



世界人口幸福指數分析報告

112/07/06學員作業（授課講師：李厚均）

學員：08_吳祖亨

homeppt

目錄

DIRECTORY



01

介紹



02

前置作業



03

分析



04

結論







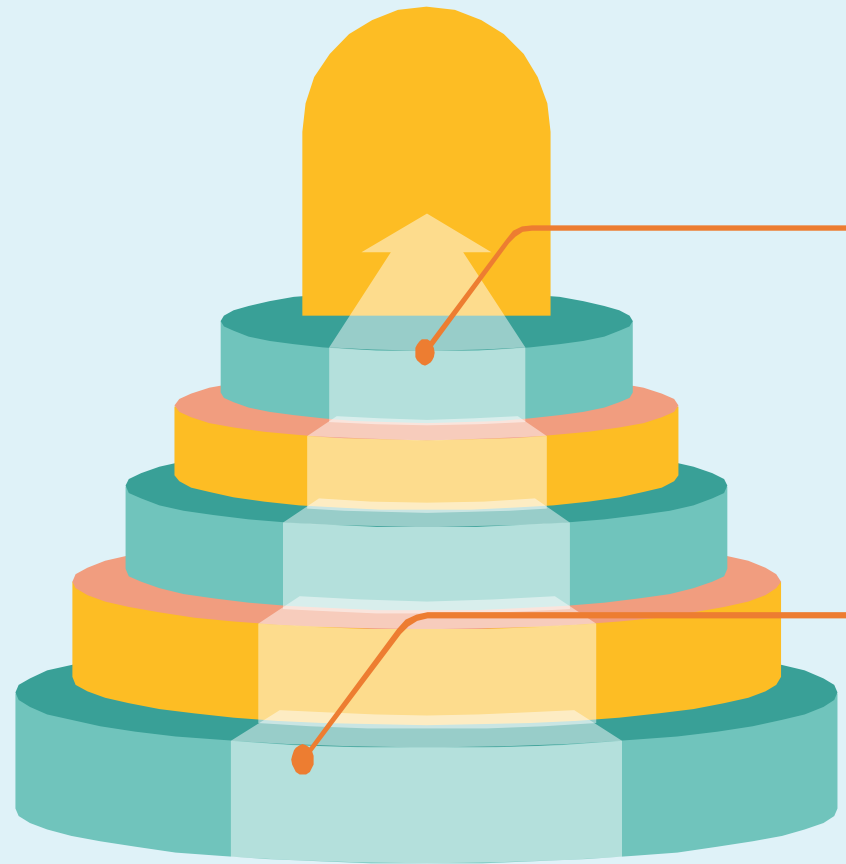
世界幸福報告

[kaggle / world-happiness](https://www.kaggle.com/world-happiness)



幸福感得分和排名使用"Gallup World Poll"民意調查的數據。
分數基於民意調查中提出的主要生活評估問題的答案。

這個問題被稱為坎特里爾階梯(Cantril ladder)，要求受訪者想像一個階梯，
其中對他們來說最好的可能生活為10，最壞的可能生活為0，
並按照該等級對自己當前的生活進行評分。



幸福分數後面的欄目基本有六個因素，在高分國家上，這些生活評價都比反烏托邦的生活評價更高。

而處於低分的國家，其六個因素中每一個因素的值均等於世界最低國家平均水平。

它們對每個國家/地區報告的總分沒有影響，但它們確實解釋了為什麼某些國家/地區排名高於其他國家/地區。





什麼是反烏托邦？

反烏托邦是一個虛構的國家，
那裡的人民是世界上最不幸福的。

建立反烏托邦的目的是為了在
六個關鍵變量中的每一個上建
立一個基準，可以對所有國家
進行有利的比較（沒有一個國
家的表現比反烏托邦更差）。

因此，觀察到六個關鍵變量的
最低分數代表了反烏托邦。

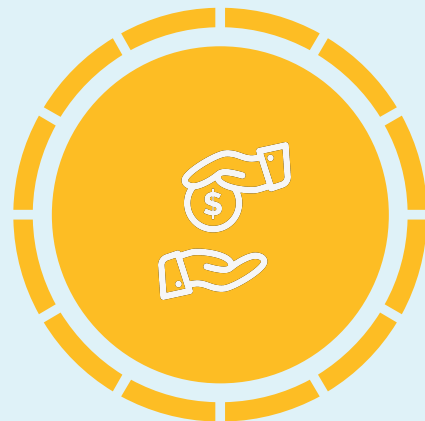
世界上收入最低
預期壽命最低
慷慨程度最低
腐敗最多
自由最少
社會支持最少

這些國家生活會非常不愉快，
與烏托邦形成鮮明對比。



自由人生選擇

Freedom to make life choices



慷慨

Generosity



對腐敗的看法

Perceptions of corruption



人均收入

GDP per capita



社會支持

Social support



健康預期壽命

Healthy life expectancy







數據清理和格式化

使用 Anaconda平台, 搭配 python 程式語言在 jupyter notebook中進行數據清理。

接下來我將檢查數據類型, 檢查數據框中的重複值與空值, 然後格式化數據。

清理數據後, 就可以進行分析了。

`describe()` -> 框架的描述性統計摘要

```
df_2019.describe()
```

	Overall rank	Score	GDP per capita	Social support	Healthy life expectancy	Freedom to make life choices	Generosity	Perceptions of corruption
count	156.000000	156.000000	156.000000	156.000000	156.000000	156.000000	156.000000	156.000000
mean	78.500000	5.407096	0.905147	1.208814	0.725244	0.392571	0.184846	0.110603
std	45.177428	1.113120	0.398389	0.299191	0.242124	0.143289	0.095254	0.094538
min	1.000000	2.853000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	39.750000	4.544500	0.602750	1.055750	0.547750	0.308000	0.108750	0.047000
50%	78.500000	5.379500	0.960000	1.271500	0.789000	0.417000	0.177500	0.085500
75%	117.250000	6.184500	1.232500	1.452500	0.881750	0.507250	0.248250	0.141250
max	156.000000	7.769000	1.684000	1.624000	1.141000	0.631000	0.566000	0.453000





```
df_2019.corr()
```

	Overall rank	Score	GDP per capita	Social support	Healthy life expectancy	Freedom to make life choices	Generosity	Perceptions of corruption
Overall rank	1.000000	-0.989096	-0.801947	-0.767465	-0.787411	-0.546606	-0.047993	-0.351959
Score	-0.989096	1.000000	0.793883	0.777058	0.779883	0.566742	0.075824	0.385613
GDP per capita	-0.801947	0.793883	1.000000	0.754906	0.835462	0.379079	-0.079662	0.298920
Social support	-0.767465	0.777058	0.754906	1.000000	0.719009	0.447333	-0.048126	0.181899
Healthy life expectancy	-0.787411	0.779883	0.835462	0.719009	1.000000	0.390395	-0.029511	0.295283
Freedom to make life choices	-0.546606	0.566742	0.379079	0.447333	0.390395	1.000000	0.269742	0.438843
Generosity	-0.047993	0.075824	-0.079662	-0.048126	-0.029511	0.269742	1.000000	0.326538
Perceptions of corruption	-0.351959	0.385613	0.298920	0.181899	0.295283	0.438843	0.326538	1.000000

```
df_2019.dtypes
```

```
Overall rank      int64
Country or region object
Score             float64
GDP per capita     float64
Social support     float64
Healthy life expectancy float64
Freedom to make life choices float64
Generosity         float64
Perceptions of corruption float64
dtype: object
```

```
df_2019.shape
```

```
(156, 9)
```

**dtypes()
shape()**
數據類型&形狀

corr()
皮爾遜積差
相關係數

**duplicated()
isnull()**
重複值&缺失值

```
df_2019.duplicated().sum()
```

```
0
```

```
df_2019.isnull().sum()
```

```
Overall rank      0
Country or region  0
Score             0
GDP per capita     0
Social support     0
Healthy life expectancy  0
Freedom to make life choices  0
Generosity         0
Perceptions of corruption  0
dtype: int64
```

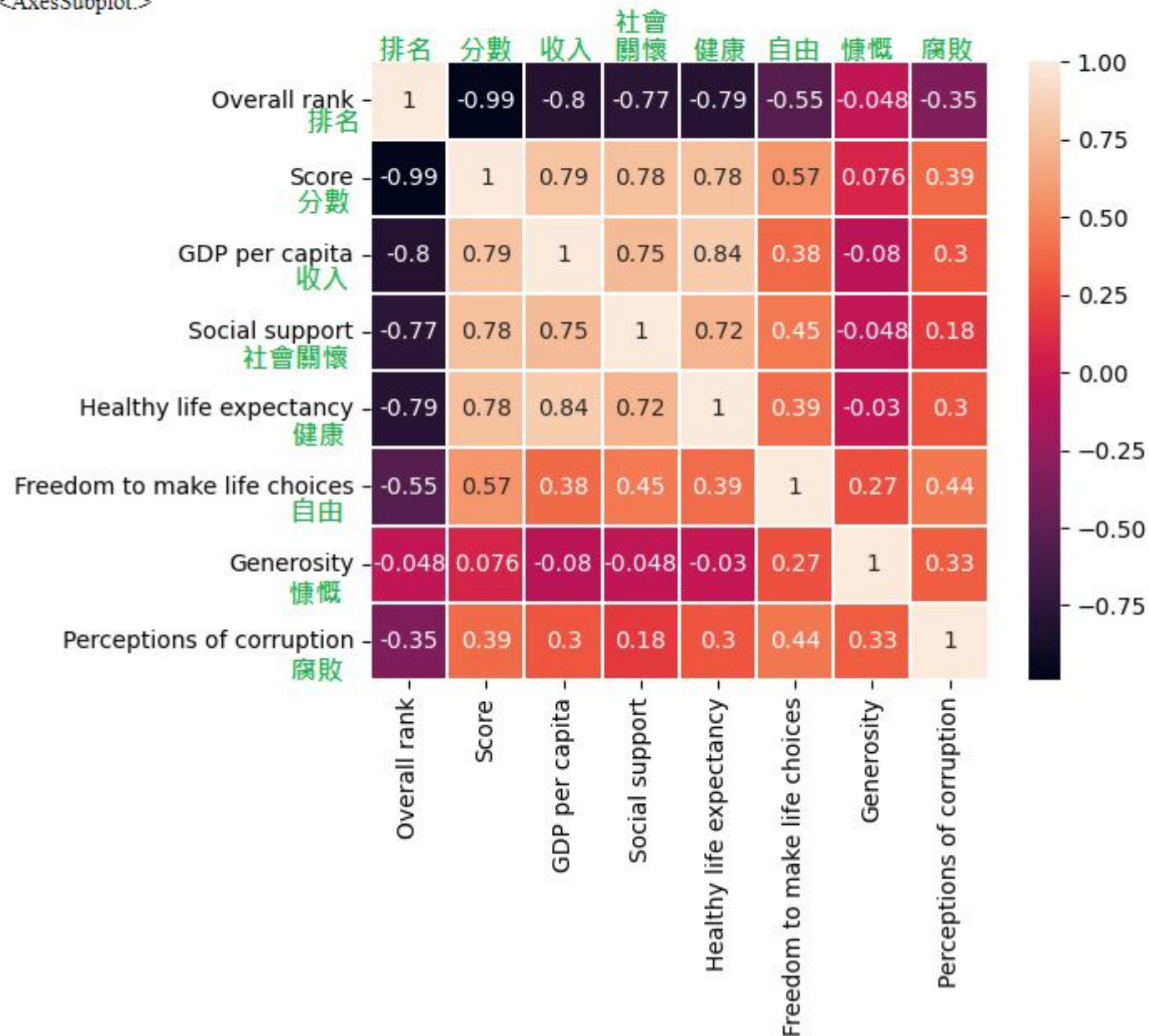




```
import seaborn as sns
```

```
sns.heatmap(df_2019.corr(),linewidths=0.1,vmax=1.0,square=True,linecolor='white',annot=True)
```

<AxesSubplot:>



熱力圖呈現

由於人類是視覺化的動物，直接查看皮爾遜積差相關係數可能會難以消化。

因此使用熱度圖來視覺化皮爾遜積差相關係數，能夠更快的理解結果。







接下來我將分析
以下內容：




了解台灣這些年來
(2015年至2019年)
的幸福狀況。



不同因素(經濟(人均GDP)、家庭、
健康、自由、信任(政府腐敗)和慷
慨)對幸福分數的影響。



尋找這些年來
(2015 年至 2019 年)
最幸福的國家 等等。



將2015~2019之間的數據做統整結合

```
df_2015=df_2015.drop(['Region', 'Standard Error', 'Dystopia Residual'],axis=1)
```

```
df_2017=df_2017.rename(columns={'Happiness.Score':'Happiness Score','Happ
```

```
df_2017=df_2017.drop(['Whisker.high', 'Whisker.low',  
 'Dystopia.Residual'],axis=1)
```

```
df_2015=df_2015.set_index("Country")
```

```
df_2016=df_2016.set_index("Country")
```

```
df_2017=df_2017.set_index("Country")
```

```
df_2018=df_2018.set_index("Country or region")
```

```
df_2019=df_2019.set_index("Country or region")
```

```
Taiwan_2015=df_2015.loc[["Taiwan"]]
```

```
Taiwan_2016=df_2016.loc[["Taiwan"]]
```

```
Taiwan_2017=df_2017.loc[["Taiwan Province of China"]]
```

```
Taiwan_2018=df_2018.loc[["Taiwan"]]
```

```
Taiwan_2019=df_2019.loc[["Taiwan"]]
```

```
Taiwan_2015["year"]=2015
```

```
Taiwan_2016["year"]=2016
```

```
Taiwan_2017["year"]=2017
```

```
Taiwan_2017=Taiwan_2017.rename({'Happiness.Rank':'Happiness Rank', 'Happiness.Score': 'Happines  
'Economy..GDP.per.Capita.': 'Economy (GDP per Capita)', 'Health..Life.Expectancy.': 'Health (Life E  
'Trust..Government.Corruption.': 'Trust (Government Corruption)' },axis=1)
```

```
df_tai_2015_17=pd.concat([Taiwan_2015,Taiwan_2016,Taiwan_2017],ignore_index=True,sort=False)
```

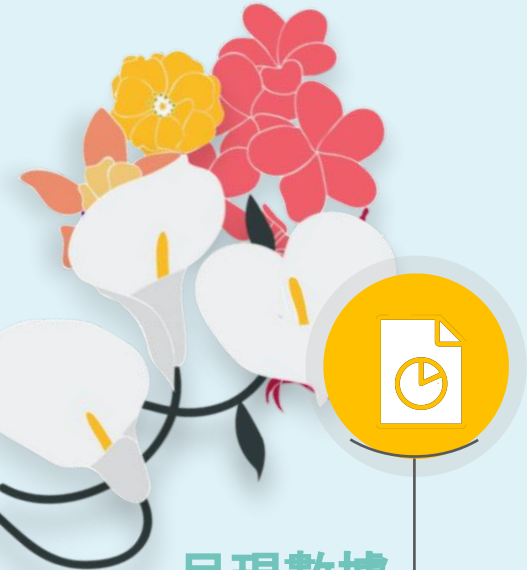
```
Taiwan_2018["year"]=2018
```

```
Taiwan_2019["year"]=2019
```

```
df_tai_2018_19=pd.concat([Taiwan_2018,Taiwan_2019],ignore_index=True,sort=False)
```

```
df_tai_2018_19.rename({'Overall rank': 'Happiness Rank', 'Score': 'Happiness Score', 'GDP per capita': 'E
```

```
df_tai_2015_19=pd.concat([df_tai_2018_19,df_tai_2015_17],ignore_index=True,sort=False)
```

呈現數據
並使其更好解釋。

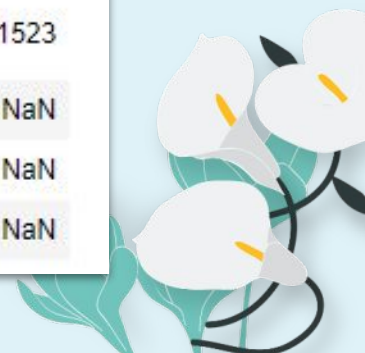
```
df_tai_2015_19.sort_values(by="year")
```

	Happiness Rank	Happiness Score	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Freedom	Generosity	Trust (Government Corruption)	year	Region	Lower Confidence Interval	Upper Confidence Interval	Dystopia Residual
2	38	6.298	1.290980	1.076170	0.875300	0.397400	0.25376	0.081290	2015	NaN	NaN	NaN	NaN
3	34	6.379	1.397290	0.926240	0.795650	0.323770	0.25495	0.066300	2016	Eastern Asia	6.305	6.453	2.61523
4	33	6.422	1.433627	1.384565	0.793984	0.361467	0.25836	0.063829	2017	NaN	NaN	NaN	NaN
0	26	6.441	1.365000	1.436000	0.857000	0.418000	0.15100	0.078000	2018	NaN	NaN	NaN	NaN
1	25	6.446	1.368000	1.430000	0.914000	0.351000	0.24200	0.097000	2019	NaN	NaN	NaN	NaN

```
df_tai_2015_19=df_tai_2015_19.set_index("year").sort_index()
```

```
df_tai_2015_19.sort_values(by="year")
```

	Happiness Rank	Happiness Score	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Freedom	Generosity	Trust (Government Corruption)	Region	Lower Confidence Interval	Upper Confidence Interval	Dystopia Residual
year												
2015	38	6.298	1.290980	1.076170	0.875300	0.397400	0.25376	0.081290	NaN	NaN	NaN	NaN
2016	34	6.379	1.397290	0.926240	0.795650	0.323770	0.25495	0.066300	Eastern Asia	6.305	6.453	2.61523
2017	33	6.422	1.433627	1.384565	0.793984	0.361467	0.25836	0.063829	NaN	NaN	NaN	NaN
2018	26	6.441	1.365000	1.436000	0.857000	0.418000	0.15100	0.078000	NaN	NaN	NaN	NaN
2019	25	6.446	1.368000	1.430000	0.914000	0.351000	0.24200	0.097000	NaN	NaN	NaN	NaN



```
#繪製散點圖 (scatter plot)
plt.scatter(df_2019["Score"],df_2019["Perceptions of corruption"],color="green")

#趨勢線的方程
x=df_2019["Score"]
z=np.polyfit(df_2019["Score"],df_2019["Perceptions of corruption"], 1)

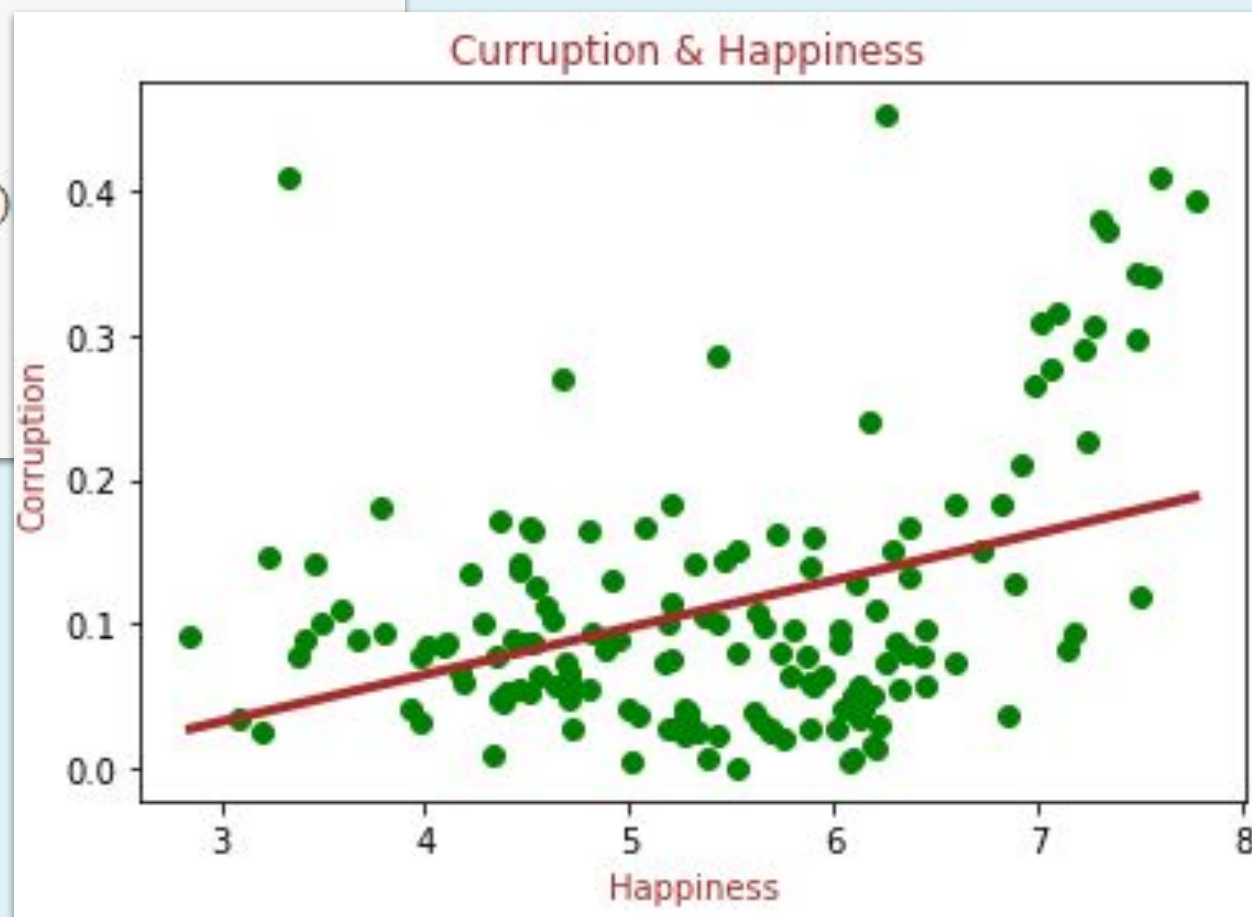
p = np.poly1d(z)

# 繪製趨勢圖 (trendline plot)
plt.plot(df_2019["Score"], p(x),color="brown",linewidth=3)
plt.title("Corruption & Happiness",color="brown")

plt.xlabel("Happiness",color="brown")
plt.ylabel("Corruption",color="brown")
```

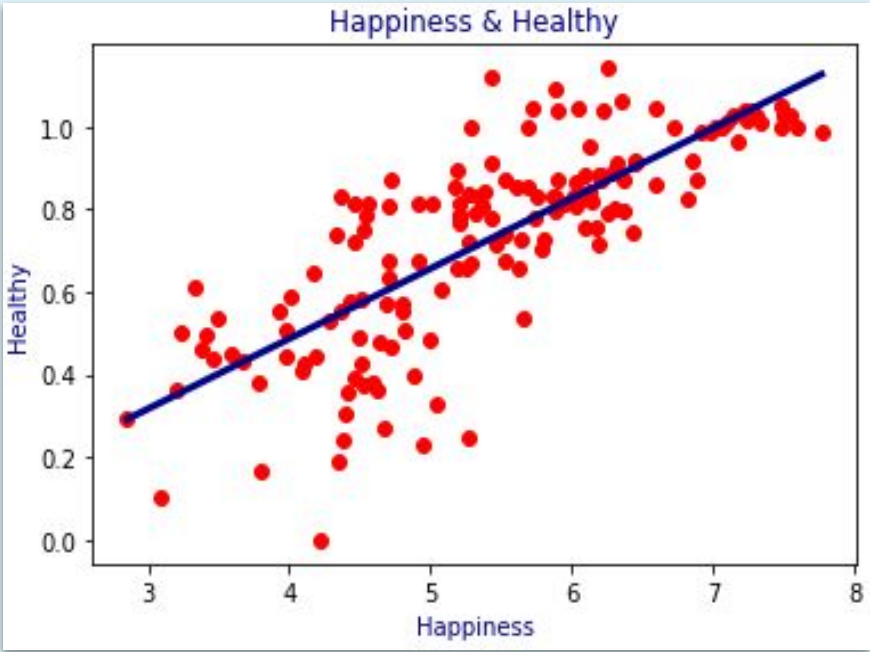
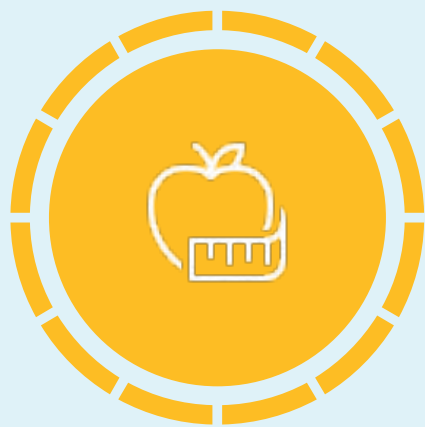
可視化+快尋要點

使用matplotlib, 我們可以為分析提供漂亮的可視化效果。



高幸福-對應六大關鍵的結果

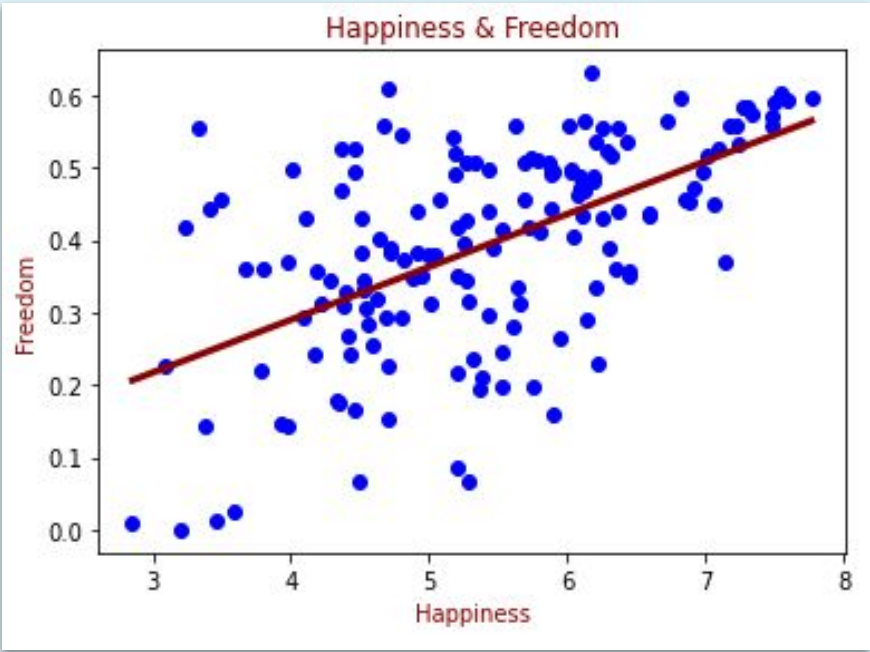
健康



社會
支持



自由



高幸福-對應六大關鍵的結果

腐敗



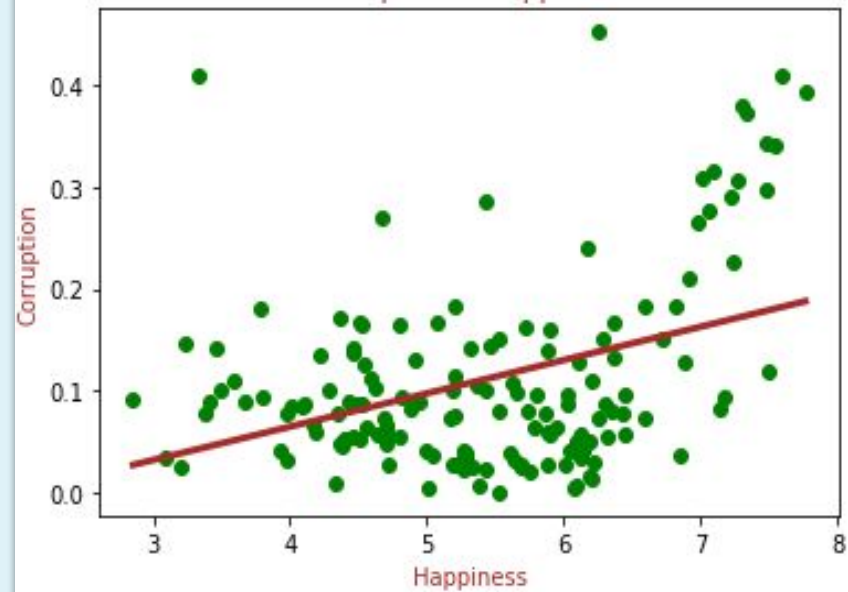
慷慨



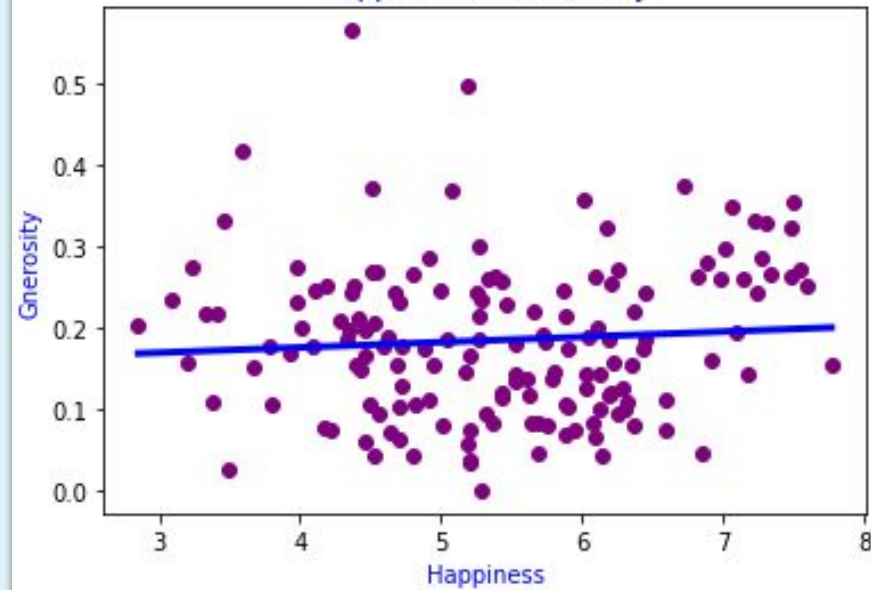
收入



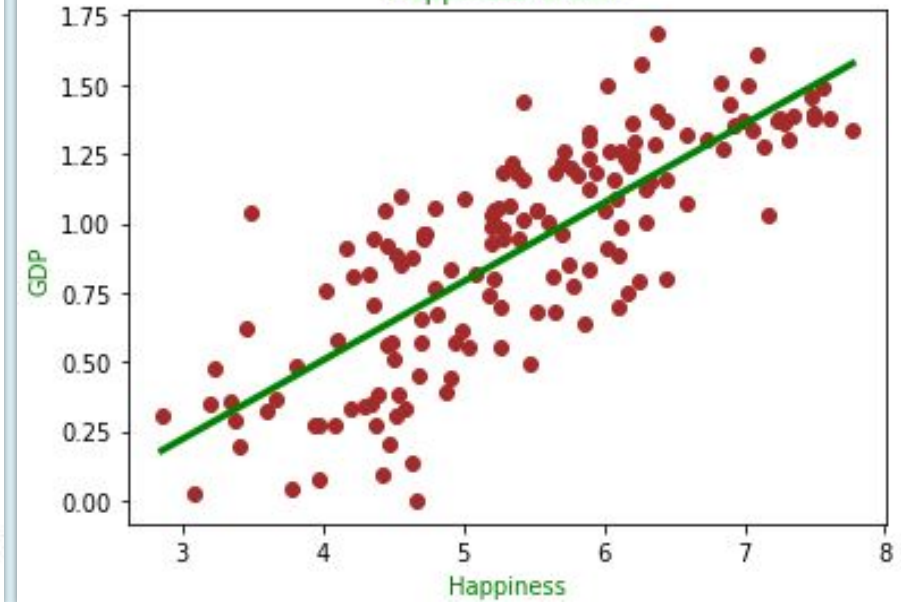
Corruption & Happiness



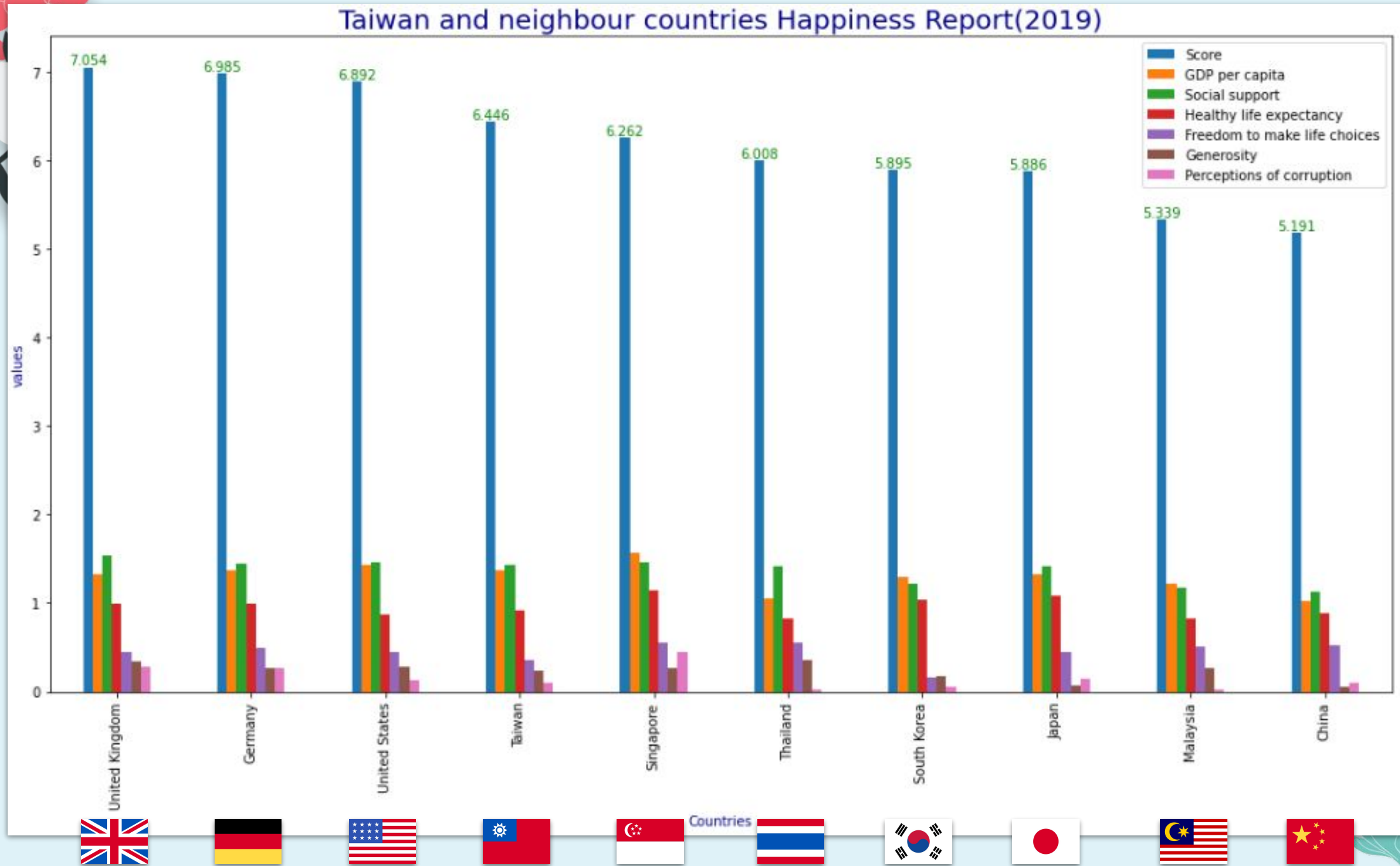
Happiness & Generosity



Happiness & GDP

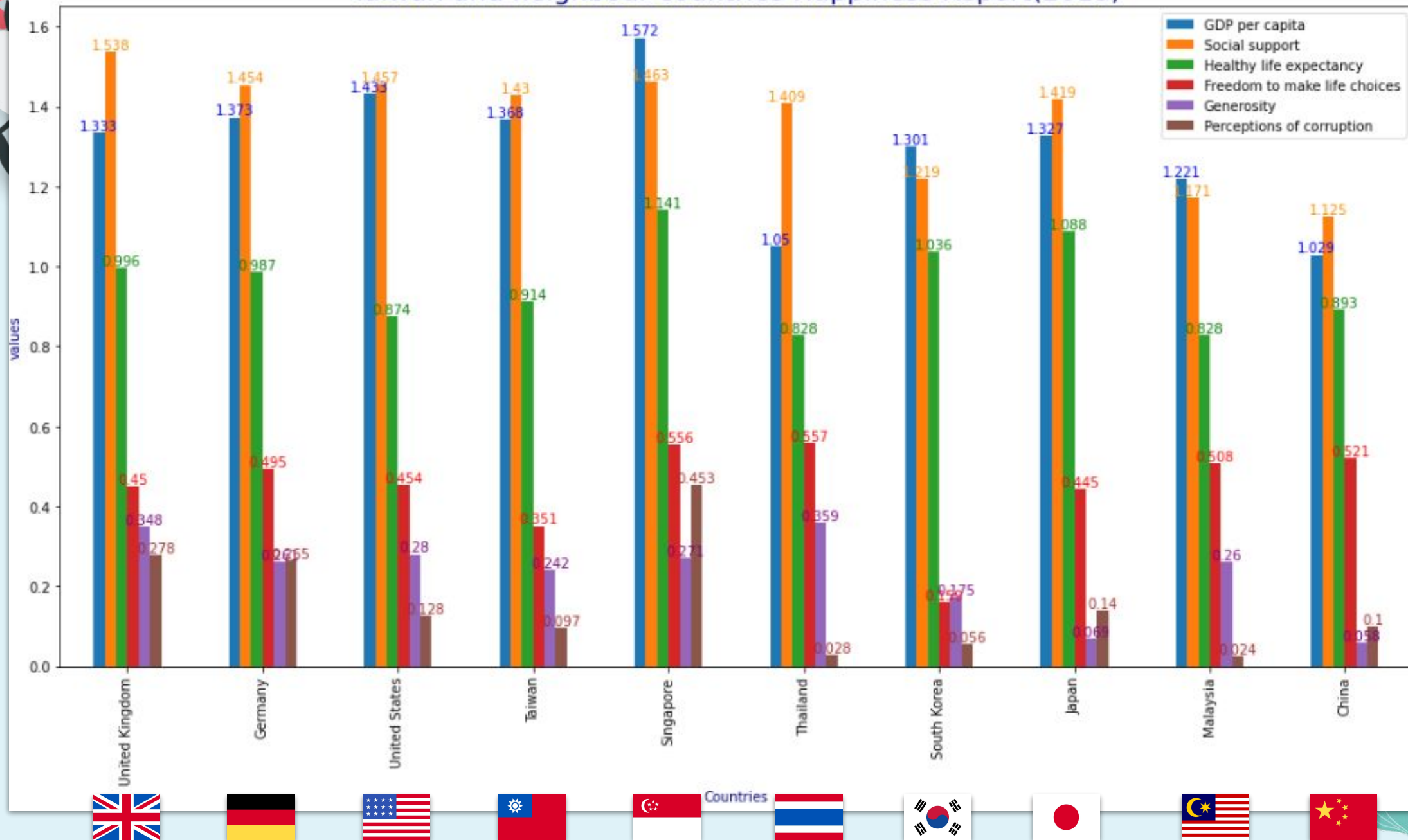


台灣及其鄰國 於(2019)相關的一些可視化繪圖



台灣及其鄰國 於(2019)相關的一些可視化繪圖

Taiwan and neighbour countries Happiness Report(2019)



台灣(2015 - 2019)相關的一些可視化繪圖

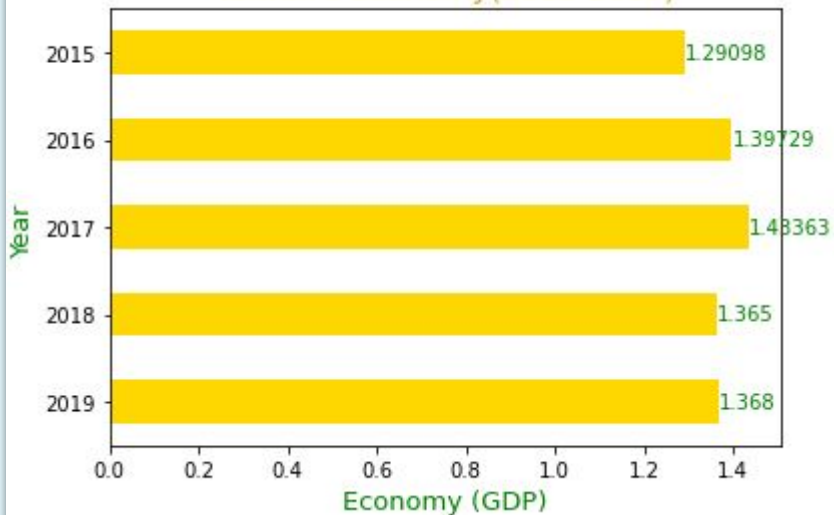
台灣幸福排名(2015-2019)



台灣幸福得分(2015-2019)

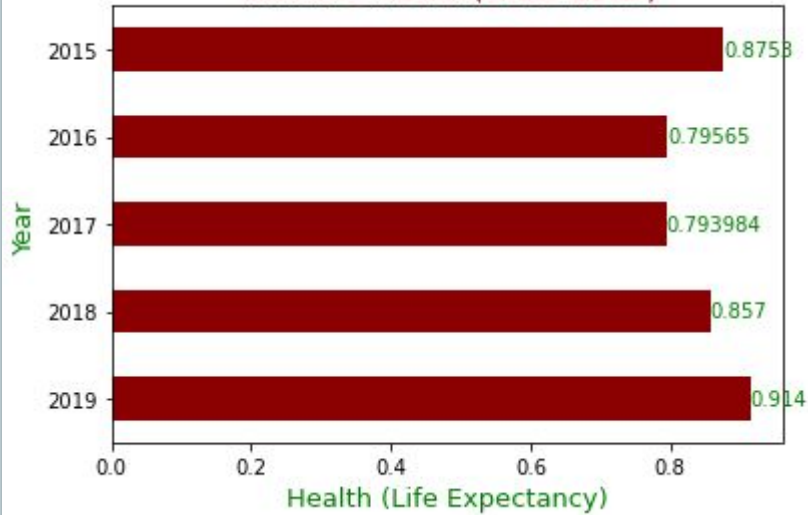


Taiwan's Economy(2015-2019)



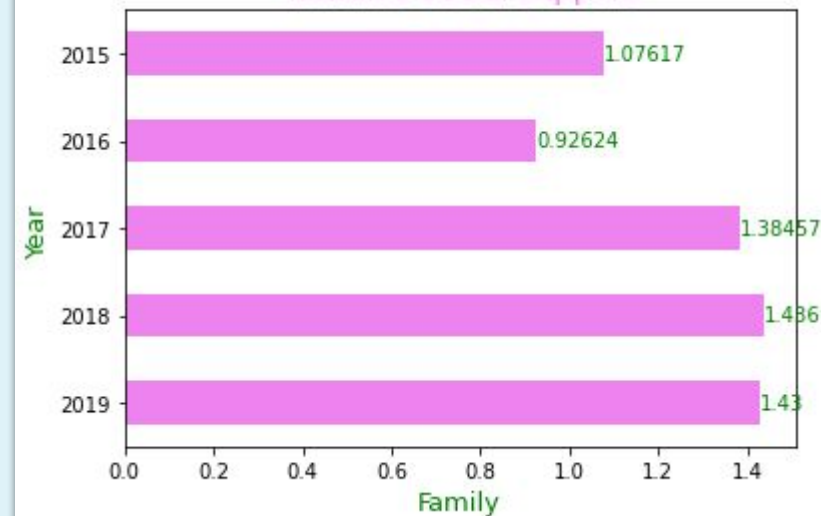
收入

Taiwan's Health(2015-2019)



健康

Taiwan's Social support



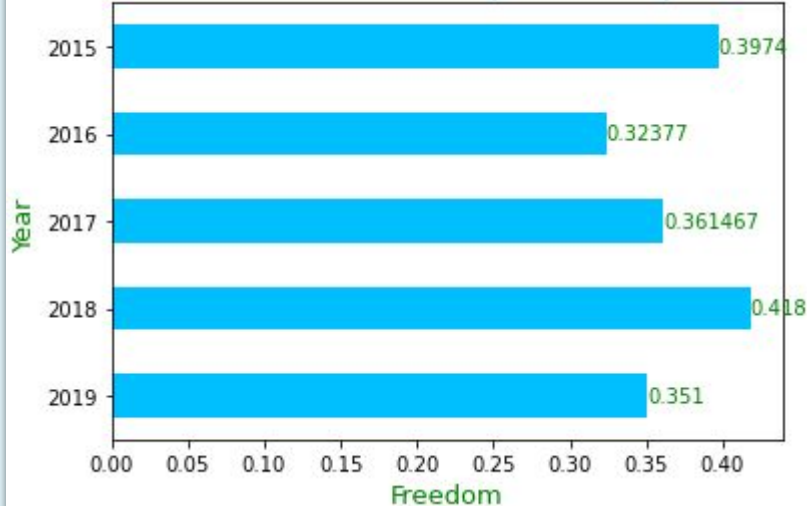
社會支持

台灣X6關鍵 結果



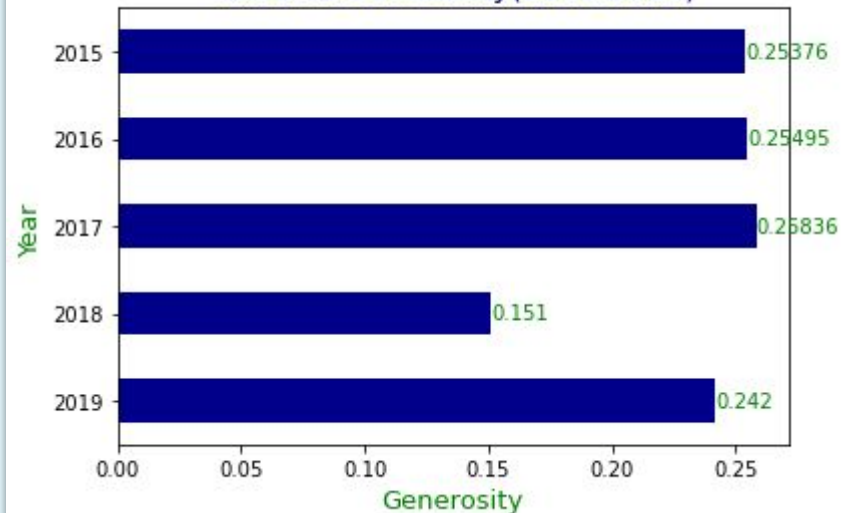
自由

Taiwan's Freedom(2015-2019)



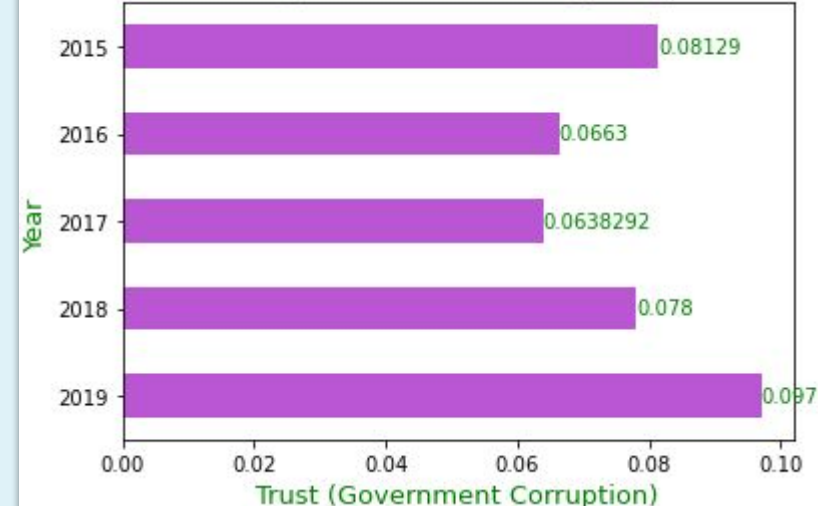
慷慨

Taiwan's Generosity(2015-2019)



腐敗

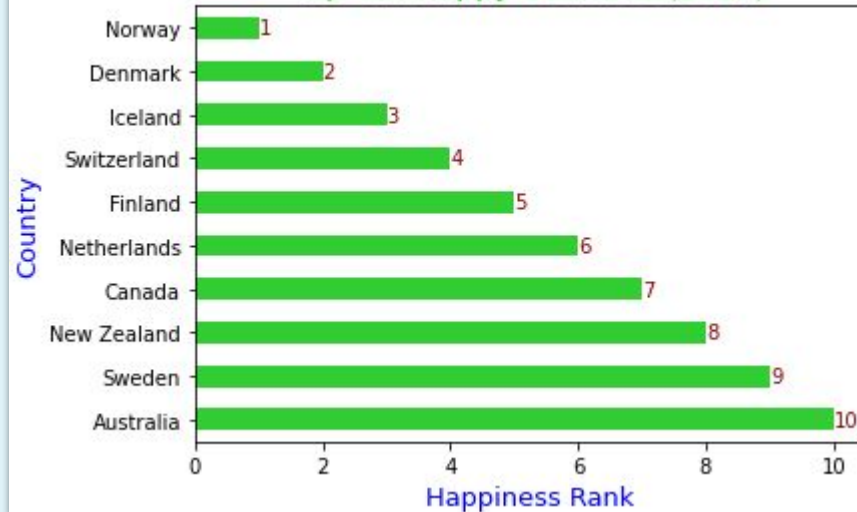
Taiwan's Trust (Government Corruption)(2015-2019)



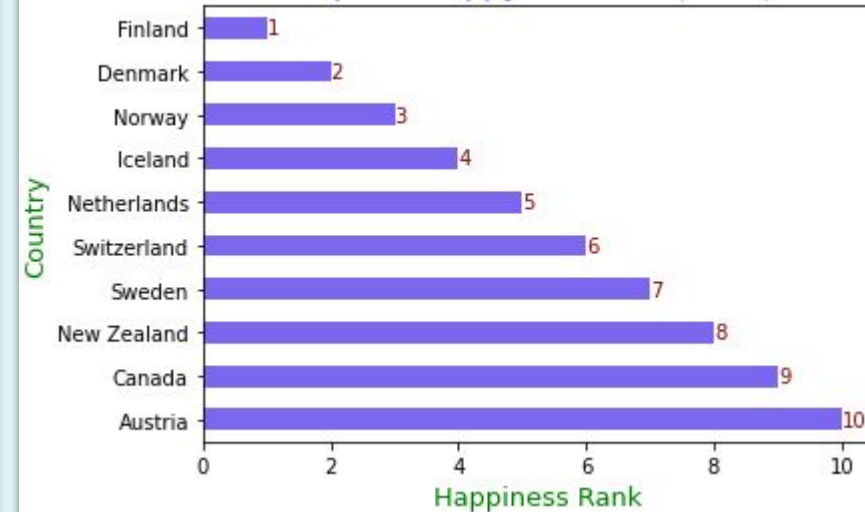
Top Ten Happy Countries(2015)



Top Ten Happy Countries(2017)



Top Ten Happy Countries(2019)



全球幸福分數排名 前10國家(2015 - 2019)

Top Ten Happy Countries(2016)



Top Ten Happy Countries(2018)



排名前3名國家 (2015-2019)

2015: 瑞士、冰島、丹麥
 2016: 丹麥、瑞士、冰島
 2017: 挪威、丹麥、冰島
 2018: 芬蘭、挪威、丹麥
 2019: 芬蘭、丹麥、挪威





我們分析了給定的數據，現在可以總結我們的分析了：

我們發現瑞士、芬蘭、挪威、丹麥和冰島是幸福指數排前的國家，而這些都是發達國家，同時他們的經濟狀況皆非常好。

在經濟(GDP)、社會支持、健康、自由和政府腐敗這六個關鍵，來跟幸福得分和幸福排名做比較後，我們可以發現到**經濟、社會支持和健康最為重要**。在圖表上，能看到趨勢和點陣是一同穩定上升的樣子。

因此，如果我們想讓我們的國家幸福，我們就應該讓我們的經濟變得更好。經濟使人們的健康變得更好，因為他們可以為家人獲得好的食物和其他生活必需品。良好的經濟和良好的健康使人們的關係變得更好，幸福的人們造就幸福的社會。

因此，我們應該努力通過增加人們的**就業機會**來改善經濟，並且提高**教育質量**。除了減少因職能不足的失業者們，也使得不同類型的人們，能夠真正做到適合他們的職位，真正的投身進社會之中。





謝謝參考

112/07/06學員作業（授課講師：李厚均）

學員：08_吳祖亨