

Assignment Visualising Data 2020

	Task	Submit	Mark
1	<p>Find a dataset on a topic that you can understand. Your final dataset should have at least eight attributes containing both qualitative and quantitative data. The dataset must have at least 200 rows. Merging will enhance your marks. If you are doing this, you must explain the merge. These attributes should include the following types of data:</p> <ul style="list-style-type: none"> • Continuous Numeric data • Categorical data • Temporal data <p>And may include</p> <ul style="list-style-type: none"> • Location data • Text data (attribute with text data, longer than 15 characters). 	<p>The URL and description of the dataset, using either MS Word or PDF. Many sources will have a description. You must add to this, explaining which attributes fit the requirements listed. If the dataset is not accessible using its URL from your programs, upload it.</p>	4
	<p>Explore the data by wrangling, and generating graphs in R, using Jupyter Notebook. Note distributions across the data and correlations between attributes. Fix errors and decide on how errors and omissions will be handled. Justify the chart types you have used in terms of the types of data in your dataset. Use markdown cells to explain what you are doing. Put your name and student number in the top of the notebook. Develop a meta-data description of your dataset.</p>	<p>Your exploratory Jupyter notebook in R, with the charts and their descriptions. This should have everything in it for the marker to run it.</p>	6
	<p>Form your Big Idea following your exploration. You should now understand what the data is telling you, so you should be able to establish a message, an audience and what you want from the story. Describe your audience, who you will approach to pass on the message, and what is at stake for them. A sentence expressing your big idea.</p>	<p>Submit your development in a word or PDF document. The complexity of your story will be taken into account.</p>	6
	<p>Develop your Storyboard in a Jupyter Notebook using R, ensuring that you are telling the truth insofar as the data tells you. If you suspect that there are factors that are not depicted in your dataset, please disclose this. The graphs that you produce should be essential to the story.</p>	<p>Your story Jupyter Notebook, in a runnable format. Put your name and student number in the top of the notebook.</p>	10

Marking:

- Data discovery, validation, enrichment, explanation, metadata.
- Development environment will be taken into account.
- Clarity, relevance and significance of visualisation.

Dataset(4 marks)

- Single dataset not meeting requirements 0
- Single dataset meeting requirements 2
- 2 or more datasets meeting requirements 4

Data discovery: (6)

- No wrangling required 0
- Very clean initial dataset but well explored with graphs 1
- Wrangling found data problems and corrected them 2
- Problematic dataset but no joining required 3
- Problematic dataset requiring joining. 4
- Dataset reshaped and deeply analysed 5.
- Meta-data description well executed (not copied from source) +1.

Marks will be lost for unclear or incorrectly submitted code.

Form your big idea (6)

- Demonstrate true understanding of what the data says using visualisation.
- Identify a strong story context. 1
- Identify a valid audience and an individual or group to which the story will be delivered. (1 mark)
- Identify risks and opportunities (1 mark).
- Story in a sentence (1 mark)

Visualisation(10 marks)

At least 4 charts in the story (2 marks each).

Chart:

1. Relevance
2. Complexity and simplicity
3. Effectiveness
4. Correctness

Story - bring it all together (2 marks)