	null	201
3	3	3	3	null	201
4	4	4	4	null	201
5	5	5	5	null	201
6	6	6	6	null	201
7	7	7	7	null	201
8	8	8	8	null	201
9	9	9	9	null	201
10	10	10	10	null	201
11	11	11	11	null	201
12	12	12	12	null	201
13	13	13	13	null	201
14	14	14	14	null	201
15	15	15	15	null	201
16	16	16	16	null	201
17	17	17	17	null	201
18	18	18	18	null	201
19	19	19	19	null	201
20	20	20	20	null	201

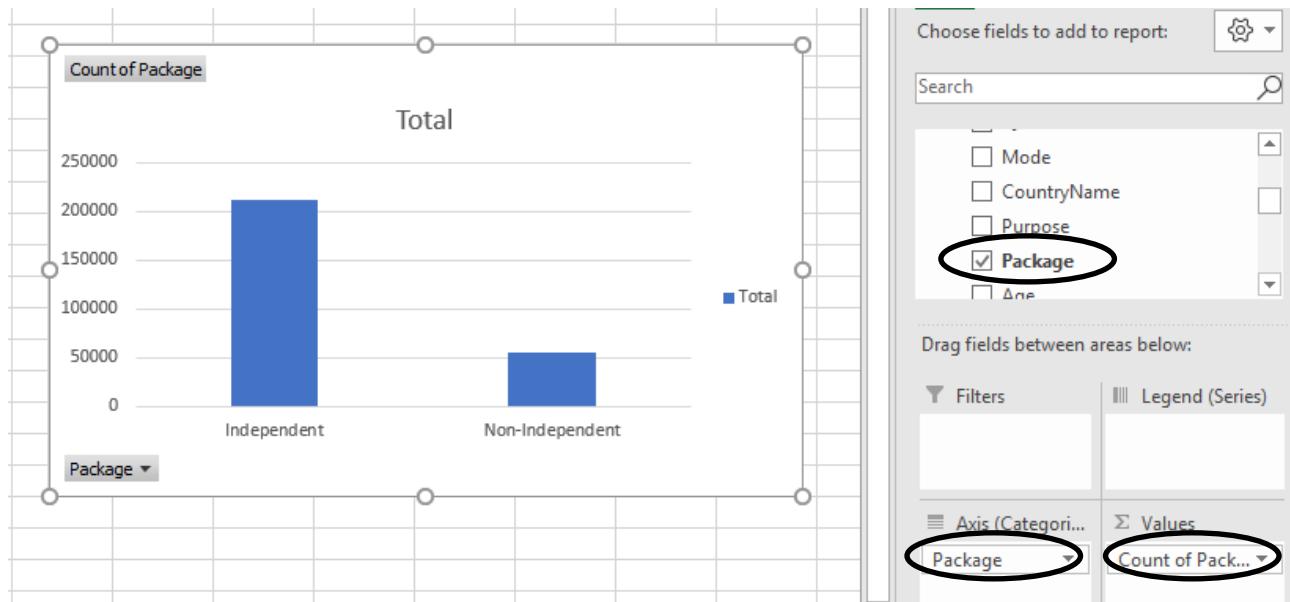
The data in the preview has been truncated due to size limits.

Load Transform Data Cancel

In this image, you can see all the data from the database system which have been imported into excel spreadsheet. Now, the pivot chart can be used to create the graphs and maps. Setting up the pivot chart can be found in **appendix C6**.

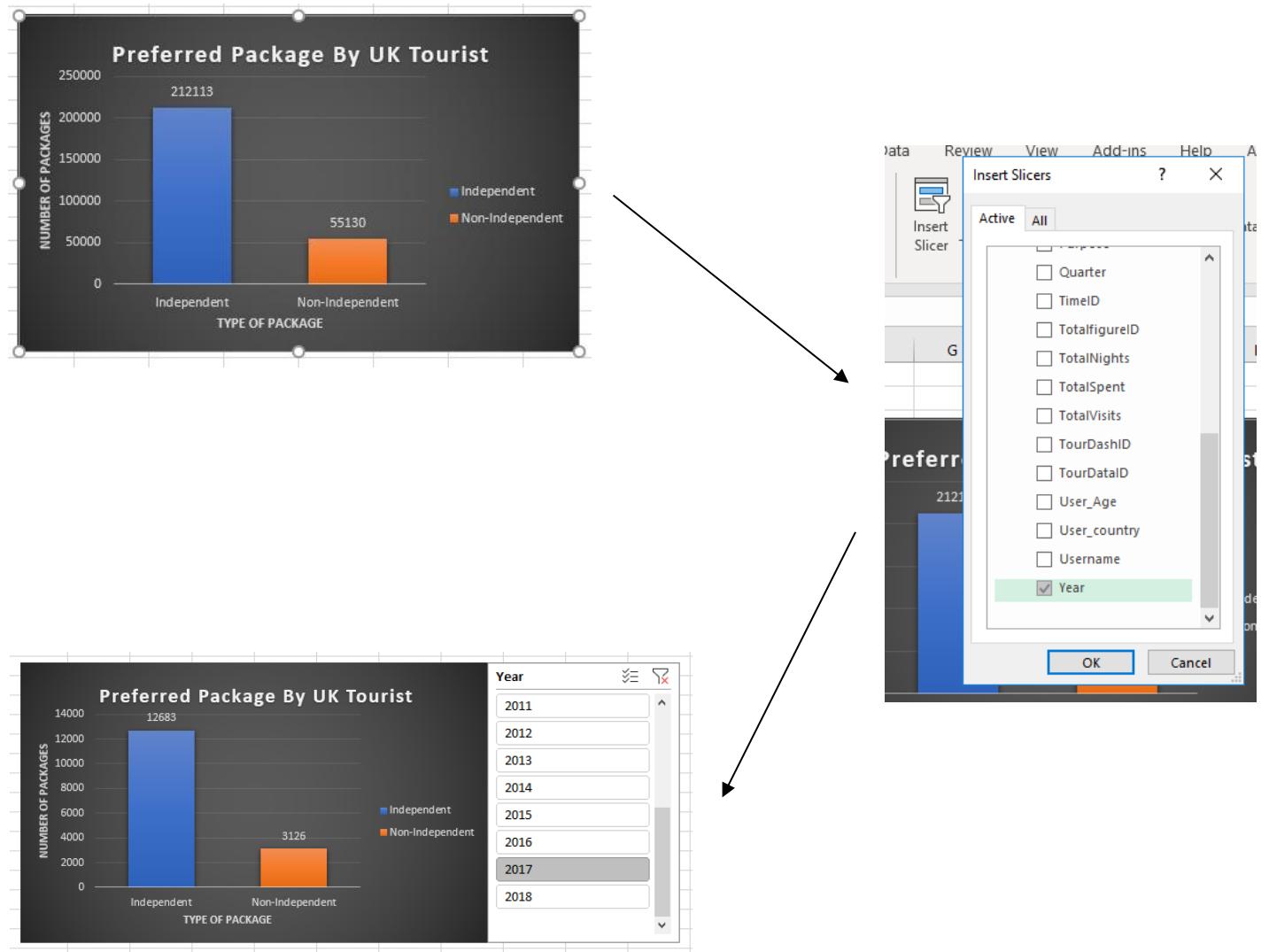
A	B	C	D	E	F	G	H	I	Purpos	
1	TourDashID	TimeID	TourDataID	TotalfigureID	D_userid	Year	Quarter	Mode	CountryName	
2	1	1	1	1		2004	Jan-Mar	Air	Austria	Holidai
3	2	2	2	2		2004	Jan-Mar	Air	Austria	Holidai
4	3	3	3	3		2004	Jan-Mar	Air	Austria	Holidai
5	4	4	4	4		2004	Jan-Mar	Air	Austria	Holidai
6	5	5	5	5		2004	Jan-Mar	Air	Austria	Holidai
7	6	6	6	6		2004	Jan-Mar	Air	Austria	Holidai
8	7	7	7	7		2004	Jan-Mar	Air	Austria	Holidai
9	8	8	8	8		2004	Jan-Mar	Air	Austria	Holidai
10	9	9	9	9		2004	Jan-Mar	Air	Austria	Holidai
11	10	10	10	10		2004	Jan-Mar	Air	Austria	Holidai
12	11	11	11	11		2004	Jan-Mar	Air	Austria	Holidai
13	12	12	12	12		2004	Jan-Mar	Air	Austria	Holidai
14	13	13	13	13		2004	Jan-Mar	Air	Austria	Holidai
15	14	14	14	14		2004	Jan-Mar	Air	Austria	Holidai
16	15	15	15	15		2004	Jan-Mar	Air	Austria	Holidai
17	16	16	16	16		2004	Jan-Mar	Air	Austria	Holidai
18	17	17	17	17		2004	Jan-Mar	Air	Austria	Holidai
19	18	18	18	18		2004	Jan-Mar	Air	Austria	Holidai
20	19	19	19	19		2004	Jan-Mar	Air	Austria	Holidai
21	20	20	20	20		2004	Jan-Mar	Air	Austria	Holidai
22	21	21	21	21		2004	Jan-Mar	Air	Austria	Holidai
23	22	22	22	22		2004	Jan-Mar	Air	Austria	Holidai

After the dataset has been selected, it is time to select the necessary field to create the graphs. Depending on the type of data chosen, different graphs can be produced, so it is important to make sure the correct data fields are selected when trying to create a valuable graph or it could produce an ineffective graph with no insight.

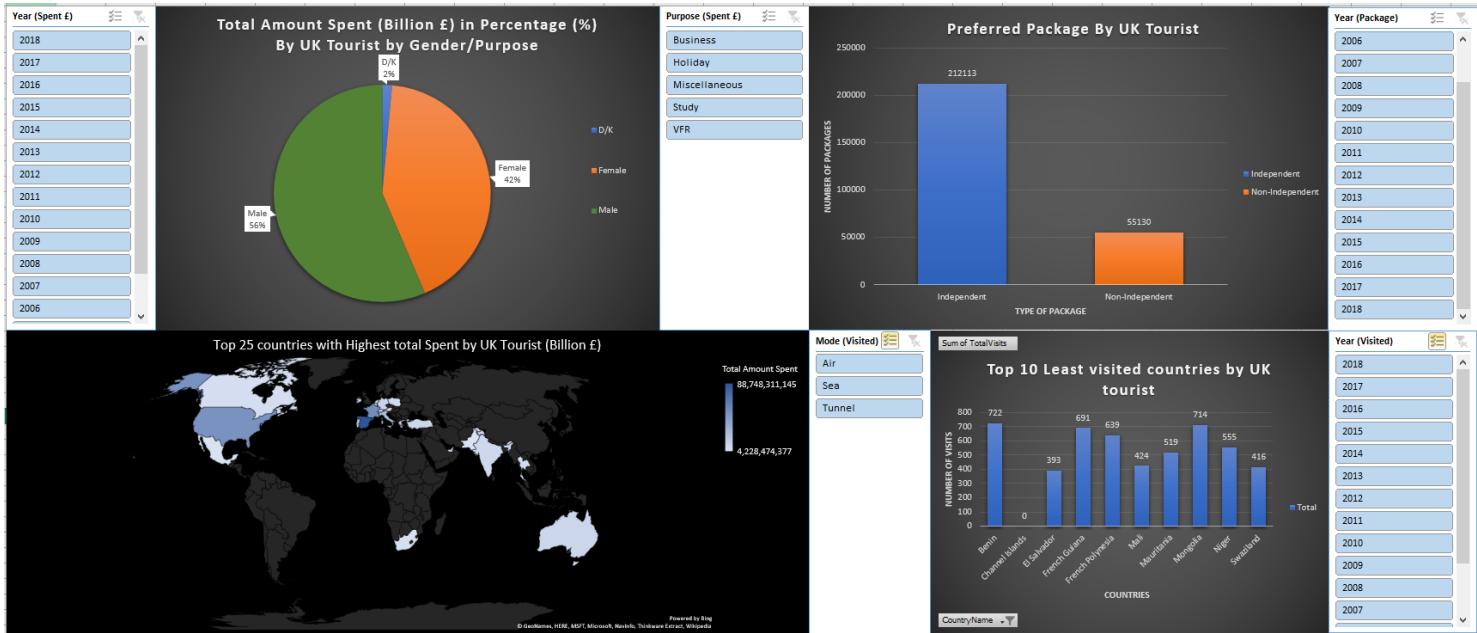


The same process will be used to create all the other graphs and charts for this prototype. Other graphs, charts and maps can be found in **appendix C9**.

Flexibility is crucial when it comes to data which are used in graphs as users may want to try different data to get different insights from the graphs. That is why 'slicer' is an important feature for the graphs in excel. It can change a static graph into a dynamic graph and can improve its effectiveness. This graph below has a 'year' slicer which lets a user to get information for each year or they can compare two or more different years. This graph contains information about the preferred package by UK tourists and it can aid in understanding if they like planning their own holiday or prefer using a travel agency. The last graph shows the preferred package by UK tourists for 2017.



After all the graphs, maps and charts are created in each spreadsheet, time has come to create the dashboard which will be the first prototype. All the pivot charts and slicers are combined in a separate spreadsheet. This is the first prototype which is a dashboard created in excel spreadsheet. This dashboard interface allows the users to view all the graphs and map in one place and can get useful insight on UK tourist.

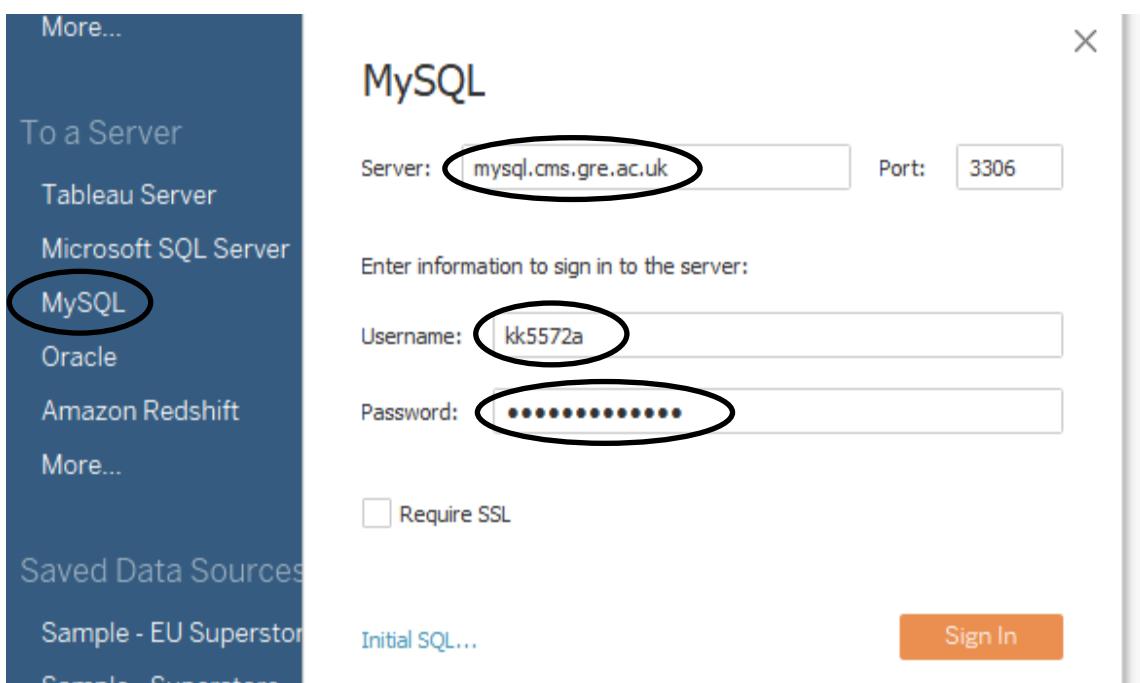


There are many features and graphs which can be added, but with limited time frame, this cannot be done for now. This prototype Excel dashboard was also proof of concept and it has proven that the graphs and dashboard can be created using the data from the database system. One of the issues encountered was related to the map and it was not flexible with what it can display with the data in the data warehouse. Any other issues and ideas found in this dashboard, will be used in the next prototype.

11.3 2nd Prototype – Tableau Dashboard

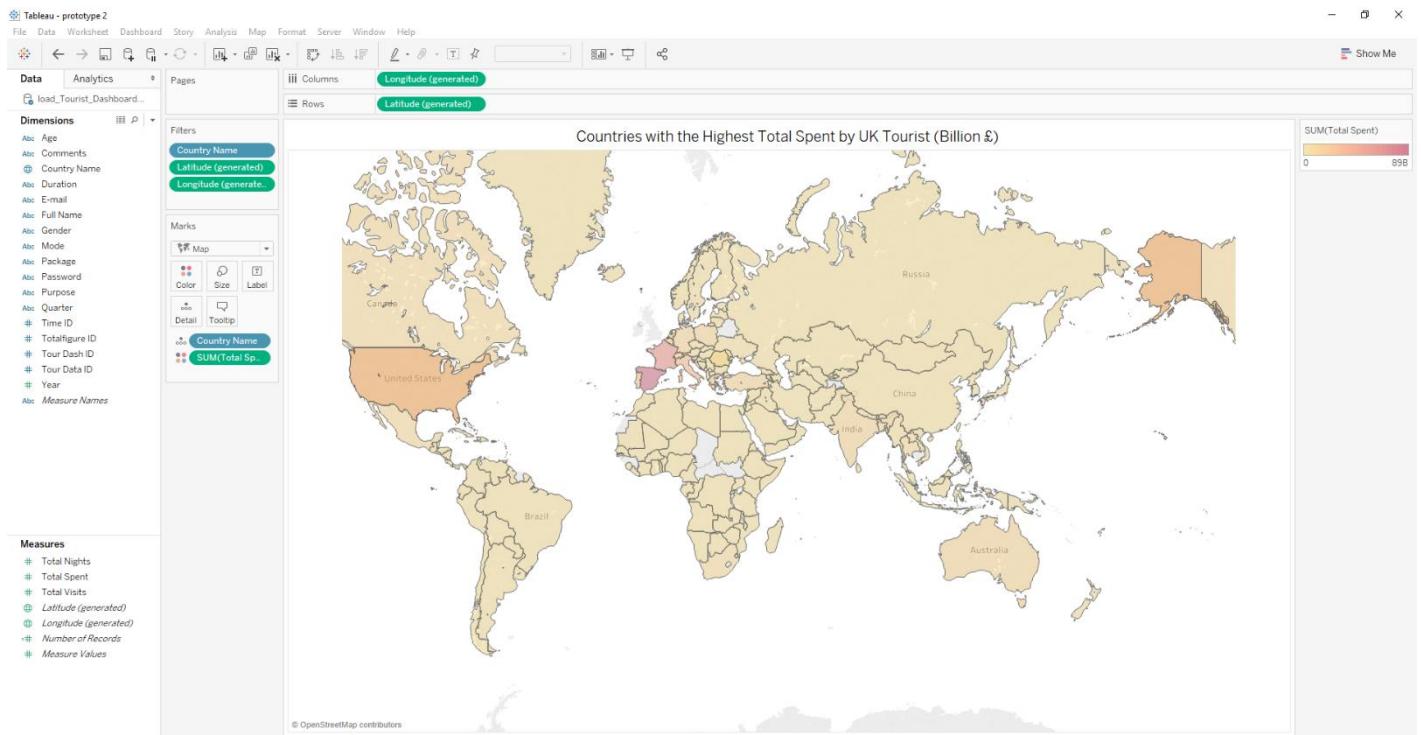
[Click here to view the appendix for Tableau Dashboard](#)

After creating the first prototype and analysing the excel dashboard, the next prototype is ready to be created which will be an improved version of the excel dashboard. Tableau software was used to create the second prototype and it contains similar functions and features from Microsoft excel. The first step is to connect the database system (phpMyAdmin MySQL) with tableau, so that the data can be accessed and used for the creation of graphs, charts and maps. Necessary server and user information must be entered in the required field to connect with the database.



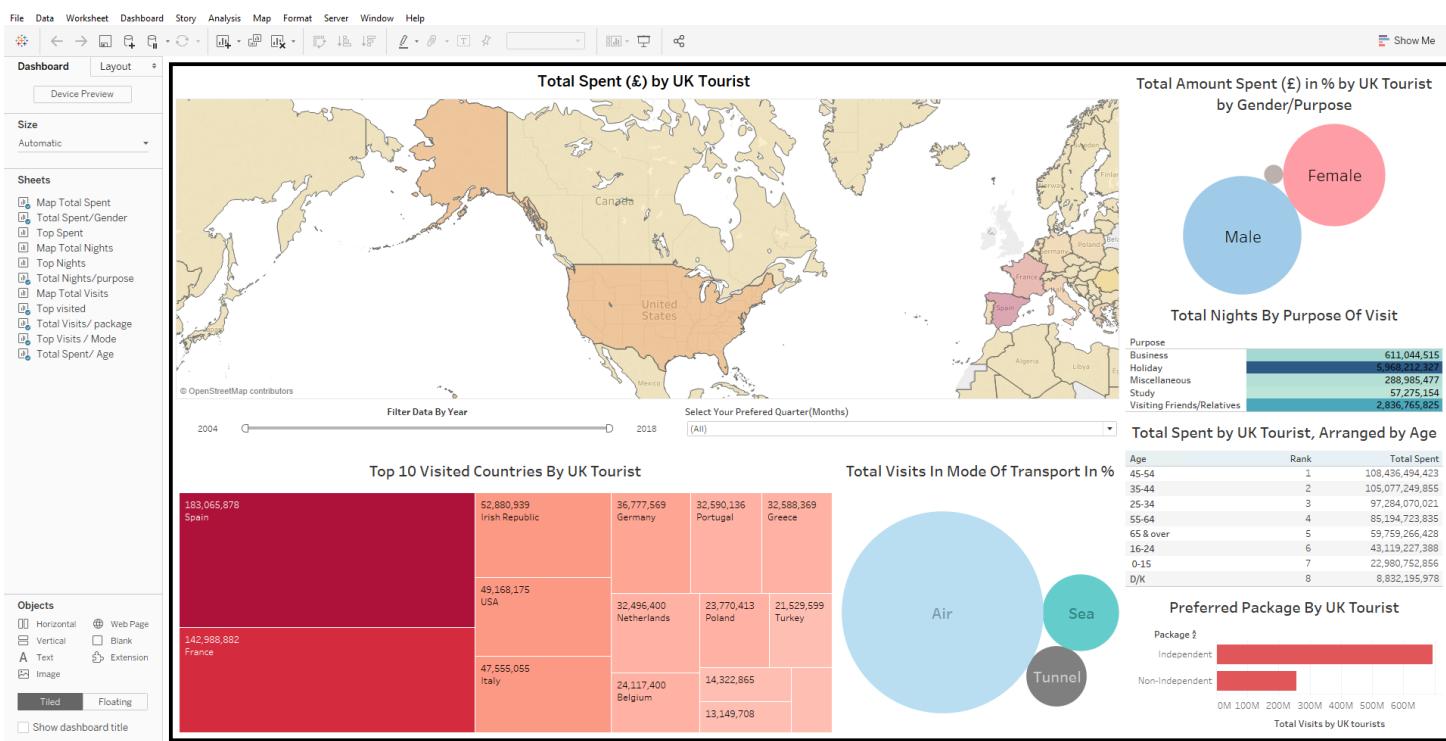
After connecting with the database, appropriate tables must be selected. Only one table from the database system is required for creating the dashboard. ‘load_Tourist_Dashboard’ is selected and ready to be used in tableau. The print screen can be found in **appendix D2**.

Combining different attributes can produce different graphs and chart. Here you can see that combining the ‘country name’ with the ‘total amount spent’ can create a map which illustrates which country has most amount spent by UK tourists. Darker colour indicates that the country has many UK tourists who spend their money. The same process will be followed to create all the graphs, charts and maps in tableau. Explanation for the other graphs and charts can be found in **appendix D4**.



After producing each individual graphs and maps, the dashboard can be created. Specific sheets (graphs) in tableau will be chosen and combined to create a dashboard. This is how it will look like after adding all the sheets together. Each individual sheet is interactive, and they can be clicked to get further information which is tailored to the option chosen by the user.

For example: Clicking 'Male' will transform all the graphs and maps in the dashboard to male and will only show information which are related to 'male'. This can be done to any other values in the dashboard and will tailor the graphs automatically.

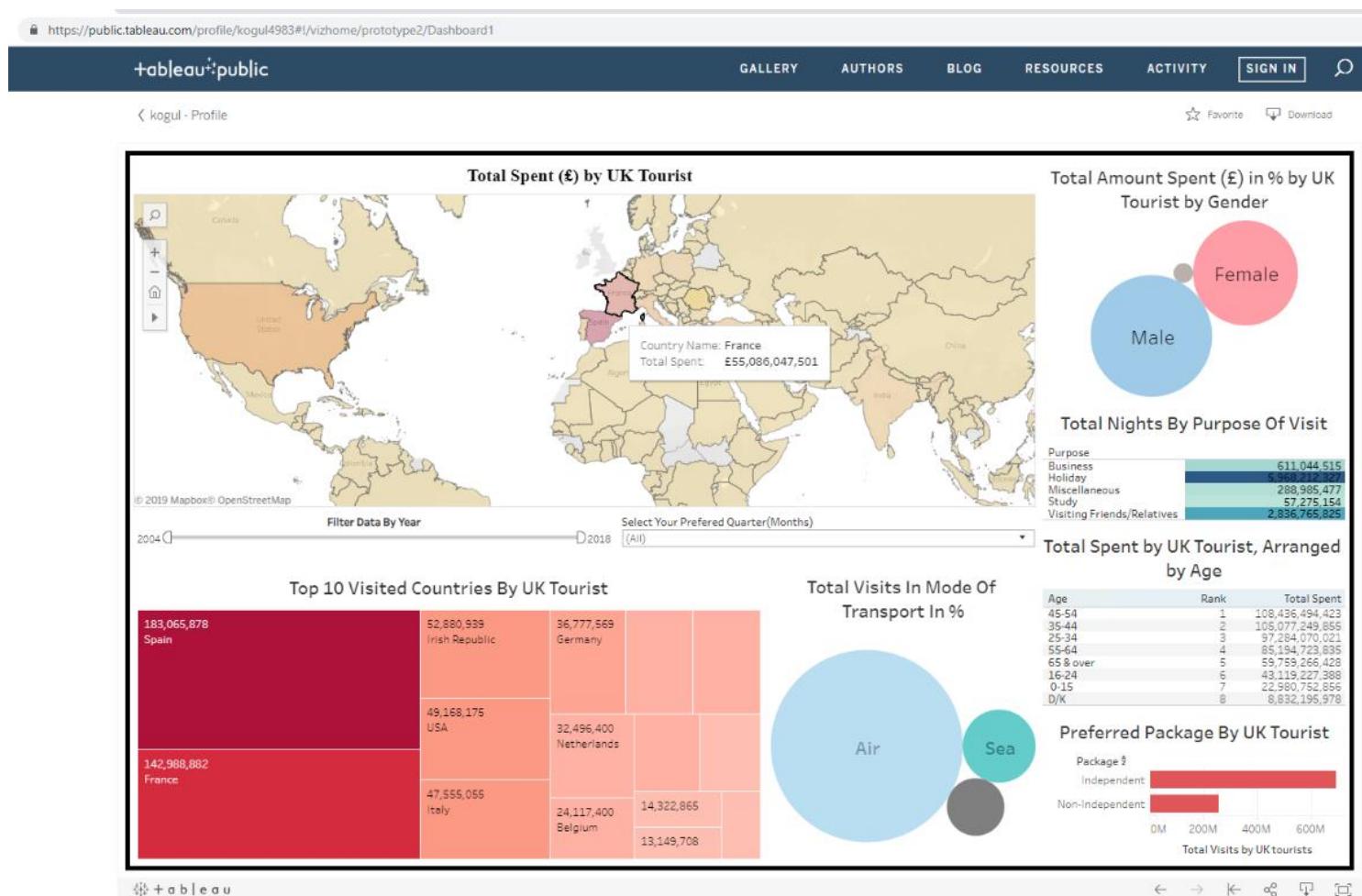


This Tableau dashboard prototype can be the finish prototype. However, it can be harder to get users to view this dashboard if it does not appear online. Tableau permits users to publish tableau dashboard if they have an account in tableau public. One must create an account before trying to publish the dashboard online. Publishing the tableau dashboard online can be found in **appendix D6 and D7**.

This is the 2nd prototype for UK tourist Dashboard. You can see here that it is published online and anyone with this link can access the dashboard.

<https://public.tableau.com/profile/kogul4983#!/vizhome/prototype2/Dashboard1>

It can also be accessed by going to the ‘Tableau public’ website and typing ‘kogul’ in the search bar to view this dashboard. All the graphs, charts and maps are interactive, and the dashboard can also be shared with anyone know who is interested in tourism industry. There are many features for using tableau public (such as, share dashboard, download dashboard, add to favourite, full screen, etc.). To get more users to view this dashboard, the website can be advertised to get more traffic to this dashboard in tableau public.



This prototype has enough features to be the final prototype. However, the tableau dashboard can be taken further by allowing the dashboard to be published online in a standalone website which can be advertised in google and other search engine easily and can also be used to earn revenue using google AdSense if needed. Currently, only the users who know about tableau public can access this dashboard which is not very effective if the purpose is to attract more users to this dashboard.

One of the issues is that the dashboard does not contain enough information for the users who are using the dashboard. Any other improvements to the dashboard will be added to the next prototype.

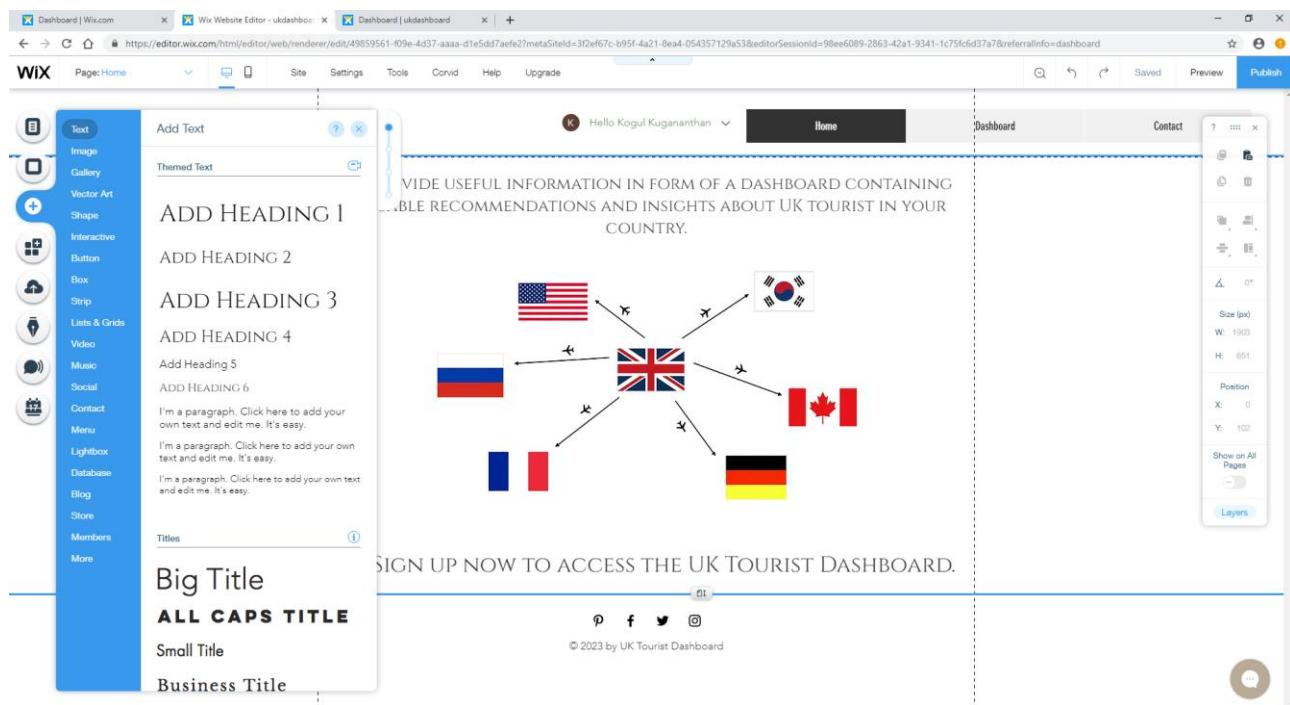
11.4 Final product – WIX Website Containing Dashboard

[Click here to view the appendix for Wix Website](#)

The current prototype is a tableau dashboard and now the 3rd prototype will incorporate the dashboard into a website using a third-party online website builder (WIX). One must create an account to use the functions and features in Wix. After creating the account, the website can be created using the template provided. Here you can see the website being created using the tools provided in Wix. The image in the middle was created to represent the UK Tourist Dashboard. The website builder also helps with hosting the website to the internet which means that anyone can access this website if they have an internet connection and a browser.

This link can be used to view the website which contains the dashboard:

<https://kk5572a.wixsite.com/ukdashboard>



This is the homepage and the first place the user will see when they enter the website. The user can login, log out and view account using the function in the top. The website has few navigation tabs and the dashboard tab contains the tableau dashboard. The user can also chat with the admin of the website if they have any question which they would like to ask about the dashboard. The user can also share this website using the social media logo below. Currently, for the prototype purpose, anyone can use the dashboard because it will be easier for the people to test the website without using an account. However, in future development the feature can be added.

UK TOURIST DASHBOARD

Hello Kogul Kugananthan

Home

Dashboard

Contact

My Account

Log Out

WE PROVIDE USEFUL INFORMATION IN VALUABLE RECOMMENDATIONS AND DASHBOARD CONTAINING ABOUT UK TOURIST IN YOUR COUNTRY.

DASHBOARD CONTAINING ABOUT UK TOURIST IN YOUR COUNTRY.

LET'S CHAT!

Hi there. I'm available to answer your questions.

SIGN UP NOW TO ACCESS THE UK TOURIST DASHBOARD.

P f t i

© 2023 by UK Tourist Dashboard

This is the contact us page which the user can use to contact the website if they have any question with the dashboard.

UK TOURIST DASHBOARD

Hello Kogul Kugananthan

Home

Dashboard

Contact

LET'S CHAT!

Hi there. I'm available to answer your questions.

GET IN TOUCH

If you have any questions about the dashboard or any other queries which you may have, please feel free to message us using the form below.

Tel: 123-456-7890 | Email: info@mysite.com

Name

Email

Subject

Message

Submit

Here you can see the dashboard in the website and each section is been explained in few words. This can help the user understand the purpose of the graphs, charts and maps in the dashboard.

SHBOARD

Hello Kogul Kugananthan ▾

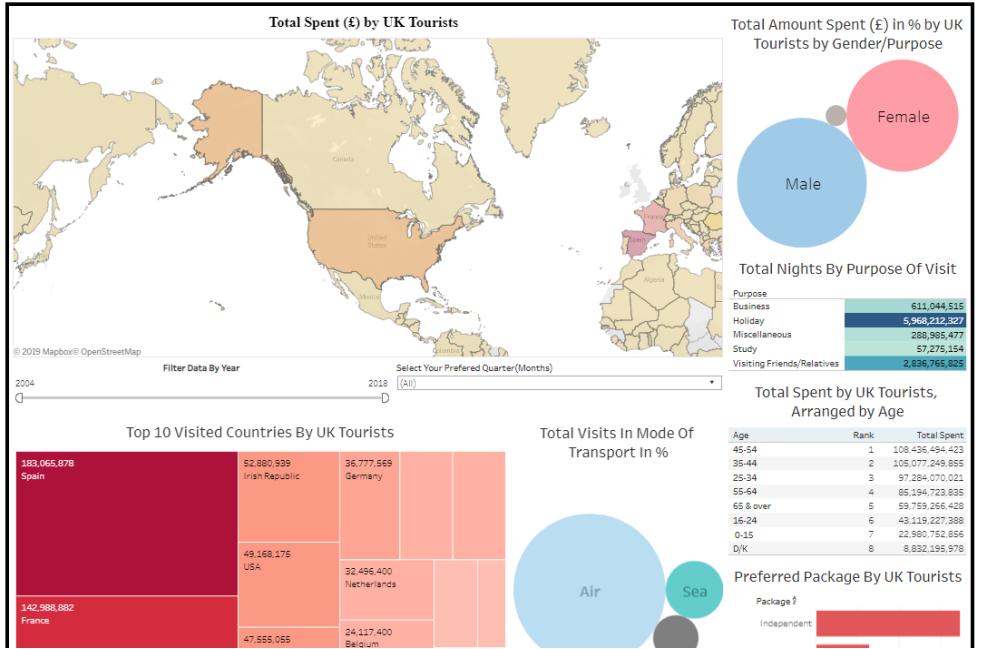
Home

Dashboard

Contact

This is a summarised version of the UK tourist dashboard and contains the most important information which you will need to gain an insight into the UK tourist who are travelling to your country.

You can interact with each of the graphs below by clicking on them and it will change all of the graphs in the dashboard to suit your selected option. All the graphs can be filtered by year and quarter.



This is the information provided to the user under the dashboard. Every quarter, this information will be updated, and it will contain key findings gathered from the UK tourist Dashboard. These findings are the most common ones which can be found by the user easily. The website also provides some recommended measures for some of the findings below. They can use these measures to help with making business decision or get a better understanding of the tourism industry in UK and worldwide.

Hello Kogul Kugananthan ▾

Home

Dashboard

Contact

KEY FINDINGS FROM DASHBOARD (FROM THE PAST 4 YEARS)

These are some of the key finding found by using the dashboard and can help with understanding the tourism industry. This section will be updated every quarter.

- UK residents favourite country to travel is Spain and all together they have spent £30.4 billion pounds in Spain over the 4-year period.
- The second and third place for the most amount spent by UK residents are France and USA.
- The least used mode of transport by UK residents around the world is the Tunnel which is equivalent to 6.77% of all mode of transport used for travelling.
- The most popular mode of transport is by Air (aeroplanes) which is 83.88% of UK residents. On the other hand, 9.35% of UK residents travel by sea (ships).
- The least visited country in the winter is Slovenia and UK residents also spent the least amount in this country which is around £13 million pounds.
- UK residents who are study abroad spend the least nights abroad and their favourite countries to study in are France, Spain, Germany and USA.
- The UK residents age group who study abroad is between 16 to 24 years old and they spent over £1.1 billion pounds over the 4 years period.
- UK residents prefer independent package over non-independent package when planning to travel abroad.
- Luxembourg is the least visited country by Female UK residents and most of them only visit the country to see their friends and family.
- Estonia is the least visited and least amount spent by Male UK residents and majority of them only visit the country to see their friends and family. They prefer to plan their own travel arrangements rather than using a travel agency.
- Slovenia has the least amount spent and least visited by business people. Only around £5.8 million pounds were spent by them over the 4-year period. Only 22% of businesswomen spend their money for business purpose over the 4-year period.
- The Top three countries visited by sea are France, Irish republic and Belgium. France and Belgium are also popular with UK residents when it comes to travelling using the tunnel.

SOME RECOMMENDED MEASURES (FROM THE PAST 4 YEARS)

These are some of the recommended measure which were produced using the key findings found using the dashboard and can help with making better business decision or have a better understanding of the tourism industry. This section will be updated every quarter.

- The popular country to travel to for UK residents is Spain and the countries has the most amount spent, most visited, and most nights spent by the UK residents. This indicates that the country gains a massive tourism from the UK and tailoring their services to accommodate the people from the UK can help increase the profit even further. Furthermore, the current



This is the end of the implementation stage and the website is ready to be used. The dashboard can be accessed by anyone from around the world and can be advertised to get more traffic to the website.

12 Testing

Black box allows one to test a software without knowing the coding of the program. This test method can be used for many parts of a program and will allow the tester to understand the software and check for any problem which may occur when using the software. Black box testing was used on all the prototypes which will help identify any problem which may occur and show all the other functions which are working properly. Refer to **Appendix G** for all the testing documentation of each prototypes.

13 Evaluation

13.1 Summarised Key findings from the project

As a result of conducting several researches in different topics, designing the prototypes and implementing them using different software. This shows that the dashboard is ready to be used for analyses. Furthermore, there are many insights which can be gathered from using different filters and graphs. Different user will want different insight into the tourism industry and will focus on different part of the dashboard only. However, there are many essential insights which can be valuable for various user groups. Some insight can only be useful for specific countries as the dashboard will tailor the information depending on the countries selected. Moreover, the total amount of data used in the dashboard equal to 15 years of tourism data from UK residents, however the only data from 2015 to 2018 will be used to find the insights and trends as it will provide latest information about the tourism industry and can be useable for the user to help with business decision. Here are a few key findings gathered from conducting this project:

- UK resident's favourite country to travel is Spain and all together they have spent £30.4 billion pounds in Spain over the 4-year period.
- The second and third place for the most amount spent by UK residents are France and USA.
- The least used mode of transport by UK resident's around the world is the Tunnel which is equivalent to 6.77% of all mode of transport used for travelling.
- The most popular mode of transport is by Air (aeroplanes) which is 83.88% of UK residents. On the other hand, 9.35% of UK residents travel by sea (ships).
- The least visited country in the winter is Slovenia and UK residents also spent the least amount in this country which is around £13 million pounds.
- UK residents who are study abroad spend the least nights abroad and their favourite countries to study in are France, Spain, Germany and USA.
- The UK resident's age group who study abroad is between 16 to 24 years old and they spent over £1.1 billion pounds over the 4 years period.
- UK residents prefer independent package over non-independent package when planning to travel abroad.
- Luxembourg is the least visited country by Female UK residents and most of them only visit the country to see their friends and family.
- Estonia is the least visited and least amount spent by Male UK residents and majority of

them only visit the country to see their friends and family. They prefer to plan their own travel arrangements rather than using a travel agency.

- Slovenia has the least amount spent and least visited by business people. Only around £5.8 million pounds were spent by them over the 4-year period. Only 22% of businesswomen spend their money for business purpose over the 4-year period.
- The Top three countries visited by sea are France, Irish republic and Belgium. France and Belgium are also popular with UK residents when it comes to travelling using the tunnel.

There are several other findings which can be found using the dashboard, however these are the most standard findings which all users can find readily while using the system. Better understanding of the tourism industry can be found when gathering the key findings using the dashboard.

13.2 Recommended measures produced using key findings

The key findings gathered using the product can be very useful for many users who are related to the tourism industry and anyone who is interested in the industry itself. Furthermore, with these key findings, one can understand the current situation within the tourism industry. However, analysing the findings can also produce measures which could be used to help with making business decision or just understanding the industry overall. The measures produced for this report are recommendation only and cannot be 100% percent accurate. Furthermore, the users/stakeholders should examine the measures to see if it can be useful for their situation in their countries. Here are some of the recommended measures produced using the key finding from the product (dashboard):

- The popular country to travel to for UK residents is Spain and the country has the most amount spent, most visited, and most nights spent by the UK residents. This indicates that the country gains a massive tourism from the UK and tailoring their services to accommodate the people from the UK can help increase the profit even further. Furthermore, the current age group which spends the least amount in Spain is 0-15 and 16-24. Developing more infrastructure to suit these age group will increase the profit and attract more tourists from the UK. The country could also invest in the education system to attract more UK residents to study in Spain because UK residents visit the least for studying.
- Over the past 4 years, 45-54 years old are the main demographic for spending the most amount when travel to other countries. 30.3 billion pounds was spent by this age group and the tourism industry around the world can tailor more tourism packages to attract more potential UK customers. Hotels and travel agency can create unique travel package or all-inclusive deals which are tailored to these age group which could lead to increased profit for the businesses.
- Currently, the preferred destination for female traveller from the UK are Spain, France and Italy. All these countries have infrastructure in place to help female travellers be safe and make their holiday more enjoyable. However, the least visited country by UK females is Luxembourg and the tourism industry in the country could focus on how to access this minority demographic. They can look at top countries which welcome female travellers from the UK and can incorporate some of the features into their tourism industry. Event

planners could create female related events and parties which could attract more females from around the world. Accommodation can also be focus on by added female only rooms to make them feel safe.

- Between October to December, the least visited country and least amount spent by UK residents is Slovenia. This could indicate that the country has not prepared enough to accommodate more tourists into their country. The language barrier could be the problem as the UK residents may feel nervous to travel to a country which does not speak their language primarily. This could be tackled by introducing more English words in the tourist parts of the city. Furthermore, the winter season could be tackled by creating more events and services which will suitable the UK travellers. The hotel industry could also create packages which combine winter sports and accommodation, and this can attract UK travellers who are interested in winter sports and want to keep it budget.

These are some of the recommended measures which was produced using the key findings. Furthermore, one can find more measures with the help of the UK tourist dashboard which could help their businesses by attracting more UK customers and increase profits by focusing on areas which are untapped or undeveloped or hidden.

13.3 Usability Evaluation - Cognitive Walkthrough

After producing the final product, it is important to evaluate using the nelson heuristics and other methods to make sure that they are followed to ensure the website and the dashboard encounters less user interface problems. There are many different types of evaluation techniques which can be used to test the User interface design (such as cognitive walkthrough, Heuristics evaluation, contextual interview, focus group, etc.) (Usability, 2019). Cognitive walkthrough shows if there is any usability issues in the product. The green colour means the user can use and understand the function, while the red colour means that the user may not know how to use the function. Below is evaluation for the website containing the dashboard and you can see that majority of the function are easier for the users to use. This confirms that the product is user-friendly and will encounter less UI problems when used by the user.

Final Product (Website with Dashboard)	Will users understand how to start the task?	Are the controls conspicuous?	Will users know the control is the correct one?	Will there be feedback to indicate the task was complete or incomplete?	Notes
Accessing dashboard					
Visit website using link	Yes	Yes	Yes	Yes	
Click login and enter details	Yes	Yes	Yes	Yes	
Click dashboard button in navigation bar	Yes	Yes	Yes	Yes	
Interact with dashboard	Yes	No	No	Yes	Dashboard can be difficult for complete beginners, so may require more information on how to use the system.
Share Dashboard					
Click on login and enter details	Yes	Yes	Yes	Yes	
Click dashboard button in navigation bar	Yes	Yes	Yes	Yes	
Click on share icon at the bottom	No	Yes	No	Yes	The share icon is a feature from tableau and cannot be modified. The user may have difficulty with using this method.
Choose appropriate sharing method	Yes	Yes	Yes	Yes	
Click submit button	Yes	Yes	Yes	Yes	
Sending a query to admin					
Click on login and enter details	Yes	Yes	Yes	Yes	
Click contact button in navigation bar	Yes	Yes	Yes	Yes	
Write message in form with contact details.	Yes	Yes	Yes	No	The user will not know if the message is sent until the click on the submit.
Click submit button to send message	Yes	Yes	Yes	Yes	
Download Dashboard					
Click on login and enter details	Yes	Yes	Yes	Yes	
Click dashboard button in navigation bar	Yes	Yes	Yes	Yes	
Click on download icon at the bottom	No	Yes	No	Yes	This icon is similar to the share button but they can hover the mouse to see if it is a download button.
Choose appropriate file format	No	Yes	Yes	Yes	The user must know what file format they want to download and understand how to use it themselves.
Click download button	Yes	Yes	Yes	Yes	

13.4 Project Evaluation

The aim of this project is to produce web-based visualisation in form of a dashboard which will provide an insight into the travel and behavioural patterns of UK residents. Recommendation will be produced for countries/continents using the visualisation tool. The project was successful as it did produce a system which was used to find insights into the tourism industry and the key findings were understandable and easier to find using the product.

When working on the project, there were many processes which performed well and some things which did work properly. The research conducted for this project was beneficial because it helped with understanding the tools and techniques which were used for the project. With the research, the project may have failed. On the other hand, the implementation could have been better because with limited time, less implementation was made, and some sections were not completed. If the time was managed properly, then more implementation could have been done and it would have improved the final product much further. To make sure the project was successful, testing was done to ensure the final products was working and usability evaluation was also done to check if there are any UI issues. The final product was improved significantly because several prototypes were created to test different sections of the product and helped in preventing new issues from rising which also reduced the time spent on fixing new issues

The project did not follow the original plan created at the start and new requirements and documentation have been added. The DSDM methodology provided the flexibility when working on the project and could not be possible if another methodology was used. With limited time, dsdm allow the project to prioritise on important section of the project which made sure that the project was finished on-time.

Overall, the project was successful and accomplished the aim which was to produce a web-based visualisation tool which anyone can use to get insights into tourism industry.

13.5 Personal Evaluation

Before starting this project, I was not sure on what topic to work on and it was difficult was me to choose as I did not know which skills I had that will be useful for the project. However, after talking with my supervisor, I was able to identify the skills which I could use for the project. I want to ensure that I incorporated many of the modules which I have learned over the 3 years of my degree. I felt confident in using the different tools and techniques which I have learnt from different courses (such as SQL, Tableau, Excel, UI design, Data warehousing, information analysis, requirement analysis, development framework and web development). Because of this, I had less difficulty with the implementation of the prototypes and the final product.

One of the issues which I have encountered is time management because after starting a new section in the project, more requirements were appearing which require more time to work on. With more time lost, less testing and documentation were made which could have be avoided if planned properly. Moreover, the remaining time were used to implement the prototypes and the final product. I also was not able to produce my own website using programming language because of the limited time I have left at the end of the project. I have learnt a lot by doing this

project and this experience will be useful in the future if I do a project.

Overall, I believe the project went well and if I had more time, I would work more on the report and implementation to improve the product even further.

13.6 Conclusion

What went well?

Majority of the sections in the report was successful and helped with finishing the project on-time. The research went well and was useful for understanding the product which was created for this project. The literature review provided the information on why this project was needed and help understand the purpose of the product (dashboard). Because of using different tools and techniques, further research was needed, and it helped solidify the knowledge which I previous had. The prototyping was successful because it helped in finding out issues which were prevented from occurring in the final product.

What did not go well?

The project would have been better if I had more time because more requirements were added at the end of the project which meant that more documentation were needed. Because of this, the remaining time left at the end of the project was dividing between documentation and implementation. This produced a product which could have been improved if more time was available. There could have been more testing done on the product to find more issues and it would have improve the usability of the product.

What have I learned from this project?

I have learnt to make planning a priority because it can have a massive impact on the project and its success rate. More time needs to be allocated to future requirements which may arise during the project. Furthermore, research is key to having a successful project because it help with understanding the purpose of the project and how to create the product with minimal error.

Future development

If I had more time to work on this project, I would add more functions and features to both the website and the dashboard. I would also create my own website using programming language (such as C# or JavaScript) because it can provide more flexibility when designing the website. I can also try adding a feature which will tailor the data depending on the country the user is from. This will make sure the data they view are useful to them because data about countries is not doing to help them to get insight into how they can improve their business in their country.

14 References

- Agaraj, X. and Murati, M. (2009). *Tourism an Important Sector of Economy Development*. [online] Ideas.repec.org. Available at: <https://ideas.repec.org/a/cbu/jrnlec/y2009v1p83-90.html> [Accessed 19 Nov. 2018].
- Alsos, G. A., Madsen, E. L. and Eide, D. (2014) Handbook of Research on Innovation in Tourism Industries. Cheltenham: Edward Elgar Publishing (Elgar Original Reference). Available at:
<http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=688133&site=ehost-live> (Accessed: 19 November 2018).
- Buhalis, D. (2000). *Strategic use of information technologies in the tourism industry*. [online] Epubs.surrey.ac.uk. Available at: <http://epubs.surrey.ac.uk/1123/1/fulltext.pdf> [Accessed 5 Jan. 2019].
- Buhalis, D. and Law, R. (2008). *Progress in information technology and tourism management: 20 years on and 10 years after the Internet*. [online] ScienceDirect. Available at: <https://www.sciencedirect.com/science/article/pii/S0261517708000162#bbib135> [Accessed 5 Jan. 2019].
- Charles, D. (2014). *Advantages & Disadvantages of Implementing Business Intelligence*. [online] Linkedin.com. Available at: <https://www.linkedin.com/pulse/advantages-disadvantages-implementing-business-charles-d-madewell> [Accessed 10 Feb. 2019].
- Cotton, J. (2014). *The importance of accuracy, integrity and real-time integration*. [online] Ww2.informationbuilders.com. Available at: <http://ww2.informationbuilders.com/blog/james-cotton/19152> [Accessed 7 Feb. 2019].
- Data.worldbank.org. (2019). *International tourism, expenditures (current US\$) / Data*. [online] Available at: <https://data.worldbank.org/indicator/ST.INT.XPND.CD> [Accessed 24 Dec. 2018].
- Dedić, N. and Stanier, C. (2016). Measuring the success of changes to existing business intelligence solutions to improve business intelligence reporting. Springer International Publishing, Vol. 268, pp. 225-236.
- E-unwto.org. (2018). *UNWTO Tourism Highlights 2018*. [online] Available at: <https://www.e-unwto.org/doi/pdf/10.18111/9789284419876> [Accessed 25 Dec. 2018].
- Few, S. (2006). *Information Dashboard Design*. [online] Thali.ch. Available at: https://www.thali.ch/files/Shop/Documents/018161_Chapter_1_Clarifying_The_Vision.pdf [Accessed 8 Jan. 2019].
- Fuchs, M., Höpken, W. and Lexhagen, M. (2014). *Big data analytics for knowledge generation in tourism destinations*. [online] Science Direct. Available at: <https://www.sciencedirect.com/science/article/pii/S2212571X14000353#bibliog0005> [Accessed 5 Jan. 2019].
- Hall, C., Timothy, D. and Duval, D. (2012). *Safety and Security in Tourism*. Hoboken: Taylor and Francis, p.2.
- Hassan, B. (2003). *Examining data accuracy and authenticity with leading digit frequency analysis*. [online] Emeraldinsight.com. Available at: <https://www.emeraldinsight.com/doi/full/10.1108/02635570310463447> [Accessed 7 Feb. 2019].
- Kaye, D. (1995). *The importance of information*. [online] Emeraldinsight.com. Available at: <https://www.emeraldinsight.com/doi/full/10.1108/EUM0000000003897> [Accessed 5 Jan. 2019].
- Klipfolio. (2019). *Ultimate Guide to Business Dashboards*. [online] Available at: <https://www.klipfolio.com/guide-to-business-dashboards> [Accessed 1 Mar. 2019].
- Kővári, I. and Zimányi, K. (2010). *SAFETY AND SECURITY IN THE AGE OF GLOBAL TOURISM*.

[online] Ageconsearch.umn.edu. Available at:
http://ageconsearch.umn.edu/bitstream/104672/2/10_Kovari_Safety_Apstract.pdf
[Accessed 5 Jan. 2019].

- Lucidchart. (2019). *UML Use Case Diagram Tutorial*. [online] Available at: https://www.lucidchart.com/pages/uml-use-case-diagram?a=1#discovery_top [Accessed 25 Feb. 2019].
- Mariani, M., Baggio, R., Fuchs, M. and Höepken, W. (2018). *Business intelligence and big data in hospitality and tourism: a systematic literature review*. [online] Emeraldinsight.com. Available at: <https://www.emeraldinsight.com/doi/full/10.1108/IJCHM-07-2017-0461> [Accessed 7 Jan. 2019].
- Nielsen, J. (1994). *10 Heuristics for User Interface Design: Article by Jakob Nielsen*. [online] Nielsen Norman Group. Available at: <https://www.nngroup.com/articles/ten-usability-heuristics/> [Accessed 21 Feb. 2019].
- oracle. (2019). *Overview of Extraction, Transformation, and Loading*. [online] Available at: https://docs.oracle.com/cd/B19306_01/server.102/b14223/ettover.htm [Accessed 15 Feb. 2019].
- Pantano, E. and Pietro, L. (2013). *From e-tourism to f-tourism: emerging issues from negative tourists' online reviews*. [online] Emeraldinsight.com. Available at: <https://www.emeraldinsight.com/doi/full/10.1108/JHTT-02-2013-0005> [Accessed 29 Dec. 2018].
- Rouse, M. (2017). *What is business intelligence (BI)?*. [online] SearchBusinessAnalytics. Available at: <https://searchbusinessanalytics.techtarget.com/definition/business-intelligence-BI> [Accessed 10 Feb. 2019].
- S, S. and Naik, M. (2017). *Data visualization using Viewshare*. [online] Emeraldinsight.com. Available at: <https://www.emeraldinsight.com/doi/full/10.1108/LHTN-02-2017-0009> [Accessed 8 Jan. 2019].
- Sas.com. (2018). *Data Visualization*. [online] Available at: https://www.sas.com/en_gb/insights/big-data/data-visualization.html#dmusers [Accessed 9 Jan. 2019].
- Sisense. (n.d.). *Digital Marketing Dashboards & KPI Templates / Sisense*. [online] Available at: <https://www.sisense.com/dashboard-examples/marketing/> [Accessed 10 Jan. 2019].
- Slutsky, D. (2014). *The Effective Use of Graphs*. [online] NCBI. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4078179/> [Accessed 9 Jan. 2019].
- Song, H. and Liu, H. (2017). *Predicting Tourist Demand Using Big Data*. [online] Springer.com. Available at: https://www.springer.com/cda/content/document/cda_downloaddocument/9783319442624-c1.pdf?SGWID=0-0-45-1592857-p180195151 [Accessed 5 Jan. 2019].
- Tasmania.tourismtracer.com. (2016). *Point of Interest*. [online] Available at: <https://tasmania.tourismtracer.com/points-of-interest> [Accessed 10 Jan. 2019].

15 Bibliography

- Heerschap, N., Ortega, S., Priem, A. and Offermans, M. (2014). *Innovation of tourism statistics through the use of new big data sources*. [online] Tsf2014prague.cz. Available at: http://tsf2014prague.cz/assets/downloads/Paper%201.2_Nicolaes%20Heerschap_NL.pdf [Accessed 5 Jan. 2019].
- Janoschek, N. (2017). *Benefits of Business Intelligence*. [online] BI Survey. Available at: <https://bi-survey.com/benefits-business-intelligence> [Accessed 10 Feb. 2019].

- LibGuides: Visualisation and Interactive Media: What is Visualisation?, (2018) *Libguides.library.curtin.edu.au*, [online] Available at: <https://libguides.library.curtin.edu.au/c.php?g=388681&p=2688784> (Accessed 25 October 2018).
- Richardson, M. (2018). *7 Reasons Business Intelligence Is Vital To Business Success*. [online] Maximizer. Available at: <https://www.maximizer.com/blog/7-reasons-why-business-intelligence-is-vital-to-business-success/> [Accessed 7 Jan. 2019].
- Softwareadvice. (2019). *Top Business Intelligence Tools - 2019 Reviews & Pricing*. [online] Available at: https://www.softwareadvice.com/bi/p/all/?price_range=1#buyers-guide [Accessed 11 Feb. 2019].
- Usability. (2019). *Usability Evaluation Methods* | *Usability.gov*. [online] Available at: <https://www.usability.gov/how-to-and-tools/methods/usability-evaluation/index.html> [Accessed 14 Apr. 2019].
- Xiang, Z., Schwartz, Z., Gerdes, J. and Uysal, M. (2015). *What can big data and text analytics tell us about hotel guest experience and satisfaction?*. [online] ScienceDirect. Available at: <https://www.sciencedirect.com/science/article/pii/S0278431914001698> [Accessed 4 Jan. 2019].

16 Appendix A – Project Proposal

COMP1682 Project Proposal

Analysis on the travel and behavioural patterns of UK Residents, to produce a web-based visualisation tool

**BSc Business Information Systems
000934841**

1 Overview

The Tourism industry has been growing at a significant rate, due to the improvements in the technologies, customer values and a decrease in the cost of travelling. “In the annual analysis of the global economic impact of Travel & Tourism, the sector is shown to account for 10.4% of global GDP and 313 million jobs, or 9.9% of total employment, in 2017.”(Manzo G, 2018) There are different kinds of businesses booming in the Tourism industry such as commercial airline, cruise, travel agency, tour organiser, local tourist information Centre, etc. All the businesses in this industry prioritise one thing in common when it comes to building a successful or profitable business which is data. There are millions of data created every second and they can be used for a variety of purposes. For example, Data collected from a commercial airline business can be used for predicting what the customers’ needs/likes are and what will make them come back to use their services again.

Collecting and storing these datasets will be an easy task to accomplish. However, the biggest challenge lies in breaking these data down into useful information. The data collected cannot be understood by people who are making the business decision, and this can have a negative impact on the growth of their business. With the help of useful information, it can provide businesses to gain insight into their customer's preferences. To produce useful information such as hidden patterns or customer preferences, there are various tools/techniques which can be used but the project will focus on using business intelligence tools. “Business intelligence (BI) is a technology-driven process for analysing data and presenting actionable information to help executives, managers and other corporate end users make informed business decisions.” (Rouse M, 2017) The project will focus on using business intelligence tools and datasets collected from an online source to produce a visualisation using a dashboard interface which will provide an insight into UK residents travel and behavioural patterns. The information gathered from this visualisation tool will be used to provide the countries with some recommendation which may improve the services they provide UK customers. This dashboard interface will then be available online and can be accessed by stakeholders who are concerned in the tourism industry in their countries, to gain an insight into the UK customers who travel to their respected countries. They can then use this information to tailor their services to accommodate UK customers.

Keywords: [Tourism, Business intelligence, Visualisation, Dashboard, Data warehousing, Spreadsheet, Tableau]

2 Aim

The aim of this project is to produce visualisation in form of a dashboard which will provide an insight into the travel and behavioural patterns of UK residents. The information gathered from the visualisation will be used to issue recommendations to improve the services provided to their UK customers who travel to their countries.

3 Objectives

The timeframe is given in [Weeks. Days]. The productive working hours per day is 6 hours.

Research [3.1]

3.1 Conduct a Feasibility Study [0.3]

- 3.1.1 Research the current need for data analysis in the industry [0.3]

3.2 Research about the tourism industry [0.5]

- 3.2.1 Research why the industry is booming over the years [0.2]
- 3.2.2 Research the types of businesses in the industry [0.2]
- 3.2.3 Research the issues in the industry [0.1]
- 3.2.4 Research on data collected in this industry [0.1]

3.3 Research Business intelligence and visualisation [0.6]

- 3.3.1 Research into the advantage and disadvantage of Business intelligence and visualisation [0.2]
- 3.3.2 Research the types of visualisation [0.1]
- 3.3.3 Research types of business intelligence tools needed for the project (Data warehousing, spreadsheet and tableau) [0.1]
- 3.3.4 Compare similar work done using BI tools. [0.2]

3.4 Research relevant topic areas [1.0]

- 3.4.1 Research Jakob Nielsen heuristics [0.1]
- 3.4.2 Research Data mining and Big Data [0.2]
- 3.4.3 Research ETL Process [0.2]
- 3.4.4 Research tools, techniques and process needed to create ERD [0.1]
- 3.4.5 Research Legal, Social, Ethical and Professional issues with the project [0.1]

Analysis [2.4]

3.5 Business and Product Requirements [1.4]

- 3.5.1 Produce Business requirements [0.5]
- 3.5.2 Produce Functional requirements using MoSCoW [0.3]
- 3.5.3 Produce Non-Functional Requirements [0.3]

3.6 Design data warehouse model [0.5]

- 3.6.1 Design a conceptual ER Diagram [0.2]
- 3.6.2 Design a physical ER Diagram [0.3]

3.7 Design Use-Case diagram [0.1]

3.8 Design rich picture relating to the project [0.1]

Design & Implementation [7.2]

3.9 Create a Data warehouse [1.4]

3.9.1 Use ETL Process to create a data warehouse using MySQL **[1.3]**

3.9.2 Test the Data warehouse using SQL queries. **[0.1]**

3.10 Produce prototype using Excel Spreadsheet [2.2]

3.10.1 Produce a visual concept of the dashboard interface using low-fidelity prototyping **[0.1]**

3.10.2 Transfer data from the Data warehouse to excel spreadsheet **[0.2]**

3.10.3 Using the data, create a visualisation using pivot tables and charts to produce useful information **[1.3]**

3.10.4 Combine the visualisation together to create a dashboard interface **[0.2]**

3.10.5 Test the dashboard interface **[0.1]**

3.11 Create Final Visualisation tool using Tableau [3.3]

3.11.1 Transfer dataset to Tableau **[0.2]**

3.11.2 Produce visualisation in form of a dashboard **[2.1]**

3.11.3 Test Dashboard interface **[0.2]**

3.11.4 Publish Visualisation tool to the web **[0.5]**

Testing [1.4]

3.12 Produce Testing Documentation using Black and White box Testing [1.4]

3.12.1 Test MySQL Data warehouse **[0.3]**

3.12.2 Test Excel Spreadsheet Dashboard **[0.4]**

3.12.3 Test Tableau Dashboard **[0.4]**

Evaluation and conclusion [1.3]

3.13 Evaluation and conclusion for Project [1.3]

3.13.1 Summaries the key finding from the project **[0.3]**

3.13.2 Produce recommendation using the information gathered from the visualisation tool **[0.4]**

3.13.3 Write a conclusion for the project **[0.2]**

3.13.4 Write future development of the system **[0.1]**

4 Project Framework or Any Methodology used

The project will be using Agile DSDM Atern to allow flexibility when proceeding with the project. Techniques such as MoSCoW prioritisation, Timeboxing, Prototyping will be used to deliver the project on-time and only produce the necessary requirements which are needed to run the business. As the project does not have a fixed requirement, the project will make sure to be adaptable to new changes if any arises during the project stages. Following the DSDM atern principles will benefit this project by delivering on time and not compromise on the quality of the application because it will only contain the essential components needed to answer the queries. The project will use iterative development and timeboxing together to produce the visualisation tool (dashboard) which will contain the necessary functions and features and will deliver on time.

There are some disadvantages when it comes to using DSDM atern but tailoring the principles to suit this project will allow it to become successful and will encounter fewer problems. By using this methodology, the project will be able to produce a minimum viable system in a few months which can offer some benefits earlier and can gain an idea into the possibilities of the system.

5 Legal, Social, Ethical and Professional issues

The dataset needed for this project will be collected from an online source. The project will follow the **data protection** laws and regulations by only allowing only authorised users to gain access to the data used in this project. Also, ensuring the data cannot be accessed, modified or deleted by unauthorised users.

The project will also respect the **copyright** law by following the terms and condition, to use the dataset in their website. Before using a dataset, it is professional to request permission from the owner of the dataset to get their **consent**. Having contacted the owner of the dataset which is needed for the project, he replied by saying that the dataset can be used and adapted under the terms of (Open government license - OGL) and any dataset used in the project must need to be referenced. (**See Appendix B**)

Equality will be taken into consideration when analysing issues related to gender and age. **Culture related issues** should also be considered when providing recommendations to the countries because each country may have a different view of thinking.

Social issues can be both negative and positive and it is mostly controlled by people. There are many issues which stems from social issues such as work, health, education, age, etc. it is important to investigate **unemployment issues** for this project because the web application tool may change the industries perspective and can lay off their employees because of the results they found.

Ethical and professional issues will be less relevant, has the size of the work force for this project is smaller. However, if more LSEP issues arise during this project, it will be included in the report.

6 Planning (see appendix A)

The project will begin with waterfall methodology at the early stages of project lifecycle to provide a structure to the project, however all the other stages will use DSDM atern approach by producing the system in an incremental process and delivering on time.

The tool used for planning the project comes from traditional model which is Gantt chart. Gantt chart contains activities and tasks of a project. The Gantt chart for this project will be attached to appendix A.

7 Appendix A

1	↳ Research	21 days	Thu 01/11/18	Thu 29/11/18		
2	↳ <u>Feasibility study</u>	3 days	Thu 01/11/18	Mon 05/11/18		
3	Research the current need for data analysis in the industry	3 days	Thu 01/11/18	Mon 05/11/18		
4	↳ <u>Research about the travel & tourism industry</u>	5 days	Tue 06/11/18	Mon 12/11/18		
5	Research why the industry is booming over the years	2 days	Tue 06/11/18	Wed 07/11/18		
6	Research the types of businesses in the industry	2 days	Wed 07/11/18	Thu 08/11/18		
7	Research the issues in the industry	1 day	Mon 12/11/18	Mon 12/11/18		
8	Research on data collected in this industry	1 day	Mon 12/11/18	Mon 12/11/18		
9	↳ <u>Research Business intelligence and visualisation</u>	6 days	Wed 14/11/18	Wed 21/11/18		
10	Research into the advantage and disadvantage of Business intelligence and visualisation	2 days	Wed 14/11/18	Thu 15/11/18		
11	Research the types of visualisation	1 day	Fri 16/11/18	Fri 16/11/18		
12	Research types of business intelligence tools needed for the project (Data warehousing, spreadsheet and tableau)	1 day	Mon 19/11/18	Mon 19/11/18		
13	Compare similar work done using BI tools.	2 days	Tue 20/11/18	Wed 21/11/18		
14	↳ <u>Research relevant topic areas</u>	6 days	Thu 22/11/18	Thu 29/11/18		
15	Research Jakob Nielsen Heuristics	1 day	Thu 22/11/18	Thu 22/11/18		
16	Research Data mining and Big Data	2 days	Thu 22/11/18	Fri 23/11/18		
17	Research ETL Process	2 days	Mon 26/11/18	Tue 27/11/18		
18	Research tools, techniques and process needed to create ERD	1 day	Wed 28/11/18	Wed 28/11/18		
19	Research Legal, Social, Ethical and Professional issues with the project	1 day	Thu 29/11/18	Thu 29/11/18		
20	↳ Analysis	16 days	Fri 30/11/18	Fri 21/12/18	1	
21	↳ <u>Business and Product Requirements</u>	11 days	Fri 30/11/18	Fri 14/12/18		
22	Produce Business requirements	5 days	Fri 30/11/18	Thu 06/12/18		
23	Produce Functional requirements using MoSCoW	3 days	Fri 07/12/18	Tue 11/12/18		

24	Produce Non-Functional Requirements	3 days	Wed 12/12/18	Fri 14/12/18		
25	↳ <u>Design data warehouse model</u>	5 days	Fri 14/12/18	Thu 20/12/18		
26	Design a conceptual ER Diagram	2 days	Fri 14/12/18	Mon 17/12/18		
27	Design a physical ER Diagram	3 days	Tue 18/12/18	Thu 20/12/18		
28	Design Use-Case Diagram	1 day	Fri 14/12/18	Fri 14/12/18		
29	Design rich picture relating to the project	1 day	Tue 18/12/18	Tue 18/12/18		
30	↳ Design & Implementation	51 days	Mon 24/12/18	Mon 04/03/19	20	
31	↳ <u>Create a Data warehouse</u>	11 days	Mon 24/12/18	Mon 07/01/19		
32	Use ETL Process to create a data warehouse in MySQL	10 days	Mon 24/12/18	Fri 04/01/19		
33	Test the Database using SQL queries.	1 day	Mon 07/01/19	Mon 07/01/19		
34	↳ <u>Produce prototype using Excel Spreadsheet</u>	16 days	Tue 08/01/19	Tue 29/01/19	31	
35	Produce a visual concept of the dashboard interface using low-fidelity prototyping	1 day	Tue 08/01/19	Tue 08/01/19		
36	Transfer data from the database to excel spreadsheet	2 days	Tue 08/01/19	Wed 09/01/19		
37	Using the data, create a visualisation using pivot tables and charts to produce useful information	10 days	Fri 11/01/19	Thu 24/01/19		
38	Combine the visualisation together to create a dashboard interface	2 days	Fri 25/01/19	Mon 28/01/19		
39	Test the dashboard interface	1 day	Tue 29/01/19	Tue 29/01/19		
40	↳ <u>Create Final Visualisation tool using Tableau</u>	24 days	Wed 30/01/19	Mon 04/03/19	31	
41	Transfer dataset to Tableau	2 days	Wed 30/01/19	Thu 31/01/19		
42	Produce visualisation in form of a dashboard	15 days	Fri 01/02/19	Thu 21/02/19		
43	Test Dashboard interface	2 days	Fri 22/02/19	Mon 25/02/19		
44	Publish Visualisation tool to the web	5 days	Tue 26/02/19	Mon 04/03/19		
45	↳ Testing	11 days	Tue 05/03/19	Tue 19/03/19		
46	↳ <u>Produce Testing Documentation using Black and White box Testing</u>	11 days	Tue 05/03/19	Tue 19/03/19	30	



A Gantt chart illustrating the timeline for various project tasks. The tasks are listed in rows, with their start and end dates indicated in the columns. The chart shows a sequence of tasks starting from March 5th, 2019, and ending on April 2nd, 2019.

47	Test MySQL database	3 days	Tue 05/03/19	Thu 07/03/19		
48	Test Excel Spreadsheet Dashboard	4 days	Fri 08/03/19	Wed 13/03/19		
49	Test Tableau Dashboard	4 days	Thu 14/03/19	Tue 19/03/19		
50	▪ Evaluation and conclusion	10 days	Wed 20/03/19	Tue 02/04/19	45	
51	▪ Evaluation and conclusion for Project	10 days	Wed 20/03/19	Tue 02/04/19		
52	Summaries the key finding from the project	3 days	Wed 20/03/19	Fri 22/03/19		
53	Produce recommendation using the information gathered from visualisation tool	4 days	Mon 25/03/19	Thu 28/03/19		
54	Write conclusion for project	2 days	Fri 29/03/19	Mon 01/04/19		
55	Write future development of the system	1 day	Tue 02/04/19	Tue 02/04/19		

8 Appendix B



Kogul Kuganathan <kk5572a@greenwich.ac.uk>

Do I need permission to use Dataset for an Undergraduate Project .

1 message

Social Survey <socialsurveys@ons.gov.uk>
Reply-To: socialsurveys@ons.gov.uk
To: kk5572a@greenwich.ac.uk

18 October 2018 at 08:19

Reference - Enquiry Number 22626

Hello Kogul,

Thank you for your email. Under the terms of the Open Government Licence (OGL), you are free to use and adapt the data set you link to (or any other data available on our website). We give the following guidance:

Users reproducing ONS content without adaptation should include a source accreditation to ONS: Source: Office for National Statistics licensed under the Open Government Licence v.3.0.
Users reproducing ONS content which is adapted should include a source accreditation to ONS: Adapted from data from the Office for National Statistics licensed under the Open Government Licence v.3.0.

The reference to the OGL should link to <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/> if possible.

I hope this is helpful.

Regards,

Paul Bloomfield,
Data Advice and Relations Team,
ONS Social Statistics.

Dear Sir/Madam,

My name is Kogul Kuganathan and I have come across the data set for tourism industry which I am interested in using for my Undergraduate project. I would like to know if I am allowed to use your data set.

This is the link to the dataset that I am thinking of using: <https://www.ons.gov.uk/peoplepopulationandcommunity/leisureandtourism/datasets/overseastavelandtourism> by Giles Horsfield

Thank you,
Kogul Kuganathan

9 Initial References

- Rouse, M. (2017). *business intelligence (BI)*. [online] SearchBusinessAnalytics. Available at: <https://searchbusinessanalytics.techtarget.com/definition/business-intelligence-BI> [Accessed 24 Oct. 2018].

- Manzo, G. (2018) *TRAVEL & TOURISM ECONOMIC IMPACT 2018 WORLD*, World Travel & Tourism Council, p. 3, [online] Available at: <https://www.wttc.org/-/media/files/reports/economic-impact-research/regions-2018/world2018.pdf> (Accessed 26 October 2018).
- Horsfield, G. (2018). *Overseas travel and tourism, quarterly - Office for National Statistics*. [online] Ons.gov.uk. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/leisureandtourism/datasets/overseastravelandtourism> [Accessed 25 Oct. 2018].
- Nationalarchives.gov.uk. (2018). *Open Government Licence*. [online] Available at: <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/> [Accessed 25 Oct. 2018].

10 Initial Bibliography

- LibGuides: Visualisation and Interactive Media: What is Visualisation?, (2018) *Libguides.library.curtin.edu.au*, [online] Available at: <https://libguides.library.curtin.edu.au/c.php?g=388681&p=2688784> (Accessed 25 October 2018).
- What Are Business Intelligence (BI) Tools? | Microsoft Azure, (2018) *Azure.microsoft.com*, [online] Available at: <https://azure.microsoft.com/en-gb/overview/what-are-business-intelligence-tools/> (Accessed 25 October 2018).
- Tableau Public, (2018) *Tableau Public*, [online] Available at: <https://public.tableau.com/en-us/s/> (Accessed 25 October 2018).
- Roser, M. (n.d.) Tourism, *Our World in Data*, [online] Available at: <https://ourworldindata.org/tourism> (Accessed 31 October 2018).
- Caine, M. (2011) What is DSDM Atern? | M.C. Partners & Associates, *Mcpa.biz*, [online] Available at: <http://www.mcpa.biz/2011/08/what-is-dsdm-atern/> (Accessed 26 October 2018).
- Legal, social, ethical & professional issues (LSEPI) | BCS - The Chartered Institute for IT, (2017) *Bcs.org*, [online] Available at: <https://www.bcs.org/content/ConWebDoc/57831> (Accessed 27 October 2018).
- Pernice, K. (2016) UX Prototypes: Low Fidelity vs. High Fidelity, *Nielsen Norman Group*, [online] Available at: <https://www.nngroup.com/articles/ux-prototype-hi-lo-fidelity/> (Accessed 28 October 2018).

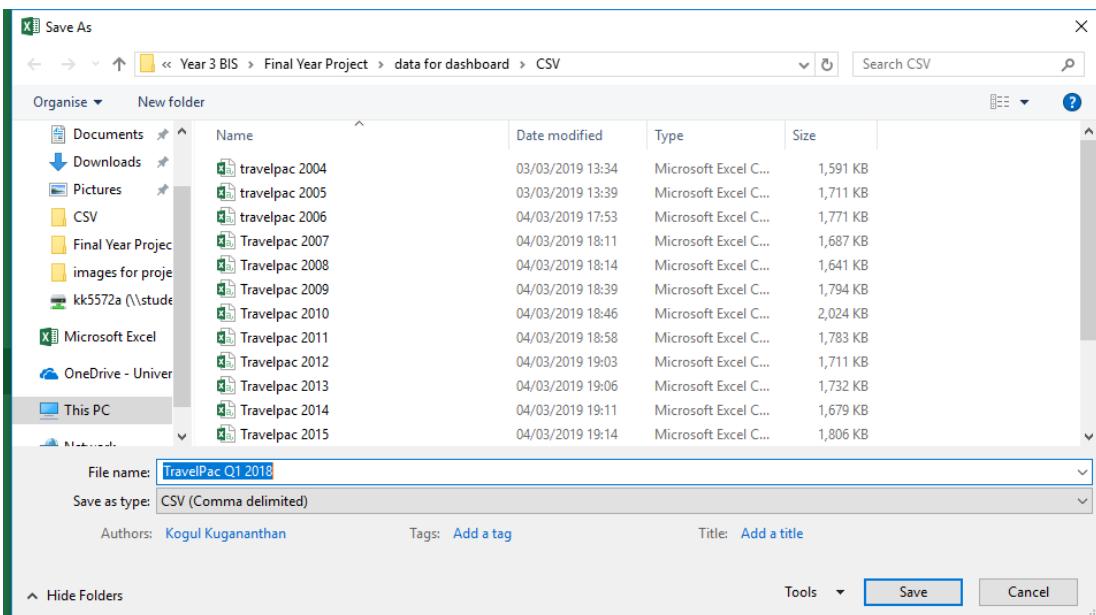
17 Appendix B (Implementation – Datawarehouse)

(1) Starting ETL process and currently in extraction phase. This is one of the datasets being used for the Datawarehouse.

1	Year	quarter	ukos	mode	country	purpose	package	Age	Sex	duration	visits	nights	spend	sample
2	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	0-15	Male	4-13 nights	4,570,400	31,899,708	2,927,838,837	4
3	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	0-15	Female	4-13 nights	2,550,820	16,970,239	2,136,828,643	4
4	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	16-24	Male	4-13 nights	3,258,280	21,451,579	2,509,682,039	3
5	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	16-24	Male	14-27 nights	2,001,483	46,034,109	1,200,889,800	1
6	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	16-24	Female	4-13 nights	1,556,020	10,892,140	1,693,727,770	2
7	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	16-24	Female	14-27 nights	960,807	14,412,105	154,689,927	1
8	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	25-34	Male	4-13 nights	4,243,375	22,148,195	1,771,885,190	5
9	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	25-34	Male	14-27 nights	960,807	13,451,298	77,825,367	1
10	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	25-34	Female	4-13 nights	3,387,346	28,420,954	3,067,399,683	4
11	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	35-44	Male	1-3 nights	5,114,328	14,071,355	2,830,271,775	4
12	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	35-44	Male	4-13 nights	6,835,964	41,548,554	5,139,596,298	9
13	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	35-44	Female	1-3 nights	2,295,051	5,536,127	953,005,800	3
14	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	35-44	Female	4-13 nights	5,386,696	36,090,760	3,647,135,606	7
15	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	45-54	Male	1-3 nights	2,164,757	6,494,271	1,271,253,044	3
16	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	45-54	Male	4-13 nights	12,031,539	73,623,248	12,293,684,767	10
17	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	45-54	Female	4-13 nights	7,153,300	56,325,052	6,224,708,081	6
18	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	45-54	Female	14-27 nights	718,059	10,052,826	644,098,923	1
19	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	55-64	Male	4-13 nights	7,165,874	50,045,771	5,984,430,021	7
20	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	55-64	Female	1-3 nights	572,454	1,144,908	171,736,200	1
21	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	55-64	Female	4-13 nights	2,227,684	15,593,788	1,238,708,651	3
22	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	55-64	Female	14-27 nights	752,855	14,304,245	506,671,415	1
23	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	65 & over	Male	4-13 nights	2,309,186	16,198,463	1,381,964,220	3
24	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	65 & over	Female	14-27 nights	485,891	8,260,147	488,806,346	1
25	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Independent	D/K	Dont know	4-13 nights	946,025	4,730,125	619,646,375	1
26	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	0-15	Male	4-13 nights	2,962,428	20,736,996	2,117,478,747	5
27	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	0-15	Female	4-13 nights	4,838,875	33,872,125	3,759,772,965	4
28	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	16-24	Male	4-13 nights	5,063,682	32,433,871	4,215,868,850	8
29	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	16-24	Female	4-13 nights	1,114,580	7,802,060	178,332,800	1
30	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	25-34	Male	4-13 nights	8,048,726	56,341,082	5,519,384,384	8
31	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	25-34	Female	4-13 nights	5,256,507	37,573,559	2,509,202,255	6
32	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	35-44	Male	4-13 nights	14,111,045	95,939,240	11,500,055,612	14
33	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	35-44	Female	4-13 nights	4,112,421	28,786,947	3,058,996,872	4
34	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	45-54	Male	4-13 nights	13,995,189	97,679,863	10,734,671,943	16
35	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	45-54	Male	14-27 nights	655,182	9,722,548	402,281,748	1
36	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	45-54	Female	4-13 nights	8,950,483	62,653,381	7,610,017,691	9
37	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	55-64	Male	4-13 nights	16,104,246	112,729,722	12,813,865,035	13
38	2013	Jan-Mar	UK residents	Air	Austria	Holiday	Non-Independent	55-64	Male	14-27 nights	2,598,924	36,384,936	1,923,203,760	1

A11271	A	B	C	D	E	F	G	H	I	J	K	L	M
1	255974	2018	Jan-Mar	Air	Austria	Holiday	Independ	0-15	Male	4-13 night	3512	23501	2874330
2	255975	2018	Jan-Mar	Air	Austria	Holiday	Independ	16-24	Male	4-13 night	713	4991	501259
3	255976	2018	Jan-Mar	Air	Austria	Holiday	Independ	16-24	Female	1-3 nights	1433	2865	152585
4	255977	2018	Jan-Mar	Air	Austria	Holiday	Independ	16-24	Female	4-13 night	4952	43571	3671850
5	255978	2018	Jan-Mar	Air	Austria	Holiday	Independ	25-34	Male	1-3 nights	2282	5710	614448
6	255979	2018	Jan-Mar	Air	Austria	Holiday	Independ	25-34	Male	4-13 night	14340	65013	8609892
7	255980	2018	Jan-Mar	Air	Austria	Holiday	Independ	25-34	Female	Nil Stay	2587	0	1552096
8	255981	2018	Jan-Mar	Air	Austria	Holiday	Independ	25-34	Female	1-3 nights	7197	14940	987962
9	255982	2018	Jan-Mar	Air	Austria	Holiday	Independ	25-34	Female	4-13 night	4362	17450	2617492
10	255983	2018	Jan-Mar	Air	Austria	Holiday	Independ	35-44	Male	4-13 night	16527	97103	13642930
11	255984	2018	Jan-Mar	Air	Austria	Holiday	Independ	35-44	Female	1-3 nights	1943	5830	1165979
12	255985	2018	Jan-Mar	Air	Austria	Holiday	Independ	35-44	Female	4-13 night	1248	8734	1188821
13	255986	2018	Jan-Mar	Air	Austria	Holiday	Independ	45-54	Male	1-3 nights	5218	12727	2335236
14	255987	2018	Jan-Mar	Air	Austria	Holiday	Independ	45-54	Male	4-13 night	14861	104770	18110180
15	255988	2018	Jan-Mar	Air	Austria	Holiday	Independ	45-54	Male	3-6 month	928	120602	1159635
16	255989	2018	Jan-Mar	Air	Austria	Holiday	Independ	45-54	Female	1-3 nights	3281	9125	3735478
17	255990	2018	Jan-Mar	Air	Austria	Holiday	Independ	45-54	Female	4-13 night	4133	18670	1674421
18	255991	2018	Jan-Mar	Air	Austria	Holiday	Independ	55-64	Male	4-13 night	4010	21810	3696172
19	255992	2018	Jan-Mar	Air	Austria	Holiday	Independ	55-64	Male	28-90 nigh	2086	66743	4692850
20	255993	2018	Jan-Mar	Air	Austria	Holiday	Independ	55-64	Female	1-3 nights	649	1298	108366
21	255994	2018	Jan-Mar	Air	Austria	Holiday	Independ	55-64	Female	4-13 night	6725	44139	6501534
22	255995	2018	Jan-Mar	Air	Austria	Holiday	Independ	65 & over	Male	4-13 night	5556	29636	4121673
23	255996	2018	Jan-Mar	Air	Austria	Holiday	Independ	D/K	Male	4-13 night	5338	34837	4640203
24	255997	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	0-15	Male	4-13 night	716	5015	980700
25	255998	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	0-15	Female	4-13 night	1149	8040	1329474
26	255999	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	16-24	Male	4-13 night	642	4497	1434468
27	256000	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	16-24	Female	4-13 night	642	4497	621838
28	256001	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	25-34	Male	1-3 nights	4628	13885	2466907
29	256002	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	25-34	Male	4-13 night	3201	22408	3944636
30	256003	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	25-34	Female	4-13 night	7144	44180	5449096
31	256004	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	35-44	Male	1-3 nights	3345	10035	1210932
32	256005	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	35-44	Male	4-13 night	1306	7365	867438
33	256006	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	35-44	Female	4-13 night	2934	20539	3144437
34	256007	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	45-54	Male	4-13 night	4928	34498	10357319
35	256008	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	45-54	Female	4-13 night	1936	12836	1546297
36	256009	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	55-64	Male	1-3 nights	593	1780	35596
37	256010	2018	Jan-Mar	Air	Austria	Holiday	Non-Inde	55-64	Male	4-13.night	5652	39564	38101780

(2) After checking each file and deleting unnecessary data, the files will be ready to be converted to CSV file.



(3) Convert excel file to csv format. I will follow this method for every file until every file is finished.

- (4) After all the files are saved as CSV files, it is time to create the table in MySQL. This image shows the table being created and it does not require any ERD design. This one table will contain all the data required for the project. 'cwe_UKtouristdata' indicates that it is in extraction phase. 'CREATE TABLE' sql command can be used but this function is easier to use to create the table.

Name	Type	Length/Values	Default	Collation	Attributes	Null	Index
UKtourdataID	INT		None				
Year	INT		None				
Quarter	VARCHAR	100	None				
Mode	VARCHAR	100	None				
Country	VARCHAR	100	None				
Purpose	VARCHAR	100	None				
Package	VARCHAR	100	None				
Age	VARCHAR	100	None				
Gender	VARCHAR	100	None				
Duration	VARCHAR	100	None				
Visits	INT		None				
Nights	INT		None				
Spent	INT		None				
Console	internal						

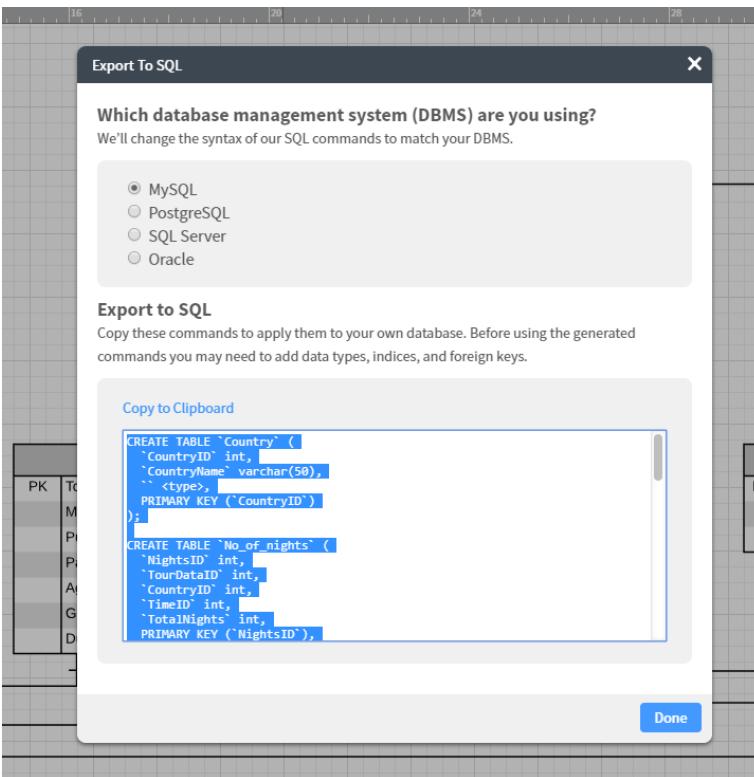
(5) After the table is created, import the csv file into the extraction table one by one.

The screenshot shows the phpMyAdmin interface with the 'File to import' dialog open. On the left, the database structure is visible, including the 'cwe_UKtouristdata' table. The 'File to import' dialog has a 'Choose file' button and a list of files from a local drive. The file 'travelpac 2004' is selected, highlighted in blue.

(6) After importing a few csv files, the table will look like this. After importing all the files into the extraction table, the total records imported was 267,243 which contains 15 years of data. The extraction phase is finished, and the next phase is transformation.

The screenshot shows the 'Structure' tab for the 'cwe_UKtouristdata' table. The table structure is displayed with 15 columns: UKtourdataID, Year, Quarter, Mode, CountryName, Purpose, Package, Age, Gender, Duration, Visits, Nights, Spent, and two additional columns at the bottom: 'Console' and 'Actions'. The table contains 17 rows of data, each representing a record from 2004 to 2018.

The screenshot shows the 'Tables' section of phpMyAdmin. It lists the 'cwe_UKtouristdata' table with the following details: 267,243 rows, 13 columns, InnoDB storage engine, latin1_swedish_ci character set, and 27.6 Mib size. The table has a primary key constraint and a sum operation defined.



(7) In the transformation phase, all the data which are no normalised should be in third normal form which means that the dataset will have less redundancy. The 3NF ERD was created in Lucidchart (diagramming tool) and it can convert the diagram into MYSQL which can be used to create the tables in the phpMyAdmin.

```

CREATE TABLE `trans_Country` (
  `CountryID` int AUTO_INCREMENT,
  `CountryName` varchar(50),
  PRIMARY KEY (`CountryID`)
);

CREATE TABLE `trans_No_of_nights` (
  `NightsID` int AUTO_INCREMENT,
  `TourDataID` int,
  `CountryID` int,
  `TimeID` int,
  `TotalNights` int,
  PRIMARY KEY (`NightsID`),
  KEY `FK` (`TourDataID`, `CountryID`, `TimeID`)
);

CREATE TABLE `trans_Time` (
  `TimeID` int AUTO_INCREMENT,
  `Quarter` varchar(50),
  `Year` int,
  PRIMARY KEY (`TimeID`)
);

CREATE TABLE `trans_Total_Spent` (
  `SpentID` int AUTO_INCREMENT,
  `TourdataID` int,
  `CountryID` int,
  `TimeID` int,
  `Totalspent` int,
  PRIMARY KEY (`SpentID`),
  KEY `FK` (`TourDataID`, `CountryID`, `TimeID`)
);

CREATE TABLE `trans_Tourist_data` (
  `TourDataID` int AUTO_INCREMENT,
  `Mode` varchar(50),
  `Purpose` varchar(50),
  `Package` varchar(50),
  `Age` varchar(50),
  `Gender` varchar(50),
  `Duration` varchar(50),
  PRIMARY KEY (`TourDataID`)
);

CREATE TABLE `trans_No_of_visits` (
  `VisitsID` int AUTO_INCREMENT,
  `TourDataID` int,
  `CountryID` int,
  `TimeID` int,
  `TotalVisits` int,
  PRIMARY KEY (`VisitsID`),
  KEY `FK` (`TourDataID`, `CountryID`, `TimeID`)
);

```

(8) Before running the sql command into phpMyAdmin, I added auto increment to primary key in the sql code. It will automatically create the ID for each record when inserting them into the table. Running this command will create all the tables for transformation phase.

- (9) You can see in the left all the tables for transformation phase. Now, it is time to populate the table with data which will come from the extraction table. ‘INSERT INTO SELECT’ is used for duplicating and transferring data from one table to another. This is for the ‘trans_Country’ table and the data is coming from ‘cwe_UKtouristdata’. The ‘DISTINCT’ is added to avoid repeated data from being transferred into the new table. After running the command, the table will contain all the crucial data.

The screenshot shows the MySQL Workbench interface. On the left, the database structure is displayed under the schema 'mdb_kk5572a'. The 'Tables' section contains the 'trans_Country' table, which is highlighted. To the right, a query editor window titled 'Run SQL query/queries on table mdb_kk5572a.trans_Country:' displays the following SQL code:

```

1 insert into trans_Country (CountryName)
2 select DISTINCT (cwe_UKtouristdata.CountryName) from cwe_UKtouristdata
  
```

Below the query are several buttons: SELECT*, SELECT, INSERT, UPDATE, and DELETE. A checkbox for 'Bind parameters' is also present. The status bar at the bottom of the query editor shows the result of the execution:

✓ 211 rows inserted.
Inserted row id: 211 (Query took 0.1625 seconds.)

insert into trans_Country (CountryName) select DISTINCT (cwe_UKtouristdata.CountryName) from cwe_UKtouristdata

At the bottom of the interface, a results grid for the 'trans_Country' table is shown. The grid has columns 'CountryID' and 'CountryName'. The data is as follows:

CountryID	CountryName
1	Austria
2	Belgium
3	Bulgaria
4	Croatia
5	Czech Republic
6	Cyprus EU
7	Denmark
8	Estonia
9	Finland
10	France
11	Germany
12	Gibraltar
13	Greece
14	Hungary
15	Iceland
16	Irish Republic
17	Italy
18	Iraq

(10) The same process was followed to populate all the transformation tables. However, different tables will require different sql commands. These images only show a sample of each transformation table.

		TimeID	Quarter	Year
<input type="checkbox"/>	Edit	Copy	Delete	1 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	2 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	3 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	4 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	5 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	6 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	7 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	8 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	9 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	10 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	11 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	12 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	13 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	14 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	15 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	16 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	17 Jan-Mar 2004
<input type="checkbox"/>	Edit	Copy	Delete	18 Jan-Mar 2004

	TourDataID	Mode	Purpose	Package	Age	Gender	Duration
<input type="checkbox"/>	1	Air	Holiday	Independent	0-15	Male	4-13 nights
<input type="checkbox"/>	2	Air	Holiday	Independent	0-15	Female	4-13 nights
<input type="checkbox"/>	3	Air	Holiday	Independent	16-24	Male	1-3 nights
<input type="checkbox"/>	4	Air	Holiday	Independent	16-24	Male	4-13 nights
<input type="checkbox"/>	5	Air	Holiday	Independent	16-24	Male	14-27 nights
<input type="checkbox"/>	6	Air	Holiday	Independent	25-34	Male	1-3 nights
<input type="checkbox"/>	7	Air	Holiday	Independent	25-34	Male	4-13 nights
<input type="checkbox"/>	8	Air	Holiday	Independent	25-34	Female	1-3 nights
<input type="checkbox"/>	9	Air	Holiday	Independent	25-34	Female	4-13 nights
<input type="checkbox"/>	10	Air	Holiday	Independent	35-44	Male	1-3 nights
<input type="checkbox"/>	11	Air	Holiday	Independent	35-44	Male	4-13 nights
<input type="checkbox"/>	12	Air	Holiday	Independent	35-44	Male	14-27 nights
<input type="checkbox"/>	13	Air	Holiday	Independent	35-44	Female	1-3 nights
<input type="checkbox"/>	14	Air	Holiday	Independent	35-44	Female	4-13 nights
<input type="checkbox"/>	15	Air	Holiday	Independent	45-54	Male	1-3 nights
<input type="checkbox"/>	16	Air	Holiday	Independent	45-54	Male	4-13 nights
<input type="checkbox"/>	17	Air	Holiday	Independent	45-54	Male	14-27 nights
<input type="checkbox"/>	18	Air	Holiday	Independent	45-54	Female	1-3 nights
<input type="checkbox"/>	19	Air	Holiday	Independent	45-54	Female	4-13 nights
<input type="checkbox"/>	20	Air	Holiday	Independent	55-64	Male	1-3 nights
<input type="checkbox"/>	21	Air	Holiday	Independent	55-64	Male	4-13 nights
<input type="checkbox"/>	22	Air	Holiday	Independent	55-64	Male	14-27 nights
<input type="checkbox"/>	23	Air	Holiday	Independent	55-64	Female	1-3 nights
<input type="checkbox"/>	24	Air	Holiday	Independent	65 & over	Male	1-3 nights
<input type="checkbox"/>	25	Air	Holiday	Independent	65 & over	Male	4-13 nights

trans_No_of_nights

	NightsID	TourDataID	CountryID	TimeID	TotalNights
Edit Copy Delete	267243	267243	59	267243	1876
Edit Copy Delete	267242	267242	58	267242	14742
Edit Copy Delete	267241	267241	53	267241	66675
Edit Copy Delete	267240	267240	35	267240	19751
Edit Copy Delete	267239	267239	35	267239	85586
Edit Copy Delete	267238	267238	35	267238	79002
Edit Copy Delete	267237	267237	35	267237	85586
Edit Copy Delete	267236	267236	33	267236	162983
Edit Copy Delete	267235	267235	32	267235	14072
Edit Copy Delete	267234	267234	32	267234	14817
Edit Copy Delete	267233	267233	32	267233	20102
Edit Copy Delete	267232	267232	32	267232	77787
Edit Copy Delete	267231	267231	32	267231	26334
Edit Copy Delete	267230	267230	31	267230	6701
Edit Copy Delete	267229	267229	28	267229	4020
Edit Copy Delete	267228	267228	28	267228	105337
Edit Copy Delete	267227	267227	28	267227	385232
Edit Copy Delete	267226	267226	28	267226	122237
Edit Copy Delete	267225	267225	28	267225	132098
Edit Copy Delete	267224	267224	28	267224	309426
Edit Copy Delete	267223	267223	28	267223	59266
Edit Copy Delete	267222	267222	28	267222	37042

trans_Overview

	VisitsID	TourDataID	CountryID	TimeID	TotalVisits
Edit Copy Delete	267243	267243	59	267243	1876
Edit Copy Delete	267242	267242	58	267242	1340
Edit Copy Delete	267241	267241	53	267241	3704
Edit Copy Delete	267240	267240	35	267240	3292
Edit Copy Delete	267239	267239	35	267239	3292
Edit Copy Delete	267238	267238	35	267238	3292
Edit Copy Delete	267237	267237	35	267237	3292
Edit Copy Delete	267236	267236	33	267236	3704
Edit Copy Delete	267235	267235	32	267235	2010
Edit Copy Delete	267234	267234	32	267234	3704
Edit Copy Delete	267233	267233	32	267233	2010
Edit Copy Delete	267232	267232	32	267232	3704
Edit Copy Delete	267231	267231	32	267231	3292
Edit Copy Delete	267230	267230	31	267230	1340
Edit Copy Delete	267229	267229	28	267229	1340
Edit Copy Delete	267228	267228	28	267228	3292

trans_Country

	SpentID	TourDataID	CountryID	TimeID	TotalSpent
Edit Copy Delete	267243	267243	59	267243	243912
Edit Copy Delete	267242	267242	58	267242	2010242
Edit Copy Delete	267241	267241	53	267241	5296939
Edit Copy Delete	267240	267240	35	267240	2221943
Edit Copy Delete	267239	267239	35	267239	3291767
Edit Copy Delete	267238	267238	35	267238	1645884
Edit Copy Delete	267237	267237	35	267237	1316707
Edit Copy Delete	267236	267236	33	267236	2222492
Edit Copy Delete	267235	267235	32	267235	319627
Edit Copy Delete	267234	267234	32	267234	6485972
Edit Copy Delete	267233	267233	32	267233	3166119
Edit Copy Delete	267232	267232	32	267232	1852077
Edit Copy Delete	267231	267231	32	267231	2742042
Edit Copy Delete	267230	267230	31	267230	645958
Edit Copy Delete	267229	267229	28	267229	1340161
Edit Copy Delete	267228	267228	28	267228	2172566
Edit Copy Delete	267227	267227	28	267227	0

(11) Here are the SQL commands which were used to populate each individual transformation tables.

- insert into trans_Country (CountryName)
select DISTINCT (cwe_UKtouristdata.CountryName) from cwe_UKtouristdata
- insert into trans_Time (Quarter, Year)
Select DISTINCT cwe_UKtouristdata.Quarter, cwe_UKtouristdata.Year
from cwe_UKtouristdata
- insert into trans_Tourist_data (trans_Tourist_data.Mode, Purpose, Package, Age, Gender, Duration)
SELECT cwe_UKtouristdata.Mode,
cwe_UKtouristdata.Purpose,cwe_UKtouristdata.Purpose,cwe_UKtouristdata.Age,
cwe_UKtouristdata.Gender,cwe_UKtouristdata.Duration from cwe_UKtouristdata
- insert into trans_No_of_nights
(trans_No_of_nights.TourDataID,trans_No_of_nights.CountryID,trans_No_of_nights.TimeID
, trans_No_of_nights.TotalNights) SELECT trans_Tourist_data.TourDataID,
trans_Country.CountryID,trans_Time.TimeID,cwe_UKtouristdata.Nights
from trans_Tourist_data, trans_Time,trans_Country, cwe_UKtouristdata
where trans_Country.CountryName = cwe_UKtouristdata.CountryName
AND trans_Time.Year = cwe_UKtouristdata.Year
AND trans_Time.Quarter = cwe_UKtouristdata.Quarter;

(12) The transformation phase is finished and starting the loading phase. Copying code from Lucidchart 2NF diagram. The 2NF ERD follows star formation.

The screenshot shows a database export interface with the following details:

- Table Schema:** A table titled "load_Tourist_Dashboard" is displayed with the following columns and data types:

TourDashID	int
TimeID	int
TourDataID	int
TotalfigureID	int
D(userID)	int
Year	int
Quarter	varchar(50)
Mode	varchar(50)
CountryName	varchar(50)
Purpose	varchar(50)
Package	varchar(50)
Age	varchar(50)
Gender	varchar(50)
Duration	varchar(50)
TotalVisits	int
TotalNights	int
TotalSpent	int
Full_Name	varchar(100)
Username	varchar(100)
Password	varchar(100)
E-mail	varchar(100)
Age	varchar(100)
User_country	varchar(100)
Comments	varchar(100)
- Generated SQL:** The interface displays two CREATE TABLE statements:


```

CREATE TABLE `load_Dash_user` (
`D(userID)` int,
`Full_Name` varchar(100),
`Username` varchar(100),
`Password` varchar(100),
`E-mail` varchar(100),
`Age` varchar(100),
`User_country` varchar(100),
`Comments` varchar(100),
PRIMARY KEY (`D(userID`)
);

CREATE TABLE `load_Total_figure` (
`TotalfigureID` int,

```
- Export Options:** A sidebar on the left asks "Which database management system (DBMS) are you using?" with options for MySQL, PostgreSQL, SQL Server, and Oracle. MySQL is selected.
- Buttons:** At the bottom right are "Done" and "Copy to Clipboard" buttons.

Run SQL query/queries on database mdb_kk5572:

```

1 CREATE TABLE `load_Dash_user` (
2   `D(userID` int AUTO_INCREMENT,
3   `Full_Name` varchar(100),
4   `Username` varchar(100),
5   `Password` varchar(100),
6   `E-mail` varchar(100),
7   `Age` varchar(100),
8   `User_country` varchar(100),
9   `Comments` varchar(100),
10  PRIMARY KEY (`D(userID`)
11 );
12
13 CREATE TABLE `load_Total_figure` (
14   `TotalfigureID` int AUTO_INCREMENT,
15   `TotalVisits` int,
16   `TotalNights` int,
17   `Totalspent` int,
18   PRIMARY KEY (`TotalfigureID`)
19 );
20
21 CREATE TABLE `load_Time` (
22   `TimeID` int AUTO_INCREMENT,
23   `Quarter` varchar(50),
24   `Year` int,
25   PRIMARY KEY (`TimeID`)
26 );
27
28 CREATE TABLE `load_Tourist_data` (
29   `TourDataID` int AUTO_INCREMENT,
30   `Mode` varchar(50),
31   `CountryName` varchar(50),
32   `Purpose` varchar(50),
33   `Package` varchar(50),
34   `Age` varchar(50),
35   `Gender` varchar(50),
36   `Duration` varchar(50),
37   PRIMARY KEY (`TourDataID`)
38 );
39

```

(13) Adding auto increment to all primary keys to allow the table to automatically create new ID for each new record inserted.

(14) After running the sql command from Lucidchart, the tables will be created in phpMyAdmin.

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra
1	TourDashID	int(11)		No	None		AUTO_INCREMENT	
2	TimeID	int(11)		Yes	NULL			
3	TourDataID	int(11)		Yes	NULL			
4	TotalfigureID	int(11)		Yes	NULL			
5	D(userID	int(11)		Yes	NULL			
6	Year	int(11)		Yes	NULL			
7	Quarter	varchar(50)	latin1_swedish_ci	Yes	NULL			
8	Mode	varchar(50)	latin1_swedish_ci	Yes	NULL			
9	CountryName	varchar(50)	latin1_swedish_ci	Yes	NULL			
10	Purpose	varchar(50)	latin1_swedish_ci	Yes	NULL			
11	Package	varchar(50)	latin1_swedish_ci	Yes	NULL			
12	Age	varchar(50)	latin1_swedish_ci	Yes	NULL			
13	Gender	varchar(50)	latin1_swedish_ci	Yes	NULL			
14	Duration	varchar(50)	latin1_swedish_ci	Yes	NULL			
15	TotalVisits	int(11)		Yes	NULL			
16	TotalNights	int(11)		Yes	NULL			
17	TotalSpent	int(11)		Yes	NULL			
18	Full_Name	varchar(100)	latin1_swedish_ci	Yes	NULL			
19	Username	varchar(100)	latin1_swedish_ci	Yes	NULL			
20	Password	varchar(100)	latin1_swedish_ci	Yes	NULL			
21	E-mail	varchar(100)	latin1_swedish_ci	Yes	NULL			
22	User_Age	varchar(100)	latin1_swedish_ci	Yes	NULL			
23	User_country	varchar(100)	latin1_swedish_ci	Yes	NULL			
24	Comments	varchar(100)	latin1_swedish_ci	Yes	NULL			

- (15) The same process used for the transformation phase will be used for the loading phase. 'INSERT INTO SELECT' command is used to duplicate and transfer the data from the transformation table to the new loading tables. Here you can see the sql command for 'load_time' and running the code will populate the table

The screenshot shows a database interface. At the top, a window titled "Run SQL query/queries on table mdb_kk5572a.load_Time:" contains the following SQL code:

```
1 insert into load_Time (load_Time.Quarter, load_Time.Year)
2 select trans_Time.Quarter,trans_Time.Year from trans_Time
```

To the left of the window, there is a sidebar listing various tables:

- + load_Time
- + load_Total_figure
- + load_Tourist_Dashboard
- + load_Tourist_data
- + trans_Country
- + trans_No_of_nights
- + trans_No_of_visits
- + trans_Time
- + trans_Total_Spent
- + trans_Tourist_data

The screenshot shows a table named "load_Time" with 19 rows of data. The table has columns: TimeID, Quarter, and Year. The data is as follows:

	TimeID	Quarter	Year
<input type="checkbox"/>	1	Jan-Mar	2004
<input type="checkbox"/>	2	Jan-Mar	2004
<input type="checkbox"/>	3	Jan-Mar	2004
<input type="checkbox"/>	4	Jan-Mar	2004
<input type="checkbox"/>	5	Jan-Mar	2004
<input type="checkbox"/>	6	Jan-Mar	2004
<input type="checkbox"/>	7	Jan-Mar	2004
<input type="checkbox"/>	8	Jan-Mar	2004
<input type="checkbox"/>	9	Jan-Mar	2004
<input type="checkbox"/>	10	Jan-Mar	2004
<input type="checkbox"/>	11	Jan-Mar	2004
<input type="checkbox"/>	12	Jan-Mar	2004
<input type="checkbox"/>	13	Jan-Mar	2004
<input type="checkbox"/>	14	Jan-Mar	2004
<input type="checkbox"/>	15	Jan-Mar	2004
<input type="checkbox"/>	16	Jan-Mar	2004
<input type="checkbox"/>	17	Jan-Mar	2004
<input type="checkbox"/>	18	Jan-Mar	2004
<input type="checkbox"/>	19	Jan-Mar	2004

(16) These are some of the other loading table which uses the same process as the previous tables.

The screenshot shows the MySQL Workbench interface with two main panes. The left pane displays the database schema with several tables listed under the 'Tables' node. The right pane shows the results of a query on the 'load_Total_figure' table.

Query Results for load_Total_figure:

```
Showing rows 0 - 24 (267243 total, Query took 0.0012 seconds.)
SELECT * FROM `load_Total_figure`
```

	TotalfigureID	TotalVisits	TotalNights	TotalSpent
1	548	3834	337412	
2	4858	36957	2073617	
3	1564	4693	242457	
4	5327	36686	2333298	
5	746	11187	37291	
6	1226	2732	278602	
7	2425	19566	1020525	
8	1948	5845	837722	
9	6108	44183	9475629	
10	1695	4803	435872	
11	4559	35147	2054804	
12	703	12660	351680	
13	2263	5564	370786	
14	6690	52008	4747390	

Query Results for TourDataID:

	TourDataID	Mode	CountryName	Purpose	Package	Age	Gender	Duration
1	Air	Austria	Holiday	Independent	0-15	Male	4-13 nights	
2	Air	Austria	Holiday	Independent	0-15	Female	4-13 nights	
3	Air	Austria	Holiday	Independent	16-24	Male	1-3 nights	
4	Air	Austria	Holiday	Independent	16-24	Male	4-13 nights	
5	Air	Austria	Holiday	Independent	16-24	Male	14-27 nights	
6	Air	Austria	Holiday	Independent	25-34	Male	1-3 nights	
7	Air	Austria	Holiday	Independent	25-34	Male	4-13 nights	
8	Air	Austria	Holiday	Independent	25-34	Female	1-3 nights	
9	Air	Austria	Holiday	Independent	25-34	Female	4-13 nights	
10	Air	Austria	Holiday	Independent	35-44	Male	1-3 nights	
11	Air	Austria	Holiday	Independent	35-44	Male	4-13 nights	
12	Air	Austria	Holiday	Independent	35-44	Male	14-27 nights	
13	Air	Austria	Holiday	Independent	35-44	Female	1-3 nights	
14	Air	Austria	Holiday	Independent	35-44	Female	4-13 nights	
15	Air	Austria	Holiday	Independent	45-54	Male	1-3 nights	
16	Air	Austria	Holiday	Independent	45-54	Male	4-13 nights	
17	Air	Austria	Holiday	Independent	45-54	Male	14-27 nights	
18	Air	Austria	Holiday	Independent	45-54	Female	1-3 nights	
19	Air	Austria	Holiday	Independent	45-54	Female	4-13 nights	
20	Air	Austria	Holiday	Independent	55-64	Male	1-3 nights	
21	Air	Austria	Holiday	Independent	55-64	Male	4-13 nights	
22	Air	Austria	Holiday	Independent	55-64	Male	14-27 nights	
23	Air	Austria	Holiday	Independent	55-64	Female	1-3 nights	
24	Air	Austria	Holiday	Independent	65 & over	Male	1-3 nights	

(17) These are the SQL codes used for the loading phase.

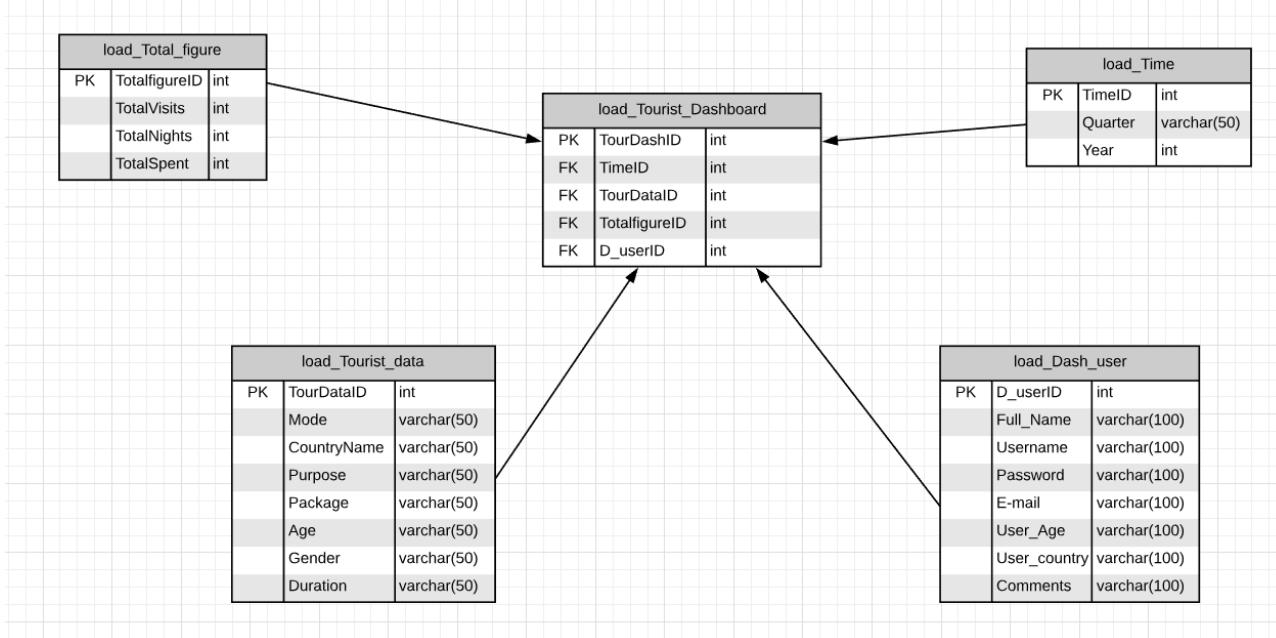
- insert into load_Tourist_data (load_Tourist_data.Mode, CountryName, Purpose, Package, Age, Gender, Duration) SELECT trans_Tourist_data.Mode, cwe_UKtouristdata.CountryName, trans_Tourist_data.Purpose, trans_Tourist_data.Package, trans_Tourist_data.Age, trans_Tourist_data.Gender, trans_Tourist_data.Duration from trans_Tourist_data, cwe_UKtouristdata where trans_Tourist_data.TourDataID = cwe_UKtouristdata.UKtourdataID;
- insert into load_Time (load_Time.Quarter, load_Time.Year) select trans_Time.Quarter, trans_Time.Year from trans_Time
- insert into load_Total_figure(TotalVisits, TotalNights, TotalSpent) select trans_No_of_visits.TotalVisits, trans_No_of_nights.TotalNights, trans_Total_Spent.TotalSpent from trans_No_of_visits, trans_No_of_nights, trans_Total_Spent where trans_No_of_visits.VisitsID = trans_No_of_nights.NightsID AND trans_No_of_nights.NightsID = trans_Total_Spent.SpentID and trans_Total_Spent.SpentID = trans_No_of_visits.VisitsID
- insert into load_Tourist_Dashboard
(TimeID, TourDataID, TotalfigureID, Year, Quarter, Mode, CountryName, Purpose, Package, Age, Gender, Duration, TotalVisits, TotalNights, TotalSpent) SELECT TimeID, TourDataID, TotalfigureID, Year, Quarter, Mode, CountryName, Purpose, Package, Age, Gender, Duration, TotalVisits, TotalNights, TotalSpent from load_Tourist_data, load_Total_figure, load_Time where load_Tourist_data.TourDataID = load_Total_figure.TotalfigureID AND load_Total_figure.TotalfigureID = load_Time.TimeID and load_Time.TimeID = load_Tourist_data.TourDataID;

- (18) This is the final loading table which contains all the data required to start the prototype. This is called the fact table and all sql queries should use this table to get the information they need. This Datawarehouse is now ready to be used for the prototypes. The ETL process is finished and the first prototype can be created.

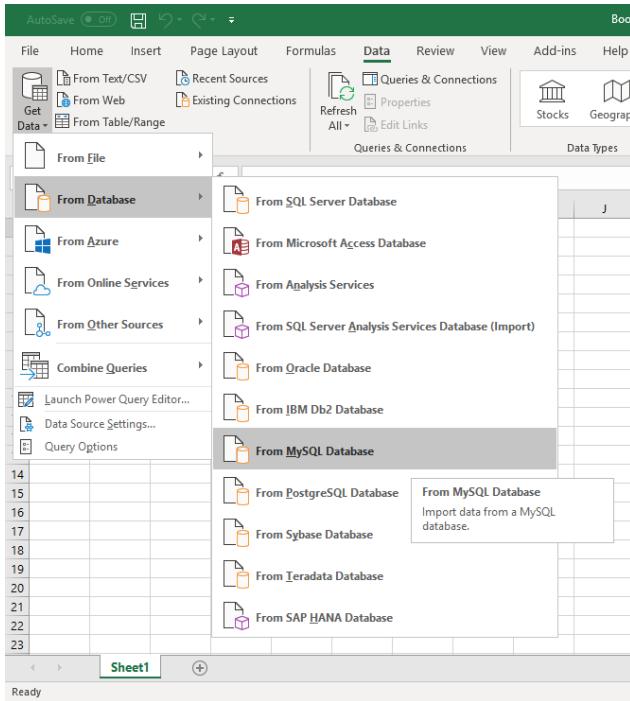
The screenshot shows the MySQL Workbench interface with the database 'mdb_kk5572a' selected. The left sidebar lists several tables, including 'load_Tourist_Dashboard', 'load_Tourist_data', 'load_Time', 'load_Total_figure', and 'load_Dash_user'. The main window displays the 'load_Total_figure' table. The table has columns: TourDashID, TimeID, TourDataID, TotalfigureID, D(userID), Year, Quarter, Mode, CountryName, Purpose, Package, Age, Gender, Duration, TotalVisits, TotalNights, and TotalSpent. The data shows 25 rows of tourist statistics from 2004, categorized by mode (Holiday, Independent), age (0-15, 16-24, 45-54, 55-64, 65+), gender (Male, Female), and duration (1-3 nights, 4-13 nights, 14-27 nights). The last row is a console entry.

TourDashID	TimeID	TourDataID	TotalfigureID	D(userID)	Year	Quarter	Mode	CountryName	Purpose	Package	Age	Gender	Duration	TotalVisits	TotalNights	TotalSpent
1	1	1	1	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	0-15	Male	4-13 nights	548	3834	337412
2	2	2	2	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	0-15	Female	4-13 nights	4858	36957	2073617
3	3	3	3	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	1-3 nights	1564	4693	242457
4	4	4	4	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	4-13 nights	5327	36686	2333298
5	5	5	5	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	14-27 nights	746	11187	37291
6	6	6	6	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Male	1-3 nights	1226	2732	278602
7	7	7	7	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Male	4-13 nights	2425	19566	1020525
8	8	8	8	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Female	1-3 nights	1948	5845	837722
9	9	9	9	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Female	4-13 nights	6108	44183	9475629
10	10	10	10	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	1-3 nights	1695	4803	435872
11	11	11	11	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	4-13 nights	4559	35147	2054804
12	12	12	12	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	14-27 nights	703	12660	351680
13	13	13	13	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Female	1-3 nights	2263	5564	370786
14	14	14	14	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Female	4-13 nights	6690	52008	4747390
15	15	15	15	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	1-3 nights	2671	6549	335918
16	16	16	16	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	4-13 nights	3956	26486	3042096
17	17	17	17	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	14-27 nights	1610	22545	1983955
18	18	18	18	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Female	1-3 nights	347	693	52937
19	19	19	19	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Female	4-13 nights	4735	30081	2547179
20	20	20	20	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Male	1-3 nights	281	561	100162
21	21	21	21	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Male	4-13 nights	1226	7740	625156
22	22	22	22	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Male	14-27 nights	1037	14523	1278041
23	23	23	23	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	55-64	Female	1-3 nights	1020	2040	179524
24	24	24	24	NULL	2004	Jan-Mar	Air	Austria	Holiday	Independent	65 & above	Male	1-3 nights	737	2211	194586

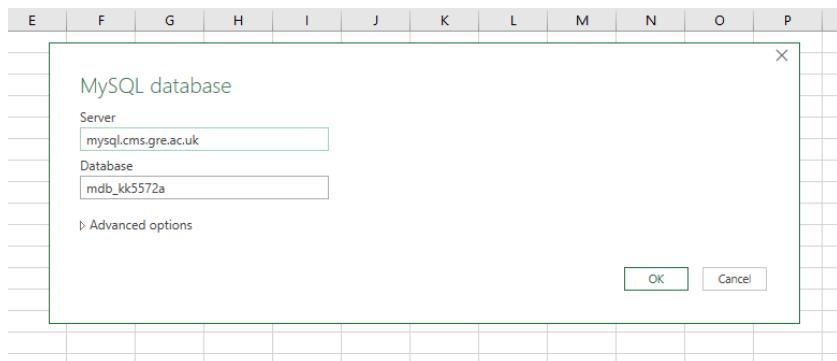
- (19) The ERD used for this Datawarehouse is different from the traditional ERD. Normally it should not contain all the data in one table. However, with limited time and getting errors from querying using the traditional ERD, the traditional design was not used. This is original ERD for the loading phase. The original ERD was modified to create a working loading tables and the modified ERD can be seen in the design section of this report.



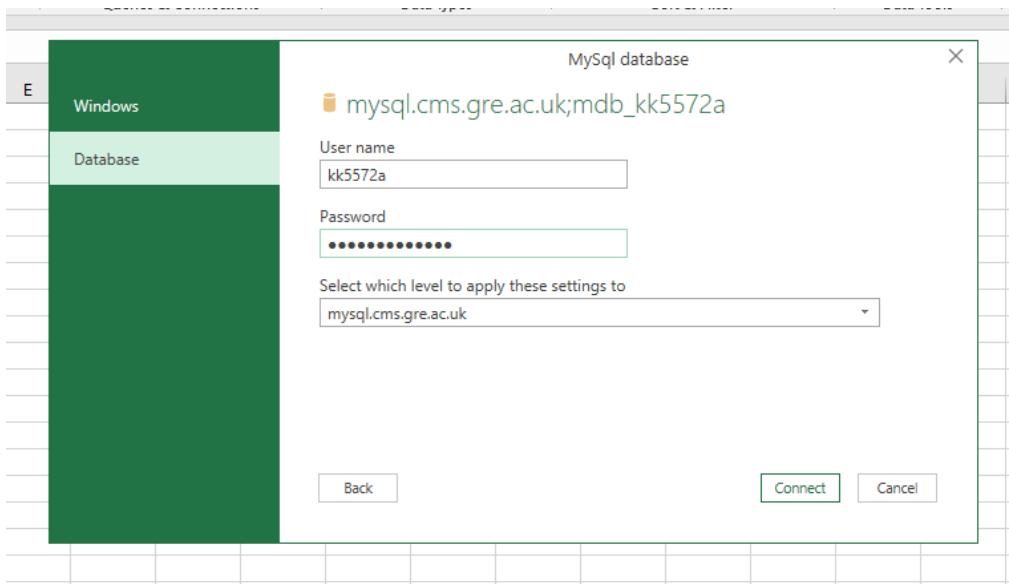
18 Appendix C (Implementation – Prototype 1 – Excel Dashboard)



(1) Datawarehouse is ready to be used and the next step is creating the first prototype in excel. First, the data from phpMyAdmin needs to be exported into excel. The appropriate database system must be chosen.



(2) Server name and database name must be entered to get access to the phpMyAdmin server.



(3) Enter username and password and click 'connect' to get access to phpMyAdmin account which contains all the sql tables.

- (4) Select the fact table from phpMyAdmin which contains all the crucial data. The table name is 'load_Tourist_Dashboard' and clicking the 'load' button will automatically import the data from the Data warehouse into the excel.

Navigator

- Select multiple items
- Display Options
- mysql.cms.gre.ac.uk: mdb_kk5572a [12]
 - mdb_kk5572a.cwe_UKtouristdata
 - mdb_kk5572a.load_Dash_user
 - mdb_kk5572a.load_Time
 - mdb_kk5572a.load_Total_figure
 - mdb_kk5572a.load_Tourist_Dashboard
 - mdb_kk5572a.load_Tourist_data
 - mdb_kk5572a.trans_Country
 - mdb_kk5572a.trans_No_of_nights
 - mdb_kk5572a.trans_No_of_visits
 - mdb_kk5572a.trans_Time
 - mdb_kk5572a.trans_Total_Spent
 - mdb_kk5572a.trans_Tourist_data

mdb_kk5572a.load_Tourist_Dashboard

TourDashID	TimeID	TourDataID	TotalfigureID	D(userID)	Year
1	1	1	1	null	201
2	2	2	2	null	201
3	3	3	3	null	201
4	4	4	4	null	201
5	5	5	5	null	201
6	6	6	6	null	201
7	7	7	7	null	201
8	8	8	8	null	201
9	9	9	9	null	201
10	10	10	10	null	201
11	11	11	11	null	201
12	12	12	12	null	201
13	13	13	13	null	201
14	14	14	14	null	201
15	15	15	15	null	201
16	16	16	16	null	201
17	17	17	17	null	201
18	18	18	18	null	201
19	19	19	19	null	201
20	20	20	20	null	201

The data in the preview has been truncated due to size limits.

Load Transform Data Cancel

- (5) In this image, you can see all the data from the Data warehouse which have been imported into excel spreadsheet.

The screenshot shows a Microsoft Excel spreadsheet with a large dataset. The columns are labeled A through I, and the rows are numbered 1 to 23. The data includes fields such as TourDashID, TimeID, TourDataID, TotalfigureID, D_userID, Year, Quarter, Mode, CountryName, and Purpose. To the right of the spreadsheet, the 'Queries & Connections' ribbon tab is selected, displaying a message about a query named 'mdb_kk5572a_load_Tourist_Dash...' which loaded 267,243 rows.

A	B	C	D	E	F	G	H	I	
TourDashID	TimeID	TourDataID	TotalfigureID	D_userID	Year	Quarter	Mode	CountryName	Purpose
1	1	1	1	1	2004	Jan-Mar	Air	Austria	Holiday
2	2	2	2	2	2004	Jan-Mar	Air	Austria	Holiday
3	3	3	3	3	2004	Jan-Mar	Air	Austria	Holiday
4	4	4	4	4	2004	Jan-Mar	Air	Austria	Holiday
5	5	5	5	5	2004	Jan-Mar	Air	Austria	Holiday
6	6	6	6	6	2004	Jan-Mar	Air	Austria	Holiday
7	7	7	7	7	2004	Jan-Mar	Air	Austria	Holiday
8	8	8	8	8	2004	Jan-Mar	Air	Austria	Holiday
9	9	9	9	9	2004	Jan-Mar	Air	Austria	Holiday
10	10	10	10	10	2004	Jan-Mar	Air	Austria	Holiday
11	11	11	11	11	2004	Jan-Mar	Air	Austria	Holiday
12	12	12	12	12	2004	Jan-Mar	Air	Austria	Holiday
13	13	13	13	13	2004	Jan-Mar	Air	Austria	Holiday
14	14	14	14	14	2004	Jan-Mar	Air	Austria	Holiday
15	15	15	15	15	2004	Jan-Mar	Air	Austria	Holiday
16	16	16	16	16	2004	Jan-Mar	Air	Austria	Holiday
17	17	17	17	17	2004	Jan-Mar	Air	Austria	Holiday
18	18	18	18	18	2004	Jan-Mar	Air	Austria	Holiday
19	19	19	19	19	2004	Jan-Mar	Air	Austria	Holiday
20	20	20	20	20	2004	Jan-Mar	Air	Austria	Holiday
21	21	21	21	21	2004	Jan-Mar	Air	Austria	Holiday
22	22	22	22	22	2004	Jan-Mar	Air	Austria	Holiday
23									

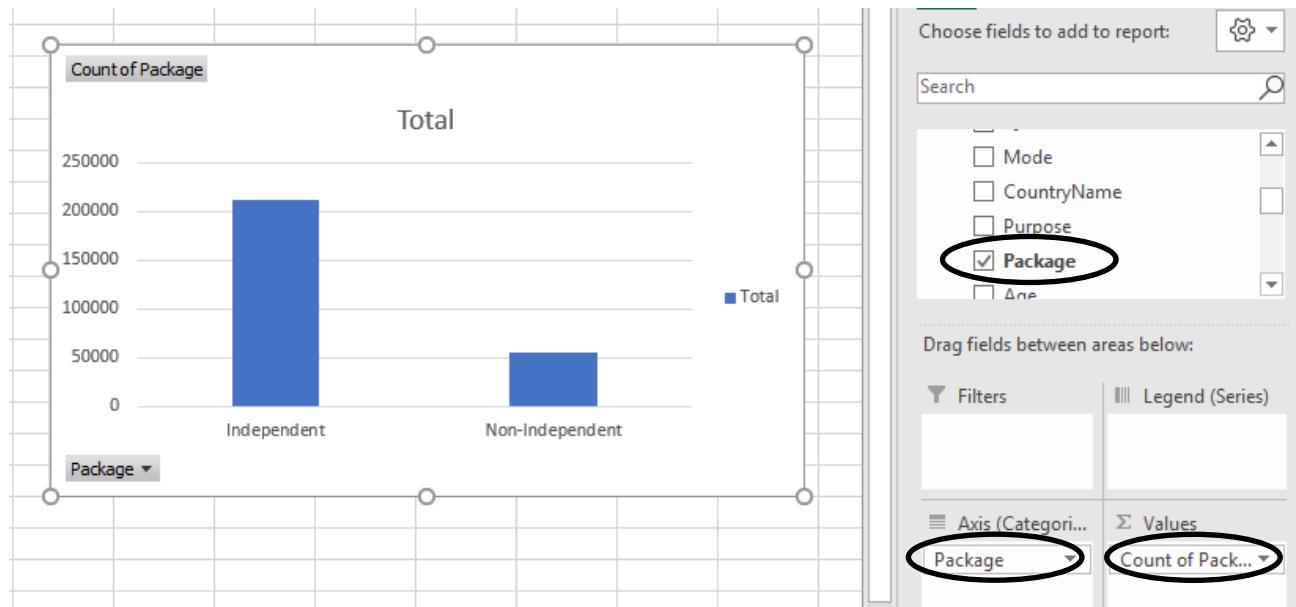
- (6) Now, it is time to create the pivot charts using the data from the excel spreadsheet. Each pivot charts should be created in separate spreadsheets and it will make it easier to create the dashboard later. The Pivot chart icon must be selected, and then appropriate dataset should be chosen. The dataset in the first spreadsheet was selected.

The screenshot shows two dialog boxes. The top one is 'Create PivotChart' with the following settings: 'Select a table or range' is set to 'Use this workbook's Data Model', 'New Worksheet' is selected for placement, and the location is 'Sheet5!\$A\$1'. The bottom one is 'Existing Connections' showing 'This Workbook Data Model' and a table named 'mdb_kk5572a_load_Tourist_Dashboard' with the range 'Sheet2!\$A\$1:\$X\$267244'. Both dialog boxes have a blue border around the table selection area.

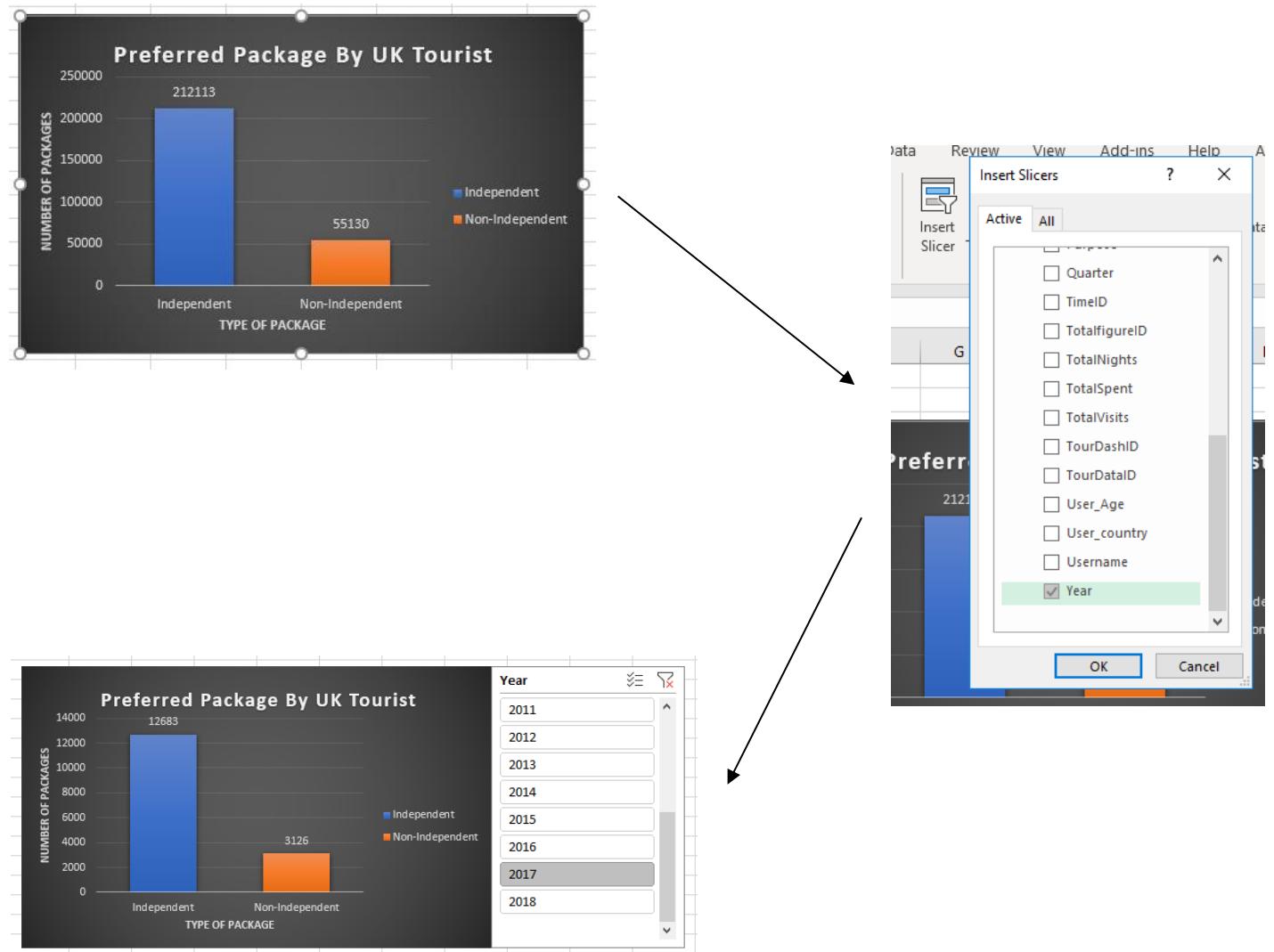
Create PivotChart
Choose the data that you want to analyze
 Select a table or range
Table/Range:
 Use an external data source
Choose Connection...
Connection name:
 Use this workbook's Data Model
Choose where you want the PivotChart to be placed
 New Worksheet
 Existing Worksheet
Location: Sheet5!\$A\$1
Choose whether you want to analyze multiple tables
 Add this data to the Data Model
OK Cancel

Existing Connections
Select a Connection or Table
Connections Tables
Show: All Tables
This Workbook Data Model
Tables in Workbook Data Model
1 Table(s)
excel prototype.xlsx (This Workbook)
mdb_kk5572a_load_Tourist_Dashboard
Sheet2!\$A\$1:\$X\$267244
Open Cancel

- (7) After the dataset has been selected, it is time to select the necessary field to create the graphs. Depending on the type of data chosen, different graphs can be produced, so it is important to make sure the correct data fields are selected when trying to create a valuable graph or it could produce an ineffective graph with no insight.



- (8) The same process will be used to create all the other graphs and charts for this prototype. However, flexibility is crucial when it comes to data used in graphs as users may want to try different data to get different insights from the graphs. That is why 'slicer' is an important feature for the graphs in excel. It can change a static graph into a dynamic graph and improving its effectiveness. This graph below has a 'year' slicer which lets a user to get information for each year or they can compare two or more different year. This graph contains information about the preferred package by UK tourists and it can aid in understanding if they like planning their own holiday or get it from a travel agent.

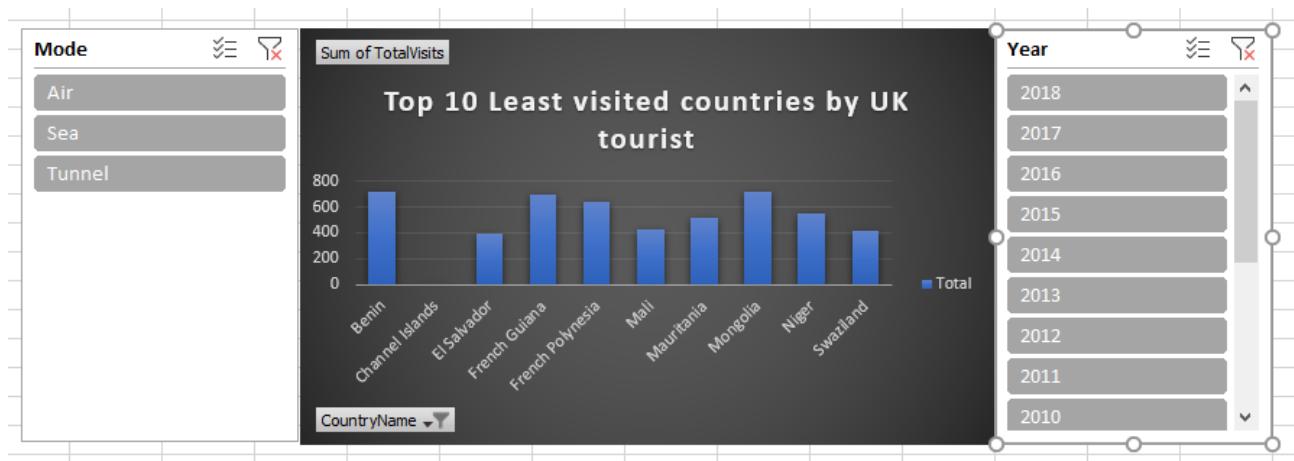


(9) These are the other graphs and charts which were created for the prototype.

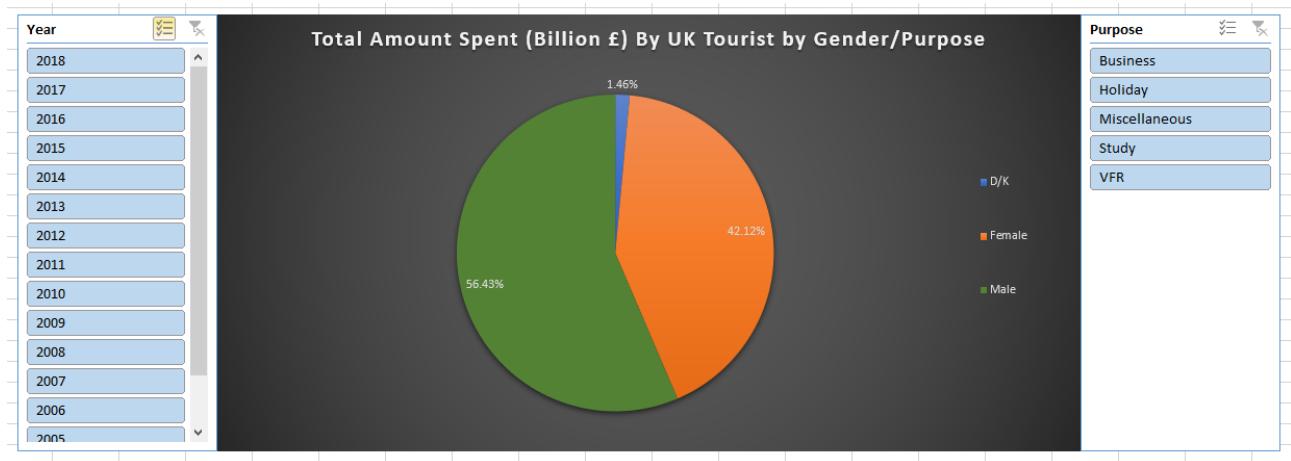
This map shows the top 25 countries with the highest amount spent by UK tourists. It can reveal the countries which UK tourist prefer the most and the UK tourism industry can capitalise on this information by tailoring their packages to their UK customers.



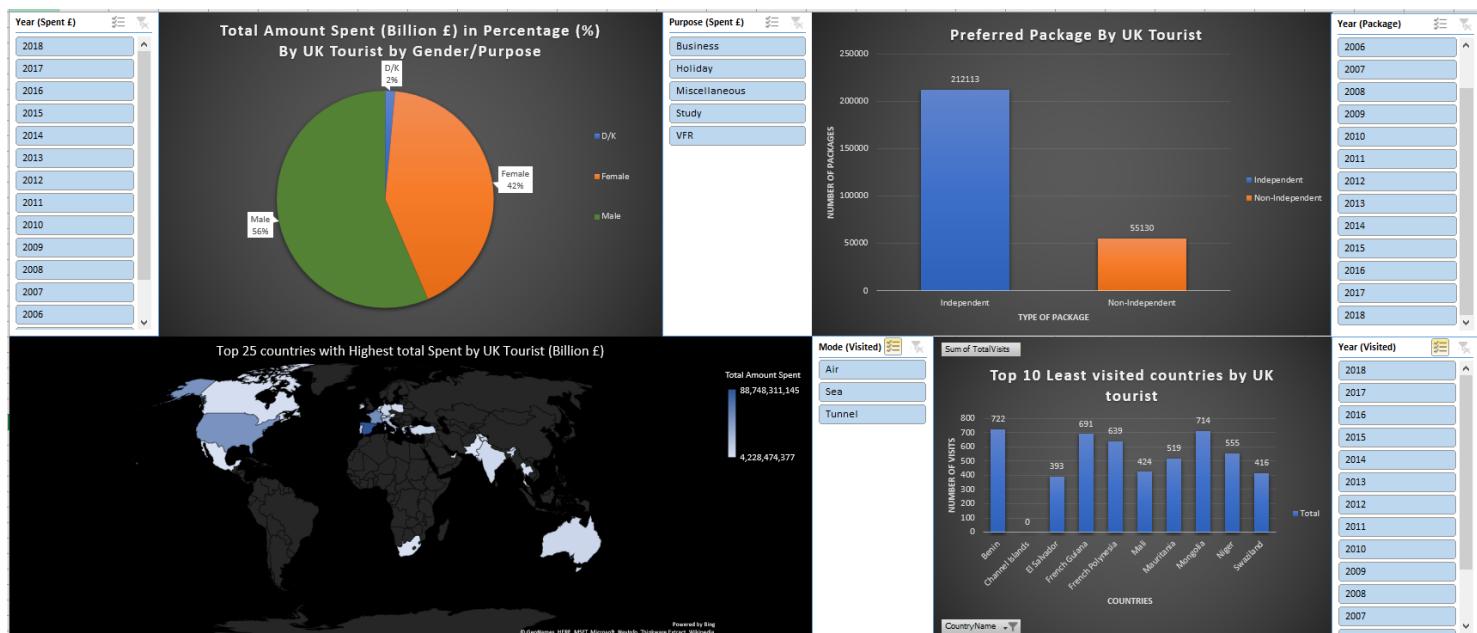
This graph shows the top 10 least visited countries by UK tourists. It can provide a small insight for the tourism industries around the world by showing which countries are least visited/interested by the UK tourists. The graph can further be modified by using the two slicers in both side (Mode, Year). The 'mode' slicer can show the least visited by mode of transportation and the 'year' slicer can be used to compare data from different years.



This graph shows the total amount spent in billions by UK tourist by gender/purpose. The pie chart shows the total percent of amount spent by UK tourist and it can help the tourism industry in understanding each gender's preference when it comes to the purpose of travel. Each purpose of travel can be viewed by selecting different options in the 'purpose' slicer. Data from different years can also be viewed using the 'year' slicer.



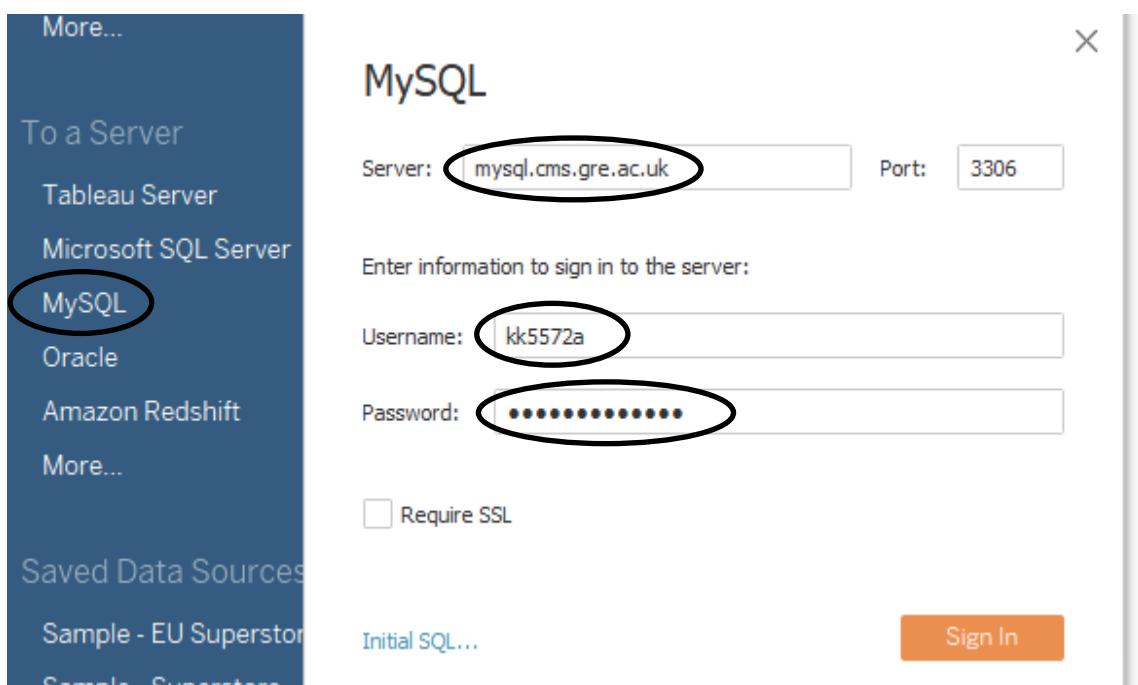
(10) After all the graphs, maps and charts are created in each spreadsheet, time has come to create the dashboard which will be the first prototype. All the pivot charts and slicers are combined in a separate spreadsheet. This is the first prototype which is a dashboard created in excel spreadsheet. This dashboard interface allows the users to view all the graphs and map in one place and can get useful insight on UK tourist.



There are many features and graphs which can be added, but with limited time frame, this cannot be done for now. This excel dashboard prototype is a proof of concept and it confirmed that the graphs and the dashboard can be created using the data from the Data warehouse. Any ideas gotten from this will be used in the next prototype.

19 Appendix D (Implementation – Prototype 2 –Tableau Dashboard)

- (1) After creating the first prototype and analysing the excel dashboard, the next prototype is ready to be created which will be an improved version of the excel dashboard. Tableau software will be used to create the second prototype and it contains many similar functions and features from Microsoft excel. The first step is to connect the database system (phpMyAdmin MySQL) with tableau, so that the data can be accessed and used for the creation of graphs, charts and maps. Necessary server and user information must be entered in the required field to connect with phpMyAdmin MySQL server.

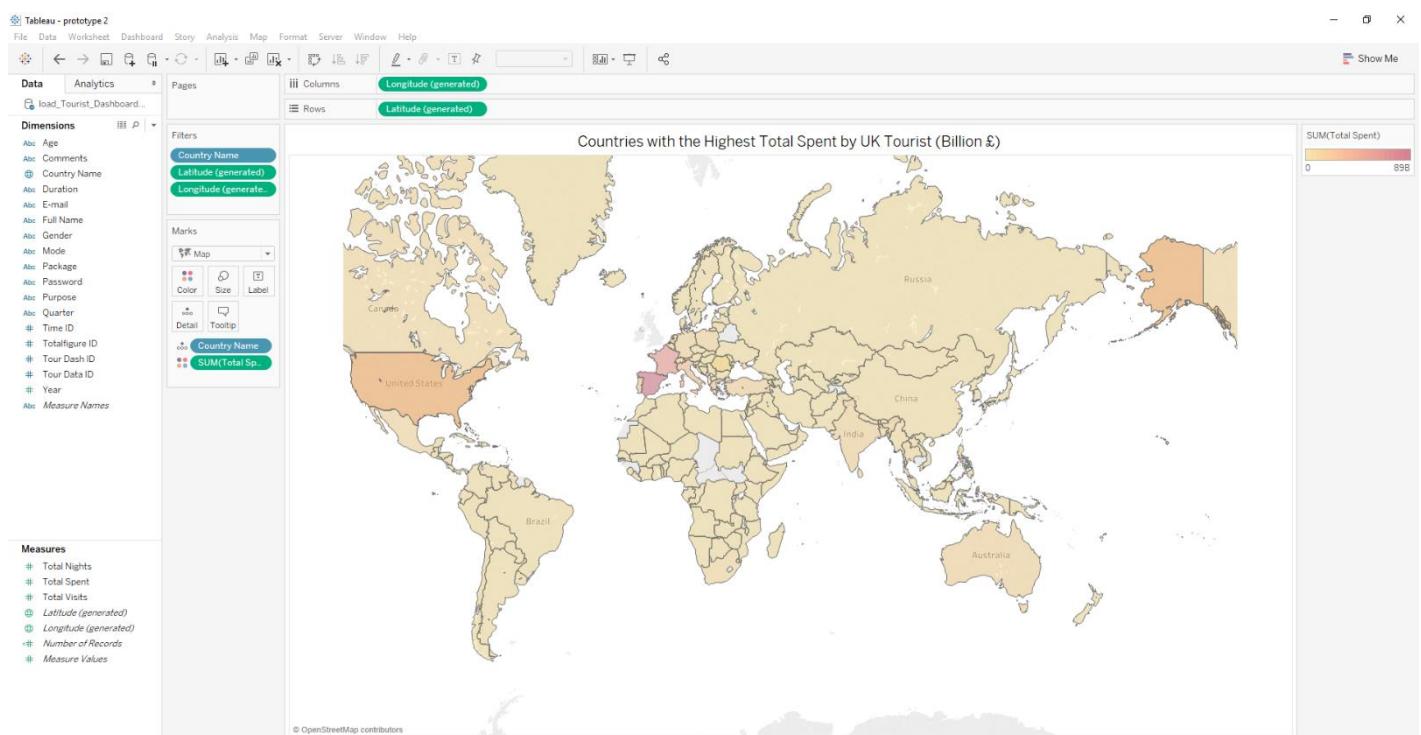


(2) After connecting and selecting the correct database name, appropriate tables must be selected. Only one table from the Data warehouse is required for creating the dashboard. ‘load_Tourist_Dashboard’ is selected and ready to be used in tableau.

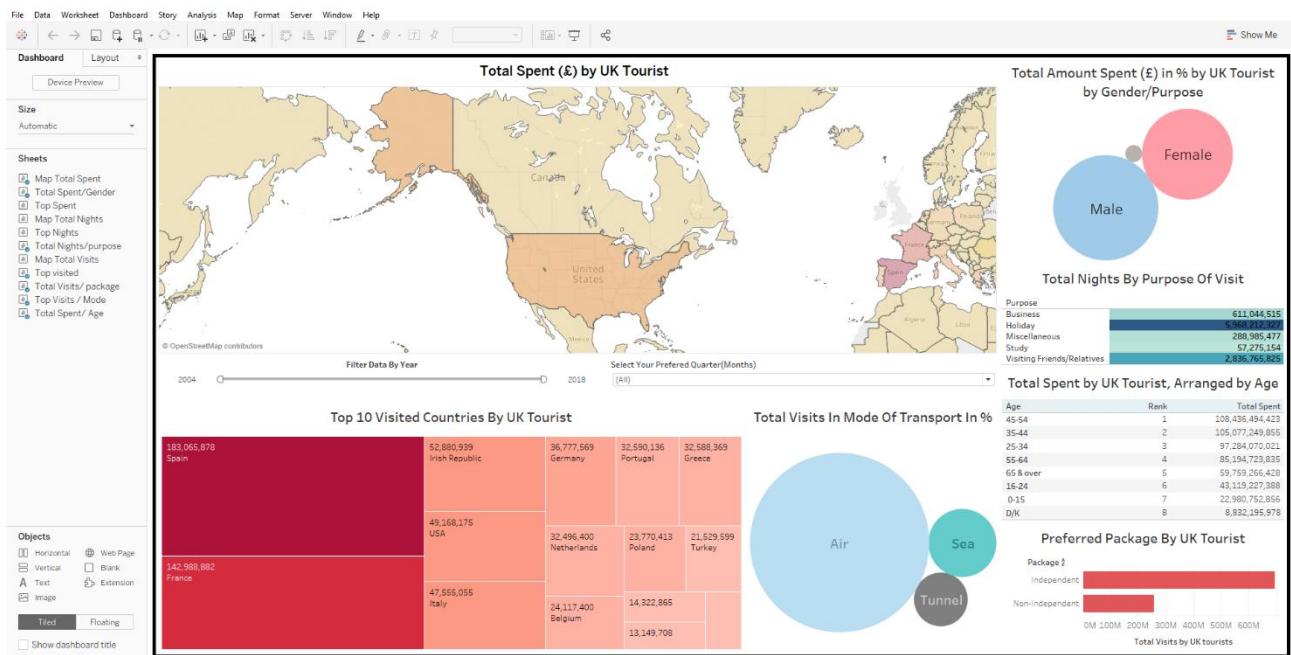
The screenshot shows the Tableau Data Source view. On the left, there are sections for Connections, Database, and Table. Under Table, the 'load_Tourist_Dashboard' table is selected and highlighted with a red oval. The main area displays the schema of the 'load_Tourist_Dashboard' table, which includes columns such as Tour Dash ID, Time ID, Tour Data ID, Totalfigure ID, Year, Quarter, Mode, Country Name, Purpose, Package, Age, Gender, Duration, Total Visits, and Total Nights. The data preview shows 15 rows of sample data.

Tour Dash ID	Time ID	Tour Data ID	Totalfigure ID	Year	Quarter	Mode	Country Name	Purpose	Package	Age	Gender	Duration	Total Visits	Total Nights
1	1	1	1	2004	Jan-Mar	Air	Austria	Holiday	Independent	0-15	Male	4-13 nights	548	3.8
2	2	2	2	2004	Jan-Mar	Air	Austria	Holiday	Independent	0-15	Female	4-13 nights	4,858	36.9
3	3	3	3	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	1-3 nights	1,564	4.6
4	4	4	4	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	4-13 nights	5,327	36.6
5	5	5	5	2004	Jan-Mar	Air	Austria	Holiday	Independent	16-24	Male	14-27 nights	746	11.1
6	6	6	6	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Male	1-3 nights	1,226	2.7
7	7	7	7	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Male	4-13 nights	2,425	19.5
8	8	8	8	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Female	1-3 nights	1,948	5.8
9	9	9	9	2004	Jan-Mar	Air	Austria	Holiday	Independent	25-34	Female	4-13 nights	6,108	44.1
10	10	10	10	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	1-3 nights	1,695	4.8
11	11	11	11	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	4-13 nights	4,559	35.1
12	12	12	12	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Male	14-27 nights	703	12.6
13	13	13	13	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Female	1-3 nights	2,263	5.5
14	14	14	14	2004	Jan-Mar	Air	Austria	Holiday	Independent	35-44	Female	4-13 nights	6,690	52.0
15	15	15	15	2004	Jan-Mar	Air	Austria	Holiday	Independent	45-54	Male	1-3 nights	2,671	6.5

(3) Combining different attributes can produce different graphs and chart. Here you can see that combining the ‘country name’ with the ‘total amount spent’ can create a map which illustrates which country has most amount spent by UK tourists. Darker colour indicates that the country has many UK tourists who spend their money. The same process will be followed to create all the graphs, charts and maps in tableau.

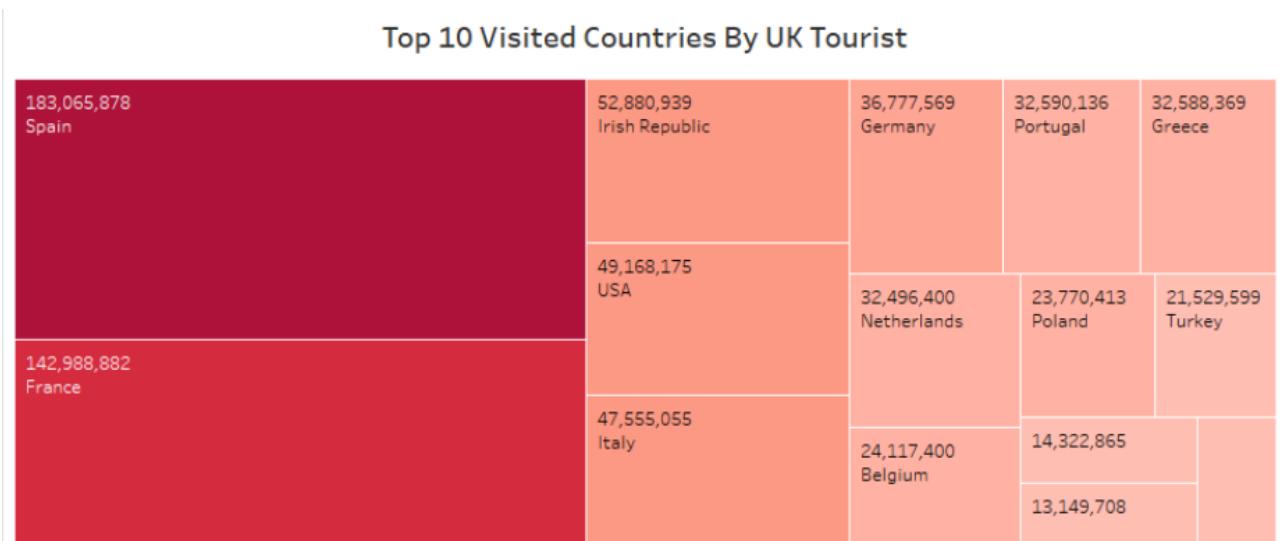


(4) After producing each individual graphs and maps, the dashboard can be created. Specific sheets (graphs) in tableau will be chosen and combined to create a dashboard. This is how it will look like after adding all the sheets together. Each individual sheet is interactive, and they can be clicked to get further information which is tailored to the option chosen by the user. For example: Clicking 'Male' will transform all the graphs and maps in the dashboard to male and will only show information which are related to male. This can be done to any other values in the dashboard and will tailor the graphs automatically.

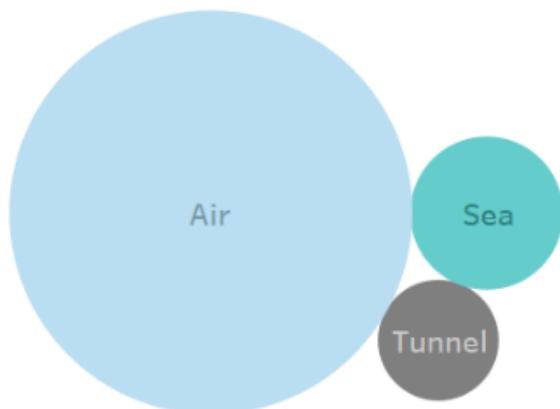


(5) These are all the sheets which are being used in the tableau dashboard prototype

This sheet shows the top 10 countries which are visited by UK tourists. It's like ranking and can give an insight by showing the number of people who are visiting the country. Each country can be clicked to view more information about that country.

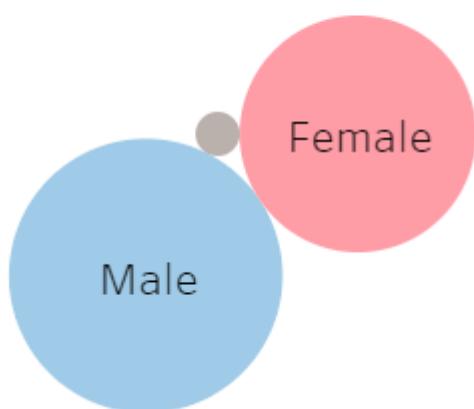


Total Visits In Mode Of Transport In %



This graph will show the total visit in % by UK tourists around the world and it is arranged in mode of transport. The value can be seen when the mouse is hovered over the circles. The tourism industry can use this information to find out the preferred mode of transportation of UK tourists and can tailor the package depending on the information from the dashboard.

Total Amount Spent (£) in % by UK Tourist by Gender



This graph shows the total amount spent in % by UK tourists and it is arranged by gender. The tourism industry can tailor their service depending on the gender. Selecting 'male' or 'female' or 'I don't know' can transform all the other graphs to only show data which are related to the chosen option.

Total Nights By Purpose Of Visit

Purpose	Nights
Business	611,044,515
Holiday	5,968,212,327
Miscellaneous	288,985,477
Study	57,275,154
Visiting Friends/Relatives	2,836,765,825

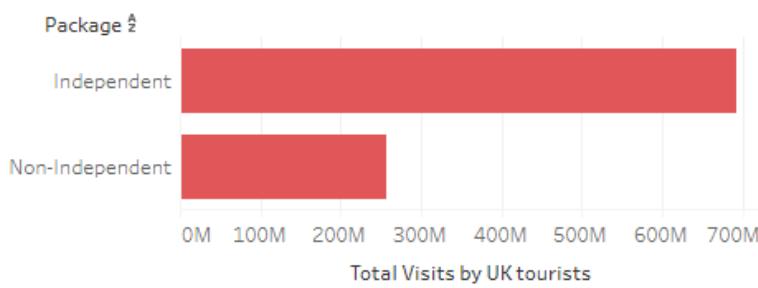
This table shows the total nights spent by UK tourist around the world and it is arranged by purpose of visit. This table will help the purpose of their travel and the tourism industry can tailor their package depending on the insight they gather using this table. The hotels can also have different type of rooms depending on the customer type.

Total Spent by UK Tourist, Arranged by Age

Age	Rank	Total Spent
45-54	1	108,436,494,423
35-44	2	105,077,249,855
25-34	3	97,284,070,021
55-64	4	85,194,723,835
65 & over	5	59,759,266,428
16-24	6	43,119,227,388
0-15	7	22,980,752,856
D/K	8	8,832,195,978

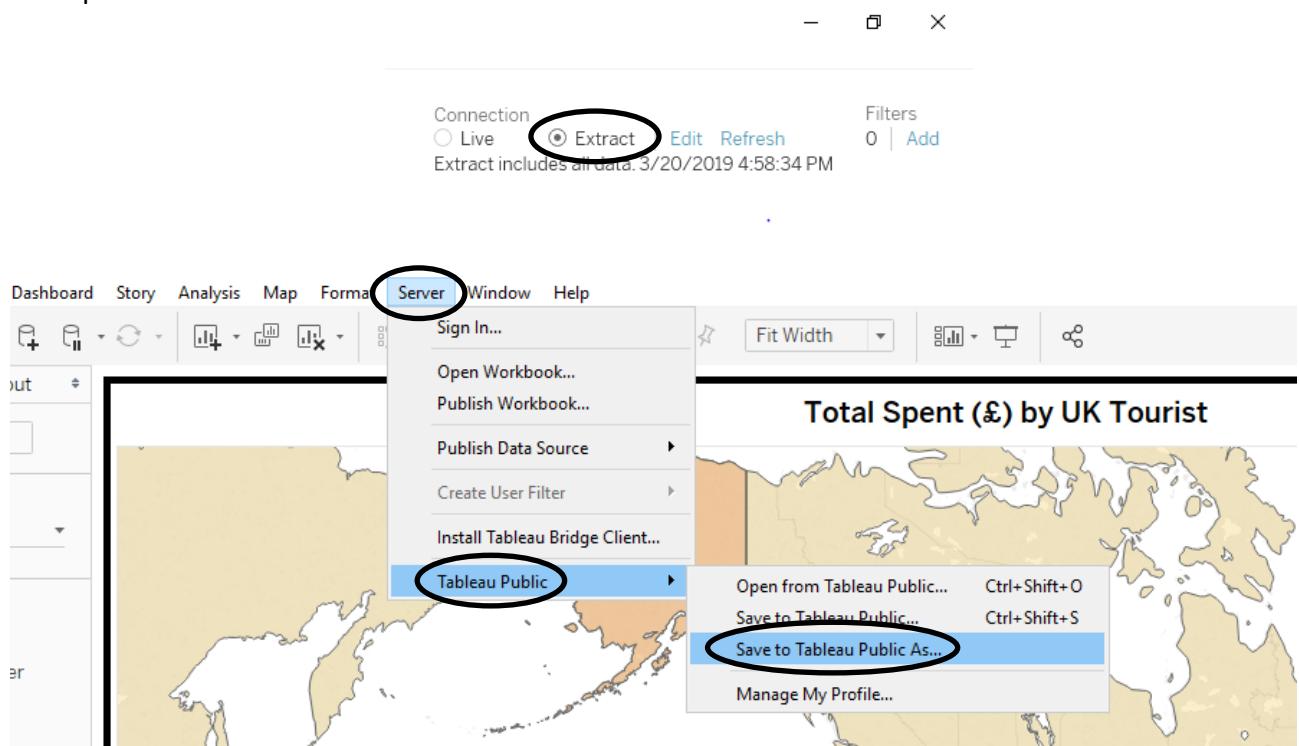
This table shows the total amount spent by UK tourist and it is arranged by age. It can help identify the age group of the travellers who are coming to the country. By looking at total spent, you can see which age group spends the most and which group spends the least. The tourism industry can use different tactics to target this age groups.

Preferred Package By UK Tourist

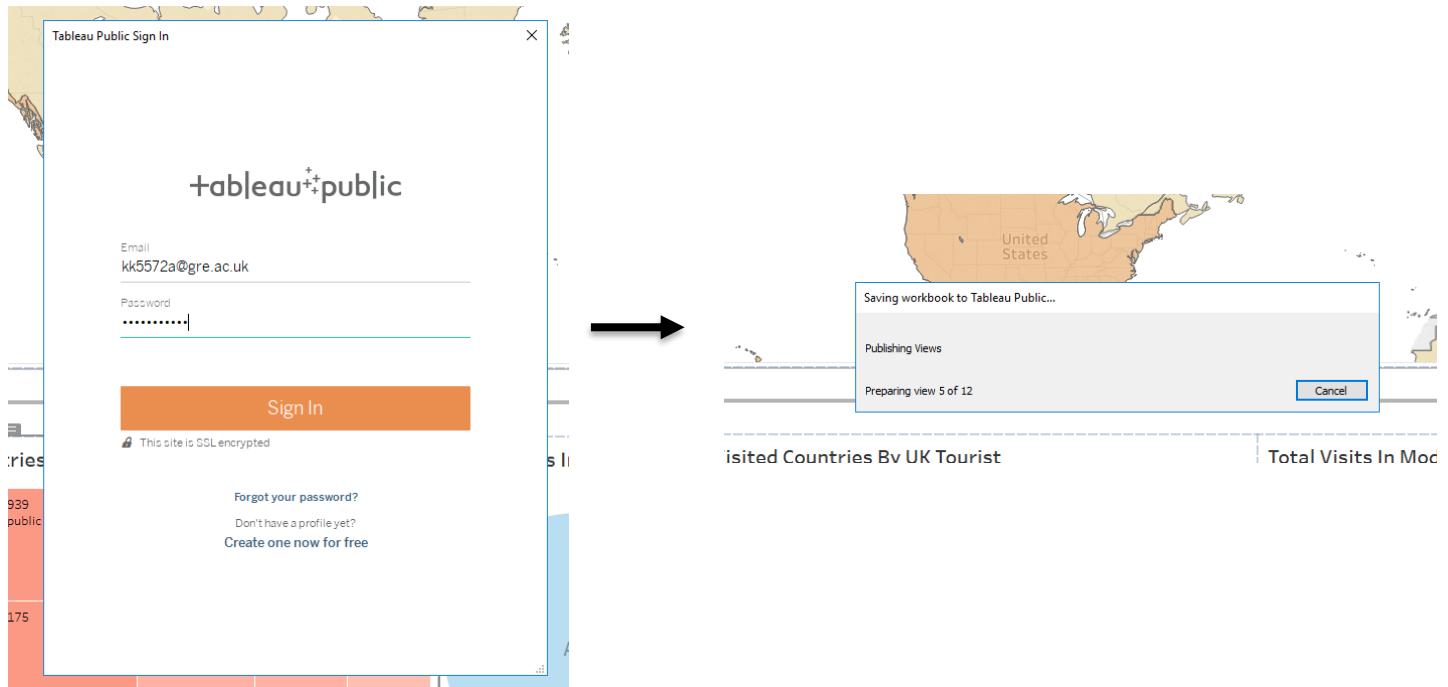


This graph shows the preferred package by UK tourist. Independent means that they would plan their own holiday using the internet. The non-independent means that they will use travel agency to plan their travel abroad. With this information, the tourism industry can understand which country the UK tourists go using which package and tailor the service according to the insight.

- (6) Now it is time to start the publishing the tableau online. First, the dataset used in the tableau must be extracted and saved in the local drive. Then select ‘server’ → ‘tableau public’ → ‘Save to tableau public as’. This will allow the tableau to be saved in tableau public online server.

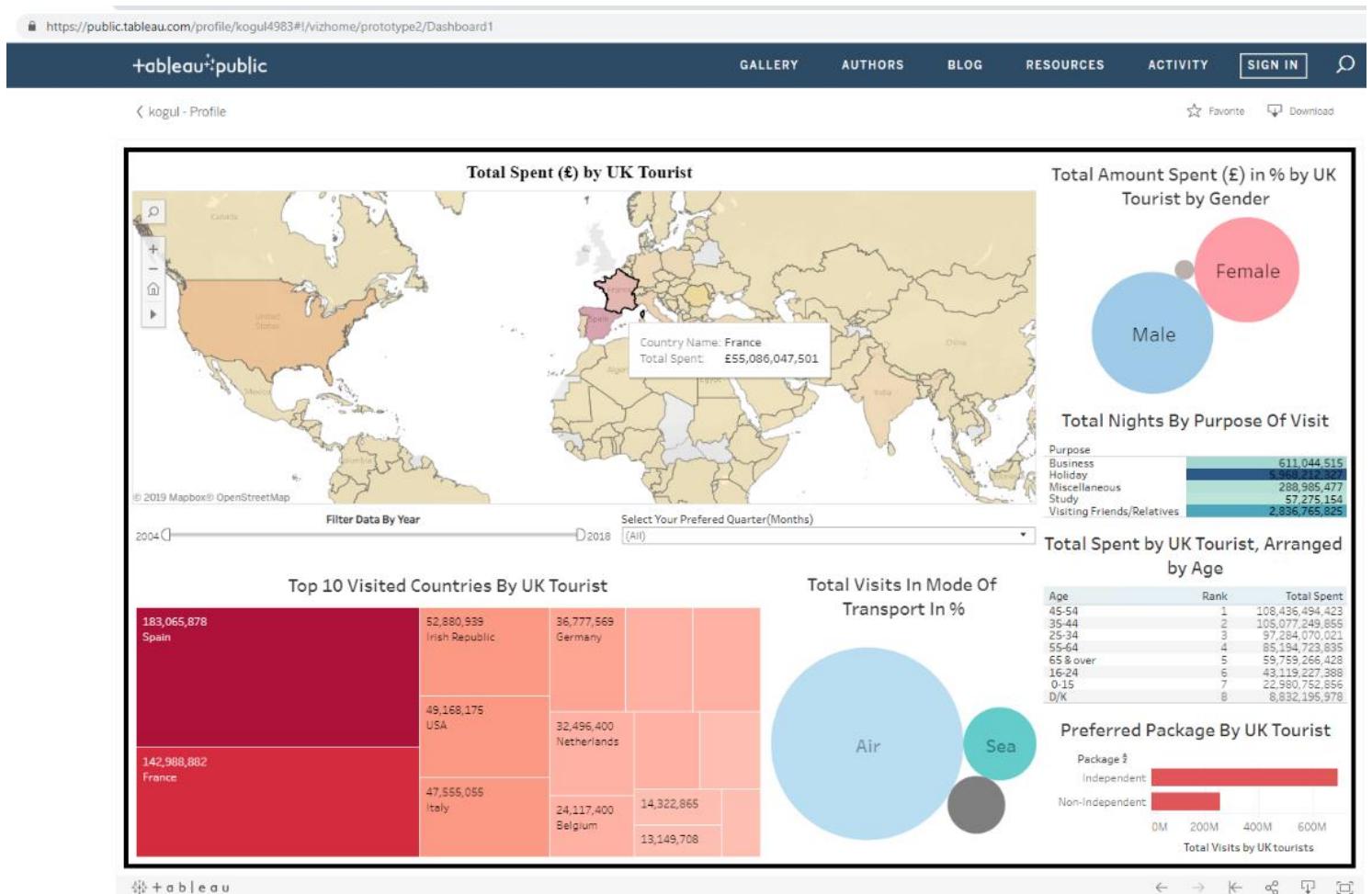


- (7) To publish the tableau dashboard, one must have an account in tableau public which is free to create. After entering the email and password into the fill, click sign in to start publishing the dashboard.



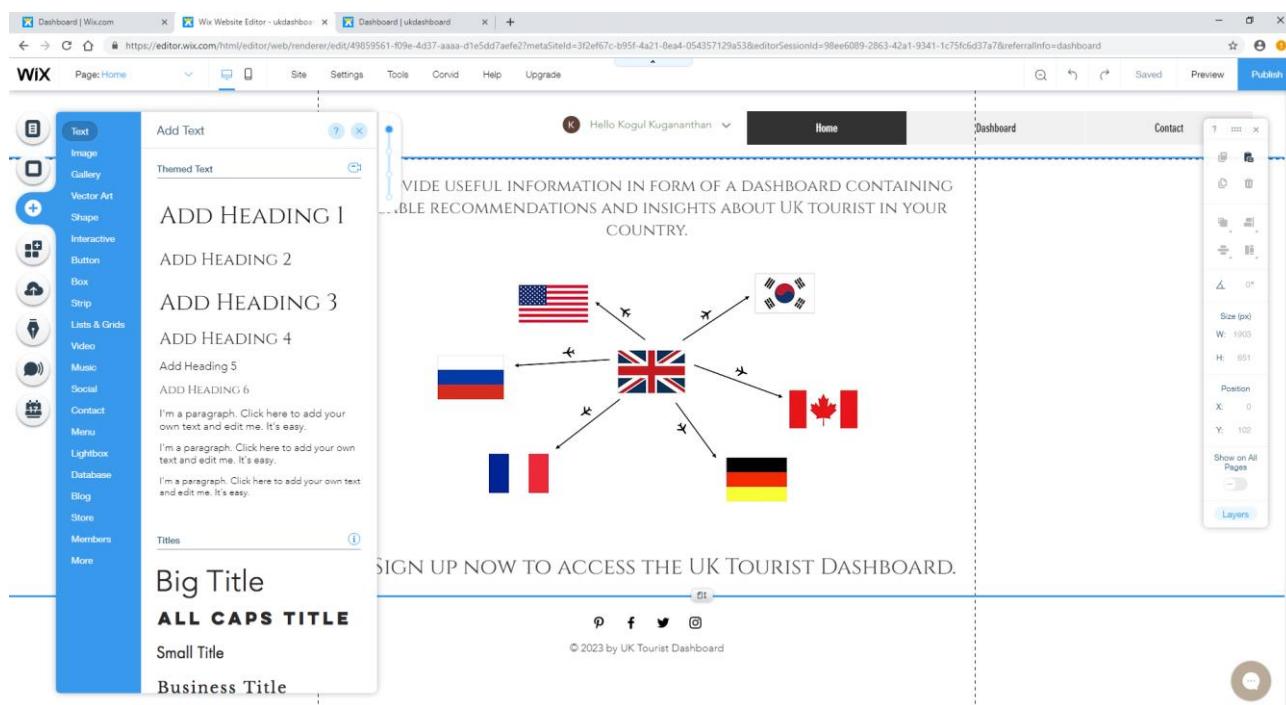
(8) This is the 2nd prototype for UK tourist Dashboard. You can see here that it is publish online and anyone with this link can access the dashboard. You can also go to the tableau public website and type 'kogul' to view this dashboard. All the graphs, charts and maps are interactive, and you can also share this dashboard with anyone you know who is interested in tourism industry. There are many features for tableau (such as, share dashboard, download dashboard, add to favourite, full screen, etc.). To get more users to view this dashboard, the website can be advertised to get more traffic to this dashboard in tableau public.

This prototype has enough features to be the final prototype. However, the dashboard can be taken further by allowing the dashboard to be publish online in a standalone website which can be advertised in google and other search engine easily and can also be used to earn revenue using google AdSense if needed.



20 Appendix E (Implementation – Final product – Wix Website containing Dashboard)

- (1) The current prototype is a tableau dashboard and now the 2nd prototype will be incorporated into a website using a third party online website builder (WIX). One must create an account to use the functions and features in Wix. After creating the account, the website can be created using the template provided. Here you can see the website being created using the tools provided in Wix. The image in the middle was created to represent the UK Tourist Dashboard.



- (2) This is the embed code from tableau dashboard and it will automatically create this code when the dashboard is published to tableau public server. This code can be used to show the dashboard in other platforms.

```
<div class='tableauPlaceholder' id='viz1553101587687' style='position: relative'><noscript><a href='><img alt='src='https://public.tableau.com/static/images/pr/prototype2/Dashboard1/1_rss.png' style='border: none' /></a></noscript><object class='tableauViz' style='display:none;'><param name='host_url' value='https%3A%2F%2Fpublic.tableau.com%2F' /><param name='embed_code_version' value='3' /><param name='site_root' value='/'><param name='name' value='prototype2/Dashboard1' /><param name='tabs' value='no' /><param name='toolbar' value='yes' /><param name='static_image' value='https://public.tableau.com/static/images/pr/prototype2/Prototype2/1.png' /><param name='animate_transition' value='yes' /><param name='display_static_image' value='yes' /><param name='display_spinner' value='yes' /><param name='display_overlay' value='yes' /><param name='display_count' value='yes' /><param name='filter' value='publish=yes' /></object></div> <script type='text/javascript'>var divElement = document.getElementById('viz1553101587687'); var vizElement = divElement.getElementsByTagName('object')[0]; vizElement.style.width='100%';vizElement.style.height=(divElement.offsetWidth*0.75)+'px'; var scriptElement = document.createElement('script'); scriptElement.src = 'https://public.tableau.com/javascripts/api/viz_v1.js'; vizElement.parentNode.insertBefore(scriptElement, vizElement); </script>
```

- (3) The dashboard can be added using the embed code from tableau dashboard. You can see here the code is being added and the tableau dashboard is being displayed on the website.

This is a summarised version of the UK tourist dashboard and contains the most important information which you will need to gain an insight into the UK tourist who are travelling to your country.

You can interact with each of the graphs below by clicking on them and it will change all of the graphs in the dashboard to suit your selected option. All the graphs can be filtered by year and.

Total Spent (£) by UK Tourist

Total Amount Spent (£) in % by Tourist by Gender

Total Nights By Purpose Of Visit

Purpose	Total Nights
Business	611,044,515
Holiday	5,966,212,327
Miscellaneous	288,995,477
Study	57,275,154
Visiting Friends/Relatives	2,836,765,825

Total Spent by UK Tourist, Arranged by Age

Age	Rank	Total Spent
45-54	1	108,436,494,423
35-44	2	105,077,249,855

Top 10 Visited Countries By UK Tourist

Country	Visits
United Kingdom	183,065,678
United States	52,880,939
Canada	36,777,569

Total Visits In Mode Of Transport In %

The map contains information about how much UK tourist have spent in each countries. You can hover your mouse over the country to view the value or click to change all graphs to only show that country's information

Choose your preferred year and quarter here

Filter Data By Year: 2004 - 2018
Select Your Preferred Quarter(Months): (All)

Information about UK tourist spending habits by purpose of visit can help understand where people spend most of their time when visiting a country.

- (4) This is the homepage and the first place the user will see when they enter the website. The user can login, log out and view account using the function in the top. The website has few navigation tabs and the dashboard tab contains the tableau dashboard. The user can also chat with the admin of the website if they have any question which they would like to ask about the dashboard. The user can also share this website using the social media logo below. Currently, for the prototype purpose, anyone can use the dashboard because it will be easier for the people to test the website without using an account. However, in future development the feature can be added.

WE PROVIDE USEFUL INFORMATION IN VALUABLE RECOMMENDATIONS AND I DASHBOARD CONTAINING ABOUT UK TOURIST IN YOUR COUNTRY.

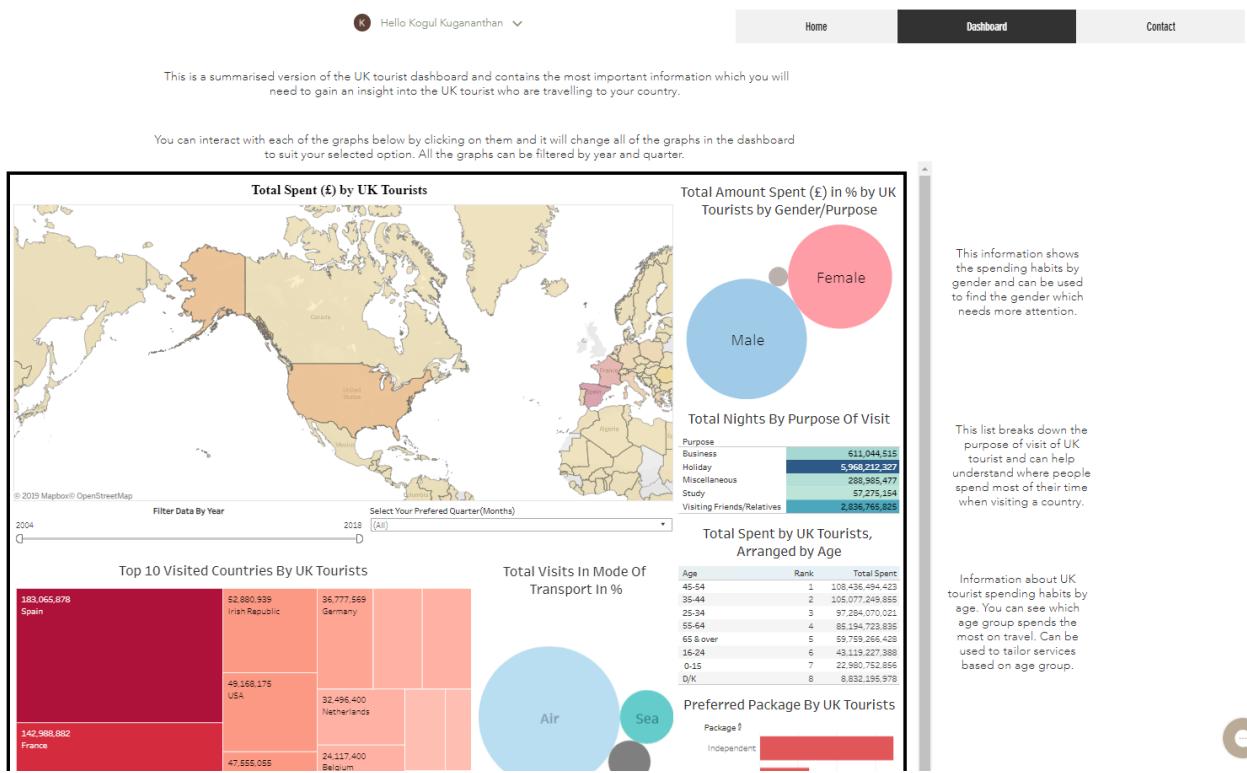
SIGN UP NOW TO ACCESS THE UK TOURIST DASHBOARD.

Let's Chat!

Hi there. I'm available to answer your questions.

- (5) Here you can see the dashboard in the website and each section is been explained in few words. This can help the user understand the purpose of the graphs, charts and maps in the dashboard.

UK TOURIST DASHBOARD



- (6) This is the contact us page which the user can use to contact the website if they have any question with the dashboard.

UK TOURIST DASHBOARD

The page has a header with the title "GET IN TOUCH" and a message: "If you have any questions about the dashboard or any other queries which you may have, please feel free to message us using the form below."

Below the message, there is a contact form with fields for Name, Email, Subject, and Message, along with a "Submit" button.

Footer information includes: Tel: 123-456-7890 | Email: info@mysite.com

21 Appendix F (Testing)

SQL Code Testing - Data warehouse																																																																																				
Test No.	SQL Code	Query	Actual result	Pass /Fail																																																																																
1	<pre>select CountryName, SUM(TotalSpent) as TotalSpent, SUM(TotalVisits) as TotalVisited, SUM(TotalNights) as TotalStayed from load_Tourist_Dashboard where year = '2017' group by CountryName ORDER BY `TotalSpent` DESC</pre>	which country has the most amount spent(£) by UK tourist for 2017 and how many times have they visited and stayed in that country?	<p>Showing rows 0 - 24 (62 total, Query took 0.0841 seconds.)</p> <pre>select CountryName, SUM(TotalSpent) as TotalSpent, SUM(TotalVisits) as TotalVisited CountryName ORDER BY `TotalSpent` DESC</pre> <table border="1"> <thead> <tr> <th>CountryName</th> <th>TotalSpent</th> <th>TotalVisited</th> <th>TotalStayed</th> </tr> </thead> <tbody> <tr><td>Spain</td><td>8683030523</td><td>15871872</td><td>145869761</td></tr> <tr><td>USA</td><td>4279222407</td><td>3401107</td><td>44182150</td></tr> <tr><td>France</td><td>3984141502</td><td>8861644</td><td>66283281</td></tr> <tr><td>Italy</td><td>2378037814</td><td>4159924</td><td>32112217</td></tr> <tr><td>Greece</td><td>1664050506</td><td>2382735</td><td>23255850</td></tr> <tr><td>Portugal</td><td>1552384904</td><td>2875598</td><td>26586459</td></tr> <tr><td>United Arab Emirates</td><td>1105440694</td><td>1015885</td><td>10910260</td></tr> <tr><td>Germany</td><td>1083709889</td><td>2909338</td><td>16484508</td></tr> <tr><td>Other Asia</td><td>1055744633</td><td>931998</td><td>22234924</td></tr> <tr><td>Irish Republic</td><td>993229378</td><td>3422119</td><td>18838173</td></tr> <tr><td>Netherlands</td><td>973560404</td><td>2659833</td><td>10767558</td></tr> <tr><td>Poland</td><td>950130504</td><td>2672368</td><td>26205224</td></tr> <tr><td>Other countries</td><td>901243881</td><td>685476</td><td>8282654</td></tr> <tr><td>Other Africa</td><td>877823654</td><td>947115</td><td>20824351</td></tr> <tr><td>India</td><td>820906396</td><td>1006423</td><td>24911348</td></tr> <tr><td>Australia</td><td>819939391</td><td>452368</td><td>14223339</td></tr> <tr><td>Turkey</td><td>712476696</td><td>1228144</td><td>14710793</td></tr> <tr><td>Canada</td><td>646968666</td><td>514862</td><td>8540803</td></tr> <tr><td>Thailand</td><td>595904494</td><td>411992</td><td>9112146</td></tr> </tbody> </table> <p>Answer: Spain.</p>	CountryName	TotalSpent	TotalVisited	TotalStayed	Spain	8683030523	15871872	145869761	USA	4279222407	3401107	44182150	France	3984141502	8861644	66283281	Italy	2378037814	4159924	32112217	Greece	1664050506	2382735	23255850	Portugal	1552384904	2875598	26586459	United Arab Emirates	1105440694	1015885	10910260	Germany	1083709889	2909338	16484508	Other Asia	1055744633	931998	22234924	Irish Republic	993229378	3422119	18838173	Netherlands	973560404	2659833	10767558	Poland	950130504	2672368	26205224	Other countries	901243881	685476	8282654	Other Africa	877823654	947115	20824351	India	820906396	1006423	24911348	Australia	819939391	452368	14223339	Turkey	712476696	1228144	14710793	Canada	646968666	514862	8540803	Thailand	595904494	411992	9112146	Pass
CountryName	TotalSpent	TotalVisited	TotalStayed																																																																																	
Spain	8683030523	15871872	145869761																																																																																	
USA	4279222407	3401107	44182150																																																																																	
France	3984141502	8861644	66283281																																																																																	
Italy	2378037814	4159924	32112217																																																																																	
Greece	1664050506	2382735	23255850																																																																																	
Portugal	1552384904	2875598	26586459																																																																																	
United Arab Emirates	1105440694	1015885	10910260																																																																																	
Germany	1083709889	2909338	16484508																																																																																	
Other Asia	1055744633	931998	22234924																																																																																	
Irish Republic	993229378	3422119	18838173																																																																																	
Netherlands	973560404	2659833	10767558																																																																																	
Poland	950130504	2672368	26205224																																																																																	
Other countries	901243881	685476	8282654																																																																																	
Other Africa	877823654	947115	20824351																																																																																	
India	820906396	1006423	24911348																																																																																	
Australia	819939391	452368	14223339																																																																																	
Turkey	712476696	1228144	14710793																																																																																	
Canada	646968666	514862	8540803																																																																																	
Thailand	595904494	411992	9112146																																																																																	
2	<pre>select COUNT(Package), Package from load_Tourist_Dashboard where year = '2018' GROUP BY Package ORDER BY `COUNT(Package)` DESC</pre>	In 2018, which package is preferred by UK tourists: Independent or Non-Independent?	<p>Showing rows 0 - 1 (2 total, Query took 0.0001 seconds.)</p> <pre>select COUNT(Package), Package from load_Tourist_Dashboard where year = '2018' GROUP BY Package ORDER BY `COUNT(Package)` DESC</pre> <table border="1"> <thead> <tr> <th>COUNT(Package)</th> <th>Package</th> </tr> </thead> <tbody> <tr><td>8798</td><td>Independent</td></tr> <tr><td>2472</td><td>Non-Independent</td></tr> </tbody> </table> <p>Answer: Independent.</p>	COUNT(Package)	Package	8798	Independent	2472	Non-Independent	Pass																																																																										
COUNT(Package)	Package																																																																																			
8798	Independent																																																																																			
2472	Non-Independent																																																																																			

3	<pre>select CountryName, COUNT(CountryName) from load_Tourist_Dashboard where Mode = 'Air' group by CountryName ORDER BY `COUNT(CountryName)` ASC LIMIT 5</pre>	<p>In the last 15 years, which 5 countries have been least visited by air by UK tourist?</p>	<p>Showing rows 0 - 4 (5 total, Query took 0.1253 s)</p> <pre>select CountryName, COUNT(CountryName) from load_Tourist_Dashboard where Mode = 'Air' group by CountryName ORDER BY `COUNT(CountryName)` ASC LIMIT 5</pre> <p>+ Options</p> <table border="1"> <thead> <tr> <th>CountryName</th> <th>COUNT(CountryName)</th> </tr> </thead> <tbody> <tr><td>Congo Brazzaville</td><td>1</td></tr> <tr><td>Mali</td><td>1</td></tr> <tr><td>Paraguay</td><td>1</td></tr> <tr><td>French Guiana</td><td>1</td></tr> <tr><td>Burkina Faso</td><td>1</td></tr> </tbody> </table>	CountryName	COUNT(CountryName)	Congo Brazzaville	1	Mali	1	Paraguay	1	French Guiana	1	Burkina Faso	1							
CountryName	COUNT(CountryName)																					
Congo Brazzaville	1																					
Mali	1																					
Paraguay	1																					
French Guiana	1																					
Burkina Faso	1																					
4	<pre>select Gender, SUM(TotalSpent) from load_Tourist_Dashboard where year = '2018' and '2017' and '2016' and '2015' and '2014' and Purpose = 'Holiday' group by Gender ORDER BY `SUM(TotalSpent)` DESC</pre>	<p>In the last 5 years, which gender have spent the most amount(£) on holiday?</p>	<p>Showing rows 0 - 2 (3 total, Que</p> <pre>select Gender, SUM(TotalSpent) from load_Tourist_Dashboard where year = '2018' and '2017' and '2016' and '2015' and '2014' and Purpose = 'Holiday' group by Gender ORDER BY `SUM(TotalSpent)` DESC</pre> <p>Show all Number of rows:</p> <p>+ Options</p> <table border="1"> <thead> <tr> <th>Gender</th> <th>SUM(TotalSpent)</th> </tr> </thead> <tbody> <tr><td>Male</td><td>14916001421</td></tr> <tr><td>Female</td><td>10762496306</td></tr> <tr><td>D/K</td><td>708453130</td></tr> </tbody> </table>	Gender	SUM(TotalSpent)	Male	14916001421	Female	10762496306	D/K	708453130											
Gender	SUM(TotalSpent)																					
Male	14916001421																					
Female	10762496306																					
D/K	708453130																					
5	<pre>select Age, SUM(load_Tourist_Dashboard.TotalNights) as TotalNights from load_Tourist_Dashboard GROUP by Age ORDER BY `TotalNights` DESC</pre>	<p>Which age group spent the most nights in 15 years?</p>	<p>Showing rows 0 - 8 (9 total, Que</p> <pre>select Age, SUM(load_Tourist_Dashboard.TotalNights) as TotalNights from load_Tourist_Dashboard GROUP by Age ORDER BY `TotalNights` DESC</pre> <p>Show all Number of rows:</p> <p>+ Options</p> <table border="1"> <thead> <tr> <th>Age</th> <th>TotalNights</th> </tr> </thead> <tbody> <tr><td>25-34</td><td>1789506968</td></tr> <tr><td>35-44</td><td>1740351898</td></tr> <tr><td>45-54</td><td>1639530668</td></tr> <tr><td>55-64</td><td>1486396507</td></tr> <tr><td>65 & over</td><td>1253287267</td></tr> <tr><td>16-24</td><td>1152295817</td></tr> <tr><td>0-15</td><td>600287428</td></tr> <tr><td>D/K</td><td>100626745</td></tr> </tbody> </table>	Age	TotalNights	25-34	1789506968	35-44	1740351898	45-54	1639530668	55-64	1486396507	65 & over	1253287267	16-24	1152295817	0-15	600287428	D/K	100626745	
Age	TotalNights																					
25-34	1789506968																					
35-44	1740351898																					
45-54	1639530668																					
55-64	1486396507																					
65 & over	1253287267																					
16-24	1152295817																					
0-15	600287428																					
D/K	100626745																					

Black Box Testing – Microsoft Excel Dashboard				
Test No.	What is being test?	Expected result	Actual result	Pass /Fail
1	Does the information appear when hovering over the graphs/map?	The tooltip must appear when hovering over the graphs/map	The tooltip appeared and displayed the necessary information	Green
2	The map should display the country with the highest amount spent by UK tourists	The highest amount spent country by UK tourist will need to be a darker colour.	The map did show the countries in darker colour which had highest amount spent by UK tourist	Green
3	The map should display the country with the least amount spent by UK tourists	The map will highlight the countries with the least amount spent by UK tourists	The map did not show the countries with the least amount spent by UK tourist. Possible solution: add filter which allow the map to have both highest and least amount spent by UK tourists.	Red
4	The filter used for the pie chart.	The pie chart should change its result when selecting a different value in the filter	The pie chart did update with new information when the filter was used.	Green
5	The bar graph in dashboard	The bar graph should display the value for each country when clicked.	The individual value did appear when clicked on the bar chart.	Green
6	Accessing excel dashboard online	The dashboard should appear online when opening the file using the internet	The excel dashboard did not work when trying to open in a browser because of the file size being too big.	Red
7	Using two filter for a graph in dashboard	The graphs should change the information it displays when two different filter are used.	The graph did change the information when both filters where used.	Green

Black Box Testing – Tableau Dashboard				
Test No.	What is being test?	Expected result	Actual result	Pass /Fail
1	Total visits in mode of transport in percentage.	The chart needs to display the information for each mode of transport.	The chart did display the required information	Green
2	Total nights by purpose of visit	The table will highlight the value with the highest figure.	The table did highlight the column with a darker colour.	Green
3	Total Spent (£) by UK tourists	The map should display the total amount spent by UK tourists	The map did show the information about the total amount	Green
4	Hovering over the map	When hovering over a country, the information should show up in a tooltip.	The tooltip did show up when hovered over a country.	Green
5	All graphs, chart, table and map.	All the information should change when a country is selected from the map	All the information did change when a country 'Spain' was selected.	Green
6	The dropdown list for quarter (months)	When click a month in the dropdown list, all information in the dashboard should change.	The information did change when selecting the months from dropdown list.	Green
7	The 'year' filter	The filter should allow to choose two different year to compare	The filter did not allow to choose two different year. Possible Solution: Change to checkbox if possible.	Red

8	Total spent by UK tourist, arranged by Age	When a individual age group is selected, change all information in dashboard. All information should be related to the selected age group.	The information did change and was related to the selected age group.	
9	Share Dashboard	Give option to share dashboard and allow user to enter login information. The dashboard should be shared when completed with login info.	The share dashboard function did work and shared the dashboard to others.	
10	Download Dashboard in PDF	The dashboard should allow the user to download the whole dashboard into PDF	The PDF was provided when clicking the download button.	

Black Box Testing – Wix Website containing dashboard				
Test No.	What is being test?	Expected result	Actual result	Pass /Fail
1	Login account	The user should be able to login into the website with a username and password	The login was successful and redirect back to the homepage when successful.	Green
2	Chat function should work	When typing into the chat, the message to be sent to the admin.	The chat function did work.	Green
3	Social media function	When the social media icon is clicked, it should redirect to the appropriate webpage.	This function did work and redirected to the appropriate webpage.	Green
4	The navigation button	When the navigation button is click, it should redirect to a different webpage.	The button did redirect to a different webpage	Green
5	The dashboard in the webpage	The dashboard needs to be interactive	The dashboard was interactive, and all the graphs and map can be clicked in the website.	Green
6	The contact page	When writing a message and clicking submit, the message should be sent to the admin	The message did not work. Possible Solution: May need to investigate the setting of the message form.	Red
7	The dashboard should scroll	The dashboard should be scrollable so that hidden content can be seen.	Yes, the dashboard can be scrolled, and the hidden information can be seen.	Green

22 Appendix G (Interaction with Supervisor and Blog Entries)

eSupervisor Communications - Add a Communication

26/04/2019	By Aditi Rawal	Met to provide feedback on the 23/4/19. The appendix does not count towards the word count.
15/04/2019	By Me	Does appendix count towards the word count in the report?
19/03/2019	By Me	Currently, I am trying to create a website to host my dashboard from tableau. Do i get more marks if i create the website myself using code or can I use a third party application like 'WIX' to create the website?
26/02/2019	By Me	When can I meet you this week about my project report?
19/02/2019	By Aditi Rawal	Data protection is a legal requirement and can translate into your non-functional requirements. So you would provision for you system security to enable this legal requirement by way of non-functional requirement.
16/02/2019	By Me	Is following data protection a functional requirement if my system works with data?
01/02/2019	By Aditi Rawal	You will still need to add a bit more on Visualisation as that seems less and requires more research. Make sure you don't have questions as headings. The product research starts very abruptly and needs an introductory passage.
01/02/2019	By Me	Thank you for the feedback. However, is there anything wrong with the Literature Review?
30/01/2019	By Aditi Rawal	Hi Kogul, You have feedback for your report, read and amend. Also make sure you don't have any repetition in your literature review with your sub headings.
14/01/2019	By Aditi Rawal	Dear All, Kindly note my office hours for this term are, Tuesday 12 -1pm and Friday 11.30- 12.30pm. As there aren't any scheduled project tutorials for this term, I would be expecting you to use these hours to drop in without an appointment to discuss your project. Regards, Aditi
08/01/2019	By Aditi Rawal	It has been signed off and now you can upload your report by the 11 Jan
07/01/2019	By Me	I submitted the file.
07/01/2019	By Aditi Rawal	Then upload your proposal here which you had uploaded to Moodle and had got feedback for and I will then sign it off
07/01/2019	By Me	How do I upload the contextual report? When I click the link, it is telling me to get my project proposal approved.
01/11/2018	By Aditi Rawal	Dear student, As part of the employability week as well as a requirement to meet up, I would like to invite you to come and see me on Tuesday 6 November between 11am -1pm at QM362. Kindly note an attendance register will be taken for this meeting and one to one feedback will be provided. With regards, Aditi

Supervisor Meetings

05/03/2019	Aditi Rawal	Met up to discuss the progress. Feedback regarding the report and the product provided. Add a comment
06/11/2018	Aditi Rawal	Met to provide formative feedback Add a comment

Diary/Blog entries - Add an entry

23/04/2019	Had a meeting with Aditi about my project and got feedback.
23/04/2019	Working on the report
15/04/2019	Documenting the implementation stage in the report.
12/04/2019	Working on prototype
29/03/2019	Met with Aditi on 26/3/2019 and got feedback on project
25/03/2019	Still working on website and tableau
18/03/2019	Finishing the second prototype and starting the website creation.
11/03/2019	Finalising my first prototype and moving on to my second prototype.
05/03/2019	Starting the implementation stage by creating the database and loading necessary data. Moving on to my first prototype.
25/02/2019	Currently working on designing the Use-case Diagram for my project.
15/02/2019	Starting the rich picture and on the side, working on prioritisation for functional requirement.
04/02/2019	Working on the report using the feedback from the contextual report submission.
29/01/2019	Met with Aditi Rawal and looked through the report and clarified some doubts I had with the report
23/01/2019	Working on requirements and researching and writing about the topics relating to my project.
07/01/2019	Currently working on literature review by researching and comparing relevant academic articles and journals.

This project record was created on 20/09/2018