

Write down your **regression equation** in basic part

$$y[i] = w[0] + w[1]*Temperature[i] + w[2]*y[i-1] + w[3]*y[i-2] + w[4]*y[i-3]$$

Briefly describe the **variables** you used in the advanced part

挪用 basic 的 variables 但是加上了 Precipitation，使用了溫度、前 3 期的感染人數和 Precipitation。

這是 basic 用來 training 的 data:

```
traindataA=pd.DataFrame({0:TemperatureA, 1:CityAshift1, 2:CityAshift2,3:CityAshift3, 4:CityA })
```

這是 advanced 用來 training 的 data:

```
traindataA_advanced=pd.DataFrame({0:TemperatureA, 1:CityAshift1, 2:CityAshift2,3:CityAshift3,4:RainA, 5:CityA })
```

Briefly describe the difficulty you encountered

第一點：沒學過 python

第二點：一開始使用 Gradient descent，要加入新的 variable 比較麻煩

第三點：去除 outlier 時不知道該去除到何種程度

第四點：不知如何實施 Linear regression 到程式上

Summarize how you solve the difficulty and your reflections

沒學過 python 導致要邊學邊寫程式，pandas,numpy 怎麼使用都要花時間去了解，一開始將全部寫在一個 .py 檔案裡，看其他人都用 .ipynb 才知道自己又浪費了很多時間。Gradient descent 寫到有點不知道在做什麼後就改用 OLS 導致大部分要重打。Outlier 的問題就一直試，因為結果也不一定。不知道如何實施 linear regression 所以看了很多網路上的教學，要手刻 linear regression 的原本就不多，其中大多都還只有使用一個 variable，導致我在 Gradient descent 和 Multivariate Regression 來回花了很多時間。所以我解決辦法的方式就是一直花時間和重來，其中覺得學到了很多東西畢竟學了很久。討論區大多很像不喜歡這項作業，應該是因為結果有點不確定性，我是滿喜歡的因為感覺學了很多。我自己使用 mape 測 validation data 都是 20 上下，希望結果不要爆炸就好。