DATA VISUALISATION 2

Semester 2, 2024. Version 1.1

Version history
- 1.0 (1 September 2024): Initial release
- 1.0.1 (2 September 2024): Removed references to Tableau in "Expectations".
- 1.1 (2 September 2024): Added submission of five design sheets.

Due date: Sunday of Week 11, 13 October, 11:55 PM.

Introduction

Similar to the Data Visualisation 1 assignment, you will design and build an effective data visualisation for a specific domain of your choice. This task involves critically evaluating information within your chosen domain, developing your own visualisation, and telling a story. The domain can be broad or specific, depending on the availability of datasets and your interests. The major difference for Data Visualisation 2 is that you will use the Vega-Lite library to create maps and diagrams.

The aim of the assignment is to apply the data visualisation techniques discussed throughout the semester and demonstrate their use in an innovative context.

Requirements

The following requirements apply specifically to the Data Visualisation 2 assignment:

- The selected domain must be distinctly different from the one used in your Data Visualisation 1 submission.
- You are required to use the Vega-Lite library for creating maps and diagrams. You may use other libraries for creating diagrams that are not feasible with Vega-Lite, however, you must obtain approval from your tutor before doing so.
- The final deliverable consists of a publicly accessible web page hosted on your GitHub account. The JSON description for each Vega-Lite diagram or map must be easily accessible in the same GitHub repository and formatted for human readability.
- You may use the Pure.css library or any other JavaScript/CSS/HTML library or authoring software to create your web page.
- Your Vega-Lite visualisation must load quickly, meaning that the total size of downloadable data should be less than 1 megabyte. If this limit is too restrictive for your visualisation (for example, due to video files or large zoomable raster images) you must seek approval from your tutor before submission.
- Your web page must include at least one geographic map. If including a map is not relevant to your chosen domain, you can request an exemption from your tutor for this requirement.

The following requirements are copied from the Data Visualisation 1 assignment and also apply for this assignment:

- Why? The visualization addresses a particular need within a specific domain of your choosing. It must be targeted to users with specific information needs in this domain or be useful and relevant to individuals interested in the chosen domain.
- What? It visualises data relevant to the domain. Data can be of any kind.
- Who? Design your visualisation for the average Australian or Malaysian, depending on whether you are a student at the Clayton or Malaysia campus.
- The visualisation must transform data into meaningful information and provide insight that would otherwise be difficult or impossible to obtain without it.
- It must provide an appropriate level of interactive exploration.
- It needs to show some innovation. While it does not have to be wholly original, it cannot be a replica of an existing visualisation. It could include an innovative visualisation idiom, or an innovative exploration of an interesting dataset.
- Your submission must demonstrate the use of the Five Design Sheet methodology for sketching and planning the design of your visualisation.
- Your submission must demonstrate the use of Munzer's What/Why/How framework discussed in lectures.
- It must apply design principles discussed throughout the unit, such as data-ink ratio, storytelling, layout, typography and visualisation idioms with appropriate use of marks and channels.

Task Description and Schedule

1. **Plan and design**: By the end of Week 8:
   a. Select a **domain** that you would like to explore.
   b. Find relevant **data** that is publicly available.
   c. Discuss your domain, design ideas, and datasets with your tutor.
   d. Design your visualisation using the **Five Design Sheet Methodology**.
2. **Create the visualisation**: Weeks 9 to 11:
   a. **Homeworks of Weeks 9 and 10** consist of creating a map and a diagram with Vega-Lite. You may include improved versions of these in your Data Visualisation 2 assignment. However, note that Monash University policy prohibits submitting the same work for two assessments. Therefore, you must significantly enhance your homework map and diagram if you include them in this assignment.
   b. Create a web page and embed your **Vega-Lite** diagrams and maps. Use CSS to style typography and design the layout. Make your visualisation publicly accessible as a **GitHub** page.
   c. Take a picture of each of your Five Design Sheets and place them in a PDF document. Make the PDF document publicly accessible on your GitHub account.
3. **Submit** the following on Moodle:
   a. A URL to your visualisation on your GitHub account.
   b. A URL to a PDF document with images of your Five Design Sheets. Upload to GitHub too
   c. A maximum of 500 words about your visualisation via a form on Moodle, covering the following points:
      i. The **domain,** the **why** and the **who** of the visualisation.
      ii. **What**: The **data** (sources, authors, relevance, creation process, etc.).

iii. **How**: Give a rationale for choosing the specific idioms and explain how they help the users to achieve their tasks. Describe any special features of your visualisation, such as custom-built elements.

Expectations

Note the following is a verbatim copy from the Data Visualisation 1 assignment, except for the first paragraph about maps and a few references to Tableau.

Maps: The visualisation includes at least one map that uses an appropriate map projection and shows data with an appropriate idiom.

Data: The visualization must use accurate and reliable real-world data; it must not use fabricated or tampered data. The most recent available dataset should be used. Kaggle datasets must be combined with an additional source. Use existing and reasonably clean data; creating datasets or performing data wrangling are not topics covered in this unit and are not included in the marking rubric.

Topic: The following visualization topics are not allowed: World happiness, COVID-19, Airbnb, crime in Victoria, road accidents in Victoria, data science salaries, house prices in Melbourne, university rankings, car sales, and electric vehicle sales. In general, the topic must be related to Australia (for students in Australia) or Malaysia (for students in Malaysia). For example, tobacco consumption in the US is not an acceptable topic, but tobacco consumption in Australia or Malaysia is acceptable. Exceptions are possible, particularly if you have a specific area of interest. If in doubt, discuss your topic with your tutor.

Format: The entire visualization must be accessible from a single URL on your GitHub account. It must be viewable on a single web page that can be scrolled, and the full width of the visualization must be visible on a small laptop screen without horizontal scrolling. There should be no buttons (or other web links) that swap major sections of the web page, though you may use buttons to show and hide visualization elements.

Presentation not exploration: The goal of this assignment is to create a visualisation that communicates interesting information in an easily accessible and graphically engaging way using storytelling elements, layout principles, typography, and graphical design. The goal is not to create an expert tool for exploring and analysing a dataset.

Audience: The topic must be relevant to and easily understandable by an audience in Malaysia (for students in Malaysia) or Australia (for students in Australia). Avoid jargon and technical terms when possible. If specialized terminology is necessary, introduce and explain these terms. Avoid visualizations that require knowledge of statistics (e.g., understanding standard deviation or confidence intervals), as the general audience is not expected to have such expertise. If statistical concepts are necessary, introduce and explain them.

Quality, not quantity: Your visualization will likely contain between three and ten charts or diagrams. There is no set minimum or maximum number of charts. Instead, we are looking for carefully designed and annotated charts that, in combination with text, icons, and pictures, guide the user through an interesting story using layout principles, typography, and graphic design principles. Avoid pixelated images, non-informative graphical elements, or trivial text.

Complement your visualization with concise, informative, and grammatically correct narrative text that tells a story.

Interactivity: Integrate interactivity where it makes sense, but avoid adding interactive elements merely for their own sake.

Copyright: You are encouraged to use icons and other simple graphical elements where appropriate. As your visualization will be publicly accessible, it is your responsibility to ensure that you have the right to use data, icons, images, etc. Consult with your tutor if in doubt. You need to indicate the source and URL (if available online) for any external elements you use, such as datasets, photos, articles, academic papers, and other sources of information. There is no need to indicate the source and URL for simple icons, such as emojis, flags, or company logos.

Authorship: As your visualisation will be publicly accessible, you need to indicate your authorship and date your work. You may choose to release your work under a [copyleft license](), but this is not a requirement and doing so does not influence your mark.

Plagiarism: We will follow up on any kind of academic misconduct. For this assignment, you cannot integrate non-trivial graphics (such as diagrams, charts, information graphics, etc.) created by others. The submission must be your own genuine work. Simply plugging a dataset into an existing visualisation template and making tweaks to layout, typography, colour design, etc. is considered plagiarism.

Use of Artificial Intelligence

Generative artificial intelligence (AI) tools can be used for this assignment. Any use of generative AI must be appropriately acknowledged (see Learn HQ). Since grammar and style of text are part of the assessment, you are encouraged to use AI tools to ensure that the text included in the visualisation is grammatically correct and stylistically appropriate.

Marking Rubric

This assignment accounts for **25%** of your final unit mark. A detailed marking rubric is provided on the next page. A **late penalty of 5%** per day and a **cut-off of seven calendar days** apply. You must pass an interview hurdle during the studio in week 12. The mark will be 0 if the domain of your Data Visualisation 2 is not clearly different from the domain of Data Visualisation 1, or if the web page or any Vega-Lite elements are not publicly accessible on your GitHub repository.

Note: Only maps and diagrams created with Vega-Lite will be marked unless you have received prior approval from your tutor for an exception. If you include the map from Week 9 and the diagram from Week 10 in this assignment, they will only be assessed if they have been substantially improved. If no map is included, the maximum possible mark for "Visualisation – Idioms and Complexity" will be reduced from 10% to 5% unless an exemption has been granted by your tutor.

	HD (80–100)	D (70–79)	C (60–69)	P (50–59)	N (0–49)
Five Design Sheets 2%, 0 if created with digital tools	Five sheets completed, large variety of detailed sketches, creative and useful outcome.	Five sheets completed, large variety of sketches, useful outcome.	Five sheets completed, some variety of sketches, some useful outcome.	Five sheets completed, limited variety of sketches, limited outcome.	Incomplete submission, repetition in sketches, outcome difficult to implement.
Visualisation (a) Idioms and complexity 10% , 5% if no maps are included	A substantial number of appropriate standard and creative custom-built idioms. Demonstrate a high-level of understanding of the use of visual marks and channels.	A substantial number of appropriate idioms. Demonstrate a good understanding of the use of visual marks and channels.	Standard idioms (e.g., bar chart, line graph). Visual marks and channels are applied correctly.	A small number of standard idioms (e.g., bar chart, line graph). Slightly incorrect use of visual marks and channels resulting in difficult to read visualisation.	Inappropriate idioms, small number of standard idioms, incomplete visualisation. Visual marks and channels not applied correctly (e.g., hue channel for ordered attribute).
Visualisation (b) Layout, colour, figure-ground 4%	Balanced and symmetric layout clearly structured in columns/rows with good use of white space. All elements aligned with sight lines. Very clear visual hierarchy by using consistent colour and figure-ground.	Balanced and symmetric layout mostly structured in columns/rows with use of white space. Most elements aligned with sight lines. Visual hierarchy by using colour and figure-ground.	Somewhat balanced and symmetric layout not consistently structured in columns/rows. Arbitrary use of white space. Some elements aligned with sight lines. Some visual hierarchy.	Layout not balanced or not symmetric, no apparent layout structure. Some elements aligned with sight lines. Limited visual hierarchy.	Layout not balanced and not symmetric, no apparent layout structure. Most elements are randomly placed. No visual hierarchy by using colour and figure-ground.
Visualisation (c) Typography, 2%	Advanced typography, (non-standard typeface matching the topic, excellent readability of main text, appropriate line height, size, weight, colour, line length, alignment, spacing of text elements, etc.)	Consistent typography (standard typeface and attention to typography).	Standard typography with minor issues, but still easy to read.	Typography with some issues that impact readability of text (e.g., overuse of highlighting, poor or inconsistent spacing of text, centred text blocks, long text lines, etc.)	Inconsistency in fonts, sizes and weights. Typography with poor readability (e.g., inappropriate typeface, poor or inconsistent font size, weight, line lengths, etc.)
Visualisation (d) Storytelling, annotations, grammar and style 5%	• Clearly structured and engaging storytelling guides the reader through the visualisation. • Extensive use of informative annotations on diagrams and text outside of diagrams. • Correct grammar. • Writing is easy to follow and understand. • Metadata (data sources, author, date, etc.) is complete and well formatted.	• Storytelling guides the reader through the visualisation. • Use of informative annotations on diagrams and text outside of diagrams. • Correct grammar. • Writing is mostly easy to follow and understand. • Metadata is complete and well formatted.	• Some guidance of the reader through the visualisation. • Use of coherent annotations on diagrams and text outside of diagrams. • Some grammar issues. • Writing is not always easy to follow and understand, jargon or obscure abbreviations are used. • Metadata is complete and well formatted.	• Limited guidance of the reader through the visualisation. • Limited amount of annotations on diagrams and text outside of diagrams, with some incoherence. • Grammar issues throughout. • Writing is not easy to follow and understand, jargon or obscure abbreviations are used. • Some Metadata is missing or not well formatted.	• No apparent guidance of the reader through the visualisation. • Small number of incoherent annotations on diagrams and text outside of diagrams. • Major grammar issues throughout. • Writing is difficult to follow and understand, jargon or obscure abbreviations are used. • Metadata is incomplete or poorly formatted.
Description of domain, who, what, why, how 2%	Domain, who, what, why, how are appropriately and succinctly described.	Domain, who, what, why, how are described with minor issues.	Domain, who, what, why, how are described with major issues.	Domain, who, what, why, how are partially described.	Substantial aspects of domain, who, what (e.g., data source), why, how are not described.