SIT112 | Data Science Concepts

Lecturer: Dr Sergiy Shelyag
Sergiy.shelyag@deakin.edu.au

DATA SCIENCE PROJECT

Due: 5pm, Friday 5 June 2020

This **project** contributes 25% to your final SIT112 mark. The project must be completed individually and submitted to *Unit Site* by **5pm, Friday 5 June 2020.**

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 your skills acquired in describing the data, exploratory data analysis and prediction.

1. Data and Resources

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

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

2. Task Description

A python notebook 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 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 Assessments/Data Science Project in SIT112 site:

1. Solution notebook: **[YourID]_project_solution.ipynb** : your jupyter notebook solution file;
2. An html exported version of your output: **[YourID]_project_solution.html.**

END OF ASSIGNMENT DESCRIPTION