

Monash University
FIT5147 Data Exploration and Visualisation
Semester 1, 2024

FIT5147 Data Visualisation Project

In this project, you are asked to create an **interactive narrative visualisation** that communicates some of your findings from the *Data Exploration Project*.

It is an **individual assignment** and **worth 40%** of your total mark for FIT5147.

The presentation described below is 3% of the 40%. The final submission of the code and report constitutes the final 37%.

Relevant Learning Outcomes

- Choose appropriate data visualisations.
- Implement interactive data visualisations using **R (Shiny)** or **JavaScript (D3)**.

Overview of the Tasks

1. Identify which **findings** from the Data Exploration Project you wish to communicate. You can be selective, and you do not need to share everything you have found. The visualisations and accompanying narration should reflect the answers to (one or more of) the questions in your Project Proposal.
2. Clearly define your **intended audience**. The audience might be the elderly, young children, your classmates, the general public, politicians or whoever you like, although the choice should make sense for the data and topic of choice. Your subsequent interactive narrative visualisation submission should be designed specifically for the intended audience.
3. Design an **interactive narrative visualisation** using the Five Design Sheet methodology.
4. Prepare a short **presentation** based on your five design sheets (one sheet per slide).
5. **Submit** the slides for your presentation in **Week 11**.
6. **Present** your presentation to your Applied Session in **Week 11 or 12**.
7. **Implement** your visualisation **using either R (Shiny) or JavaScript (D3)**. The use of other visualisation libraries and packages is subject to approval by your Teaching Associate (see the section “Notes on Implementation”). Note that you are not allowed to use R Markdown.
8. Write a **report** and export it to PDF.
9. **Submit your report and your source code** (see the section “How to Submit”) at **the start of the exam period**.

Presentation Details

The presentation is an opportunity to gain feedback on your designs from your teaching associates and peers. Prepare a **three minute** presentation based on your five design sheets. Your presentation should consist of **6 slides** covering:

1. An **Introduction**: Name, project title, aims and motivation (one slide)
2. Each of your **Five Design Sheets** (i.e., one sheet per slide).

The five design sheet methodology must be completely followed. All components should be legible and readable. For your sheets you may use pen and paper, and then digitise them (at high quality to ensure readability), or use digital pen technology. If you do import any images from external sources to your sheets you must reference the source appropriately in the relevant design sheet.

The design slides you present in person in your Applied session must match those submitted on Moodle. If you do not manage your presentation and go overtime, your Teaching Associate may stop your presentation, which may restrict the feedback you get on elements of your design.

Report Structure

Write a **report of up to 15 pages** (excluding cover page, table of contents, bibliography, and appendix - which together have a limit of 10 pages) that consists of the following sections:

1. *Project title*
Title of your narrative visualisation. This can be included in the cover page.
2. *Your identity*
Your full name, student ID, Applied session number, and Teaching Associate's name. This can be included in the cover page.
3. *Introduction [Approx. ½-1 page]*
A precise and succinct description of what findings and messages you wanted your narrative visualisation to convey, and who the intended audience is.
4. *Design Process [Approx. 3-5 pages]*
A description and justification of your narrative visualisation design process. This should briefly refer to each of your five design sheets (you must provide the 5 design sheets in the Appendix), and **justify your design choices based on the theoretical content of the unit** (throughout Weeks 1-12), for instance: describing consistency in design and interaction; reasons for a particular colour palette; referring to layout, typography, choice of visual variables, aspects of Munzner's what-why-how framework, genres of narration style, or aspects of the human visual system; etc. It is important that this section is not simply a description of which charts you chose in the different sheets, but must justify your ideas and the process that lead you to your final design choices. All visualisations in your design process should be relevant for your data and there should be a clear indication of how your choice of audience has influenced your design process.
5. *Implementation*
 - 5.1. Technical Implementation [Approx. ½ - 1 page]*
This section contains a high-level description of your implementation, including libraries used, references to external code sources such as templates, and reasons for any differences between your final design (that was justified and explained in Design Process) and the implemented design (described in the following section), if applicable. You are not required to explain the code in detail. You should briefly explain the reasons why your project was challenging (e.g., extensive wrangling was required, advanced use of D3, etc. - see Marking Criteria 4 for further information).
 - 5.2. Interactive Narrative Visualisation Implementation [Approx. 2-5 pages]*
Describe your final implemented submission, including screenshots where necessary. This section should clearly explain your final implementation and how its narrative portrays the data insights to your audience.
 - 5.3. Using the Implementation [Approx. 1 page]*

This section contains instructions for how to run, view and use your interactive narrative visualisation. This should emphasise any parts of your visualisation that may be easily missed by a reader (e.g., some interaction you have implemented that might not be immediately visible).

6. Conclusion [Approx. ½ page]

Summarise your findings and what you have achieved with your narrative visualisation. Reflect on what you have learnt in this project, including what in hindsight you might have done differently to improve the result and any future work that you would like to do.

7. Bibliography

Appropriate references of all resources that have influenced your work in IEEE or APA style (refer to the [Monash Uni library's guide](#) page for Monash citing and reference style). This should include any code templates, design influences and sources on theory, as well as references which influence any data insights.

8. Appendix

Include your five design sheets in the Appendix. Make sure you provide clear images and any handwriting is readable.

Your report should contain high-quality images of your narrative visualisation and five design sheets. If possible, avoid using a single screenshot of the entire page since the resolution might be low; instead, crop and explain individual sections of the page. It is also recommended that you export your PDF using a local word processor (e.g., Microsoft Word), as exporting your document as a PDF directly from Google Docs will result in low-quality images. Make sure you can read and understand the PDF document and its images at A4 size without requiring further enlargement.

Notes on the Design

- **The visualisation must be a narrative.** Elements of the design must tell the data story, using text and visualisation techniques to narrate how the data and the findings of the exploration process enable the questions about the topic of interest to be answered.
- **Your design must follow the Five Design Sheet process** and provide all the required information according to the 5dS template. The designs for Sheets 2-4 must be distinct from each other. The final design on Sheet 5 is expected to be a refinement of one of those sheets (or a combination of components from each).
- **Each design for Sheets 2-4 is expected to complete the narrative independently of the other designs.** No design sheet should just respond to one of the questions or findings narrated by the visualisation. No design sheet should consist of a single visualisation technique, e.g., one graph. Every design should resemble a complete solution with a clear layout that follows a particular narration style.
- In **Sheet 5** (and your final implementation) your final design should include a location to provide information about the data used in this work and a link to the original data source. It should also include how the interface will provide guidance to the audience on using the interactive narrative visualisation.

Notes on the Implementation

- Your implemented narrative visualisation should be based on the result of your Five Design Sheet process. It does not need to follow it exactly, however it should resemble the final design in Sheet 5. Small changes to your final design are allowed (e.g., layout, visualisation choices, navigation method, colour) but any such differences between your design and how it was implemented must be explained **and justified** in the *Implementation* section of your

report. Likewise, any differences between the final design in your presentation and that in your report in light of feedback to your presentation must be explained and justified.

- As a rule of thumb, all *visualisation* packages and libraries that are included in this unit are allowed for your implementation. **This includes, but is not limited to:**
 - For R Shiny: *ggplot2*, *ggmap*, *ggraph*, *Leaflet*, *Plotly*, *igraph*, *wordcloud*, etc.
 - For D3: *D3* itself, *Leaflet*, *MapBox*, etc. **Libraries which act as high-level wrappers for D3 are NOT allowed (e.g., *C3.js*, *dimple*).**

If you are unsure if a particular visualisation package or library is allowed, please discuss it with your Teaching Associate, or ask on the Ed discussion board.

- Tools or packages used for data wrangling, data cleaning, Shiny theming, HTML5 templating, CSS styling, etc. are not subject to these rules and can be used freely (i.e., for anything other than the visualisations themselves). However, you should not use server-side code, like *Django* or *node.js*, when implementing your design. Any data used for your DVP must be read from the files submitted with your code.
- For performance reasons, it is recommended that you pre-format all of your data files before loading them into R Shiny or D3. In other words, all data wrangling and cleaning steps (if any) should be performed outside of your narrative visualisation code. You are not required to include the code for data wrangling and cleaning as part of your submission. However, if you have done considerable work since your Data Exploration Project, then you should describe these steps in your DVP report (see Marking Criteria 3).

Marking Criteria

Data Visualisation Project: Presentation [3%]

- Quality of oral presentation (confidence, speed, voice) and quality of slides (legibility, design, images) [1%].
- Logical structure [1%].
- Choice of content (completeness, appropriate level, discussion of design and implementation alternatives) [1%].

Data Visualisation Project: Report and Source Code [37%]

When grading your submission, **all components** (i.e., the quality of your narrative visualisation design, technical implementation, and the written report) are taken into account:

1. Project Continuity [2%]

The degree to which the narrative visualisation presents data insights from questions proposed in your submitted Project Proposal and discovered during your Data Exploration Project. This explanation should be clearly outlined in your *Introduction*. Further exploration or improvements can be done post DEP, when necessary, but must be described and justified within the *Technical Implementation* section.

2. Design Process and Justifications [8%]

- a. Appropriate use of the Five Design Sheet methodology and evaluation of your alternative designs [5%].
- b. Relevance and quality of your justification of the final design based on data visualisation theory taught in the unit (Weeks 1-11). [3%].

3. Implementation of Final Visualisation Design [12%]

- a. Quality of implemented narrative: clear and appropriate signposting of messages and intended narrative [2%].

- b. Quality and appropriateness of implemented interactive visualisation: provision of appropriate context, clean and appropriate layout, suitable typography, attention to detail, good use of colour, appropriate description and reference to data source(s), and appropriateness for the intended audience [5%]
- c. Correctness and robustness, performance and usability of the implementation [3%].
- d. Code comments and code quality [2%].

4. Project Difficulty [10%]

The degree to which the visualisation project demonstrates sophistication and complexity in terms of its technical, theoretical and design implementation. Marks for this section will be allocated for the following:

- a. Sophisticated use of different data sources, in particular non-tabular data [2%].
- b. Dealing with very large datasets [2%].
- c. Advanced implementation of D3 / R (Shiny) [3%]
- d. Sophisticated user interaction (e.g., animation, linked interaction) [3%]

Note: Other technical, theoretical and/or design aspects will be considered for marks in this difficulty section. It is therefore crucial to make the marker aware of the complexity of your project by ensuring you mention and justify all elements in your written report.

5. Project Report [5%]

- a. Quality of writing, logical structure, quality and suitability of images, grammar/spelling, appropriate figure and table use with clear captions and in-text referencing (e.g. Figure 1, Table 1), appropriate academic referencing (APA or IEEE) and citations [1.5%].

Note: Please pay particular attention to your visualisations in your report. They must be clear and legible with readable font and font size, clear titles, legends and labelled axes in written English with no special characters such as underscores.

- b. Completeness (i.e., all the above sections should be submitted and complete) [3.5%].

Check Your Code!

Please be sure to check that your code runs correctly. If possible, check if your code works on other computers and operating systems. If you do not have access to another computer you can try checking via the Monash [MoVE platform](#).

If your code requires some steps for it to run, then be sure to make these very clear in *readme* notes for your marker and describe this in the *User Guide* section of your report. Your code must run on your marker's computer on the first attempt for us to be able to mark your submission. **If your submission does not run correctly, 5% (from the *Implementation* mark) will be instantly deducted from your grade.** If after some troubleshooting your grader is still unable to get the code to run, further deductions will occur as we will not be able to fully grade your interactive narrative visualisation.

Your code must also contain meaningful comments and be formatted and designed in such a way that it is easily readable and understandable.

Originality

As this is academic work, it must be original and must clearly indicate what elements were your work and what are based on someone-else's work. If you are including facts, data, opinions or any other written or graphical information from another source, you must cite the source and reference the bibliographic details for the source, using the APA or IEEE style guide. **This includes any third-party programming code or software you use in your data exploration and analysis.** If you directly quote or

replicate any material from a reference, you must do so in a manner appropriate to the APA or IEEE style guide. **Be sure to acknowledge sources that influence your code through your code comments and references in your bibliography. Do not copy complete designs from any external sources.**

If you are retaking this unit from a previous semester, please ensure you choose a completely new design and new visualisation code. The text, design and code cannot have been used in any other unit. **Likewise, you cannot reuse any code or written content that you have used in any previous assessment tasks for any units.** The only self-plagiarism that is allowed is the questions you set in your Project Proposal this semester and reusing some R code from your Data Exploration Project this semester (if you wish). **Any other written content from your Proposal or DEP may not be reused. It must be rewritten for your DVP.**

Generative Artificial Intelligence (Generative AI) software or systems like ChatGPT or Midjourney cannot be used for any part of this assessment task, including (but not limited to) generating written or visual components of your submitted work.

If your work is believed to not be original, due to potential instances of plagiarism, collusion with other students or contract cheating, your academic integrity will be reviewed. If any breaches of the academic integrity are confirmed, penalties may be applied to your assignment, the unit and/or even your enrolment in your course.

Submission Due Dates

- Submit your **presentation slides** to Moodle, due **Tuesday Week 11 (see Moodle for date and time)**. Presentations will take place during Week 11 & 12 in your Applied session. Attendance for both weeks is **mandatory**.
- Submit a **PDF report** and a **zip file containing your code and the data to Moodle**, due **the first MONDAY of exam period (see Moodle for date and time)**.

NOTE: All submission times are in Melbourne, Australia local time.

How to Submit

Presentation

1. **Prepare a PDF file** containing all five of your design sheets.
2. Name the file *StudentName_StudentID_Presentation.pdf*
3. Submit it via Moodle under Assessments/Presentation (3%).

Report and Source Code

1. **Prepare a PDF report** (max 15 pages) and a **ZIP file** containing the *source code* for your narrative visualisation and *any data files* that are required to run it.
2. Name the files using the following format:
 - a. *StudentName_StudentID_Report.pdf*
 - b. *StudentName_StudentID_Code.zip*
3. Submit both files via Moodle under Assessments/Data Exploration Project Submission (33%). These **must** be two separate files. **Do not put your PDF inside of the ZIP archive**. Note that only .zip is recommended, and you should **not** use other extensions such as .rar or .7z.

Notes on submissions:

- We cannot mark any work submitted via email or stored on file hosts such as Google Drive or Dropbox. Please ensure that you submit correctly via Moodle since it is only in this process that you complete the required student declaration, without which your work cannot be assessed.
- Your assignment **MUST** show a status of "Submitted for grading" before it will be marked. If your submission shows a status of "Draft (not submitted)", **it will not be assessed and will incur late penalties if submitted after the due date/time**. Note that this applies even if your file was uploaded to Moodle as draft prior to the due date.
- It is your responsibility to **ENSURE** that the files you submit are the correct files. We strongly recommend after uploading a submission, and prior to actually submitting on Moodle, that you download the submission and double-check its contents.
- Turnitin is used to help staff review the academic integrity for all submissions. For this reason, it will not be shared with students unless a student's work is under review.
- There is a maximum file size of 500MB. This is rarely hit by students in the unit, but it can cause an issue if your data files are very large. If you believe the limit affects you, check your zipped folder size and look to reduce the size of your data (e.g., by removing columns you are not using). If this is not possible, then **only then** can you consider storing your data remotely, e.g., via Google Drive, but be sure to test your code and provide access. Be sure to note this restriction in your code comments and any instructions, if needed. If access and instructions are not provided, your mark will be penalised.
- You **do not** need to publish your app on the web.

Late Submissions and Special Consideration

Presentation

- We encourage everyone to submit their presentation slides on time.
All Presentation Slides submitted late will receive zero marks.

Report and Source Code

- Assessments received after the submission deadline, or after the extended submission date for those with special consideration, will be **penalised 10% of the available total marks per day up to a maximum of seven days**. Submissions seven days after the due date will receive a mark of zero, and may not receive feedback.
- For information on eligibility for **Extensions** and **Special Consideration**, please refer to the relevant section on the Assessment page on Moodle.