COMP1730/6730 S1 2020
- Project Assignment

University IDs and Names:

201900800110　　吴岳东（contribution:50%）　　201900800103　　蔡宇翔（contribution:10%）

201900800173　　孔德政（contribution:10%）　　201900800148　　刘家溪（contribution:10%）

201900800145　　赵晟康（contribution:10%）　　201900800161　　耿嘉　（contribution:10%）

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.
```

2. 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：

```
1. 3069.411233404914        # model error of simple_model
2. 1233.4821871759684       # model error of complex_model
```

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: