

# 探討在台灣買房 該怎麼選擇

Author：張德昱

python資料處理分析&視覺化專題

# 大綱

- ➡ 資料蒐集
- ➡ 資料處理與分析：
  - 1.房價、薪資與人口數等數據
  - 2.利用資料分析各種原因
- ➡ 結論

# 資料蒐集

## ➡ 使用工具:

- Python
- Matplotlib
- pandas
- numpy
- jupyter

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
plt.rcParams['font.sans-serif']=['Microsoft J
plt.rcParams['axes.unicode_minus']=False
df=pd.read_csv('C:/Users/ASUS/Desktop/我的專題
```

- ➡ 主要是利用Google 網路  
爬蟲 抓取資料儲存為csv檔
- ➡ 或政府資料開放平台取得

```
url='https://pip.moi.gov.tw/V3/E/SCRE0105.aspx'
```

```
datas=[]
for i,tr in enumerate(trs):
    if i<2:
        continue
    else:
        data=[]
        for td in tr.find_all('td'):
            print(td.text.strip(),end='\t')
            data.append(td.text.strip())
        print(data)
        datas.append(data)
```

datas

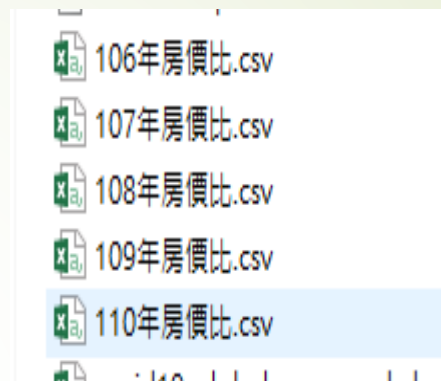
全國	36.27	-0.27	1.63	9.07	-0.06	0.41
新北市	48.48	-0.53	2.44	12.13	-0.12	0.62
臺北市	63.11	0.94	5.54	15.79	0.25	1.40

```
df.to_csv('各縣市房價比.csv',encoding='utf-8-sig')
```

# 資料蒐集處理

各縣市房價比.csv

	A	B	C	D	E	F	G	H	I	J
1		縣市	房貸負擔率	房貸負擔率	房貸負擔率	房價所得比	房價所得比	房價所得比(倍)	年變動值	
2	0	全國	36.27	-0.27	1.63	9.07	-0.06	0.41		
3	1	新北市	48.48	-0.53	2.44	12.13	-0.12	0.62		
4	2	臺北市	63.11	0.94	5.54	15.79	0.25	1.4		
5	3	桃園市	30.29	0.34	0.6	7.58	0.09	0.16		
6	4	臺中市	39.46	-1.17	1.05	9.87	-0.29	0.27		
7	5	臺南市	31.64	0.52	2.46	7.92	0.14	0.63		
8	6	高雄市	30.36	-0.04	1.28	7.6	0	0.33		
9	7	宜蘭縣	34.83	-0.23	1.86	8.71	-0.05	0.47		
10	8	新竹縣	33	1.51	0.81	8.26	0.39	0.21		
11	9	苗栗縣	31.73	0.55	2.86	7.94	0.15	0.72		
12	10	彰化縣	34.2	0.26	-0.45	8.56	0.08	-0.1		
13	11	南投縣	34.57	-0.13	4.53	8.65	-0.02	1.14		



```
In [1]: import pandas as pd

In [31]: df1=pd.read_csv('C://Users//ASUS//Desktop//我的專題報告//106年房價比.csv',encoding='utf-8')

In [12]: df2=pd.read_csv('C://Users//ASUS//Desktop//我的專題報告//107年房價比.csv',encoding='utf-8')

In [13]: df3=pd.read_csv('C://Users//ASUS//Desktop//我的專題報告//108年房價比.csv',encoding='utf-8')

In [21]: df4=pd.read_csv('C://Users//ASUS//Desktop//我的專題報告//109年房價比.csv',encoding='utf-8')

In [22]: df5=pd.read_csv('C://Users//ASUS//Desktop//我的專題報告//110年房價比.csv',encoding='utf-8')
```

```
In [25]: df11=pd.concat([df1,df2,df3,df4,df5])
df11

Out[25]:
```

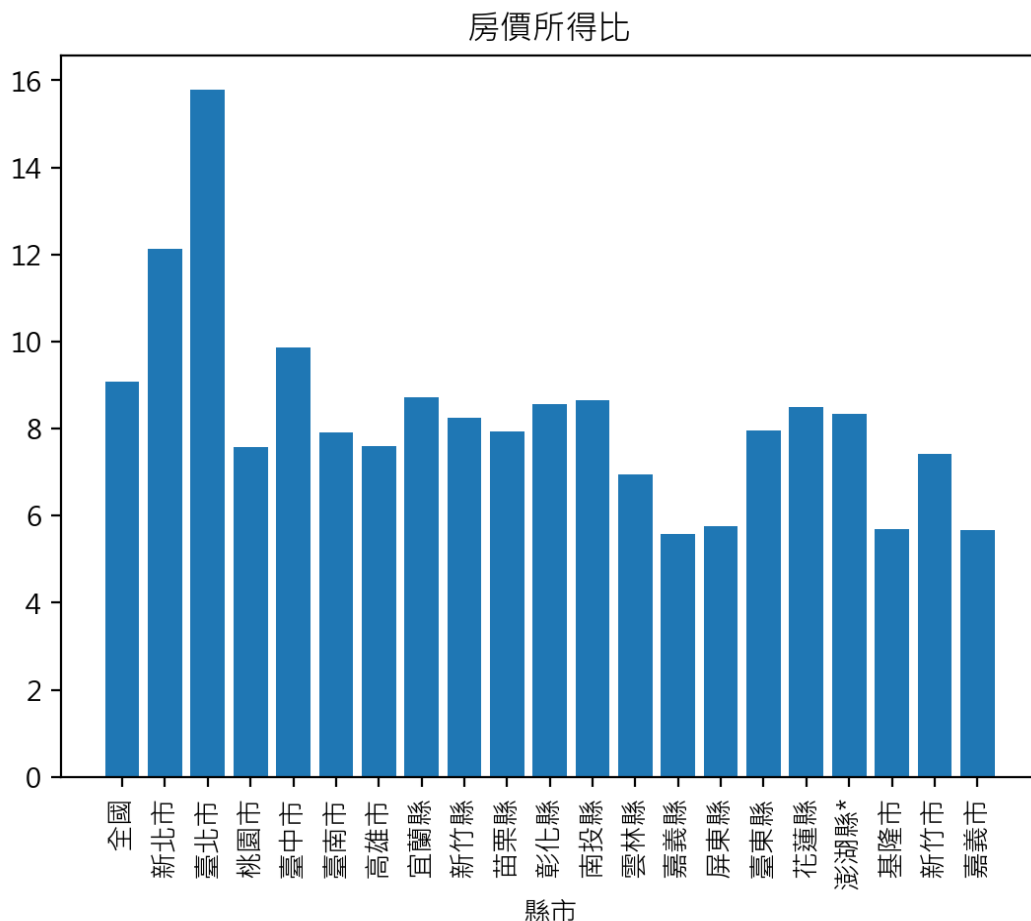
Unnamed: 0	年分	縣市	房貸負擔率(%)	房貸負擔率(百分點)	年變動值	房貸負擔率(百分點)	年變動值	房價所得比(倍)
0	0	2017	全國	38.90	0.88	1.76	9.48	
1	1	2017	臺北市	64.29	0.39	1.90	15.64	
2	2	2017	新北市	52.19	0.74	0.38	12.70	
3	3	2017	臺中市	39.60	1.0	0.92	9.63	
4	4	2017	新竹市	36.14	-0.4	3.71	8.79	
...	...	...	...	...	...	...	...	
16	16	2021	花蓮縣	33.98	0.39	2.84	8.50	

# 資料分析

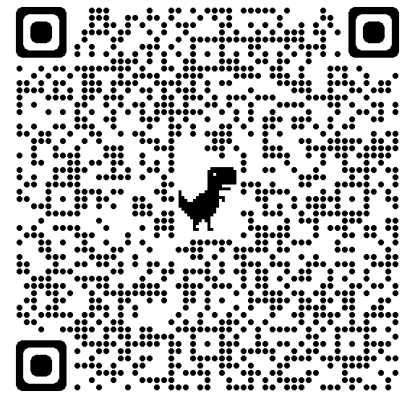
- 根據下圖所示
- 台北市買房難度位居第一
- 購買房子需要工作奮鬥15年

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
plt.rcParams['font.sans-serif'] = ['Microsoft JhengHei']
plt.rcParams['axes.unicode_minus'] = False
df = pd.read_csv('C:/Users/ASUS/21號張德昱python/各縣市房價比.csv', encoding='utf-8')
print(df.to_string())
df.info()
house = df.drop("Unnamed: 0", axis='columns')

fig = plt.figure(dpi=200)
plt.bar(data[:, 0], data[:, 4])
plt.xticks(data[:, 0], rotation='vertical')
plt.title('房價所得比')
plt.xlabel('縣市')
# fig.savefig('房價所得比圖表.png', dpi=fig.dpi, bbox_inches='tight')
plt.show()
```

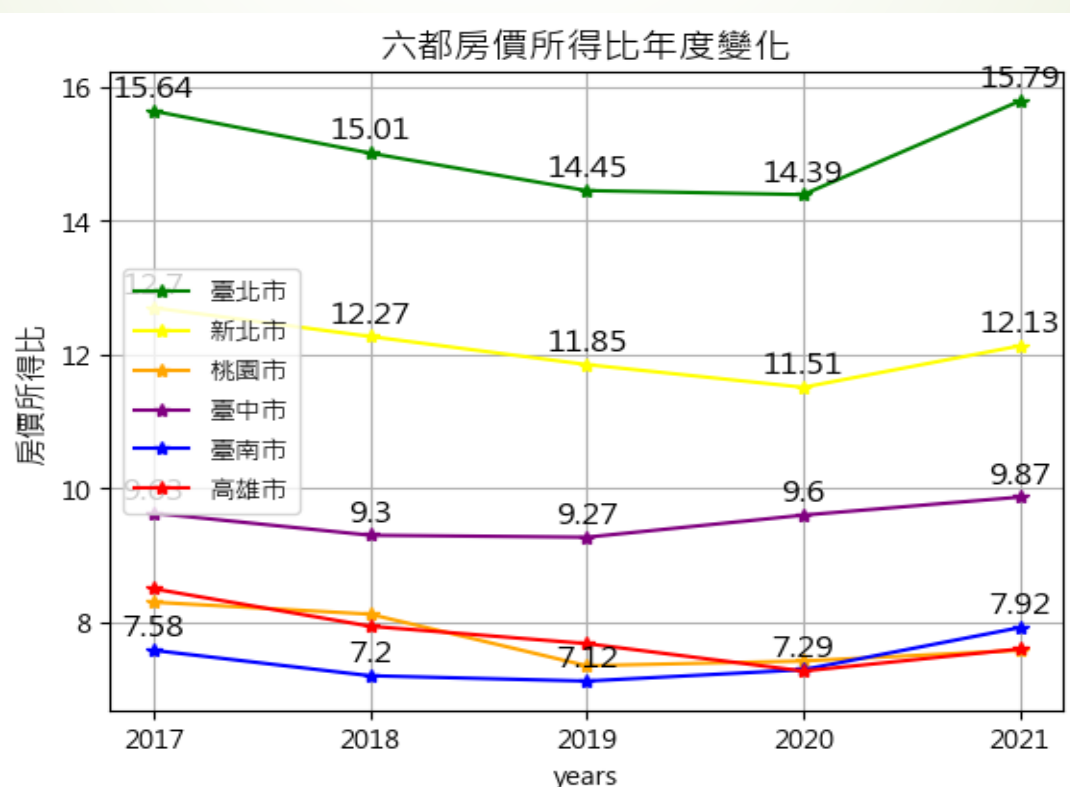


# 資料分析



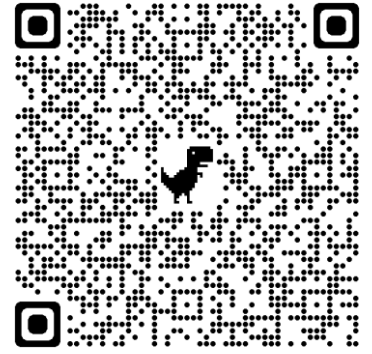
- 根據下圖所示
- 最近的房價所得比也是慢慢地往上升，儘管中間有小跌，但過了幾年很快又回升到原本的高度，顯示房價要跌下來是有一定難度的。
- 台北浮動較大，則台中則是緩緩爬升，台南也有小幅增加的現象。

```
fig = plt.figure()
plt.plot(taipei[:, 0], taipei[:, 5], '-*', color='green')
plt.plot(Xinbei[:, 0], Xinbei[:, 5], '-*', color='yellow')
plt.plot(Taoyuan[:, 0], Taoyuan[:, 5], '-*', color='orange')
plt.plot(Taichung[:, 0], Taichung[:, 5], '-*', color='purple')
plt.plot(Tainan[:, 0], Tainan[:, 5], '-*', color='blue')
plt.plot(Kaohsiung[:, 0], Kaohsiung[:, 5], '-*', color='red')
plt.legend(labels=['臺北市', '新北市', '桃園市', '臺中市', '臺南市', '高雄市'], loc='center left')
for a, b in zip([x for x in taipei[:, 0]], taipei[:, 5]):
    plt.text(a, b+0.1, b, ha='center', va='bottom', fontsize=12)
for a, b in zip([x for x in Taichung[:, 0]], Taichung[:, 5]):
    plt.text(a, b+0.1, b, ha='center', va='bottom', fontsize=12)
for a, b in zip([x for x in Tainan[:, 0]], Tainan[:, 5]):
    plt.text(a, b+0.1, b, ha='center', va='bottom', fontsize=12)
for a, b in zip([x for x in Xinbei[:, 0]], Xinbei[:, 5]):
    plt.text(a, b+0.1, b, ha='center', va='bottom', fontsize=12)
plt.gca().xaxis.set_major_locator(MaxNLocator(integer=True))
plt.title('六都房價所得比年度變化', fontsize=14)
plt.xlabel('years', fontsize=10)
plt.ylabel('房價所得比', fontsize=12)
plt.grid()
fig.savefig('六都房價所得比年度變化.png', dpi=fig.dpi, bbox_inches='tight')
plt.show()
```

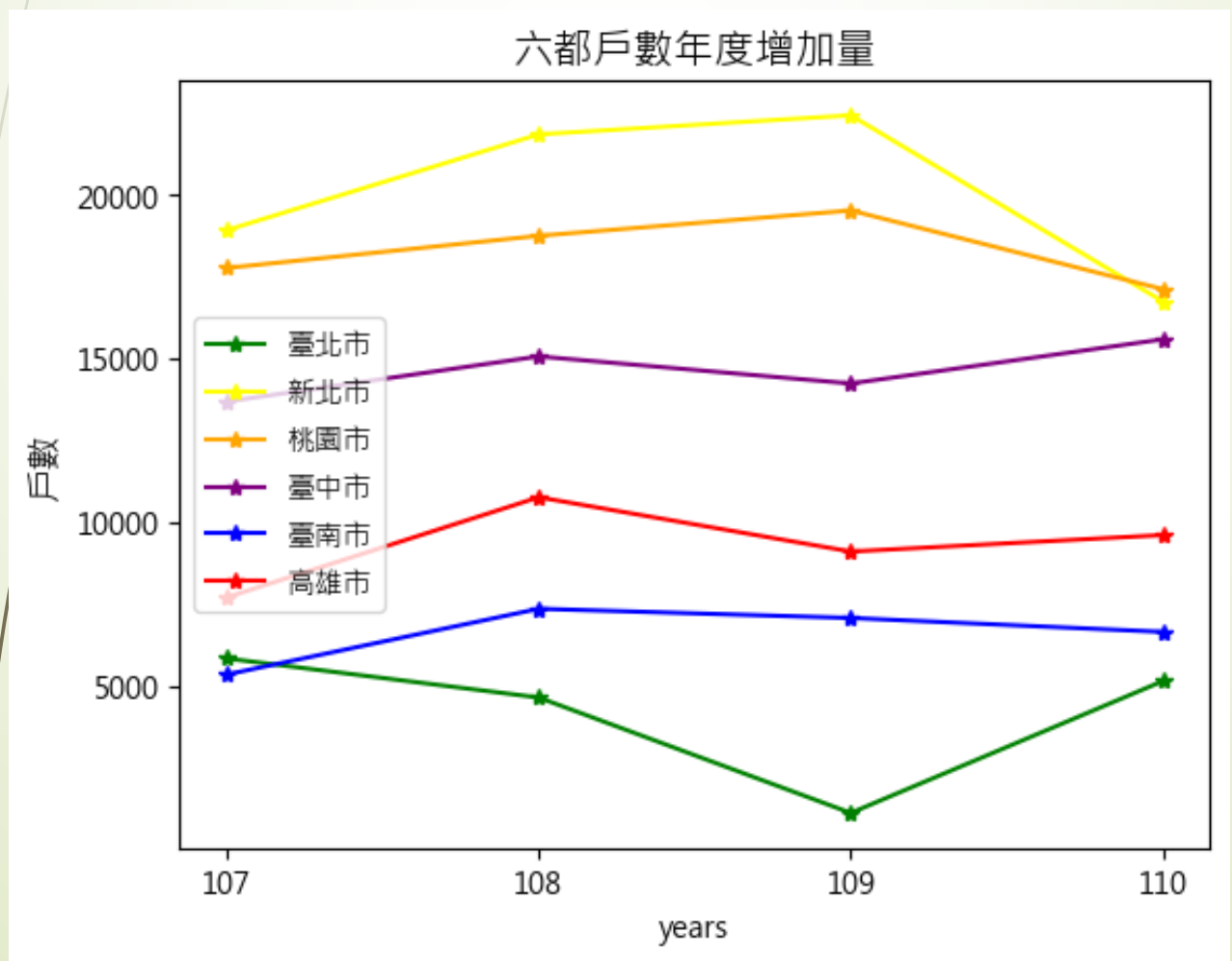




# 資料分析



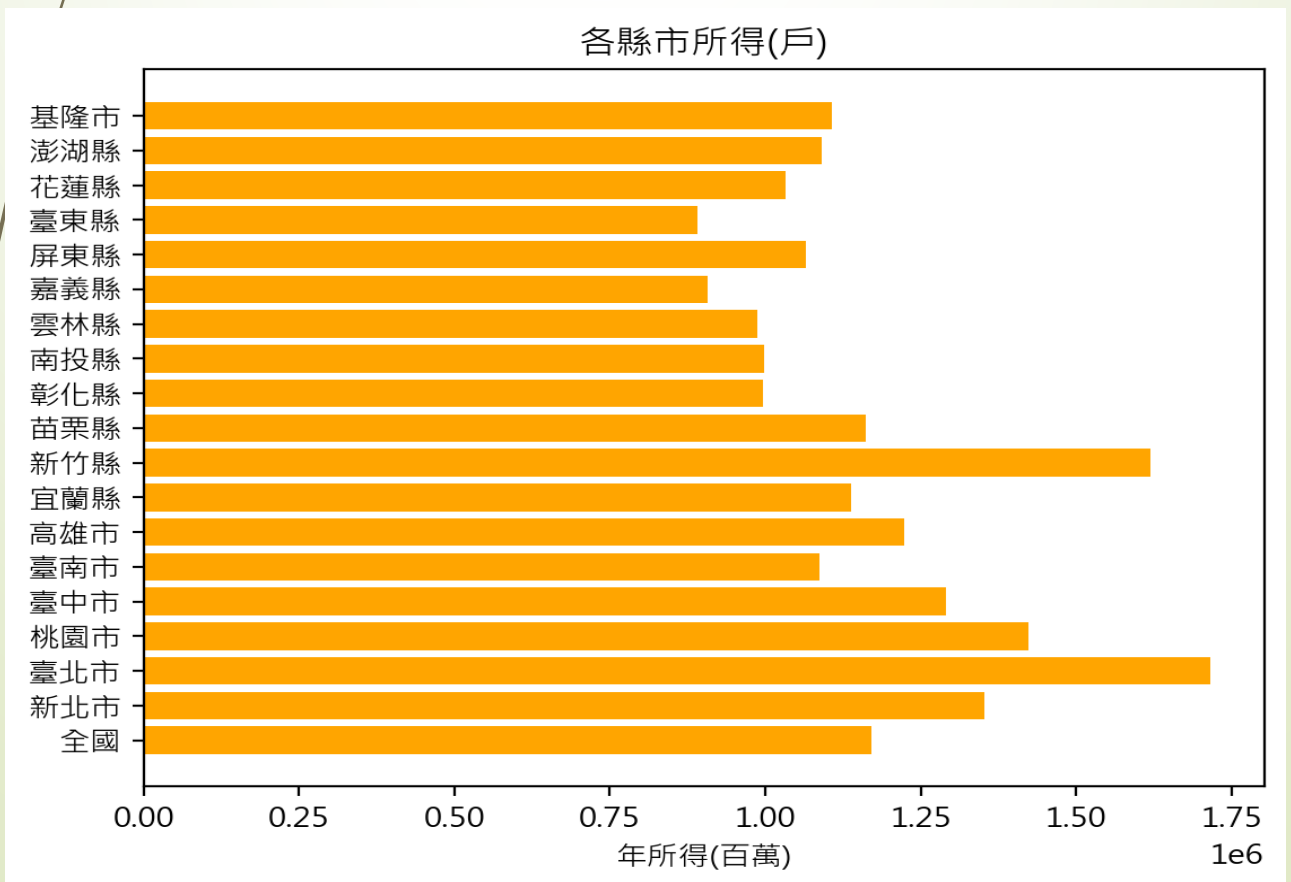
- 下圖可看出新北市每年的戶口增加量長期平均都位居第一，反觀台北市每年的增加量一直是不太多的情況下，可看出台北市的房價比過高，且疑似有過度哄抬的現象，導致大部分的民眾寧可購買鄰近的新北房子。
- 則排第二名的桃園市在未來也有可能像新北市房價一樣悄悄地往上升，就像樹林、龜山和林口。



# 資料分析

- 數據顯示台北市平均所得是位居第一，不過以現在房價高得如此嚇人且成長率也位居第一的台北，也許你能考慮在新北市買房並在台北工作。
- 另外可以看出桃園市的平均所得位居第三名，看出儘管房價所得比低於新北的桃園也能有這麼高的所得，那麼相較起來如果7年能在桃園買房，或許比12年才能在新北買房來得更好一點呢。

```
print(df.to_string())
data=np.array(df)
fig=plt.figure(dpi=200)
plt.barh(data[:,0],data[:,1],color='orange')
plt.title('各縣市所得(戶)')
plt.xlabel('年所得(百萬)')
fig.savefig('各縣市所得圖表.png',dpi=fig.dpi,bbox_inches='tight')
plt.show()
```

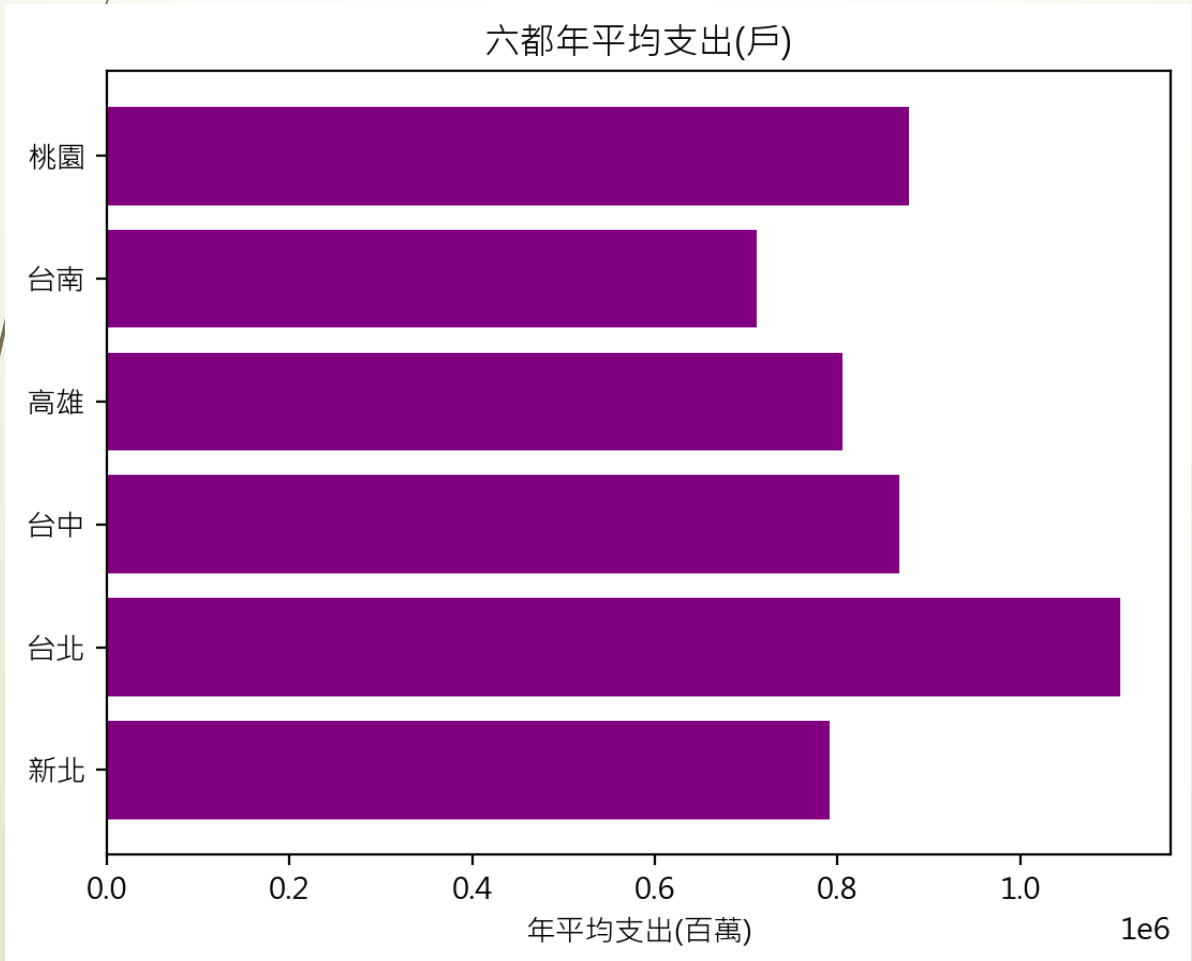




# 資料分析

- 在台北也反映出了跟其他縣市相比，支出的量遠遠高於大家，其中可能也包含著高額的房租費。

```
print(df.to_string())
data=np.array(df)
fig=plt.figure(dpi=200)
plt.barh(data[:,0],data[:,1],color='purple')
plt.title('六都年平均支出(戶)')
plt.xlabel('年平均支出(百萬)')
fig.savefig('六都年所得支出圖表.png',dpi=fig.dpi,bbox_inches='tight')
plt.show()
```



# 結論

- 台北市的確是一個高薪的地區，不過相對支出的量比其他縣市更多，購買房子也是一個問題，因為你得工作15年才能夠在台北買棟房子，如果價格一直居高不下，台北的房子也只能望塵莫及或努力賺錢。
- 如果沒有非要在台北買房的話，新北跟桃園會是你的好選擇，畢竟不像台北可能是空有價格但實質上需求量並不大，能從戶數年度變化量看出。
- 而台中在各數據裡面也有不俗的表現，在六都中屬中高所得，且每年戶數增加量大，現已為第二大直轄市
- 所以如果要買房：
- 依上述房價所得比、縣市戶數增加量分析  
我認為桃園市首選、第二新北、第三台中
- 如果依工作年份來選擇的話  
想工作 8 年買房：桃園  
想工作 1 0 年買房：台中  
想工作 1 2 年買房：新北

如考量其他方面的因素，如交通、教育、生活機能..等，就會影響在何處買房子的選擇。