FIT1043: Assignment 1

Due: Sunday, 26 August 2018

Aim:

The aim of this assignment is to investigate and visualize data. In particular, we will identify trends, gain insights from how those trends change over time.

Data:

The data we will use contains Airport traffic data and comes from the Australian Bureau of Infrastructure, Transport and Regional Economics.

- The CSV file contains 30 years worth of monthly data listing international air traffic volumes in terms of passengers, freight and mail, to and from Australian airports.
- The file is available on Moodle and publicly available from data.gov.au

Hand-in Requirements

Please hand in a **PDF file** containing your answers and a **Jupyter notebook file (.ipynb)** containing your Python code to all the questions respectively.

- You can use Word or other word processing software to format your submission. Just save the final copy to a PDF before submitting.
- Make sure to include screenshots/images of the graphs you generate in order to **justify your answers** to all the questions. (You will need to use screen-capture functionality to create appropriate images.)
- We would like to see the Python code you write to format the data. Please **include a copy of your Python code** in your submission (submitting your Jupyter notebooks).

Python Availability

You will need to use Python to complete the assignment. You can do this by either:

1) running a Jupyter Notebook on a computer in the labs; or
2) installing Python (we recommend Anaconda) on your own machine.

Assignment Tasks

This assignment involves building a visualisation using Motion Chart and then answering a series of questions about the data.

Visualising Airport Traffic

The aim of the first part of the assignment is to build a Motion Chart showing passenger, freight and mail quantities over time for five different airports in Australia (**Melbourne**, **Sydney**, **Brisbane**, **Adelaide** and **Perth**). More specifically, you should:

- Generate a Motion chart with Passengers_Total on the x-axis, Freight_Total_(tonnes) on the y-axis and where the colour of the bubbles is given by the AustralianPort and the size of the bubbles is given by the Mail_Total_(tonnes). Please copy the generated motion chart and include it in your PDF file.
- Note that the CSV file is both too big and not in the right format to visualise directly, so you will first need to **aggregate** the data **using Python** at both the 'Year' and 'AustralianPort' level. Code to help you do this was discussed during the lectures or tutorials. You need to **select** only **five** Australian airports before the aggregation.
- You can then output the data as a CSV file (please note you do not need to submit your output CSV file), which you can use in order to create the appropriate visualisation (see Activity: Motion Chart – Motion Chart in Python in Alexandria).

Having graphed the data as a Motion Chart, answer the following questions (please provide a snapshot of your motion chart to support your answers in your PDF file if required):

1. Why are the values lower in 2016 than they were in 2015?
2. In which year did Brisbane have almost the same number of passenger numbers as Melbourne?
3. Which city has the largest number of international air passengers traveling through it in 2015?
4. Name two cities that have roughly the same number of passengers and freight quantity in total in 1987 (Please make sure the scale in x and y axes are linear when you answer to this question)?
5. Which city has the least number of passengers travelling through it comparing to all other four cities? In which year this value is maximum?
6. Has the number of passengers travelling through Sydney airport ever decreased from one year to the next? If so, when did it happen? Any idea why that might have occurred?

Good luck!