Data Science Laboratory 1

Assignment 1

In this assessment it is your turn to do some research! You are expected to generate a one page report that presents a basic analysis and visualisation of a data set of your choosing (if you need some inspiration, some hints can be found below). The goal is to tell a story using data. You are expected to identify an open access (i.e. public) data set and provide the reader with an easily digestible summary of its main properties, interesting features, trends etc. The report should contain explanatory text, tables and figures. basic stats + visualisation (not models yet)

You will be expected to use the visualisation methods described in the notes, but you are also free to explore other techniques if you wish. You will be assessed on the following:

1. The clarity of your report including spelling and grammar, use of captions and explanatory text.

2. Appropriate use of visualisation techniques and statistical summaries.

3. Quality of figures (e.g. legibility, proper labelling).

4. Use of references where appropriate (e.g. data sources, any unusual techniques employed).

The exact format and layout of the report is up to you, but you should follow these guidelines:

- It should be no more than one A4 page (PDF format please) with 2 cm margins all around.

- The minimum font size should be 11 pt, with the exception of figure labels, footnotes and references.

- The report should have a clear title and short abstract or introductory paragraph explaining the topic.

Some ideas for data sources:

- Government: `https://data.gov.uk`

- MET Office: `https://www.metoffice.gov.uk/research/climate/maps-and-data/data/index`

- World Bank Open Data: `https://data.worldbank.org`

- WHO Global Health Observatory: `https://www.who.int/data/gho`

- Google Public Data Explorer: `https://www.google.com/publicdata/directory`

- FiveThirtyEight: `https://data.fivethirtyeight.com`

- UCI Machine Learning Repository: `https://archive.ics.uci.edu/ml/index.php`