

COMP1730/6730 S1 2020

- Project Assignment

# Answers to the following questions:

1. Choose part of the code you have written for this assignment and explain it.

All functions with code reference in assignment.py:

|  |
| --- |
| 1. def read\_dataset(filepath): *#Takes the file path of the dataset as input, reads the*  *data,and returns the dataset in a suitable format.* |
| 2. def get\_column(data, header): *#Get the data under a specific header from the dataset*  *.* |
| 3. def largest\_area(data): *#Get the largest area covered by the lake from the dataset.* |
| 4. def calculate\_average(data, header): *#to calculate the avarage of specific header da*  *ta* |
| 5. def average\_volume(data): *#Get the average volume of the lake from the dataset.* |
| 6. def most\_average\_rainfall(data): *#Get the time whose rainfall is closest to the aver*  *age from the dataset.* |
| 7. def hottest\_month(data): *#Get the hottest month on average from the dataset.* |
| 8. def area\_vs\_volume(data): *#Get a plot of the area of Lake George against its volume.* |
| 9. def lake\_george\_simple\_model(data, evaporation\_rate): *#A simple model to show how La*  *ke George fills and ebbs over time.This model is based on the assumption that the ev aporation of the lake is constant.* |
| 10.def lake\_george\_complex\_model(data): *# A complex model to show how Lake George fills*  *and ebbs over time.This model is based on the assumption that the evaporation is ch anged based on the environment.* |
| 11.def evaluate\_model(data, volumes): *#To check the veracity of a model. By compare the*  *model to the actual data.* |

1. Which model (simple or complex) is the best? Why?

Complex model is best. Run assignment\_test.py in IDE, results related to Model Evaluation are below：

2. 1233.4821871759684

# model error of complex\_model

# model error of simple\_model

1. 3069.411233404914

These floats indicate how bad the model is at estimating the real values. It’s called the model error and it should be equal to zero if and only if the model volumes are all exactly the same as the real volumes. Accuracy is inversely proportional to this value. Thus, complex model’s

return value is smaller than another. And We prefer higher accuracy ,so the complex model is best. 3. What assumptions did you have to make in order to solve Q4? Are they realistic? How would you improve it?

We assume that the catchment area and surface area of Lake George are both equal to the maximum surface area of the lake (which is calculated in Q2) and then refining them. Exactly, they are not realistic. We can write another function to accurately estimate these areas.

# The copy of the plot generated in Question3:

