**書面報告**

1. **簡介**

利用爬蟲技術將電腦中的Youtube首頁的影片抓下來，顯示你想看到前\_名觀看人數的影片，再利用長條圖做觀看次數的比較

1. **動機**

現代人每天有大部分時間都拿來看影片，於是我選了一個最大的影片平台來分析首頁中，也就是你比較有興趣的影片中其中觀看人數更高的影片，利用這個分析就可以了解最新Youtube的趨勢，簡單的增加與他人的話題也不失看自己喜歡影片的原意

1. **使用到的技術**
   1. 將爬到的內容刪去不必要的資訊

( Youtube自己的版面配置、重複的影片、頻道連結 …等等 )

* 1. 觀看數將數字重新拼接，並判斷是否大於6位數使其分割次數不同，

例如：123,456 => 123456 和 123,456,789 => 123456789

* 1. 抓取所有影片的並分析出需要數量的超連結
  2. 擷取需要的影片後和流量輸出成長條圖
  3. 依照觀看量決定圖表大小

1. **使用到的package**
   1. BeautifulSoup
   2. request
   3. matplotlib
   4. re (re.split)
2. **Reference**
   1. 參考了助教的程式碼
   2. Lab中的長條圖
   3. <https://blog.csdn.net/programmer_at/article/details/77409507>

(Python中使用多个分隔符分隔字符串re.split)

* 1. 網路上關於beautiful soup的教學

如<https://kknews.cc/tech/3jjpj8g.html>

<https://ithelp.ithome.com.tw/articles/10196817>

1. **與Reference差異**
   1. Beautiful soup大部分結構跟助教的程式碼沒有差太多

但是都是有自己嘗試理解和再打的，資料的處理也都是用上課學到的方法加以運用

* 1. 示範好像沒有關於超連結的部分