DEAKIN COLLEGE | **DEAKIN UNIVERSITY** in association with

Title: Data Science Project

Group:☐   Individual:☒

Graded out of: 25   Weight: 25%

Due date:	Week 11 Sunday	10 pm, 20th September 2020

Submission: Online:☒   Hardcopy:☐

Instructions: Students submit using the submission tool located under the Resources and Assessment section of Moodle.

Summary:

This Data Science Project aims to apply machine learning techniques to understand and visualize relationships in data. Our task specifications for this project will focus on two supervised learning tasks: linear regression and classification. You must demonstrate through skills acquired in describing the data, exploratory data analysis and prediction.

This project contributes 25% to your final SIT112 mark. It must be completed individually and submitted before the due date: **10pm, 20/09/2020, Sunday of Week 11.**

Plus, **compulsory presentation assessment (in-class) in week 12**. You are required to present based on your project in our ZOOM based online class, either on Wednesday or Saturday class. The 25 marks will be divided to 25*50% = 12.5% for your project solution, and another 12.5% for your presentation.

1. Data and Resources

In the Data Science Project folder, you will find the following files:

Filename	Description
Data Science_Project_instructions.pdf	This is the file contains the instruction to complete your project.
DC_project_notebook.ipynb	This is the Jupyter notebook which has been prepared and pre-filled for you to complete the programming task.

Relevant Content Weeks & ULOs

☐ Week 1	☐ Week 2	☐ Week 3	☐ Week 4	☐ Week 5	☐ Week 6
☐ Week 7	☒ Week 8	☒ Week 9	☒ Week 10	☐ Week 11	☐ Week 12

☒ ULO1   ☒ ULO2   ☒ ULO3

Please read the full assignment details that follow.

2. Task Description

A python note book file **project_notebook.ipynb** has been prepared for you to complete this task. Download this notebook, load it up and follow instructions inside the notebook to complete the task. *You are required to submit your solution in an Jupyter Notebook format as well as its exported version in html*.

3. Summary for submission

This project is to be completed individually and submitted online by the due date. You are required to submit the following files to the corresponding *Moodle page* in *SIT112 Data Science Concepts site*:

1. **[YourID]_project_solution.ipynp** : your IPython notebook solution source file.
2. **[YourID]_project_output.html:** the output of your IPython notebook solution in html.

END OF ASSIGNMENT DESCRIPTION