

# **Project Management & Data Analysis**

**2810ICT— Software Technologies**

School of ICT

Griffith University

Trimester 2, 2020

Part A due at the end of Week 6, Sunday 30th August, midnight.

Part B due at the end of Week 12, Sunday 17th October, midnight.

## **Assignment Description**

The assignment work for this course is a group-based (max 3 people) data analysis project that is split into 2 parts – design and implementation. There are several different datasets listed at the bottom of this document – you will need to choose 1 that your team will work on for both Part A and B. The overall goal for this assignment is to develop a simple data analysis and visualisation tool for a dataset. You will need to design and implement this tool, and it must provide a graphical user interface that will handle the analysis and visualisation of the data. Each dataset has a number of analysis and visualisation tasks that your interface must provide. In addition to the listed tasks, you must come up with 1 additional analysis/visualisation/insight that your software provides (you will need to come up with an appropriate task based on the dataset you pick).

### ***Part A – Project Management and Software Design (Due week 6, Sunday 30<sup>th</sup> August)***

In the first part, you will need to prepare a project plan that includes a Project Overview, Work-Breakdown Structure, Activity Definition and estimation and a Gantt chart for displaying scheduling & time estimation. This project plan should include sensible estimates for the various tasks required for both Part A and B (including estimates on preparing the project planning documentation). As you complete various components, you should put the actual completion time/dates on your Gantt chart to track how close you were to your estimates. For the subsequent stage of the assignment, you should revise your project plan with any additional details and continue to track your work.

You will then need to prepare a Software Design Document for one of the given data sets and related questions. You should start with a System Vision Document (this can be included in your design document). You should then produce a formal list of requirements that will need to be satisfied, some use cases for your software, a listing of system components and the related software design, and an early user interface design / wireframe (to be implemented in part B). There are templates provided for these documents.

### ***Part B – Data Analysis and Visualisation (Due week 12, Friday 9<sup>th</sup> October)***

For Part B, you will need to implement the software that you designed in Part A. Before you begin writing code, you should come up with a testing plan that will thoroughly test the various components of your system. As you begin your implementation, you can start testing these components as they are finished. You must present the results of your testing in a testing report. You should prepare a brief user manual that explains how to use your software and shows its features. Finally, you should use your software to prepare a report

that analyses the data over a 12 month period and presents the results in an executive summary. Your report should present the results from all of your required features for your chosen dataset, and should contain images from your software as well as some analysis and comments about the data.

*It is important to note that submission of this assignment is a requirement for passing the course. Late submissions will be marked according to Griffith University's assessment policy. 5% of the overall mark will be deducted for each business day late. After 5 days, no submissions will be accepted.*

## **Group Requirements**

This assignment should be completed in groups of up to 3 people. If you are unable to work in a group or have any specific issues, you should discuss this with your campus convenor. Make sure you include the name and student number of all group members in your documentation.

## **Submission Requirements**

This assignment must be submitted online via L@G under the assessment page. Only 1 submission per group is needed. Your submission should include;

### **Part A**

- A project plan (doc/pdf). You should use the provided template.
- A Gantt chart. You should embed this in your project plan (as an image or other), but also provide the original file.
- A software design document (doc/pdf). You should use the provided template.

### **Part B**

- All required .py files containing the data analysis program
- An instruction document/user manual (doc/pdf)
- A software testing report containing a brief testing plan and unit tests & coverage test results (doc/pdf). You should use the provided template.
- Updated project plan, design document and Gantt chart from Part A (doc/pdf).
- An executive summary covering a 12 month (doc/pdf). You should use the provided template.

## Marking

Each part of this assignment is worth 20% of your final grade (40% total). All group members will be given the same mark, so it is important that you all contribute and work together to complete the assignment. A full rubric for each part will be made available on the course website. Please note that all submitted assignments will be analysed by a plagiarism detector that is specifically designed for assignment submissions containing program source code. We will attempt to have all assignments marked within 2 weeks of the due date.

## Data Sets – Choose 1

### Victoria State Accident DataSet:

<https://www.kaggle.com/gaurav896/victoria-state-accident-dataset/>

#### Required Features:

For a user-selected period, display the information of all accidents that happened in the period.

For a user-selected period, produce a chart to show the number of accidents in each hour of the day (on average).

For a user-selected period, retrieve all accidents that contains a keyword (user entered) in the DCA\_CODE, e.g. animal, ped.

Allow the user to analyse the impact of alcohol time in accidents – ie: trends over days, the relationship with dark street lights, etc.

One other 'insight' or analysis tool of your choice

### Sydney Airbnb Data:

<https://www.kaggle.com/tylerx/sydney-airbnb-open-data>

#### Required Features:

For a user-selected period, report the information of all listings in a suburb

For a user-selected period, produce a chart to show the distribution of prices of

For a user-selected period, retrieve all records that contain a keyword (user entered), e.g. pool, pet.

Analysing how many customers commented on factors related to cleanliness (multiple key words may be associated with cleanliness).

One other 'insight' or analysis tool of your choice

**NSW Traffic Penalty:**

<https://www.kaggle.com/lilian/australia-nsw-traffic-penalty-data-20112017>

**Required Features:**

For a user-selected period, report the information of all penalty cases.

For a user-selected period, produce a chart to show the distribution of cases in each offence code

For a user-selected period, retrieve all cases captured by radar or camera based on offence description

Analysing the cases caused by mobile phone usage - ie: trend over time, offence code, and so on.

One other 'insight' or analysis tool of your choice

**New York Restaurant Inspection results:**

<https://www.kaggle.com/a16575659140/nycrestaurantinspectionresults-from-dohmh>

**Required Features:**

For a user-selected period, retrieve all inspection details.

For a user-selected period, plot the distribution of violations over the different suburbs

For a user-selected period, retrieve all violations that contain a keyword (user entered).

Analyse the cases related to animals, e.g., rats, mice or others, and their trend over time and distribution over suburbs

One other 'insight' or analysis tool of your choice