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:

這是advanced用來training的data:



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。