

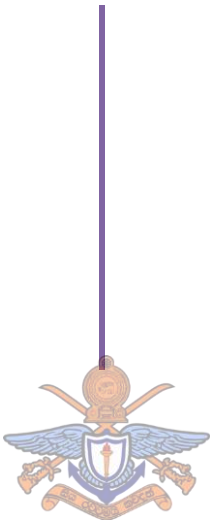


Advanced SQL and Cloud Databases

Assignment 2

Assignment Information

Assessment Title : Greater Manchester Road Safety Analyzer & Greater Manchester Higher Education Analyzer	
Programme Code(s): NA	Issue date¹: 1/10/2023
Weighting : 25% of the total module mark	Submission date²: 30/11/2023
Assessor(s) : Dr Charith Silva	Return date³: The final mark will be available in a maximum of 4 weeks after submission.



Learning Outcomes of this Assessment

The learning outcomes covered by this assignment are you will be able to,

- Develop skills in importing, cleaning, transforming, and shaping raw data for analysis.
- Gain hands-on experience with creating Power BI reports and dashboards.
- Provide a broad overview of the general field of report and dashboard designing.
- Provide an overview of the general field of 'Cloud databases'.
- Learn to deal with common data issues like missing values, duplicates, and inconsistent formats.

Key Skills to be Assessed

This assignment assesses your skills in:

- Designing and implementing a SQL Database.
- Designing and implementing Reports
- Use of Azure PaaS Database
- Use of T-SQL
- Use of Power BI Report Builder & Power BI Desktop

Recommended Reading

The module notes are sufficient literature for completing this assignment successfully.

Equipment and Facilities to be Used

The assignment has to be implemented using Azure PaaS SQL Database (Preferred) or SQL Server Database, Power BI Report Builder and Power BI Desktop. Ideally, data should be stored in an Azure PaaS Database, but if you have technical difficulties, you can use SQL Server Database. SQL statements should be stored in plain text files.

Team Composition

This is a team assignment. Each team may freely choose **two, three or a maximum of four members** from the class. Group assignments allow you to develop key transferable skills such as teamwork, organisation, communication, leadership and problem-solving. In addition to this, it can be highly beneficial to both your academic performance and long-term personal development, particularly as the ability to work well with others is a key requirement for many graduate jobs.

Workload

For the successful completion of this assignment, a total of 120 hours should be budgeted.

Task 1 (50%)

Imagine you are employed as a Power BI developer for a Manchester-based (UK) road safety-related research company. They are currently in the process of developing a Power BI based reporting tool called **Greater Manchester Road Safety Analyzer** using Azure SQL PaaS/SQL Server Databases, Power BI Builder and Power BI Desktop to analyse historical **fatal and serious** roadside accidents in the Greater Manchester area.

Greater Manchester Road Safety Analyzer is a simple, low-cost reporting solution that can analyse and evaluate fatal and serious roadside accidents. It allows users to extract and present data in charts, tables, and other visualisations so users can find useful information. It can also allow you to build paginated reports that are ideal for printing. The purpose of this tool is to translate data into actionable information to understand **Fatal** and **Serious** roadside accidents only in Local Authority Districts in Greater Manchester, which are Bolton, Bury, Manchester, Oldham, Rochdale, Salford, Stockport, Tameside, Trafford, Wigan in the past five years.

Link for the Datasets:

<https://www.data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data>

Road Safety - Casualties last 5 years	.CSV	14 October 2021
Road Safety - Vehicles last 5 years	.CSV	14 October 2021
Road Safety - Accidents last 5 years	.CSV	14 October 2021

Link for the Metadata :

<https://www.data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data>

Supporting documents

Link to the document	Format	Date added
Road Safety - Digital Breath Test Data Guide	XLS	05 October 2015
Published statistics and supporting documents	HTML	05 October 2015
Understanding historical road safety data	.docx	14 October 2021
Severity Adjustment Figure Guidance 2020	.docx	14 October 2021
Road Safety Open Dataset Data Guide	.xlsx	15 October 2021

A road traffic accident is one of the undesirable events that is uncertain. Road safety is a major societal issue worldwide; road accidents are one of the main causes of unnatural fatalities, disability, and property harm. Along with fatalities, road traffic crashes cause thousands of severe and slight injuries annually. European roads remain the safest in the world. In 2017, the EU counted 49 road fatalities per one million inhabitants, against 174 deaths per million globally. Sweden, the Netherlands, the UK, and Malta have the lowest reported road fatality rate, below 30 dead per million Inhabitants. According to data from the World Health Organisation, about 1.3 million people die each year on the world's roads, of which 25,300 lost their lives in the EU last year. For every person killed in traffic crashes, about five more suffer serious injuries with life-changing consequences. Serious injuries are often costly to society because of long-time rehabilitation and healthcare needs.



Millions of citizens use road infrastructure daily for walking, riding or driving. Any kind of serious accident on our roads is a disaster. As a responsible member of the community, we must take the road safety measures more seriously to protect the community. In some cases, victims of a roadside accident can fully recover from their injuries within a short period of time, while other victims become permanently disabled. Road traffic injuries are a majorly ignored public health problem, requiring rigorous efforts to provide effective and sustainable protection. One way to reduce the number of traffic accidents is to analyse traffic accidents recorded in-depth and understand the causes. Understanding where traffic accidents most happen and identifying peak time windows are crucial for improving road safety and traffic enforcement officers. One of the most common methods of analysing road traffic accidents is descriptive statistics and time series analysis.

This Research organisation seeks to evaluate existing data and interpret roadside accidents from multiple angles, at various levels and more broadly, to discover potentially useful rules to improve road safety. Improving road safety is the main goal of intelligent transport systems and leads to the importance of road infrastructure safety. Accident severity prediction information can provide rescuers with valuable

information in response to the severity of accidents and other associated risks.

This study provides details about road traffic accidents in Great Britain that were reported by the police during the last five years using the STATS19 reporting system. The data used in this project was collected by police forces, either through officials who take part in the scene or from members of the public reporting the incident in police stations after the incident. Accidents that happened on private land or car parks are not included in the study. Damage-only accidents not resulting in personal injury are also excluded from this study. A fatal accident is defined as a roadside accident in which at least one person is killed. A serious accident is defined as that at least one person is seriously injured. However, no one involved in the accident is killed. Slight accident is defined as when at least one person is injured but does not require medical treatment.



The client requires your knowledge and expertise to design, implement and test a reporting system for this project using Azure PaaS Database/ SQL Server Database, Power BI Report Builder and Power BI Desktop. Use Power BI Report Builder and Power BI Desktop as a presentation layer.

You will provide a report detailing your approach and proposed front-end design in Power BI Report Builder and Power BI Desktop. You will also provide your SQL statements with appropriate comments. In addition to the front-end design, the client also requires working T-SQL statements with meaningful comments to achieve the following:

- Insert the last 5 years' dataset from the <https://www.data.gov.uk> to Azure PaaS SQL Database or SQL Server Database.
- Create all the tables using T-SQL.
- Create more than **five** meaningful summarised reports using Power BI Report Builder that allow users to explore.
- Include filtering, sorting and grouping functionality in the reports.
- Include various search facilities in the reports.
- Also, create a Power BI Dashboard using data in the database that allows users to evaluate the factors that correlate with the Fatal and serious road traffic accidents in Greater Manchester.
- Interpret the findings from the dashboard.
- Discuss any patterns, anomalies, or areas of interest.
- Discuss the implications of your findings

Extra features to be implemented

To get more than a "Satisfactory" mark, a number of extra features should be implemented. Features include, but are not limited to:

- Use of Views, CTEs and stored procedures
- Use of system functions, including aggregate and ranking functions
- Use of Maps in Power BI
- Use of Advanced Power BI Visualizations such as create custom visuals with R or Python

Report

A 2000-2500-word report that documents your database, report and dashboard design.

Requirements / Marking Scheme

Requirement	Assessment Method	Weight (%)
Power BI Report design	Report	15%
Power BI Dashboard design	Report	15%
Report	Report	20%

Task 2 (50%)

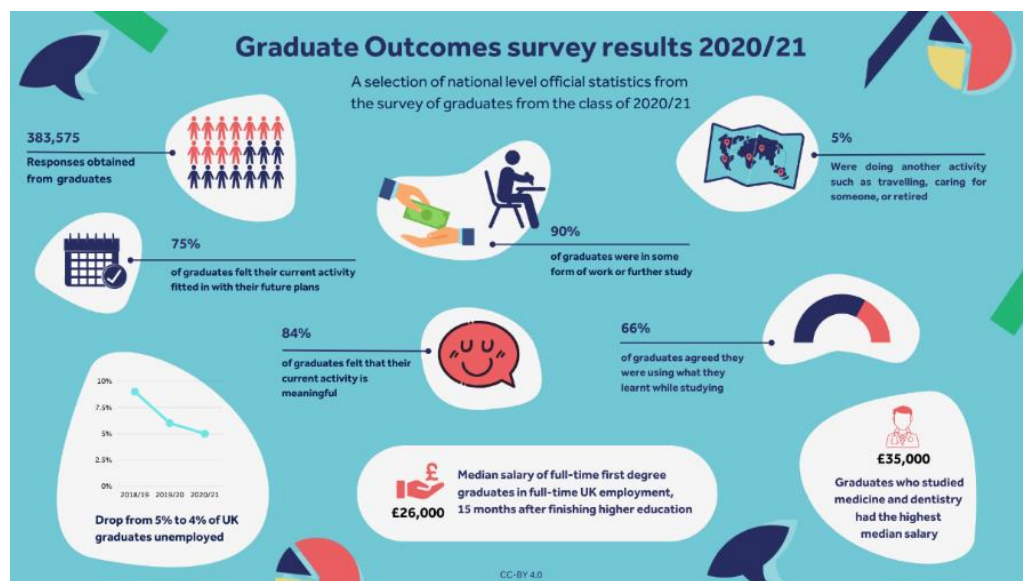
Imagine you are employed as a research assistant for a Manchester-based Higher Education organisation. They are currently in the process of developing a comprehensive data visualisation platform called **Greater Manchester Higher Education Analyser** using the latest UK Higher Education Graduate Outcomes Data. You are free to use any technology you like for this task e.g., SQL Server Databases, Azure Databases, Power BI Builder, Microsoft Report Builder, Power BI Desktop, Excel, etc., to analyse performance in higher education institutes in Greater Manchester. Greater Manchester has several world-class universities and higher education organisations (<https://www.studyinmanchester.com/universities>).

This task of the assignment aims to help you develop your data analysis, visualization, and interpretation skills using real-world datasets. Ensure that your dashboard is not just informative but also insightful, offering valuable perspectives on the outcomes of the graduates studied from Higher Education institutes located in Greater Manchester. .



Graduate Outcomes is the biggest UK annual social survey and captures recent graduates' perspectives and current status. All graduates who completed a course will be asked to take part in the survey 15 months after they finish their studies. The survey aims to help current and future students gain insight into career destinations and development.

This survey will help Universities or colleges evaluate and promote their courses. The Higher Education Statistics Agency (HESA) owns the Graduate Outcomes survey, which is the higher education sector's designated data body. The administration of the Graduate Outcomes survey is undertaken by a third-party survey contractor who works in collaboration with HESA and universities. Participation in the Graduate Outcomes survey is a statutory requirement for all higher education providers, and most university graduates are required to participate. Anyone who graduates after the 1st of August 2017 will be asked to complete the survey. All eligible students will be contacted to ensure consistency and accuracy across all universities.



Link for the Datasets and Metadata:

<https://www.data.gov.uk/dataset/37b401c3-1689-4f3c-bac4-b6cc39cdefa7/higher-education-graduate-outcomes-data>

There are several different kinds of data sets available in this data location. Free to use any of the dataset or multiple datasets for your project. However, make sure you filter out only the data related to higher education institutes in Greater Manchester.

Supporting documents

Link to the document	Format	Date added
Higher Education Graduate Outcomes Data	HTML	17 July 2020
Definitions	HTML	17 July 2020
Data intelligence	HTML	17 July 2020

Data links

Link to the data	Format	File added	Data preview
Table 30- UK domiciled graduates in full-time paid employment in the UK by subject area of degree and salary band 2019/20 onward	CSV	21 July 2023	Preview
Table 29 - Response rates by subject of study 2019/20 onward	CSV	21 July 2023	Preview
Table 28 - Graduate activities by provider and subject area of degree 2017/18 onward	CSV	21 July 2023	Preview
Table 27 - Graduates in further study by subject area of former degree and level of qualification aimed for in further study during census week 2017/18 onward	CSV	21 July 2023	Preview
Table 26 - UK domiciled graduates who obtained first degree qualifications and entered full-time paid employment in the UK by provider and salary band 2017/18 onward	CSV	21 July 2023	Preview
Table 25 - UK domiciled graduates in full-time paid employment in the UK by subject area of degree and salary band 2017/18 to 2018/19	CSV	21 July 2023	Preview
Table 24 - UK domiciled graduates in full-time paid employment in the UK by personal characteristics and salary band 2017/18 onward	CSV	21 July 2023	Preview
Table 23 - Standard industrial classification of graduates entering work in the UK by standard occupational classification 2017/18 onward	CSV	21 July 2023	Preview
Table 22 - Standard occupational classification of graduates entering work in the UK by provider 2017/18 onward	CSV	21 July 2023	Preview
Table 21 - Response rates by subject of study 2017/18 to 2018/19	CSV	21 July 2023	Preview
Table 20 - Graduate activities by provider and subject area of degree 2017/18 to 2018/19	CSV	21 July 2023	Preview
Table 19 - Standard industrial classification of graduates entering work in the UK by provider 2017/18 onward	CSV	21 July 2023	Preview
Chart 17 - Graduates' subjective wellbeing by subject area of degree 2019/20 onward	CSV	21 July 2023	Preview
Chart 16 - Graduates' reflection on activity by subject area of degree 2019/20 onward	CSV	21 July 2023	Preview

This is a broad assignment that gives students the flexibility to showcase their skills in data visualization, analysis, and presentation. Depending on the actual complexity of the dataset, you can further refine or expand upon these requirements. Develop an interactive dashboard to analyze and visualize the outcomes of recent UK graduates, shedding light on key metrics such as employment rates, salary figures, further education pursuits, etc.

Tips:

- Focus on ensuring that the dashboard is user-friendly and interactive.
- Clearly label all visualizations and include a brief description or observation for each.
- Validate your interpretations with relevant external sources or literature, if applicable

Extra features to be implemented

To get more than a "Satisfactory" mark, a number of extra features should be implemented. Features include, but are not limited to:

- Use of Maps in Power BI
- Use of Advanced Power BI Visualizations such as create custom visuals with R or Python
- Create a video that explains the functionality of the dashboard (Kind of a user manual)

Report

A 2000-2500-word report that documents your database, report and dashboard design.

Requirements / Marking Scheme

Requirement	Assessment Method	Weight (%)
Visualisation Platform Design	Report	30%
Report	Report	20%

Notes

- The assignment must be completed with your team members. Any code snippets you or your team members do not directly write must be referenced within your code. You must directly comment on the code to explain its source. Failure to reference a code that is not yours will be treated as plagiarism.
- The assignment must be completed on time.

If you submit work late, it will be marked according to the University's late submission policy.

Submission

Your submission should be a single ZIP file uploaded via Teams. The file should be named as:

<<Your surname>>_<<Your name>>.zip — for example: **Smith_John.zip**.

All items in the zip file should also be prepended by your surname. (Ensure you replace "Smith" by your surname in the names below).

The following items must be included in your submission for each task.

1. A folder named **Smith_Reports** contains the Report Builder reports (Task 1).
2. A folder named **Smith_Dashboard** contains the Power BI Dashboard (Task 1 and Task 2).
3. A PDF document named **Smith_report.pdf** containing your report (Task 1 and Task 2)

It is assumed that you will also address any social/legal and ethical issues surrounding the design and implementation of the project, such as copyright, references, and web law.