SIT 112 | Data Science Concepts
Lecturer: Dr Truyen Tran | truyen.tran@deakin.edu.au

ASSIGNMENT TWO Due:
Friday, 5pm 11th May, 2018

This is a **group assignment** which contributes 30% to your final SIT112 mark. The assignment must be completed and submitted to *Unit Site* by **5pm, 11th May 2018**.

The theme for this assignment, and the subsequent data science project, is to continue to explore data problems related to Australia. In this assignment, you are provided with sets of keywords. Each set contains three keywords. Each group will need to select **ONE** set of three keywords and perform two tasks:

- Crawling data from Twitter within Australia based your selected keywords

- Performing analytics tasks as instructed in the IPython notebook supplied.

The sets of keywords are as follows:

- computer, animal, movie
- game, bike, fiction
- food, holiday, song
- restaurant, singer, fun
- gift, science, actor
- study, actor, fashion

1. Assignment Resources

In the Assignment 2 folder, you will find the following files:

Filename	Description
Assignment2_notebook.ipynb	This is the IPython notebook which has been prepared and partially completed for this assignment. You will need this notebook to complete your assigned tasks.
Assignment2_instructions.pdf	This is the assignment instruction file which contains information for submission instructions.

2. Task Description

The description of assigned tasks has been given in details in the assignment notebook. **You are required to follow the instructions in the notebook to complete your tasks**.

3. Submission Instruction

Each group needs to submit only one solution. You are required to submit **five** files:

- Solution notebook: **[groupID]_assignment2_solution.ipynb.
  [groupID] is your group ID or group name.**

- An exported version of your output: **[groupID]_assignment2_output.html**

- Three json data files storing tweets collected for three keywords,
  named **[groupID]_[w1].json**, **[groupID]_[w2].json**, **[groupID]_[w3].json**
  where **w1, w2, w3** are three keywords you selected. For example: 10_game.json
  if your group ID/name is 10 and you choose the keyword "game".

END OF ASSIGNMENT DESCRIPTION