



團隊測驗報告

報名序號：113147

團隊名稱：我不知道為什麼我在這

目次

01 - 分析摘要

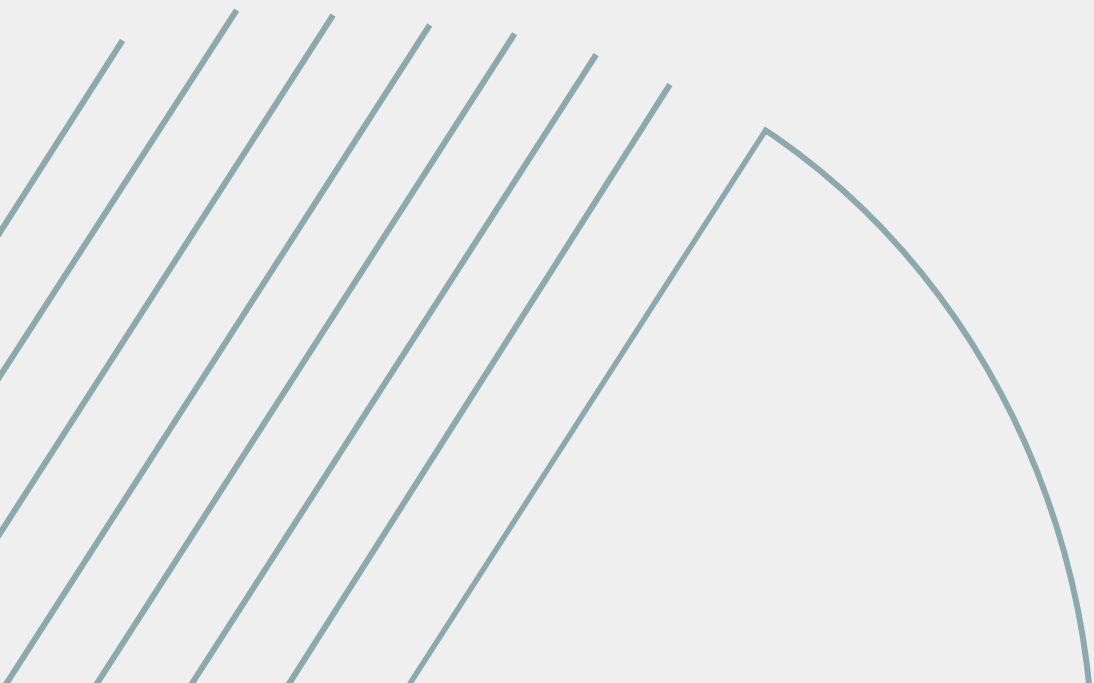
02 - 資料前處理

03 - 模型介紹

03 - 預測結果



分析摘要





分析摘要



經過資料觀察後，發現實驗數值隨著檢測次數增加呈現遞減狀態

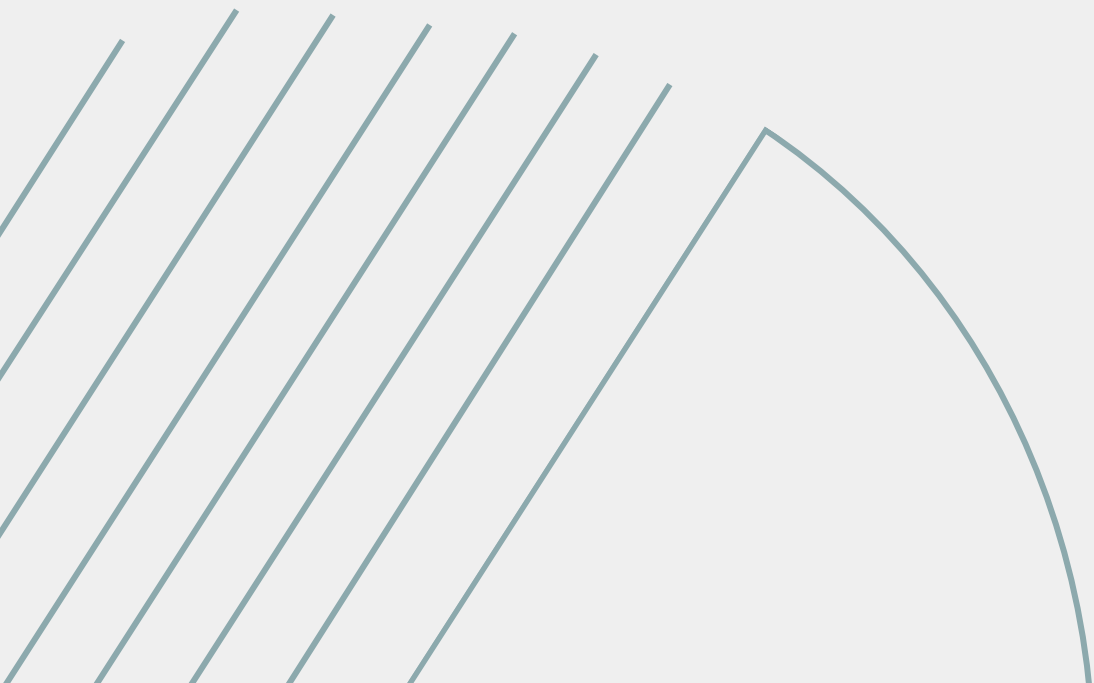
推斷數值下降的原因與時間變數相關

採用時間序列ARIMA進行模型訓練及預測





資料前處理



資料前處理

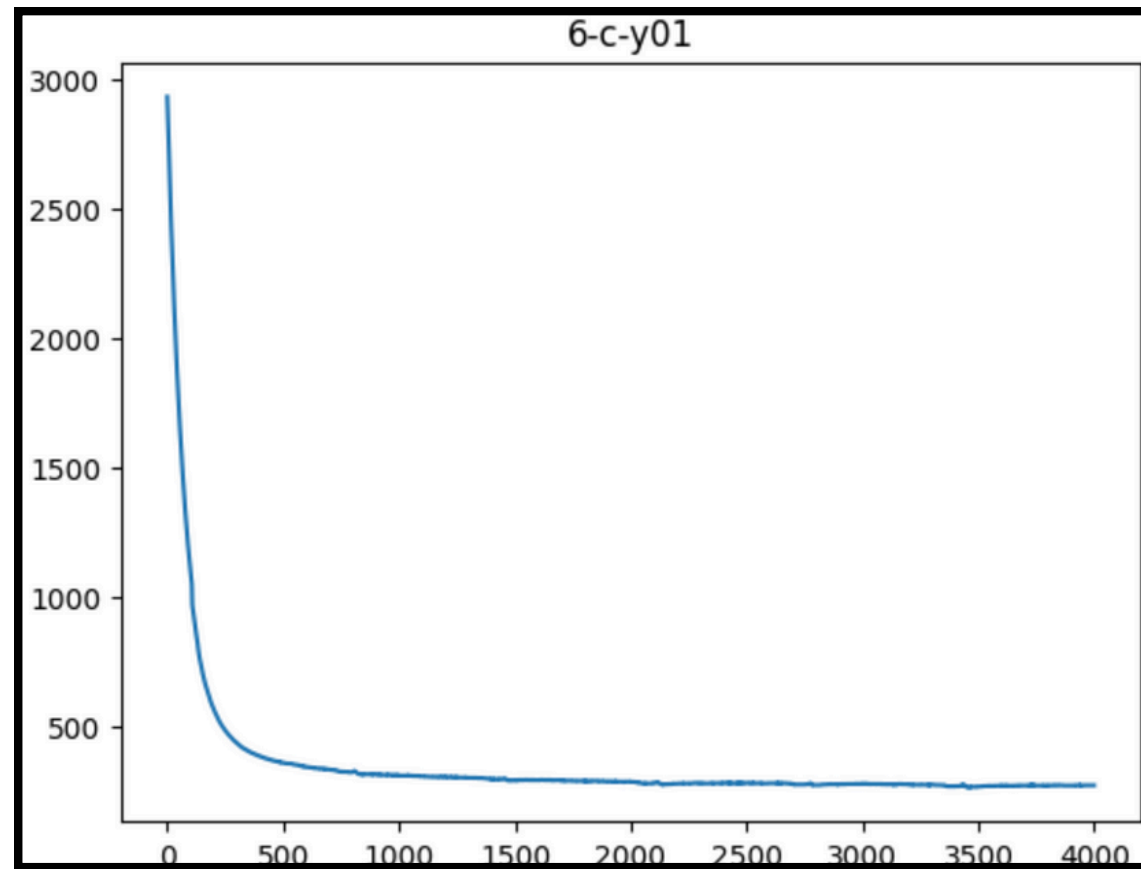
- 資料視覺化

- 隨著觀測次數增加，數值呈現指數遞減。
- 在使用a、b 電壓下，電流起始值較平穩。
- 在使用c、d電壓下，電流起始值起伏較大。
- 觀察起始值的盒狀圖，6-c 與 測驗集的統計量較為相似。

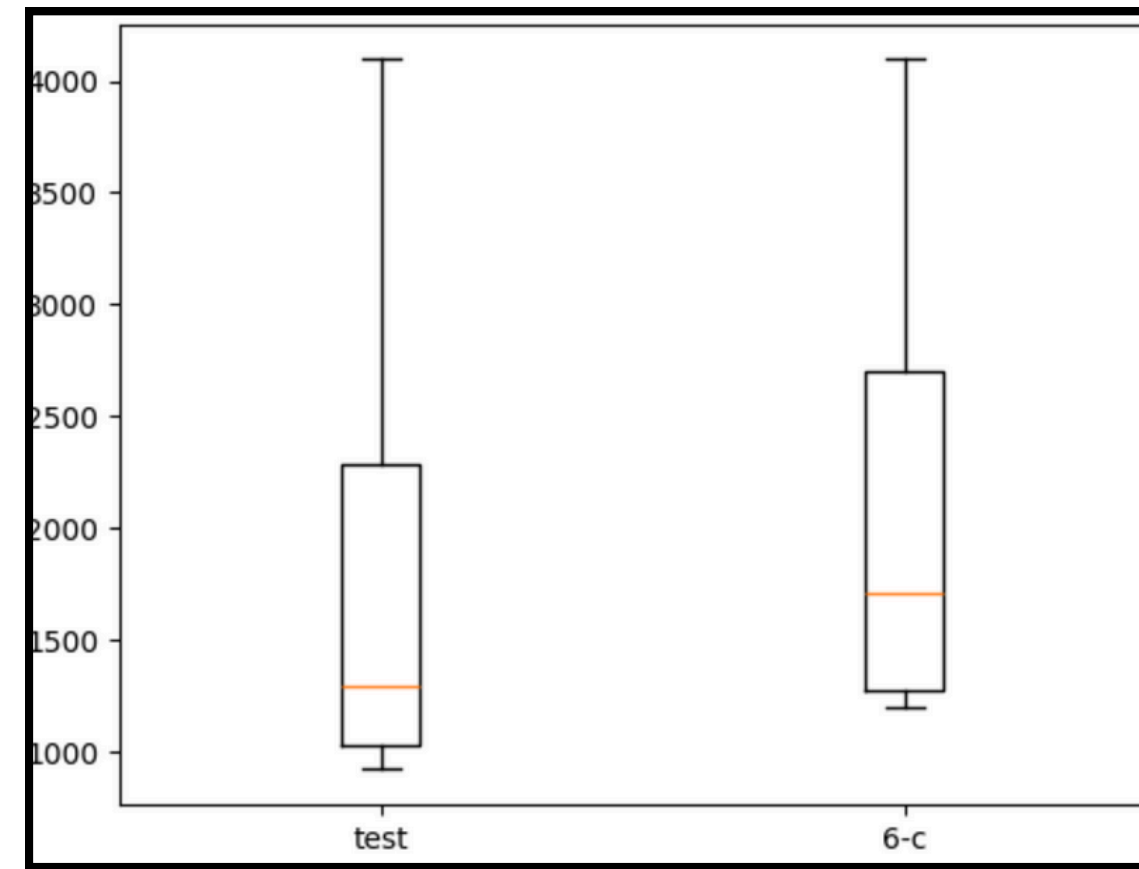
- 資料合併

- 將 y01~y10 欄位合併成一個欄位，進行時間序列分析。

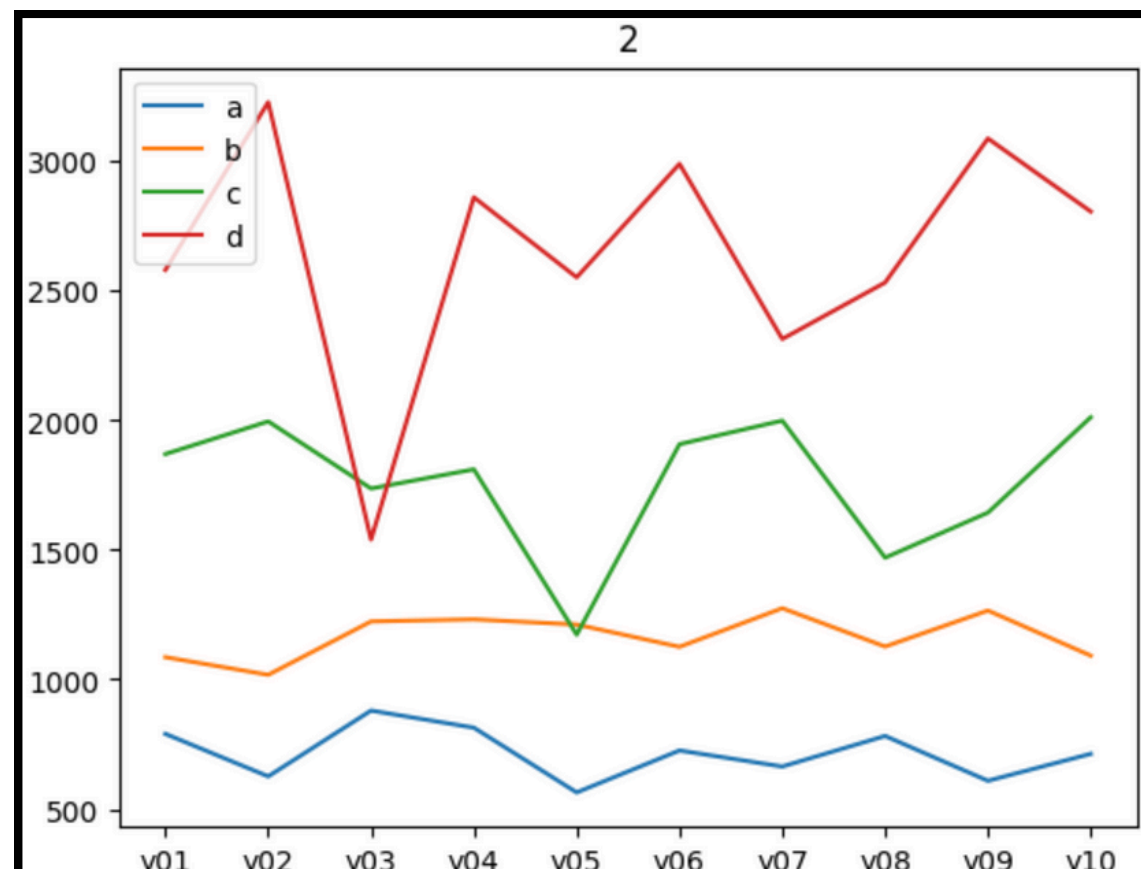
資料前處理



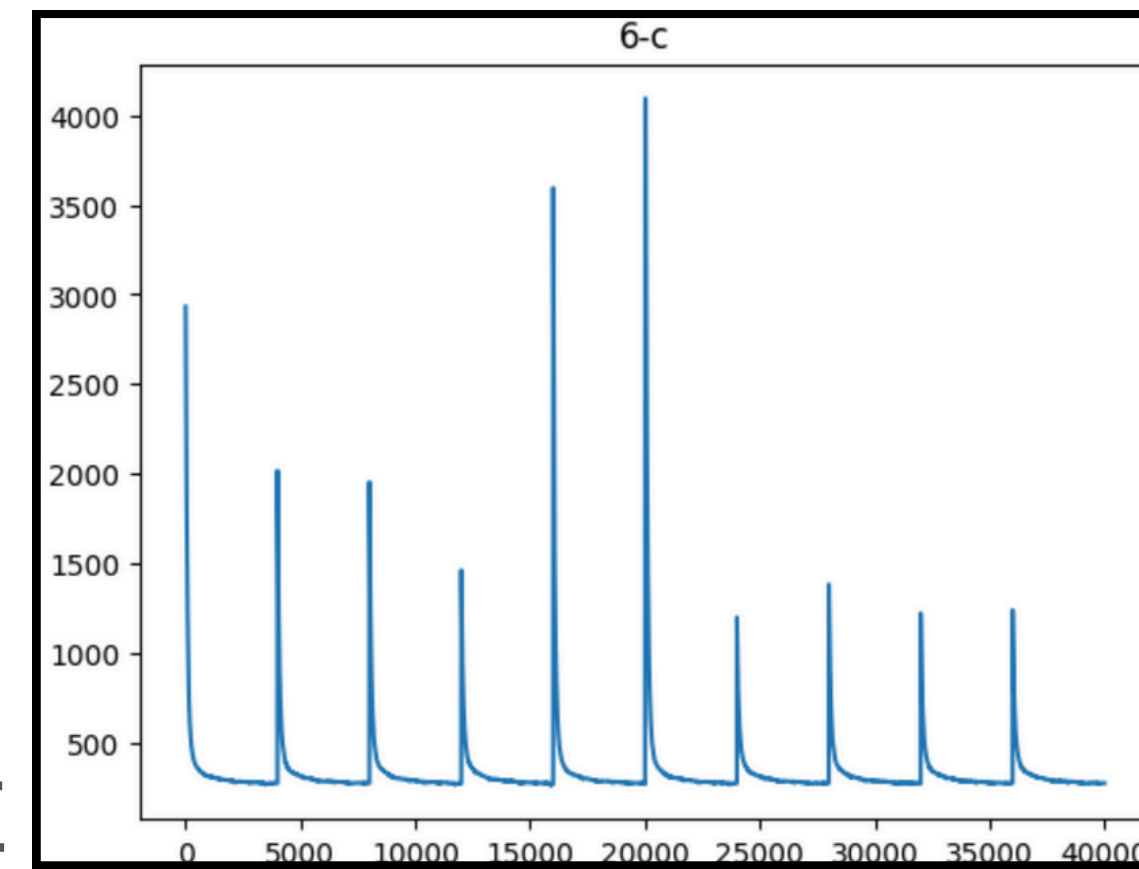
- 遞減趨勢



- 相似盒狀圖



- 不同電壓值



- 合併欄位



模型介紹







模型介紹：ARIMA

- ARIMA (AutoRegressive Integrated Moving Average)
 - 一種時間序列數據的統計模型，用於分析和預測未來數據趨勢。
 - 三個主要參數：
 - 自回歸項數 (p):


自回歸部分的滯後觀測值數量，反映序列當前值與前幾個值之間的線性依賴性。
 - 差分次數 (d)


使時間序列數據變為平穩的差分次數，為了消除趨勢。
 - 移動平均項數 (q)


移動平均部分的滯後誤差項數量，當前值與前幾個誤差之間的關係。
- 




模型介紹：參數選擇



- Augmented Dickey Fuller Test 進行平穩性檢定
 - 決定差分次數 (d)
 - 畫出 PACF 與 ACF 圖
 - 決定自回歸項數 (p)、移動平均項數範圍 (q)
 - 比較模型 AIC 與 BIC 數值
 - 決定參數組合 (p, d, q)
- 




模型介紹：參數選擇

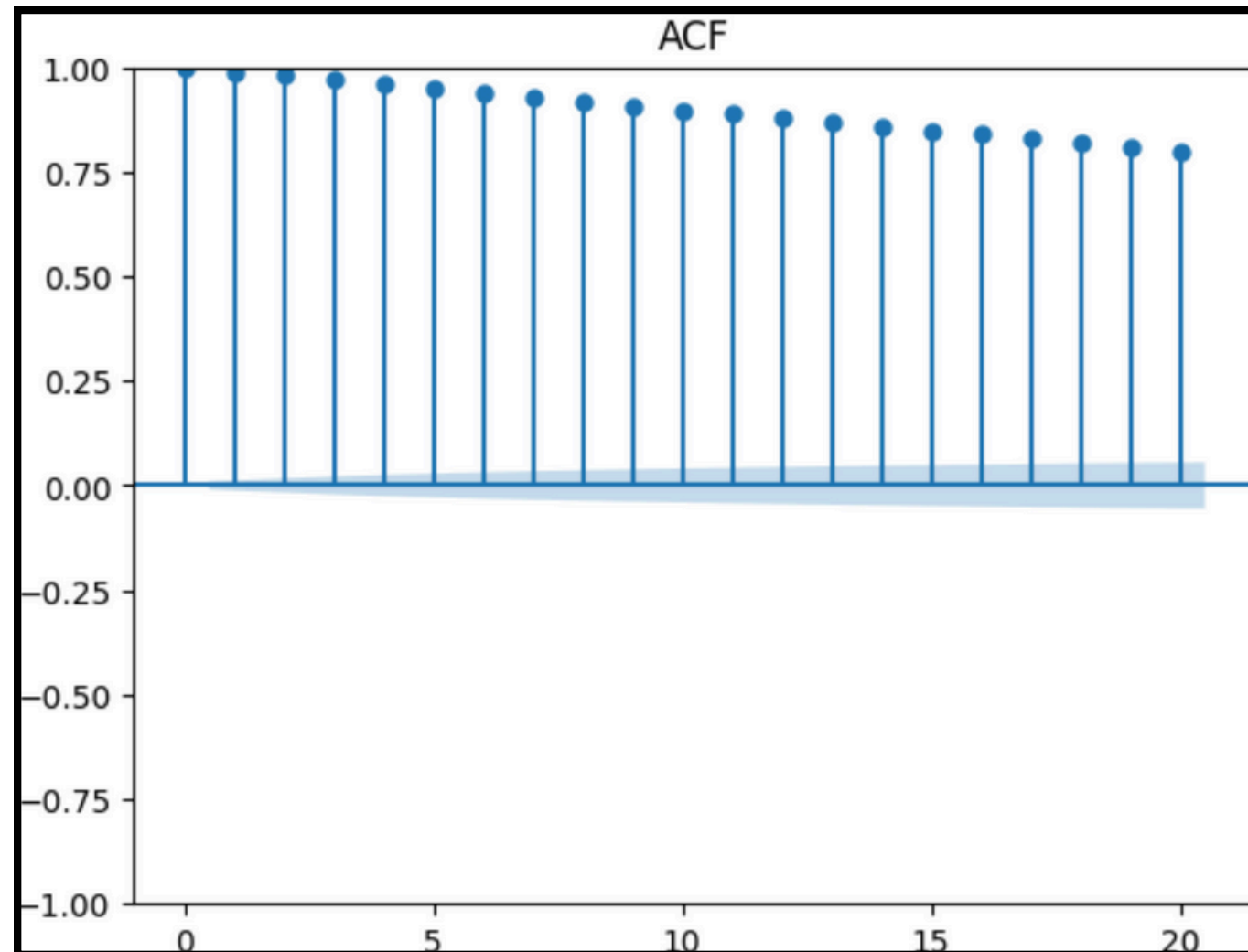


```
ADF Statistic: -15.68580062762325  
p-value: 1.453962048338028e-28  
Critical Values:  
    1%: -3.4305134970801805  
    5%: -2.861612261953491  
   10%: -2.566808462717237
```

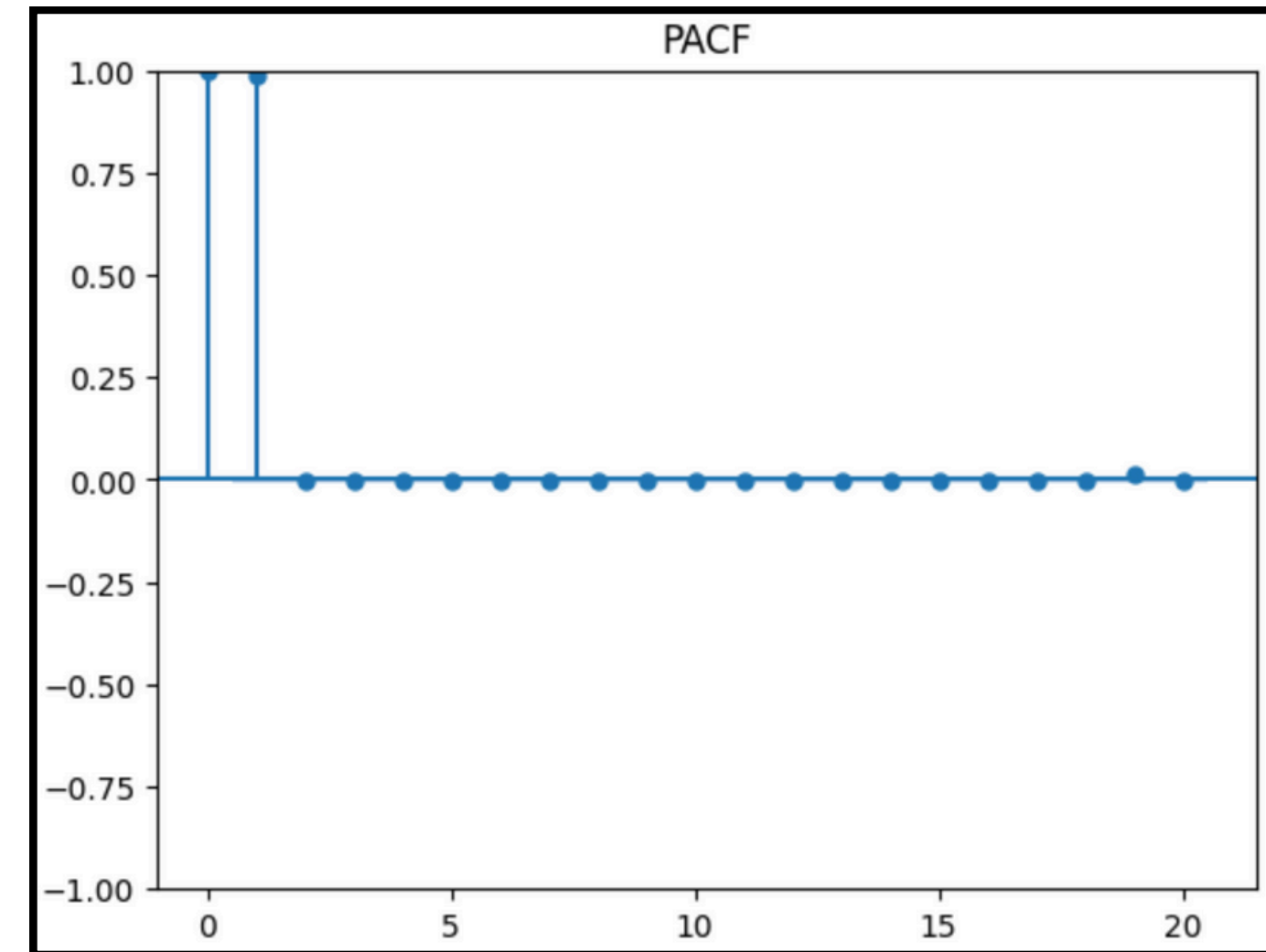
拒絕虛無假設：足夠證據顯示資料平穩，不需進行差分
($d=0$)



模型介紹：參數選擇



無明顯截尾，移動平均項範圍大
($q=0\sim 4$)



有明顯截尾，自迴歸項範圍小
($p=1\sim 2$)

模型介紹：參數選擇

```
ARIMA(1, 0, 0) AIC=29429.50, BIC=29452.83
ARIMA(1, 0, 1) AIC=28547.87, BIC=28578.98
ARIMA(1, 0, 2) AIC=28531.75, BIC=28570.63
ARIMA(1, 0, 3) AIC=28527.02, BIC=28573.67
ARIMA(1, 0, 4) AIC=28491.86, BIC=28546.29
ARIMA(2, 0, 0) AIC=28543.46, BIC=28574.57
ARIMA(2, 0, 1) AIC=28534.08, BIC=28572.96
ARIMA(2, 0, 2) AIC=28446.29, BIC=28492.95
ARIMA(2, 0, 3) AIC=28431.86, BIC=28486.29
ARIMA(2, 0, 4) AIC=28393.98, BIC=28456.18
```

```
Best ARIMA model based on AIC: ARIMA(2, 0, 4) Results: ARIMA(2, 0, 4) AIC=28393.98, BIC=28456.18
Best ARIMA model based on BIC: ARIMA(2, 0, 4) Results: ARIMA(2, 0, 4) AIC=28393.98, BIC=28456.18
```

選擇較小 AIC 與 BIC：選擇參數(2, 0, 4)模型



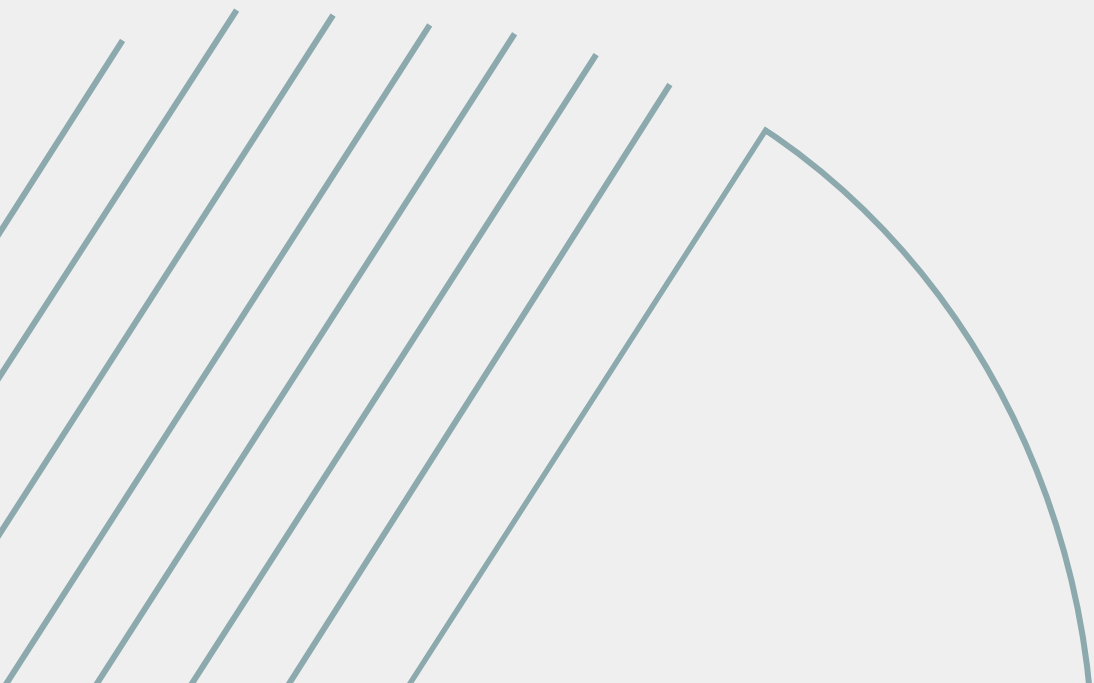
模型介紹：演算法



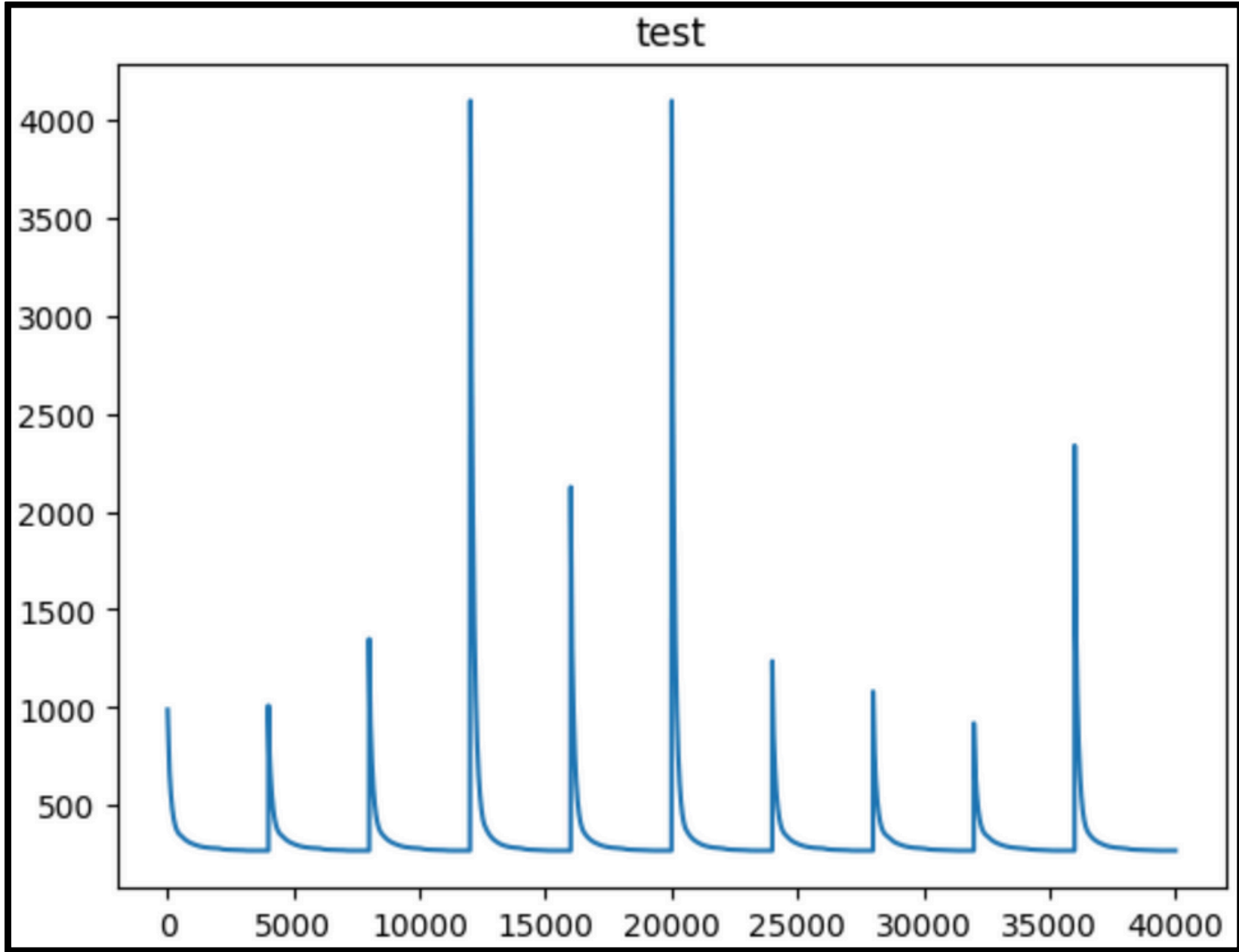
1. 將 6-c 訓練資料中 $y_{01} \sim y_{10}$ 與測試集任一欄位的前50筆資料進行合併。
2. 共進行六次預測，每次只取部分預測數值。
 - a. 第一次：參數 (1, 0, 0)，預測觀測點 50~4000，取 500 筆預測數值
 - b. 第二次：參數 (2, 0, 0)，預測觀測點 550~4000，取 500 筆預測數值
 - c. 第三次：參數 (2, 0, 0)，預測觀測點 1050~4000，取 1000 筆預測數值
 - d. 第四次：參數 (2, 0, 4)，預測觀測點 2050~4000，取 500 筆預測數值
 - e. 第五次：參數 (2, 0, 4)，預測觀測點 2550~4000，取 500 筆預測數值
 - f. 第六次：參數 (2, 0, 4)，預測觀測點 3050~4000，取 950 筆預測數值
3. 最後獲得測試集任一欄位的 3950 筆預測數值。



預測結果



預測結果



趨勢與訓練集相似

3991	270.7491951	270.7491626	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898
3992	270.7491951	270.7491626	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898
3993	270.7491951	270.7491626	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898
3994	270.7491951	270.7491626	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898
3995	270.7491951	270.7491626	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898
3996	270.7491951	270.7491626	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898
3997	270.7491951	270.7491626	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898
3998	270.7491951	270.7491625	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898
3999	270.7491951	270.7491625	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898
4000	270.7491951	270.7491625	270.748848	270.7487303	270.7488362	270.7488511	270.7491562	270.749149	270.7491918	270.7491898

由後10筆預測資料趨近於270
與訓練資料相符



END

