Monash University

FIT5147 Data Exploration and Visualisation

Semester 1, 2024

Programming Exercise 2: R (5%)

Please carefully review all the requirements below to ensure you have a good understanding of what is required for your assessment.

---

1. Due Date

Tuesday Week 6, 9:30AM

---

2. Instructions & Brief

In this assignment you will demonstrate your capability in creating an interactive visualisation page with simple narrative elements using R Shiny.

It is an individual assignment and worth 5% of your total mark for FIT5147.

Relevant learning outcomes for FIT5147:

1. Perform exploratory data analysis using a range of visualisation tools;
6. Implement interactive data visualisations using R and other tools.

3. Details of task

3.1 Dataset

The data set used in the assignment is based on the Vehicle Road Crash Data released by the Victorian Department of Transport and Planning in January 2024. **However, we have wrangled the data for the purpose of this assignment, so you must use the data that we provide on Moodle.** The name of the data set is

"**Victoria_Accident_Data_FIT5147S12024PE2v2.csv**". This is a different version of the data to that used in PE1.

3.2. Design Brief

The task is to use R Shiny, ggplot2, and Leaflet to create a **data visualisation page** using the provided dataset and following a specified layout mockup and design requirements.

For this assignment you will need to create three visualisations 1) **VIS 1**, 2) **VIS 2** and 3) **MAP** combined into one layout.

The mock-up shown in Figure 1 shows 1) the expected layout of the visualisation page; and 2) the expected content in each section. In order to create this layout you are expected to use **fixedPage** (not a fluidPage) layout and are expected to position the visualisation elements in an appropriate number of rows and columns.

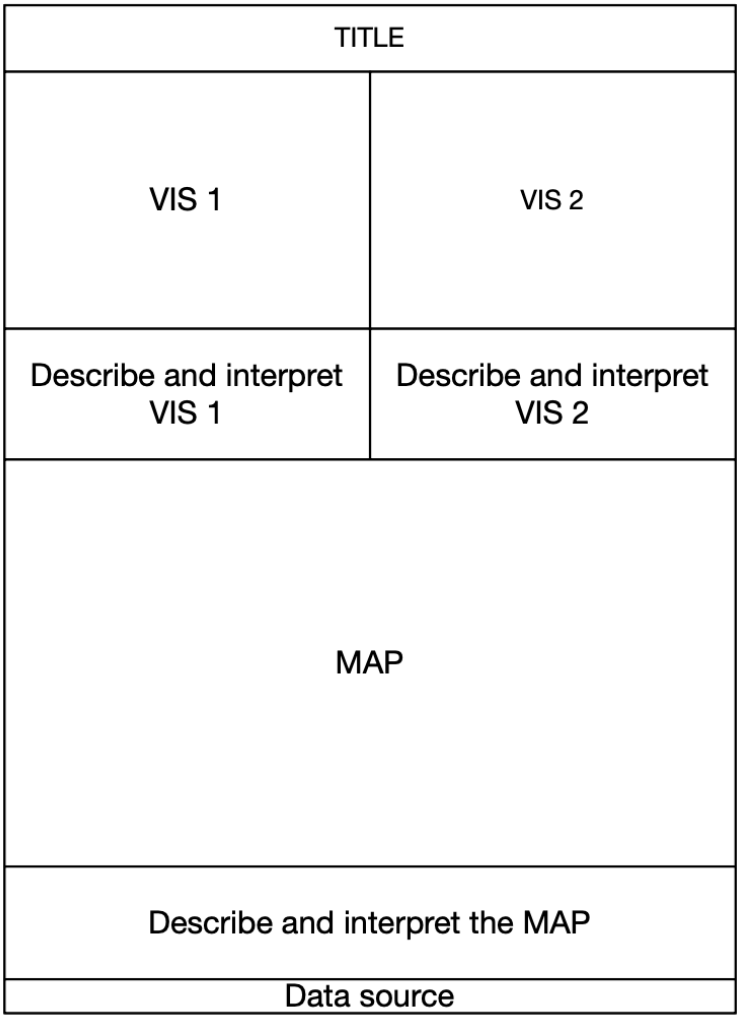

Figure 1. PE2 visualisation mock-up showing approximate position for your three visualisations (Vis 1, Vis 2 and Map) and three descriptions.

3.4 VIS 1 and VIS 2

For this assignment, these two visualisations should be **static** (not interactive) with the following requirements:

1. **VIS 1** should show the number of accidents that occurred under different *light_condition_desc* values, given different *speed_zone* categories, for example a stacked bar chart.
2. VIS 2 should show the number of accidents that occurred *each hour*, for each of the top 4 *speed_zones* in which accidents occurred, according to VIS 1*.*
3. The two charts must be created using ggplot2.

3.5 MAP

For the **map component**, you will create an **interactive** proportional map using *Leaflet*.

You must follow the following requirements:

1. Plot the accidents on the map using circle markers, with the following data mappings and design aspects:
    a. **Colour** should be mapped to a ***daynight* attribute**. You can choose any appropriate colour palette for the type of data. The *daynight* value (“day”, “night” or “dusk/dawn”) will need to correspond to certain values in the *light_condition_desc* attribute.
        “day”: “Day” light condition
        “dusk/dawn”: “Dusk/Dawn” light condition
        “night”: all other light conditions
    b. **Radius** should be mapped to the ***severity_rank* attribute** (you may need to scale the size so as to reduce the data overlap on the map, or use opacity, but some overlap is expected. The scale should be reversed to match the severity, i.e. 1 is the most severe)
    c. Provide a **colour legend** at the bottom right of the map.
2. Implement at least two of the following interactive features:
    a. Provide a **tooltip** for each circle marker on hover over that shows at least the *accident_date, accident_type_desc, light_condition_desc, road_geometry_desc* and *speed zone.*
    b. Add a **range slider** (a slider with min and max value) for filtering the circles by *severity_rank.* By default, all accident data should be shown on the map, i.e., the sliders’ values should be equal to the min and max values.
    c. Provide a **checklist** for filtering the *daynight* value (or make your colour legend interactive) such that circles corresponding to zero or more of the values may be filtered at any time.

3.6 Descriptions

The description text in the description boxes should **describe and interpret** the related visualisations. They should help the viewer see insights in the

data. The source of the data must be provided and described in the appropriate position.

NOTES:

1. You should avoid explicitly hard coding any of the data such as lists of speed zones, accident types, road geometries or times.
2. Whilst a map legend is usually expected for all data map elements, a legend for the size of the proportional symbols on the map is not a requirement for this assignment. (You may provide one if you wish)
3. Your design choices for VIS 1, VIS 2 and MAP will be included in the evaluation of your work. Make sure you meet the requirements of the assignment and choose any visual elements wisely.
4. No data checking or cleaning is required, but you may need to perform data transformations. You could use an R package such as *dplyr* (https://dplyr.tidyverse.org/) for this purpose, but it is not a requirement.
5. You may find these settings useful for creating your map (feel free to adapt them for your map needs): Centre point of -38.482461 latitude and 145.465783 longitude, zoom level 10 and map provider CartoDB.Positron.
6. Whilst there are no requirements on the use of specific colour palettes for this assignment, appropriate use of colour is essential. We recommend the use of appropriate *colorbrewer* palettes (https://ggplot2.tidyverse.org/reference/scale_brewer.html).
7. The text for your description text boxes, titles and any visualisation element content should be readable and legible to the viewer. Choose a clear font and appropriate font size for your submission and be sure to check your grammar and spelling.
8. Code is run through academic integrity checks. No collusion between students is permitted and any R code that is largely based on any third party code **must cite the original source** in comments within the R scripts(s), including webpages or social media messages. Otherwise your work may be considered to be plagiarising the code of others. No code provided by Generative AI can be used in this work. No part or description of this assessment task may be input into a generative AI system.

4. Assessment Resources

See the Assessments section on Moodle for the data:
Victoria_Accident_Data_FIT5147S12024PE2v2.csv

---

5. Assessment Criteria

The following outlines the criteria which you will be assessed against.

- Demonstrate the ability to read in and transform data using R [1%]
- Demonstrate the ability to create static visualisations in R using ggplot2 [1%]
- Demonstrate the ability to create a data map in R with Leaflet [1%]
- Demonstrate the ability to create an interactive visualisation in R with Shiny [2%]

As part of the grading process, **mandatory interviews to discuss your submission will occur during your tutorial in Week 7**. If you do not have a satisfactory interview, your mark and feedback for the assignment will be withheld.

---

6. How to Submit

Submit a zip file containing all files required to run your work. Name the zip file in this format:

PE2_[LAST NAME]_[STUDENT ID].zip.

Before submitting your assignment, please double check that your Shiny application runs correctly. To do so, clear objects from the workspace by clicking on the “Broomstick” icon on the top-right section of RStudio. Afterwards, make sure your application is still working by clicking the “Run App” button on RStudio.

The files that you need to include in your submission are:

- The one dataset supplied for this assignment
- R script(s) for the final Shiny application (you can use a single R script, or two scripts for UI and Server)
    - Have all required "library(xxx)" or "require(xxx)" statements at the beginning of your R files (you do not need the code to install the packages)
    - Use relative paths when reading your dataset (do not use absolute paths that refer to specific drives)

7. Late penalty

See the late penalty guidelines in the Assessments section on Moodle.