

Project Description Outline

Names of group members

Jack Forrester;

Chris Allen;

Rob Rodden;

Sarah(Xi) Bonham-Yang

Roles & responsibilities of each member

Everyone worked on:

- Exploring cleaning and analysis datasets;
- Designing dashboard wireframe;
- Building Shiny App;

Jack Forrester worked on:

- Cleaning and analysing data relate to “Hospital beds occupancy”;
- Creating visualisations for beds occupancy;
- Building bed occupancy section into Shiny app

Chris Allen worked on

- Cleaning and analysing data relate to “A&E Waiting time” ;
- Creating visualisations for A&E waiting time and;
- Building basic stiction and health board input section into Shiny app;
- Building visualisations to Shiny dashboard

Rob Rodden worked on:

- Cleaning and analysing data relate to “SIMD”;
- Creating visualisations for SIMD;
- Building SIMD section into Shiny app

Sarah(Xi) Bonham-Yang worked on:

- Cleaning and analysing data relate to “specialty”;
- Creating visualisations for specialty;
- Building specialty section into Shiny app

Brief description of dashboard topic

Our dashboard contains :

- Health board user input
- Map of health boards
- Trends of hospital admissions before and during covid time
- Bar chart of top 5 biggest increase of specialty in selected health boards
- Changes of hospital admissions over different genders and age groups
- Changes of hospital admissions over different SIMD groups
- Hospital KPI (Waiting time and bed occupancy)

Our topic is

- COVID impact on hospital performance and on different groups of general population

The dashboard outlines our topic in terms of ...

Stages of the project

- Planning & dashboard wireframe
- Git branching & version control
- Choosing datasets
- Data exploring, cleaning & analysis
- Drafting and choosing graphs to use
- Creating dashboard
- Presentation and report

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 (collaboration & communication)

How did you gather and synthesise requirements for the project?

We synthesised the information given in the brief by ...

We prioritised ...

Motivations for using the data you have chosen

We used all the datasets on hospital performance across the Scotland to answer 'is the winter crisis real' and 'the impact of COVID'. We picked out the information and demographic groups that we are interested in to answer our questions.

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 is good, it meet a high standard of data quality.

Therefore, the dataset may not be biased.

To clean the datasets we picked out the information that we needed.

Used date time information to get year, month, season information.

Then we joined the health board names table with the hospital data to allow comparing.

How is the data stored and structured

[**Hint:** This page offers a good starting point for understanding the data structure: <https://guides.statistics.gov.scot/article/34-understanding-the-data-structure>]

The data is in the form of linked data. Wherever possible, any data point contains a linked dataset should have a unique identifiable URL.

This means data point can be browsed using a web browser and can be referenced by other linked datasets.

Benefits of storing the data like this are that it makes it very convenient to connect datasets.

Ethical and legal considerations of the data

[**Hint:** this page contains information on the Open Government License which covers most datasets on statistics.gov.scot: <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>]

There are no personal information in the datasets. However, some years are shorter than others which makes the records fewer too. It could make our analysis to be biased.

We also don't have enough information in things like death rate and what happens after a patient is transferred which could be important for us to understand the impact of COVID.

The datasets are covered by the Open Government License, which means public can use the datasets as long as they acknowledge the source.