INFS 5116 – DATA VISUALISATION

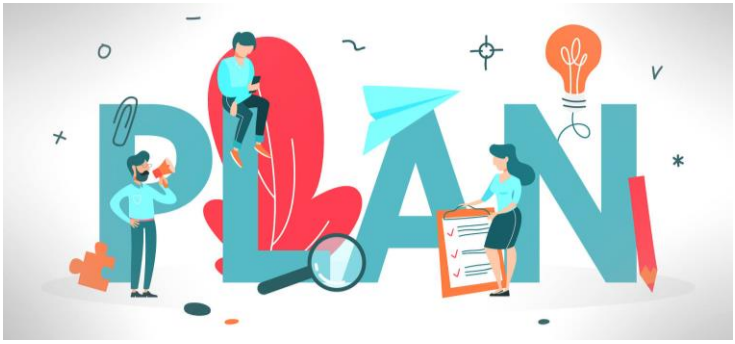
Visualisation Project Plan (SP5 2023)

Due 17 September by 11pm

General instructions:

- This assignment is worth 27% of your final grade and it is due no later than **11pm on Sunday 17 September**.
- You will need to submit your assignment via learnonline.
- The submitted assignment needs to be a **single file** in pdf file format.

Assessment aim:



The aim of this *Visualisation Project Plan* is to help you structure your proposed visualisation project and to perform some exploratory analysis of your data, focusing on graphics. It is an important preliminary work which will form a framework for your final submission.

Assessment task:

You need to choose your project topic, acquire relevant data and then go through initial stages of working with your data, in preparation for building a data story for the final assessment in this course.

Context – choosing your project:

- What are you curious about? What would you like to know? Choose a topic that is of personal interest to you.
- What is the purpose of your project going to be? What will you focus on? Who is going to be your intended audience?

[You will have already completed this step for your Padlet post.]

Data acquisition:

- What type of data will you need? What data is actually available?
- Where might you find the data you need?
- How will you obtain your data?

[You will have already completed this step for your Padlet post.]

Once you have identified the topic and data you would like to work with, you should perform data examination, transformation and exploration tasks.

Data examination:

- Physical properties (type, size, condition) of your data

Data transformation:

- Cleaning (identifying and correcting problem areas)
- Creating new variables
- Consolidating data sources

Data exploration:

- This is an important step in planning your project.
- Use visual techniques and exploratory statistical analysis (but only as appropriate) in order to 'see' the data for yourself.
- Graphics created at this stage may not be the ones you use in your final version. The idea here is to interrogate the data from many different angles to better understand it. Once you really know what you have, you may need to modify some of your plans.
- You do not have to visualise all of your data; using only some of the variables or using summarised data may be more appropriate for your purpose. Ensure you are clear on your purpose.

Submission instructions:

You will need to submit a single pdf file with the following structure:

Title page

Include the following:

- A (preliminary) title of your project based on your topic, with 'Visualisation Project Plan' as a subtitle;
- Your name and student ID number.

1. ***Introduction***

   Explain the context for the proposed project and the main question(s) you hope to be able to answer using visualisations.

   One short paragraph (1/4 page) is sufficient.

2. *Data Sources*

Describe the data source(s) that you plan on using for your visualisation project, where it comes from, how it was collected, its size, number and type of variables etc.

The data source(s) you choose should be large and rich enough to allow for visual presentation of data features from multiple perspectives and at different levels of complexity.

Note:

- The data you use for this project should not come from other courses you have completed or are currently undertaking. Rather, it should be identified specifically for the purposes of this project.
- Do not include screen dumps from Python or R, or a full data dictionary! Find a way to briefly summarise your data.

One paragraph (up to 1/2 page) is sufficient. You can use dot points.

3. *Data Preparation*

Perform and provide a summary of any data preparation tasks that are required before you embark on building your data visualisations, e.g. joining of files, identifying data problems (missing values or data errors) or creating new variables.

One short paragraph is sufficient. You can use dot points.

4. *Project Aims*

Propose some elementary, intermediate and overall level questions that may be addressed using your chosen data set(s). Identify the scope of the proposed visualisation project and possible problems that may occur along the way.

No more than one and a half page. You can use dot points.

5. *Data Exploration*

Present results of preliminary exploration of your data, which can include descriptive statistics and graphs aimed at helping you to get to know and understand your data before you begin building your main visualisations.

What do your preliminary results suggest in relation to the questions of interest identified in section 4? Comment briefly.

Four to six graphics (at least two types) plus one paragraph is sufficient. You can use dot points if you wish.

6. *References*

Provide a list of all references that you have cited in the project plan. [You need to provide in-text references as well].

Use APA referencing style.

Assessment criteria:

A grade will be allocated by applying the following criteria:

- The extent to which you demonstrate engagement with your chosen topic and provide sufficient context for what you propose to do **(25%).**
- Ability to describe your data sources and data preparation tasks concisely **(15%).**
- Quality of your graphics and how you use your graphics to show key features of the data **(40%).**
- How well you articulate your project plan and potential insights from visualising your data (visualisation for analysis) **(20%).**

Additional considerations:

In presenting your project plan, you need to strike a balance between technical details and describing the context, features of your data and its potential to address questions that you have posed for your project:

- Context is key – why did you choose your topic? What might be important or interesting about it to your audience? **Note:** Context is not about explaining why visualisation is considered useful for gaining insights from data. You will need to demonstrate this through your work and not just talk about it. Focus on your topic and be specific.
- What are you aiming to achieve with your project? What are you hoping to find out by working with your data? Again, be specific!
- What have you learned from your data exploration about key features of your data?
- Your code, full descriptions of all variables in the data set or a lengthy discussion of your data cleaning tasks is not helpful or interesting for the purposes of describing the plan for your project.
- **Less is more!** Aim to get to the point quickly and get your graphics to do most of the talking.