Project Description Outline

Names of group members

Graeme, Rashpal, Sarina, Stephanie

Roles & responsibilities of each member

Graeme worked on creating a Shiny Dashboard framework

Rashpal worked on gathering data on life expectancy and plotting the appropriate plots

Sarina worked on creating a geographical picture of mental wellness using Leaflet

Stephanie worked on gathering data and creating plots on long-term health conditions as well as mental health

Everyone worked on presentation creation and presentation delivery

Brief description of dashboard topic

Our dashboard contains information regarding mental wellness and how this changes with various demographics, geography & time.

Our health topic is mental wellness

The dashboard outlines our health topic in terms of:

- An overall picture for Scotland
- Mental wellness overtime
- Mental wellness by geographic location
- Mental wellness based on various demographics (age, gender, job status, underlying health condition & could potentially split this by region

Stages of the project

- Planning & dashboard wireframe
- Reading relevant documentation
- Git branching & version control
- Choosing datasets
- Manipulation of datasets
- Plotting data
- Creating basic Shiny dashboard framework
- Populating relevant sections with plots
- Preparing a presentation

Which tools were used in the project

- Zoom (daily stand-ups and occasional mob programming)
- Trello (planning & task allocation)
- Git/GitHub (collaboration & version control)
- Slack

How did you gather and synthesise requirements for the project?

We synthesised the information given in the brief by discussing available topics and reviewing potential data sources.

We used the Q&A session with 'the client' to gather further information on what information was best to present, what to do if we did not have the most up-to-date data for our graphs and to confirm what statistical analysis was appropriate given the data available.

We prioritised ensuring we understood the goal of our presentation and how we would appropriately communicate our findings.

Motivations for using the data you have chosen

We used the dataset on life expectancy on statistics.gov.uk to find the trend over the last 30 years to obtain an overview of general health in Scotland.

The hospital admissions dataset was from statistics.gov.uk and contained the number of admissions by long-term condition which allowed the trends of long-term condition to be visualised.

We choose various mental health indicators from statistics.gov.uk to further convey the problems associated with mental health.

We used data from the National Records of Scotland to find data for life expectancy within the UK.

Data quality and potential bias, including a brief summary of data cleaning and transformations

According to the About tab on the dataset page/dedicated page online, the data quality, according to the statistics.gov.uk:

The survey is currently designed to yield a representative sample of the general population living in private households in Scotland every year. Estimates at NHS Health Board and Local Authority level are available over a four year reporting period. Those living in institutions, who are likely to be older and, on average, in poorer health than those in private households, were outwith the scope of the survey. This should be borne in mind when interpreting the survey findings. The survey uses a clustered, stratified multi-stage sample design. In addition, weights are applied when obtaining survey estimates

The dataset may be biased because the sample size is limited and therefore we may not have an accurate

representation of the entire population.

Datasets which were based on surveys may include response bias.

To clean the datasets we used various techniques taught during the course to remove unnecessary columns, clean unusable data and get the data into ‘tidy format’.

How is the data on statistics.gov.scot stored and structured

The data on statistics.gov.scot is in the form of linked data

This means the data can be combined to produce more advanced queries

Benefits of storing the data like this are that it:

It allows more flexibility in organising the data

It is dynamic data structure.

It’s efficient with the data stored once

Ethical and legal considerations of the data

There may be ethical considerations, however the data has been stripped of all personal identifiers.

The datasets are covered by the Open Government License, which means you are free to use the data for both commercial and non-commercial purposes as long as you clearly state where the data has been obtained.

More information can be found at:

http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/]