

班級：醫資三 姓名：王殿馨 學號：409570298

```
[1] import matplotlib.pyplot as plt
import matplotlib.style
import pandas as pd
import numpy as np

matplotlib.style.use('ggplot')
```

```
[4] df = pd.read_csv('http://bit.ly/kaggletrain') #讀取檔案
df
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cummings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q

891 rows x 12 columns

0秒 完成時間: 上午10:13

首先，將會使用到的套件import進程式碼中。再來將檔案網址寫入讓程式碼讀取，並且測試讀取是否成功（成功的話資料將會output在下方）。

```
[13] df_sex_man = df[ df["Sex"] == "male"]
df_sex_man.shape[0] #計算有幾筆男性資料
```

577

```
[20] df_sex_female = df[ df["Sex"] == "female"]
df_sex_female.shape[0] #計算有幾筆女性資料
```

314

接下來讀取資料中性別的資料，並且分為男性、女性去計算在資料中總共有多少筆。



這裡我沒有使用Matplotlib繪製，因為一直顯示錯誤。就想到之前好像學過另一種繪圖的方式，就試試看了，沒想到成功把性別資料畫出來了。



年齡的分布我也用了剛剛的放法去畫，不過發現畫出來跟作業要求的完全不一樣。
但我也沒有找到其他方法，所以算是失敗了。

而最後兩個圖，我實在是不會，所以有上網找了參考資料

(<https://hackmd.io/@Go3PyC86QhypSl7kh5nA2Q/Hk4nXFYkK>)，也有試著從中找到一些方法去修正，不過沒有成功。(下有失敗的程式碼)

```
[x] cols = ['Survived', 'Pclass']
n_rows = 1
n_cols = 1
fig, axs = plt.subplots(n_rows, n_cols, figsize = (n_cols * 3.2, n_rows * 3.2))
for r in range(0, n_rows):
    for c in range(0, n_cols):
        i = r * n_cols + c
        if i < 5 :
            ax = axs[r][c]
            sns.countplot(df[cols[i]] hue = df['Pclass'], ax= ax)
            ax.set_title(cols[i])
            ax.legend(title = 'survived', loc = 'upper right')
plt.tight_layout()
```

學習心得補充：

這次的作業我個人覺得滿不容易寫的，常常腦中想的很美好覺得這樣應該行，但現實中就是寫不出來或是一直被錯誤擋住。

我也想過要直接把資料放在畫圖那裡，然後一次畫完，結果發現做不到。

```
[3] plt.subplot(2, 2, 1) # plt.subplot(列數, 行數, 圖形編號)設定第一張圖位置
plt.plot() # plt.plot(x軸資料, y軸資料,)

# 畫第二張圖，位於右上角
plt.subplot(2, 2, 2)
plt.plot()

# 畫第三張圖，位於左下角
plt.subplot(2, 2, 3)
plt.plot()

# 畫第四張圖，位於右下角
plt.subplot(2, 2, 4)
plt.plot()

# 儲存圖檔
plt.savefig("Titanic.jpg", # 儲存圖檔
            bbox_inches='tight', # 去除座標軸占用的空間
            pad_inches=0.0) # 去除所有白邊
plt.close() # 關閉圖表
```

暑假時參加的鐵人賽也有學過用python跟Matplotlib繪圖，但這次發現暑假學的東西果然不夠多，好多東西其實都還不會靈活運用。

經過這次作業，覺得應該要好好規劃時間來更加學習這一塊了！