

# world-happiness-report-eda

January 25, 2023

```
[60]: !pip install pycountry_convert --quiet
```

```
[61]: !pip install plotly --quiet
```

```
[62]: !pip install missingno --quiet
```

```
[63]: import pycountry_convert as pc
import missingno as msno
```

```
[56]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

```
[2]: data=pd.read_csv("C:\\Users\\MSI\\Desktop\\VIT DS\\Python for Data_
↪Science\\World Happiness Report 2022 CSV.csv")
data.head()
```

```
[2]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low	\
0	1	Finland	7.821	7.886	7.756	
1	2	Denmark	7.636	7.710	7.563	
2	3	Iceland	7.557	7.651	7.464	
3	4	Switzerland	7.512	7.586	7.437	
4	5	Netherlands	7.415	7.471	7.359	

	Dystopia (1.83) + residual	Explained by: GDP per capita	\
0	2.518	1.892	
1	2.226	1.953	
2	2.320	1.936	
3	2.153	2.026	
4	2.137	1.945	

	Explained by: Social support	Explained by: Healthy life expectancy	\
0	1.258	0.775	
1	1.243	0.777	
2	1.320	0.803	
3	1.226	0.822	

4	1.206	0.787
---	-------	-------

	Explained by: Freedom to make life choices	Explained by: Generosity \
0	0.736	0.109
1	0.719	0.188
2	0.718	0.270
3	0.677	0.147
4	0.651	0.271

	Explained by: Perceptions of corruption
0	0.534
1	0.532
2	0.191
3	0.461
4	0.419

```
[3]: data.shape
```

```
[3]: (146, 12)
```

```
[4]: data.columns
```

```
[4]: Index(['RANK', 'Country', 'Happiness score', 'Whisker-high', 'Whisker-low',
          'Dystopia (1.83) + residual', 'Explained by: GDP per capita',
          'Explained by: Social support', 'Explained by: Healthy life expectancy',
          'Explained by: Freedom to make life choices',
          'Explained by: Generosity', 'Explained by: Perceptions of corruption'],
          dtype='object')
```

```
[5]: data.describe()
```

	RANK	Happiness score	Whisker-high	Whisker-low \
count	146.000000	146.000000	146.000000	146.000000
mean	73.500000	5.553575	5.673589	5.433568
std	42.290661	1.086843	1.065621	1.109380
min	1.000000	2.404000	2.469000	2.339000
25%	37.250000	4.888750	5.006250	4.754750
50%	73.500000	5.568500	5.680000	5.453000
75%	109.750000	6.305000	6.448750	6.190000
max	146.000000	7.821000	7.886000	7.756000

	Dystopia (1.83) + residual	Explained by: GDP per capita \
count	146.000000	146.000000
mean	1.831808	1.410445
std	0.534994	0.421663
min	0.187000	0.000000
25%	1.555250	1.095500

50%	1.894500	1.445500
75%	2.153000	1.784750
max	2.844000	2.209000

	Explained by: Social support	Explained by: Healthy life expectancy \
count	146.000000	146.000000
mean	0.905863	0.586171
std	0.280122	0.176336
min	0.000000	0.000000
25%	0.732000	0.463250
50%	0.957500	0.621500
75%	1.114250	0.719750
max	1.320000	0.942000

	Explained by: Freedom to make life choices	Explained by: Generosity \
count	146.000000	146.000000
mean	0.517226	0.147377
std	0.145859	0.082799
min	0.000000	0.000000
25%	0.440500	0.089000
50%	0.543500	0.132500
75%	0.626000	0.197750
max	0.740000	0.468000

	Explained by: Perceptions of corruption
count	146.000000
mean	0.154781
std	0.127514
min	0.000000
25%	0.068250
50%	0.119500
75%	0.198500
max	0.587000

```
[7]: data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 146 entries, 0 to 145
```

```
Data columns (total 12 columns):
```

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	RANK	146 non-null	int64
1	Country	146 non-null	object
2	Happiness score	146 non-null	float64
3	Whisker-high	146 non-null	float64
4	Whisker-low	146 non-null	float64
5	Dystopia (1.83) + residual	146 non-null	float64

```

6   Explained by: GDP per capita          146 non-null    float64
7   Explained by: Social support          146 non-null    float64
8   Explained by: Healthy life expectancy 146 non-null    float64
9   Explained by: Freedom to make life choices 146 non-null    float64
10  Explained by: Generosity              146 non-null    float64
11  Explained by: Perceptions of corruption 146 non-null    float64
dtypes: float64(10), int64(1), object(1)
memory usage: 13.8+ KB

```

```
[8]: data[data['Country']=="India"]
```

```

[8]:      RANK Country  Happiness score  Whisker-high  Whisker-low  \
135   136   India          3.777          3.828          3.726

      Dystopia (1.83) + residual  Explained by: GDP per capita  \
135                          0.795                          1.167

      Explained by: Social support  Explained by: Healthy life expectancy  \
135                          0.376                          0.471

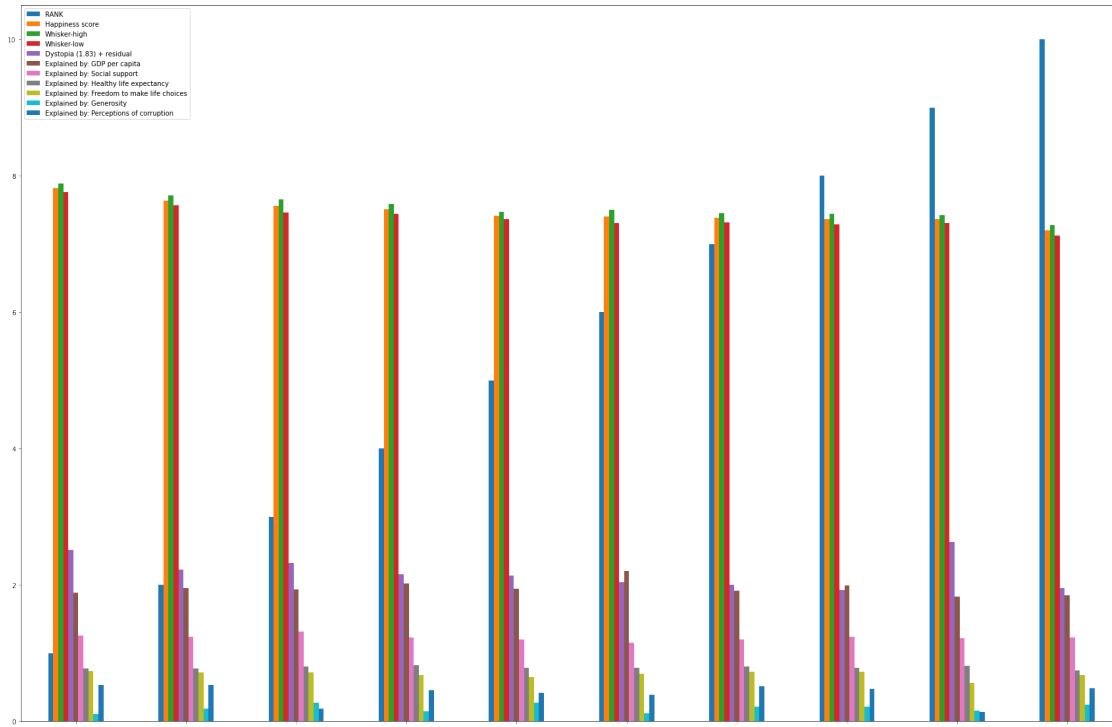
      Explained by: Freedom to make life choices  Explained by: Generosity  \
135                          0.647                          0.198

      Explained by: Perceptions of corruption
135                          0.123

```

```
[16]: data[:10].plot(kind="bar",figsize=(30,20))
```

```
[16]: <AxesSubplot:>
```



```
[156]: top= data.sort_values('Happiness score', ascending=False)[:10].
        ↪sort_values('Happiness score')
top
```

```
[156]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low	\
9	10	New Zealand	7.200	7.279	7.120	
8	9	Israel	7.364	7.426	7.301	
7	8	Norway	7.365	7.440	7.290	
6	7	Sweden	7.384	7.454	7.315	
5	6	Luxembourg	7.404	7.501	7.307	
4	5	Netherlands	7.415	7.471	7.359	
3	4	Switzerland	7.512	7.586	7.437	
2	3	Iceland	7.557	7.651	7.464	
1	2	Denmark	7.636	7.710	7.563	
0	1	Finland	7.821	7.886	7.756	

	Dystopia (1.83) + residual	Explained by: GDP per capita	\
9	1.954	1.852	
8	2.634	1.826	
7	1.925	1.997	
6	2.003	1.920	
5	2.042	2.209	
4	2.137	1.945	
3	2.153	2.026	

2	2.320	1.936
1	2.226	1.953
0	2.518	1.892

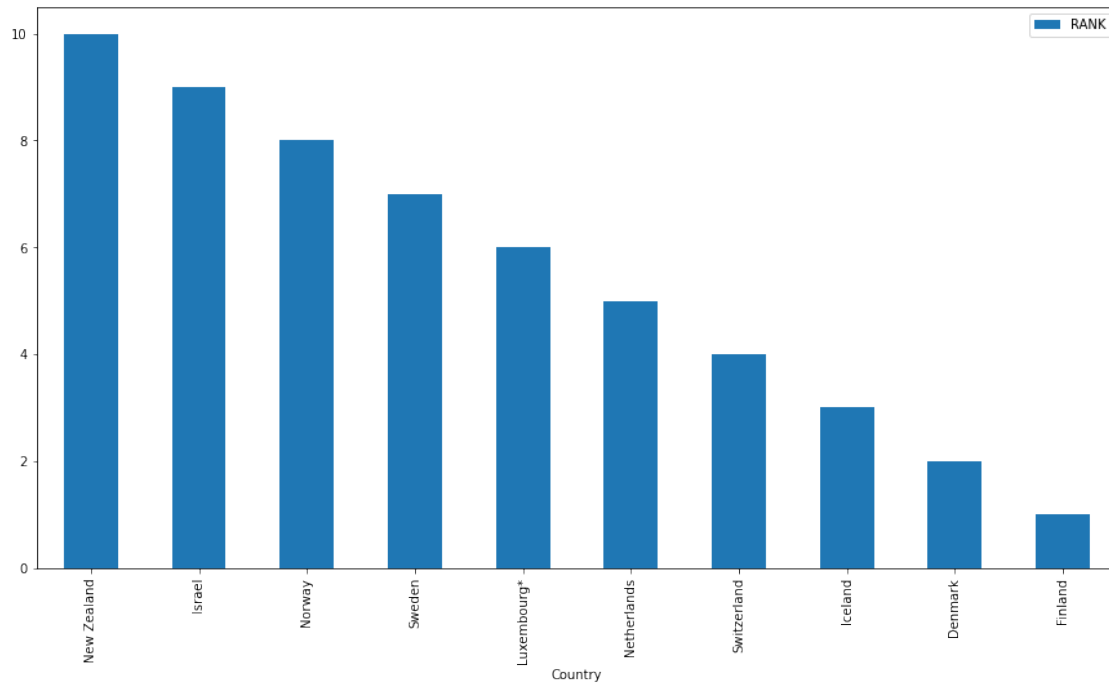
	Explained by: Social support	Explained by: Healthy life expectancy \
9	1.235	0.752
8	1.221	0.818
7	1.239	0.786
6	1.204	0.803
5	1.155	0.790
4	1.206	0.787
3	1.226	0.822
2	1.320	0.803
1	1.243	0.777
0	1.258	0.775

	Explained by: Freedom to make life choices	Explained by: Generosity \
9	0.680	0.245
8	0.568	0.155
7	0.728	0.217
6	0.724	0.218
5	0.700	0.120
4	0.651	0.271
3	0.677	0.147
2	0.718	0.270
1	0.719	0.188
0	0.736	0.109

	Explained by: Perceptions of corruption	Country code	Continent
9	0.483	NZ	OC
8	0.143	IL	AS
7	0.474	NO	EU
6	0.512	SE	EU
5	0.388	LU	EU
4	0.419	NL	EU
3	0.461	CH	EU
2	0.191	IS	EU
1	0.532	DK	EU
0	0.534	FI	EU

```
[36]: top.plot(kind="bar",x="Country",y="RANK",figsize=(15,8))
```

```
[36]: <AxesSubplot:xlabel='Country'>
```



```
[38]: last= data.sort_values('Happiness score')[:10].sort_values('Happiness score',
↪ascending=False)
last
```

```
[38]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low	\
136	137	Zambia	3.760	3.902	3.618	
137	138	Malawi	3.750	3.941	3.560	
138	139	Tanzania	3.702	3.847	3.558	
139	140	Sierra Leone	3.574	3.732	3.416	
140	141	Lesotho*	3.512	3.748	3.276	
141	142	Botswana*	3.471	3.667	3.275	
142	143	Rwanda*	3.268	3.462	3.074	
143	144	Zimbabwe	2.995	3.110	2.880	
144	145	Lebanon	2.955	3.049	2.862	
145	146	Afghanistan	2.404	2.469	2.339	

	Dystopia (1.83) + residual	Explained by: GDP per capita	\
136	1.135	0.930	
137	1.661	0.648	
138	0.735	0.848	
139	1.556	0.686	
140	1.312	0.839	
141	0.187	1.503	
142	0.536	0.785	
143	0.548	0.947	

144	0.216	1.392
145	1.263	0.758

	Explained by: Social support	Explained by: Healthy life expectancy \
136	0.577	0.306
137	0.279	0.388
138	0.597	0.425
139	0.416	0.273
140	0.848	0.000
141	0.815	0.280
142	0.133	0.462
143	0.690	0.270
144	0.498	0.631
145	0.000	0.289

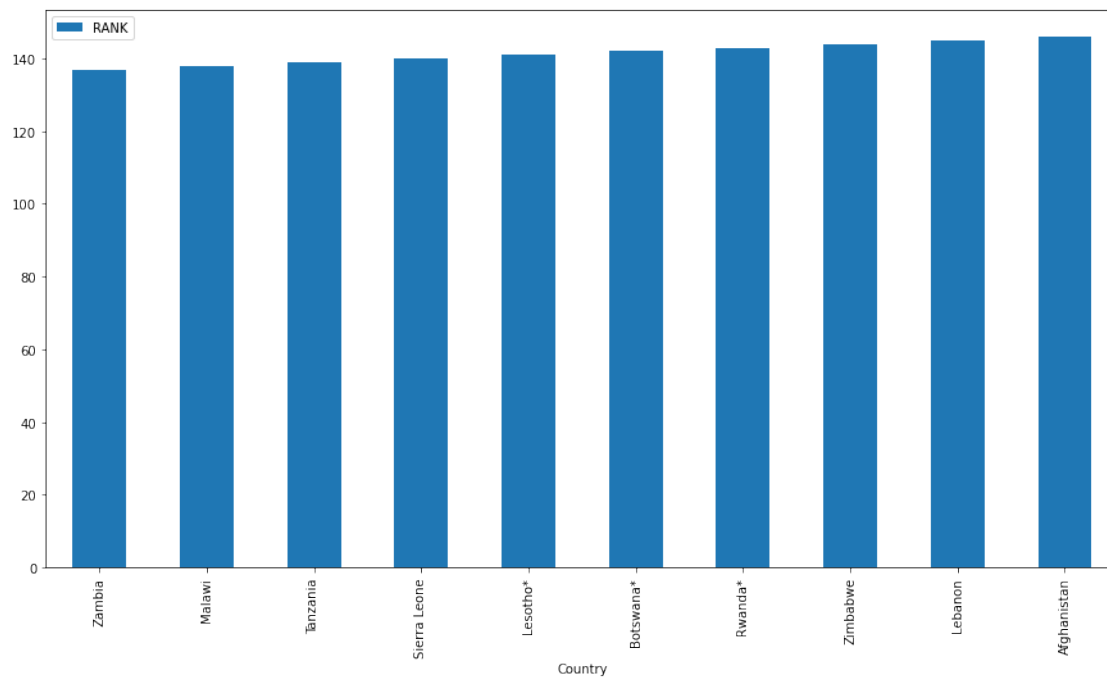
	Explained by: Freedom to make life choices	Explained by: Generosity \
136	0.525	0.203
137	0.477	0.140
138	0.578	0.248
139	0.387	0.202
140	0.419	0.076
141	0.571	0.012
142	0.621	0.187
143	0.329	0.106
144	0.103	0.082
145	0.000	0.089

	Explained by: Perceptions of corruption
136	0.083
137	0.157
138	0.270
139	0.055
140	0.018
141	0.102
142	0.544
143	0.105
144	0.034
145	0.005

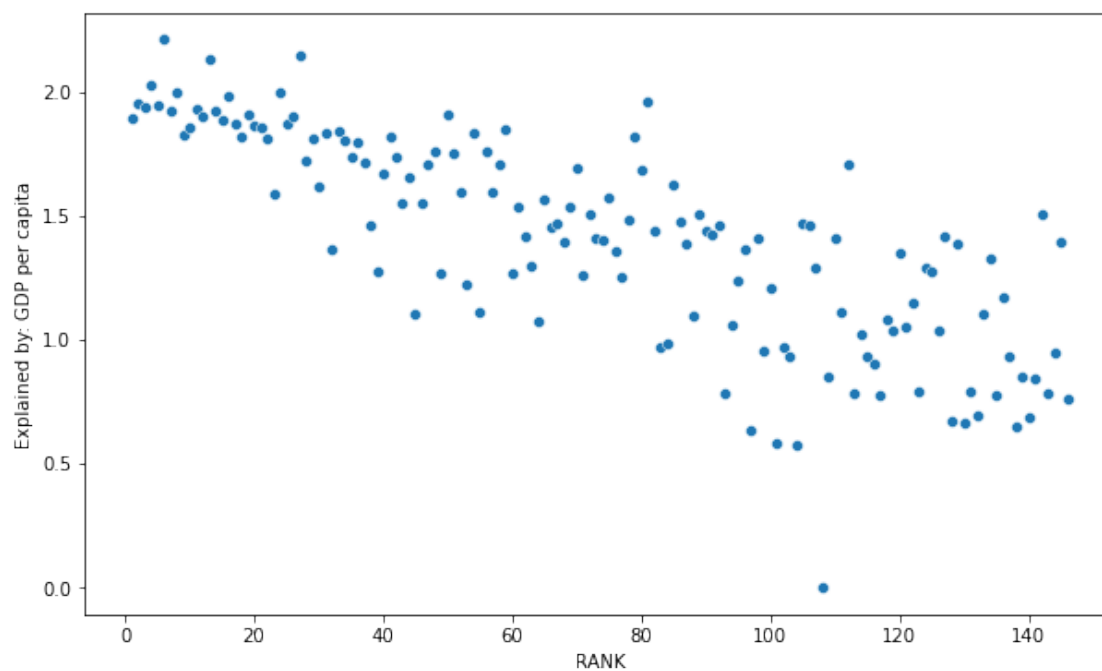
```
[39]: last.plot(kind="bar",x="Country",y="RANK",figsize=(15,8))
```

```
[39]: <AxesSubplot:xlabel='Country'>
```

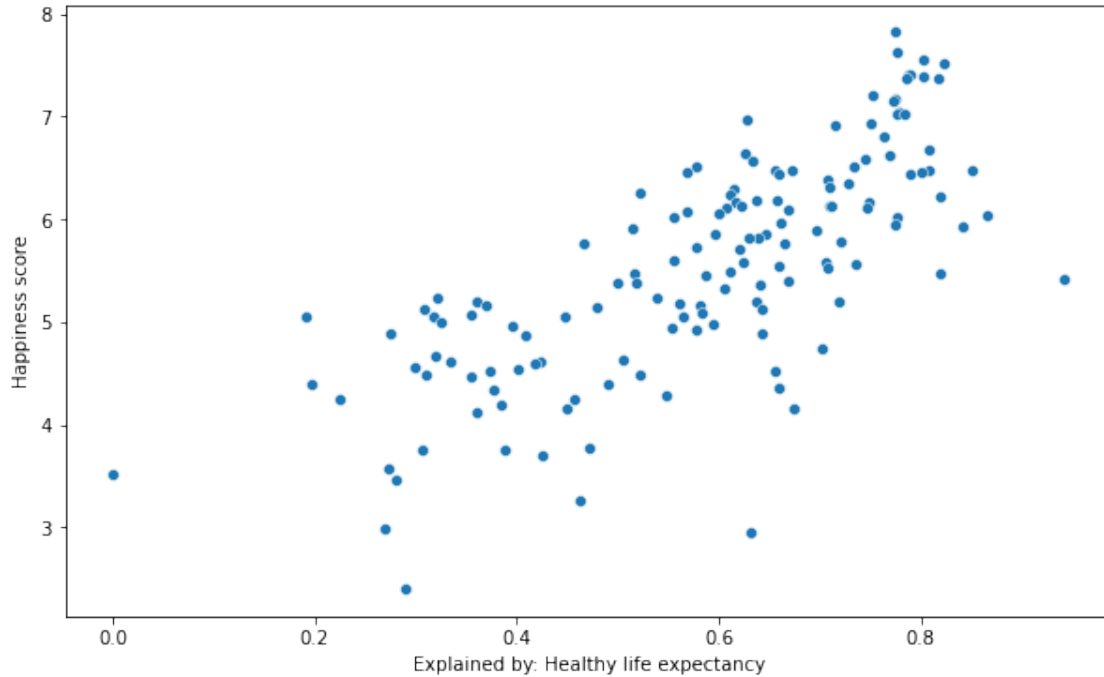




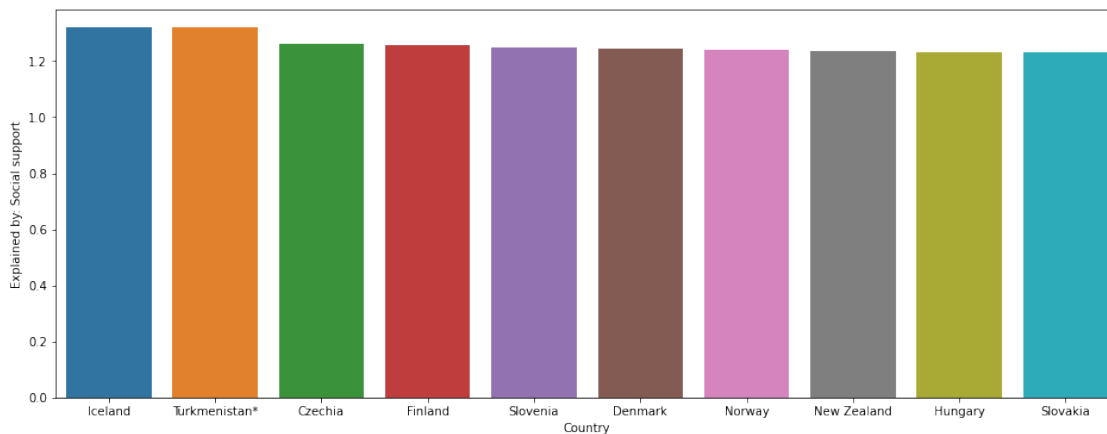
```
[44]: plt.figure(figsize=(10,6))
sns.scatterplot(x=data['RANK'],y=data['Explained by: GDP per capita'])
plt.show()
```



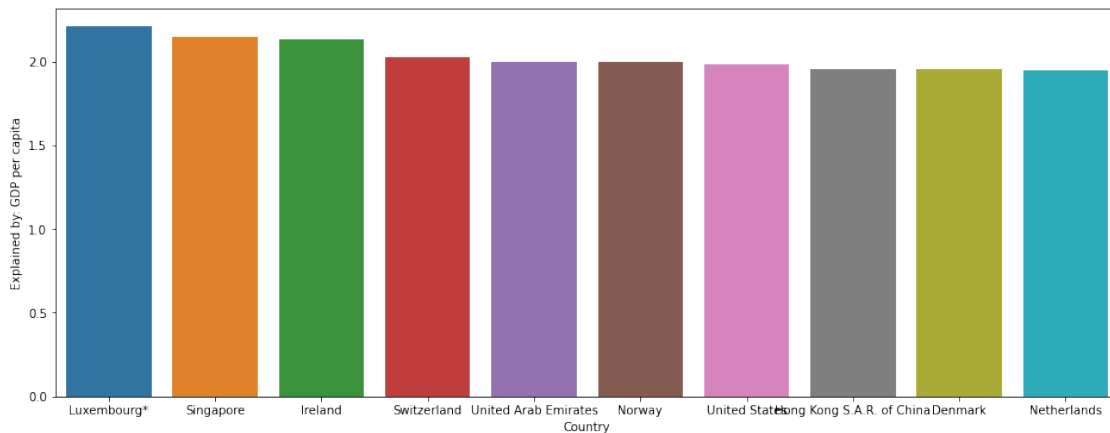
```
[46]: plt.figure(figsize=(10,6))
sns.scatterplot(x=data['Explained by: Healthy life_
↳expectancy'],y=data['Happiness score'])
plt.show()
```



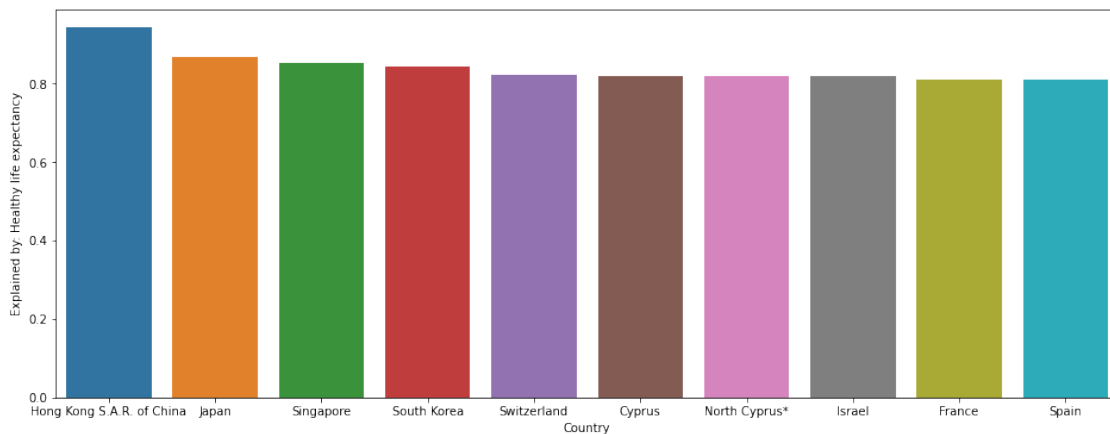
```
[45]: plt.figure(figsize=(16,6))
sns.barplot(x='Country',
            y='Explained by: Social support',
            data=data.nlargest(10, 'Explained by: Social support'))
plt.show()
```



```
[43]: plt.figure(figsize=(16,6))
sns.barplot(x='Country',
            y='Explained by: GDP per capita',
            data=data.nlargest(10, 'Explained by: GDP per capita'))
plt.show()
```

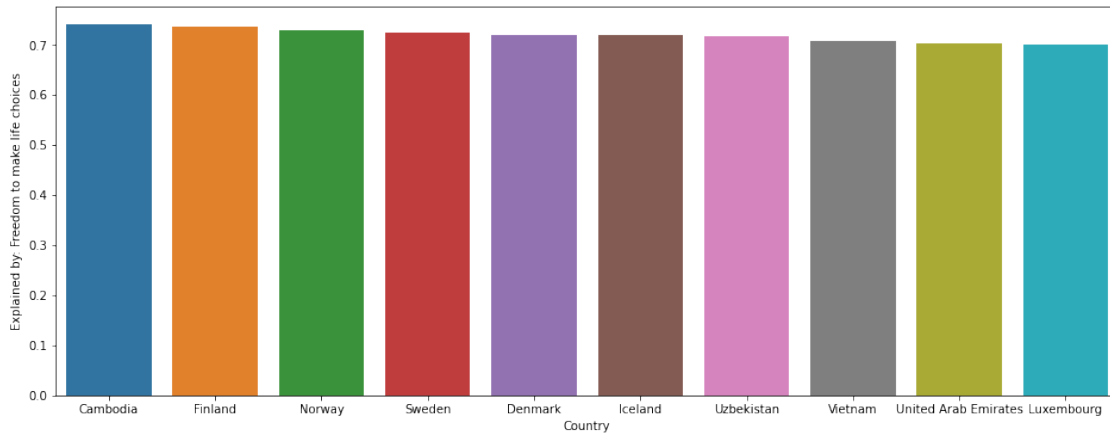


```
[47]: plt.figure(figsize=(16,6))
sns.barplot(x='Country',
            y='Explained by: Healthy life expectancy',
            data=data.nlargest(10, 'Explained by: Healthy life expectancy'))
plt.show()
```

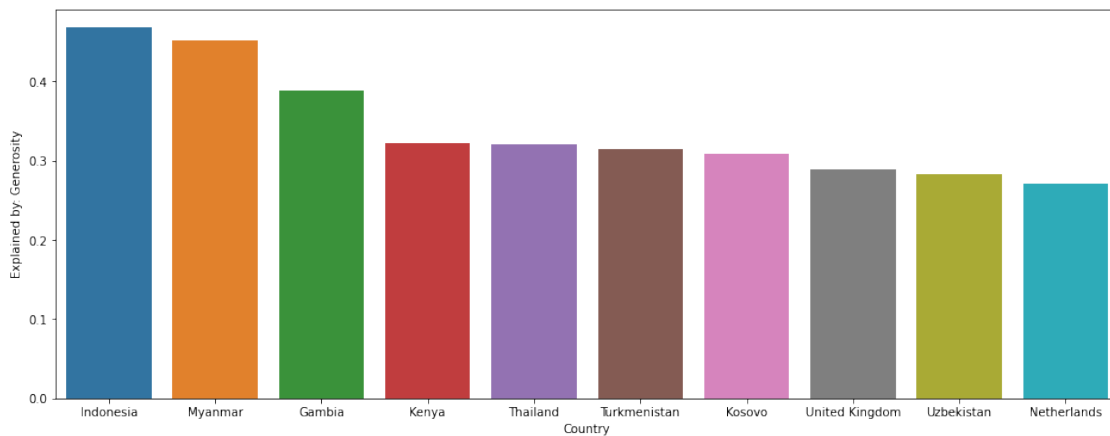


```
[164]: plt.figure(figsize=(16,6))
sns.barplot(x='Country',
            y='Explained by: Freedom to make life choices',
```

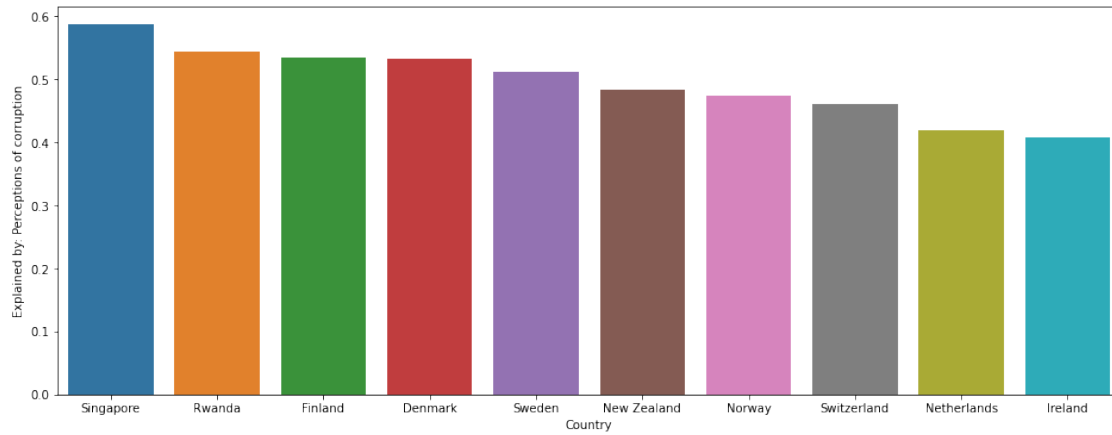
```
data=data.nlargest(10, 'Explained by: Freedom to make life_
↳choices'))
plt.show()
```



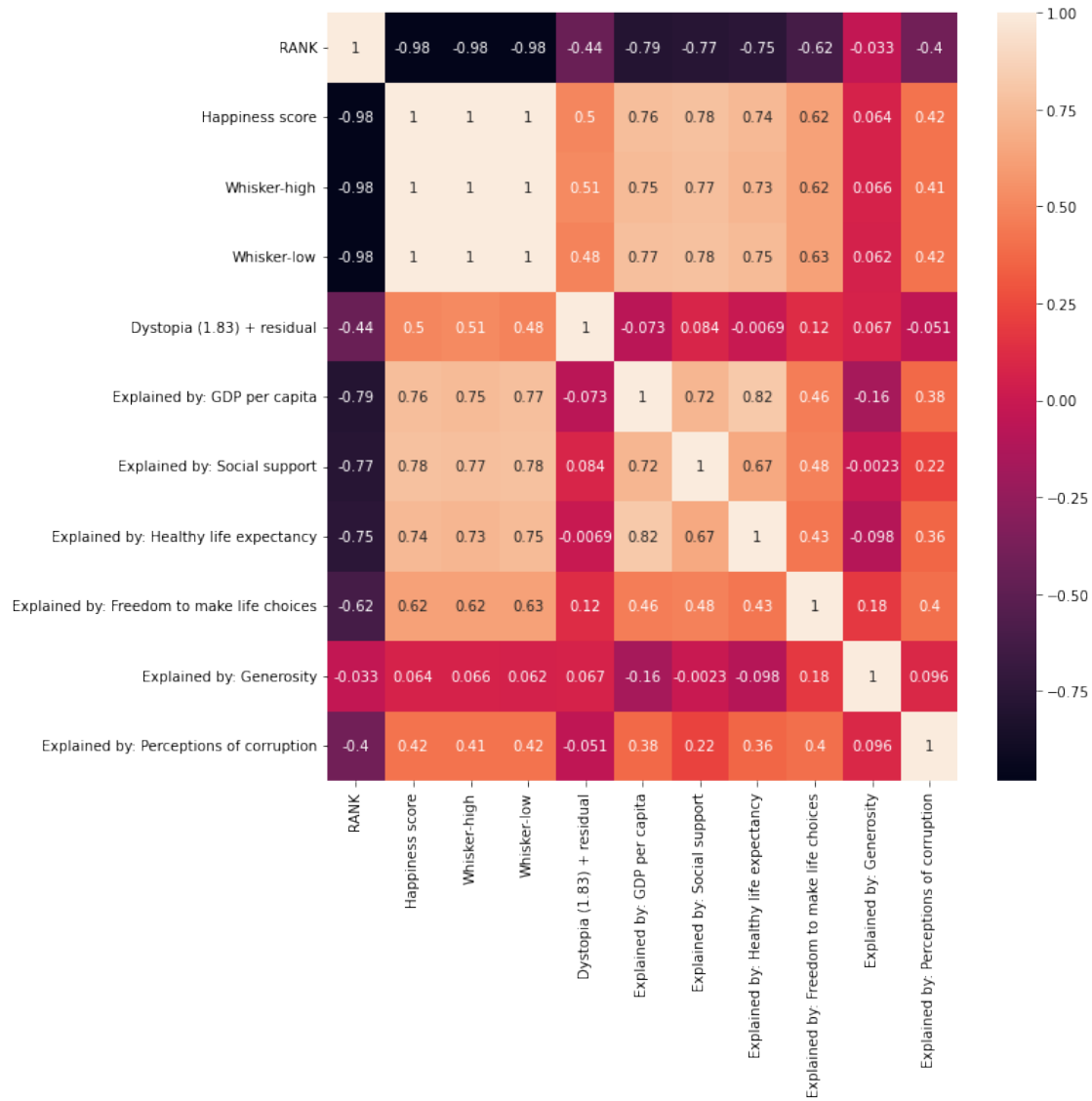
```
[165]: plt.figure(figsize=(16,6))
sns.barplot(x='Country',
            y='Explained by: Generosity',
            data=data.nlargest(10, 'Explained by: Generosity'))
plt.show()
```



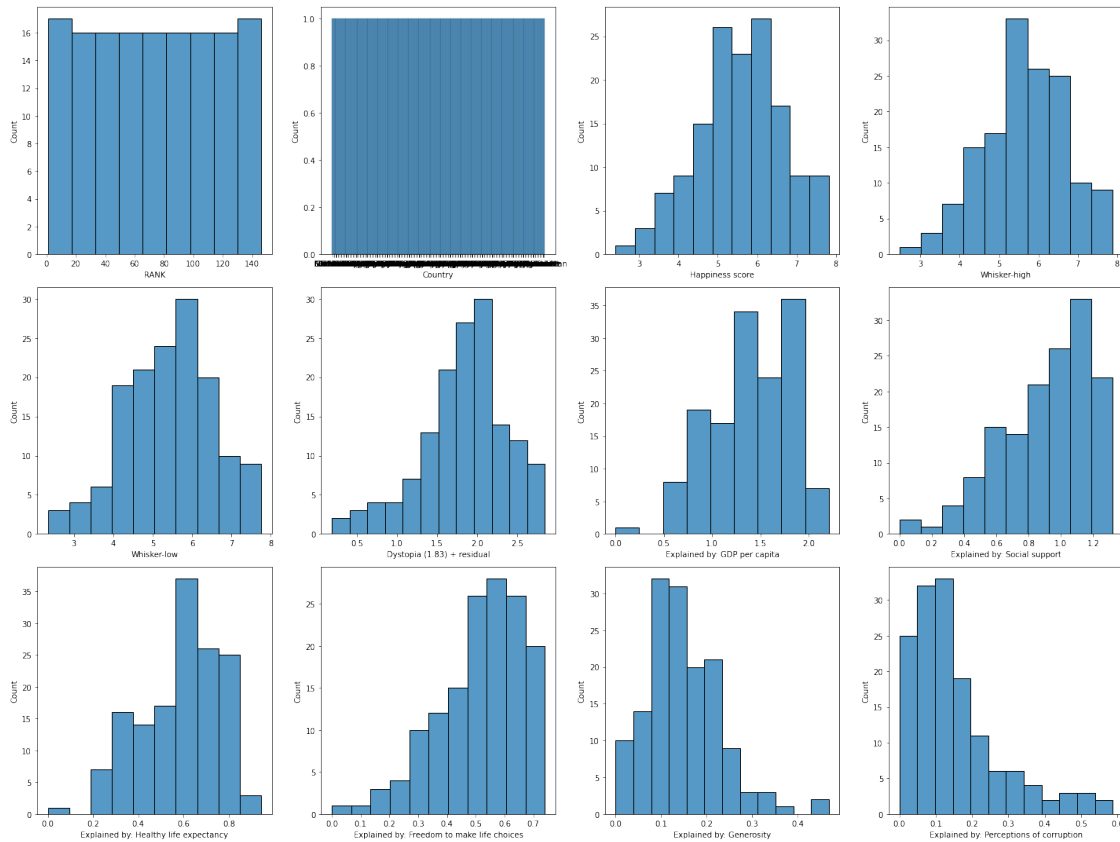
```
[166]: plt.figure(figsize=(16,6))
sns.barplot(x='Country',
            y='Explained by: Perceptions of corruption',
            data=data.nlargest(10, 'Explained by: Perceptions of corruption'))
plt.show()
```



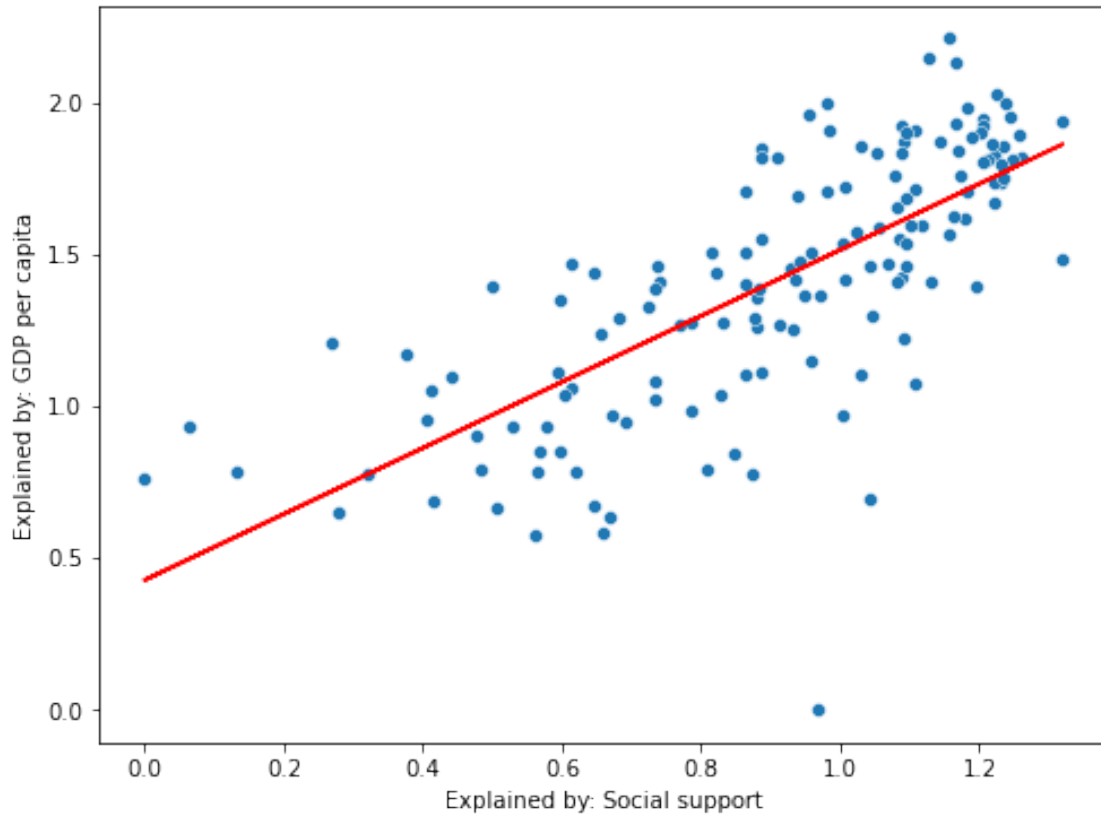
```
[51]: plt.figure(figsize=(10,10))  
sns.heatmap(data.corr(),annot=True);
```



```
[54]: num_rows = (len(data.columns) + 2) // 4
fig, axes = plt.subplots(nrows=num_rows, ncols=4, figsize=(20,15))
for i, col in enumerate(data.columns):
    sns.histplot(data=data, x=col, ax=axes[i // 4, i % 4])
plt.tight_layout()
plt.show()
```



```
[57]: x=data['Explained by: Social support']
y=data['Explained by: GDP per capita']
parameter = np.polyfit(x,y,1)
y2 = parameter[0] * x + parameter[1]
plt.figure(figsize=(8,6))
sns.scatterplot(x=x,y=y)
plt.plot(x, y2, color='r')
plt.show()
```



## 1 Data Cleaning

```
[58]: data["Country"] = data["Country"].str.replace("*", "")
```

C:\Users\MSI\AppData\Local\Temp\ipykernel\_1532\1774426780.py:1: FutureWarning:  
The default value of regex will change from True to False in a future version.  
In addition, single character regular expressions will *not* be treated as  
literal strings when regex=True.

```
data["Country"] = data["Country"].str.replace("*", "")
```

```
[59]: data.shape
```

```
[59]: (146, 12)
```

```
[65]: data.isna().sum()
```

```
[65]: RANK          0
      Country      0
      Happiness score  0
      Whisker-high  0
```



```

Whisker-low                                0
Dystopia (1.83) + residual                  0
Explained by: GDP per capita                0
Explained by: Social support                0
Explained by: Healthy life expectancy      0
Explained by: Freedom to make life choices 0
Explained by: Generosity                    0
Explained by: Perceptions of corruption    0
dtype: int64

```

```
[67]: data.duplicated().sum()
```

```
[67]: 0
```

```
[68]: data["Country"].replace("Palestinian Territories","Palestine",inplace = True)
data["Country"].replace("Hong Kong S.A.R. of China","Hong Kong",inplace = True)
data["Country"].replace("Russia","Russian Federation",inplace = True)
data["Country"].replace("Taiwan Province of China","Taiwan",inplace = True)
```

```
[70]: def countryCode (country_name):
    try:
        return pc.country_name_to_country_alpha2(country_name)
    except:
        return None

if __name__ == "__main__":
    data['Country code']= data.apply(lambda x: countryCode(x.Country), axis = 1)
```

```
[71]: def continent(country_code):
    try:
        return pc.country_alpha2_to_continent_code(country_code)
    except:
        return None

if __name__ == "__main__":
    data['Continent']= data["Country code"].apply(lambda x: continent(x))
```

```
[72]: data.head()
```

```
[72]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low	\
0	1	Finland	7.821	7.886	7.756	
1	2	Denmark	7.636	7.710	7.563	
2	3	Iceland	7.557	7.651	7.464	
3	4	Switzerland	7.512	7.586	7.437	
4	5	Netherlands	7.415	7.471	7.359	

```
Dystopia (1.83) + residual  Explained by: GDP per capita \
```

0	2.518	1.892
1	2.226	1.953
2	2.320	1.936
3	2.153	2.026
4	2.137	1.945

	Explained by: Social support	Explained by: Healthy life expectancy \
0	1.258	0.775
1	1.243	0.777
2	1.320	0.803
3	1.226	0.822
4	1.206	0.787

	Explained by: Freedom to make life choices	Explained by: Generosity \
0	0.736	0.109
1	0.719	0.188
2	0.718	0.270
3	0.677	0.147
4	0.651	0.271

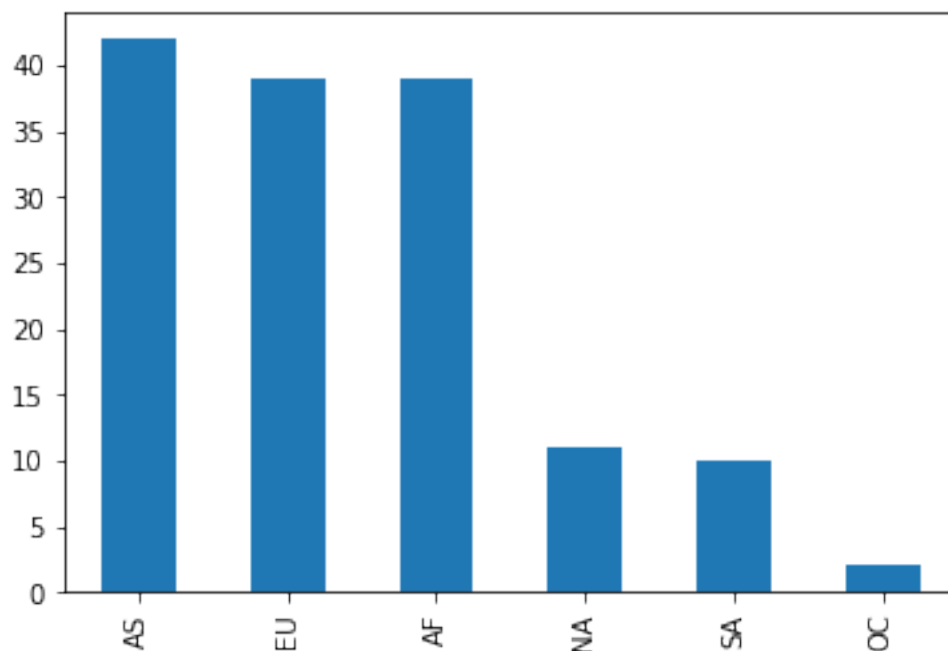
	Explained by: Perceptions of corruption	Country code	Continent
0	0.534	FI	EU
1	0.532	DK	EU
2	0.191	IS	EU
3	0.461	CH	EU
4	0.419	NL	EU

```
[76]: data['Continent'].value_counts()
```

```
[76]: AS    42
      EU    39
      AF    39
      NA    11
      SA    10
      OC     2
      Name: Continent, dtype: int64
```

```
[85]: data['Continent'].value_counts().plot(kind='bar')
```

```
[85]: <AxesSubplot:>
```



## 2 Continent Wise Data Analysis

### 2.0.1 Asia

```
[89]: ASdata=data.filter(["RANK","Country"],axis=1)
ASdata = data[data["Continent"] == "AS"]
ASdata.head()
```

```
[89]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low \
8	9	Israel	7.364	7.426	7.301
20	21	Bahrain	6.647	6.779	6.514
23	24	United Arab Emirates	6.576	6.660	6.492
24	25	Saudi Arabia	6.523	6.637	6.409
25	26	Taiwan	6.512	6.596	6.429

	Dystopia (1.83) + residual	Explained by: GDP per capita \
8	2.634	1.826
20	2.092	1.854
23	1.809	1.998
24	2.075	1.870
25	2.002	1.897

	Explained by: Social support	Explained by: Healthy life expectancy \
8	1.221	0.818
20	1.029	0.625

23	0.980	0.633
24	1.092	0.577
25	1.095	0.733

	Explained by: Freedom to make life choices	Explained by: Generosity \
8	0.568	0.155
20	0.693	0.199
23	0.702	0.204
24	0.651	0.078
25	0.542	0.075

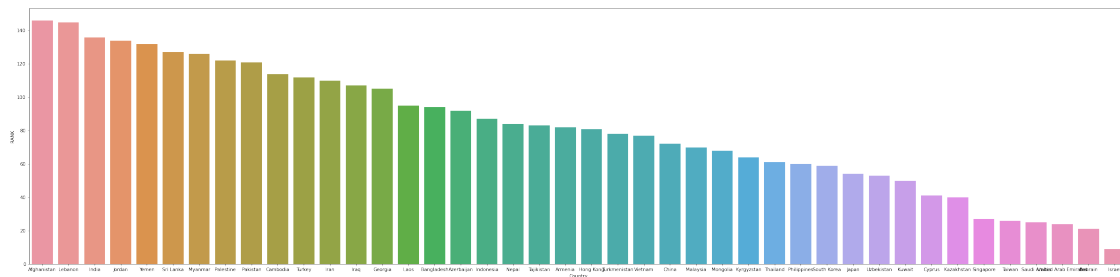
	Explained by: Perceptions of corruption	Country code	Continent
8	0.143	IL	AS
20	0.155	BH	AS
23	0.250	AE	AS
24	0.180	SA	AS
25	0.168	TW	AS

```
[90]: ASdata.shape
```

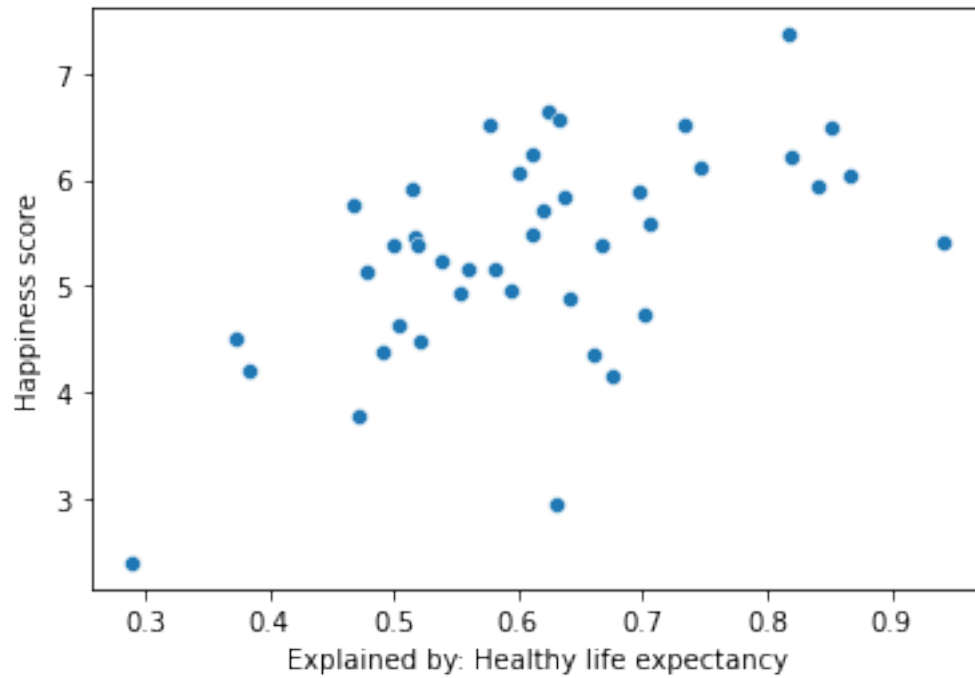
```
[90]: (42, 14)
```

```
[102]: plt.figure(figsize=(42,10))
sns.barplot(x='Country',
            y='RANK',
            data=ASdata.nlargest(42, 'RANK'))
```

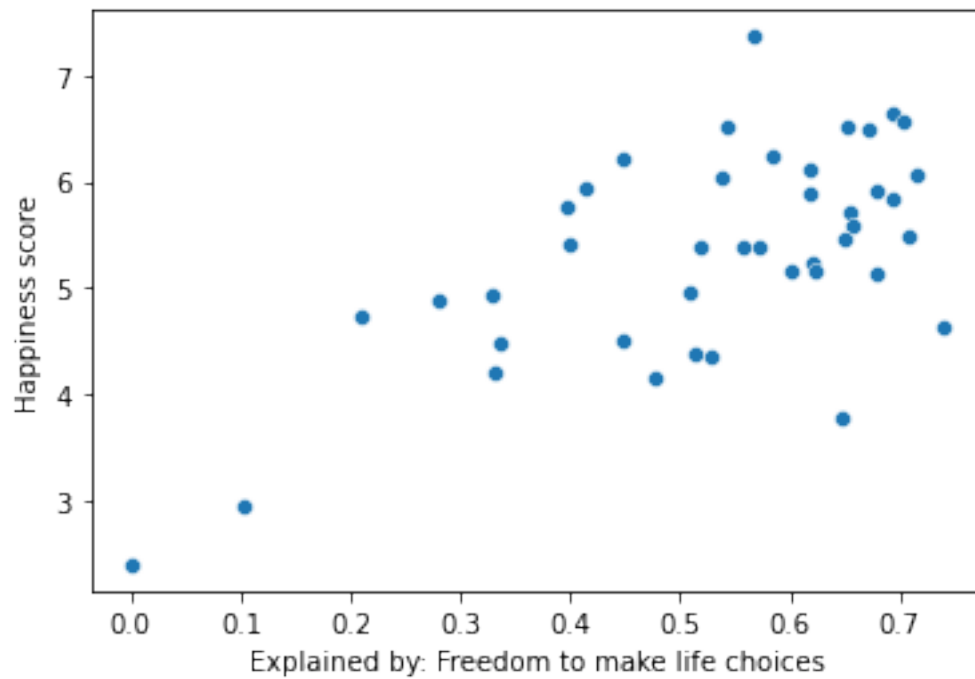
```
[102]: <AxesSubplot:xlabel='Country', ylabel='RANK'>
```



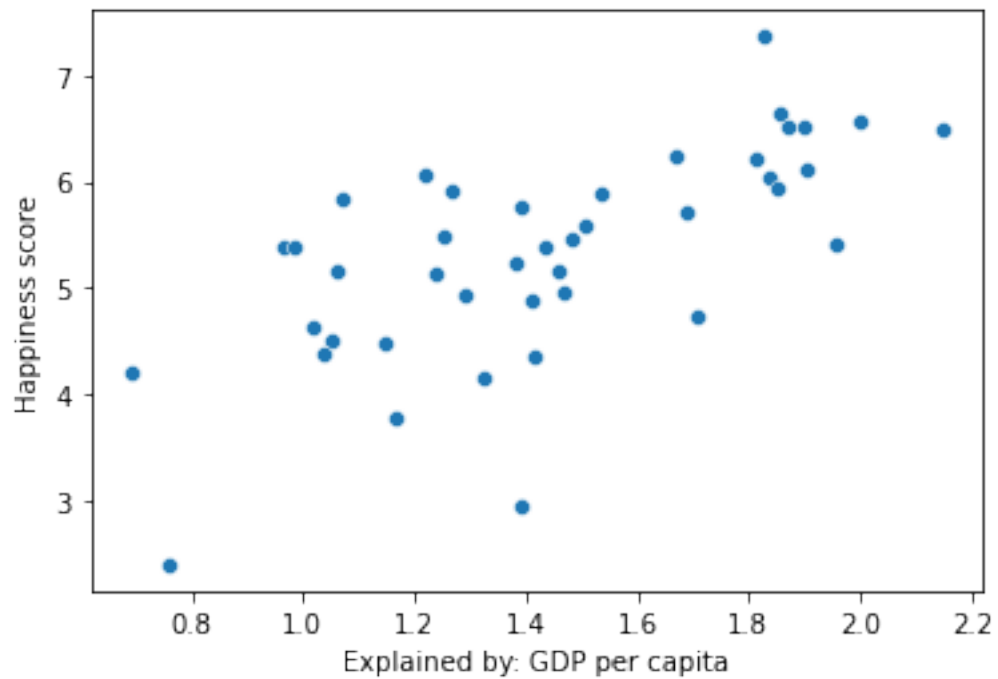
```
[104]: sns.scatterplot(x=ASdata['Explained by: Healthy life_
↪expectancy'],y=ASdata['Happiness score'])
plt.show()
```



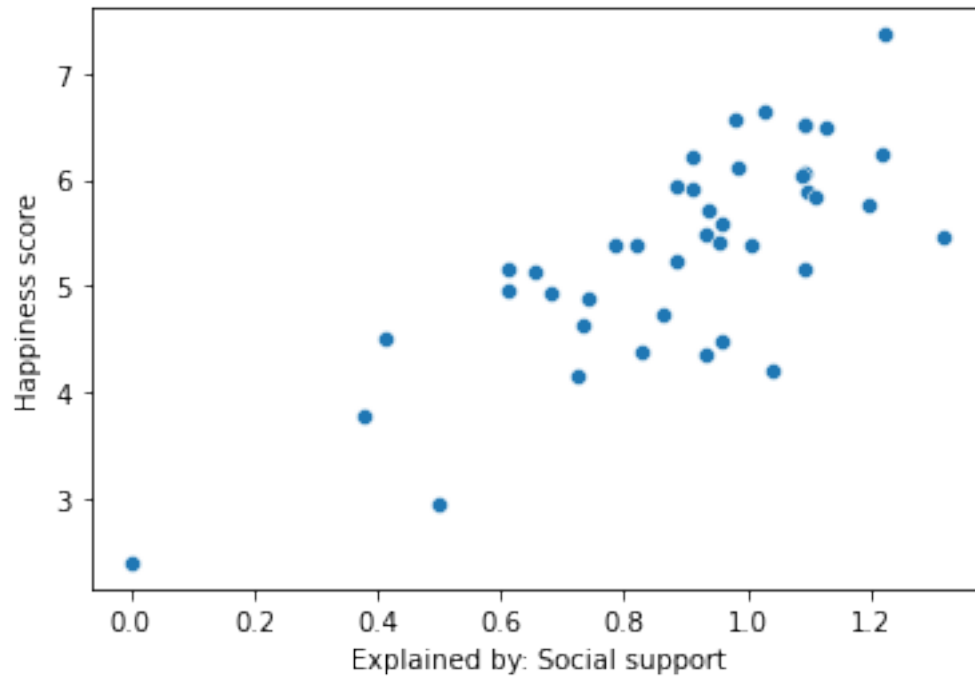
```
[105]: sns.scatterplot(x=ASdata['Explained by: Freedom to make life_
↳choices'],y=ASdata['Happiness score'])
plt.show()
```



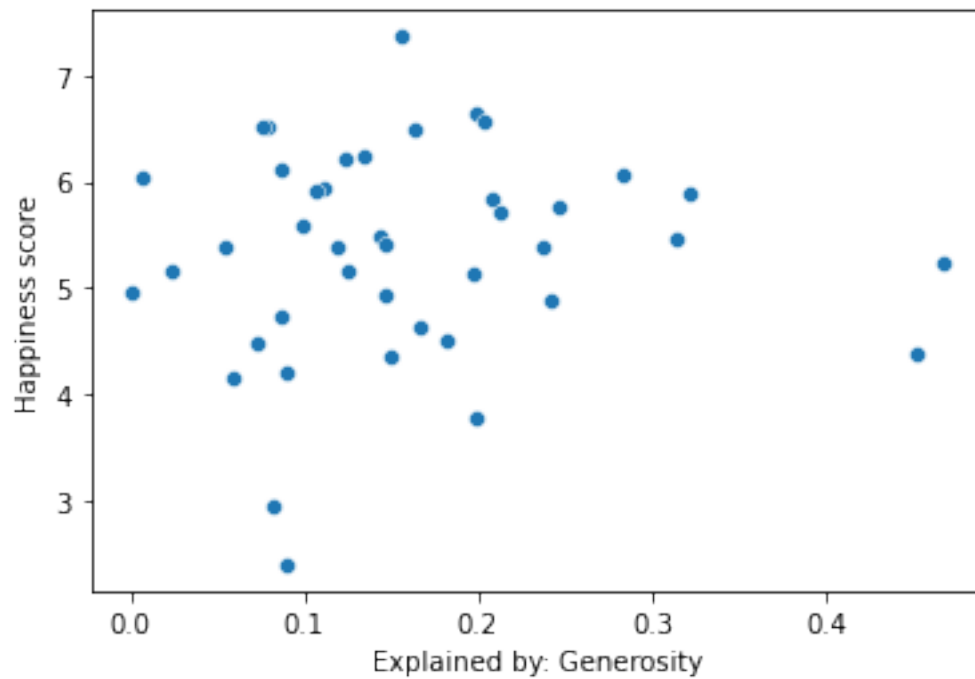
```
[106]: sns.scatterplot(x=ASdata['Explained by: GDP per capita'],y=ASdata['Happiness_↵score'])  
plt.show()
```



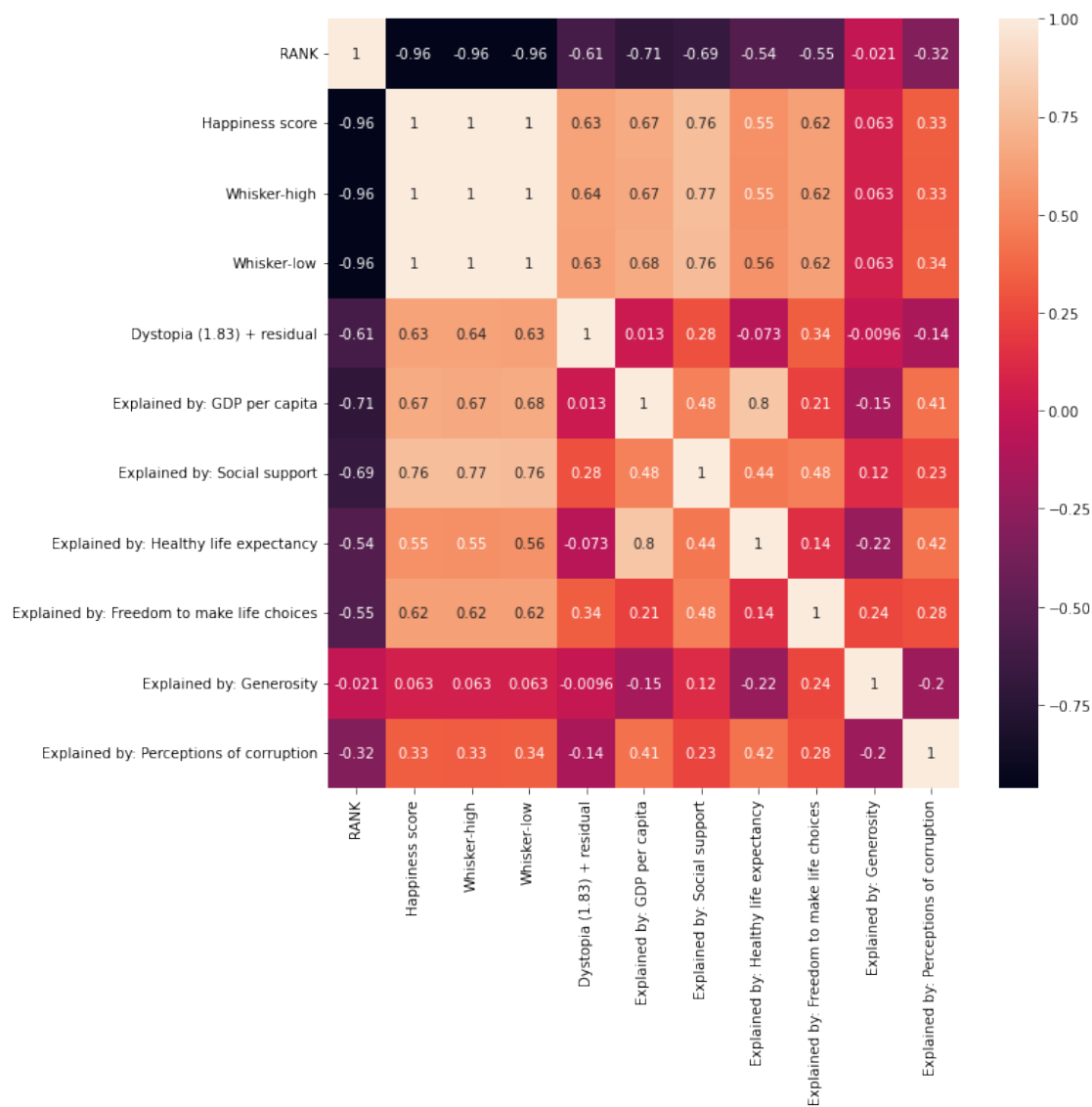
```
[107]: sns.scatterplot(x=ASdata['Explained by: Social support'],y=ASdata['Happiness_↵score'])  
plt.show()
```



```
[108]: sns.scatterplot(x=ASdata['Explained by: Generosity'],y=ASdata['Happiness_↪score'])
plt.show()
```



```
[152]: plt.figure(figsize=(10,10))
sns.heatmap(ASdata.corr(),annot=True);
```



```
[153]: ASdata.to_csv(path_or_buf="C:\\Users\\MSI\\Desktop\\VIT DS\\Python for Data_
Science\\ASdata.csv")
```



## 2.1 North America

```
[109]: NAdata=data.filter(["RANK","Country"],axis=1)
NAdata = data[data["Continent"] == "NA"]
NAdata.head()
```

```
[109]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low	\
14	15	Canada	7.025	7.107	6.943	
15	16	United States	6.977	7.065	6.888	
22	23	Costa Rica	6.582	6.683	6.481	
36	37	Panama	6.309	6.464	6.154	
38	39	Guatemala	6.262	6.460	6.064	

	Dystopia (1.83) + residual	Explained by: GDP per capita	\
14	1.924	1.886	
15	2.214	1.982	
22	2.346	1.584	
36	2.086	1.715	
38	2.746	1.274	

	Explained by: Social support	Explained by: Healthy life expectancy	\
14	1.188	0.783	
15	1.182	0.628	
22	1.054	0.744	
36	1.107	0.709	
38	0.831	0.522	

	Explained by: Freedom to make life choices	Explained by: Generosity	\
14	0.659	0.217	
15	0.574	0.220	
22	0.661	0.089	
36	0.592	0.049	
38	0.662	0.112	

	Explained by: Perceptions of corruption	Country code	Continent
14	0.368	CA	NA
15	0.177	US	NA
22	0.102	CR	NA
36	0.051	PA	NA
38	0.115	GT	NA

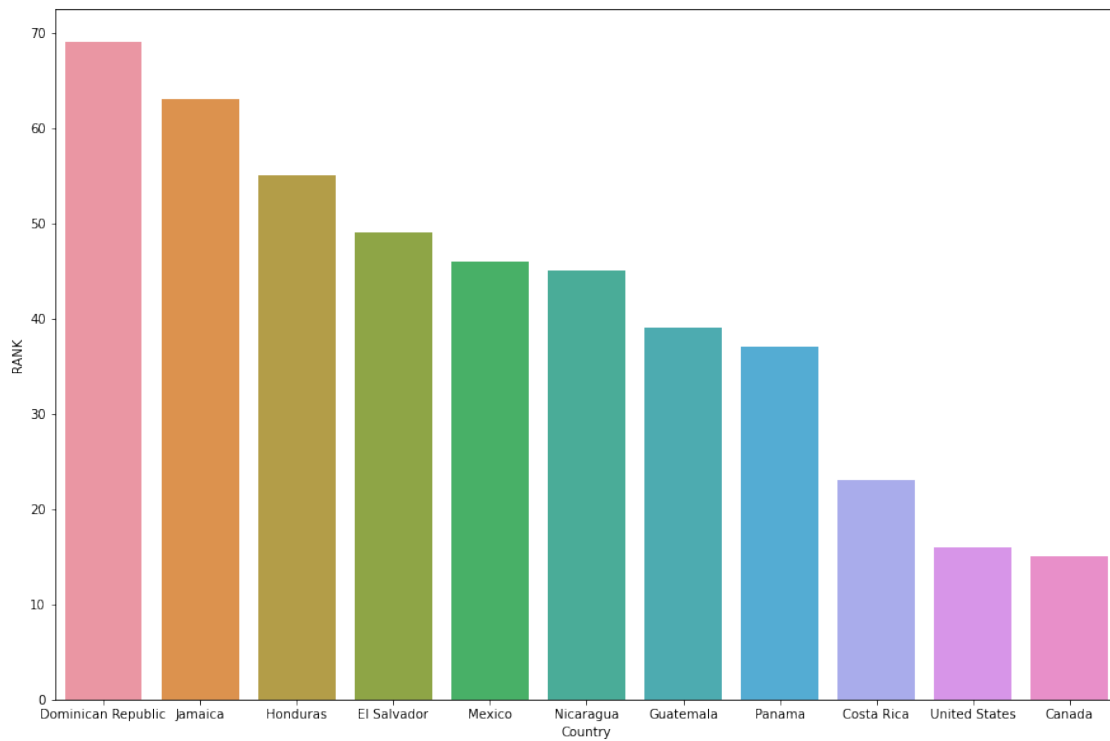
```
[110]: NAdata.shape
```

```
[110]: (11, 14)
```

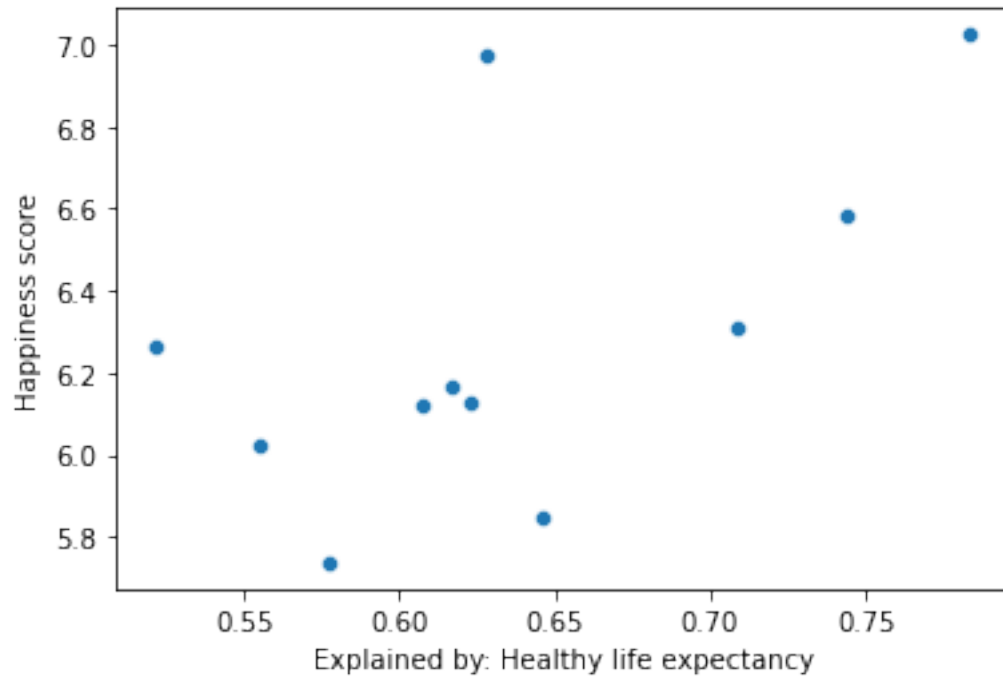
```
[113]: plt.figure(figsize=(15,10))
sns.barplot(x='Country',
            y='RANK',
```

```
data=Ndata.nlargest(11, 'RANK')
```

```
[113]: <AxesSubplot:xlabel='Country', ylabel='RANK'>
```



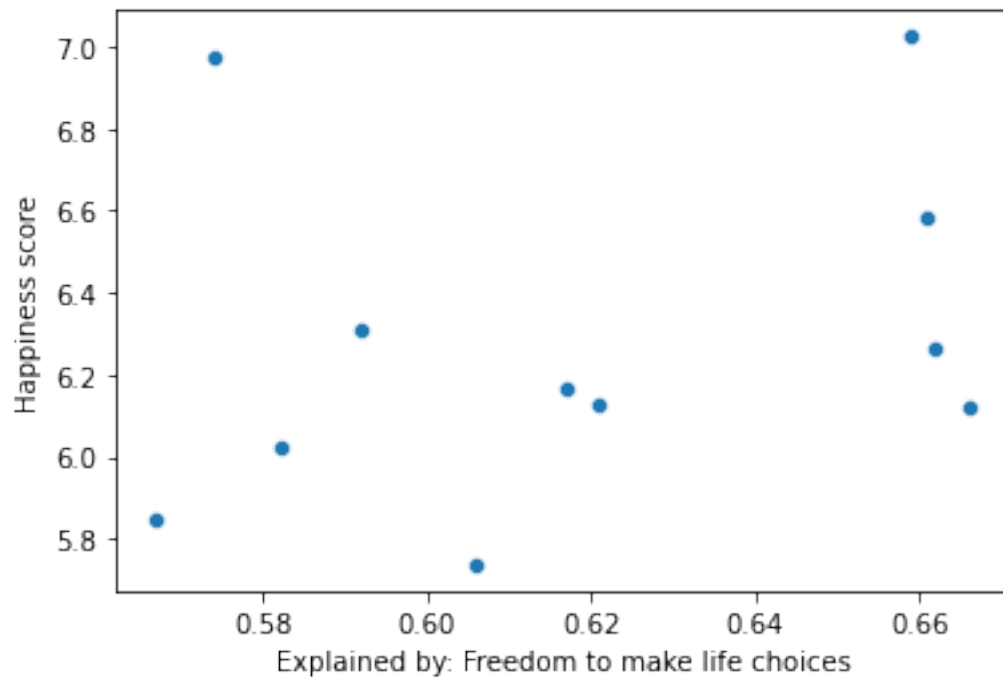
```
[114]: sns.scatterplot(x=Ndata['Explained by: Healthy life_↵  
↵expectancy'],y=Ndata['Happiness score'])  
plt.show()
```



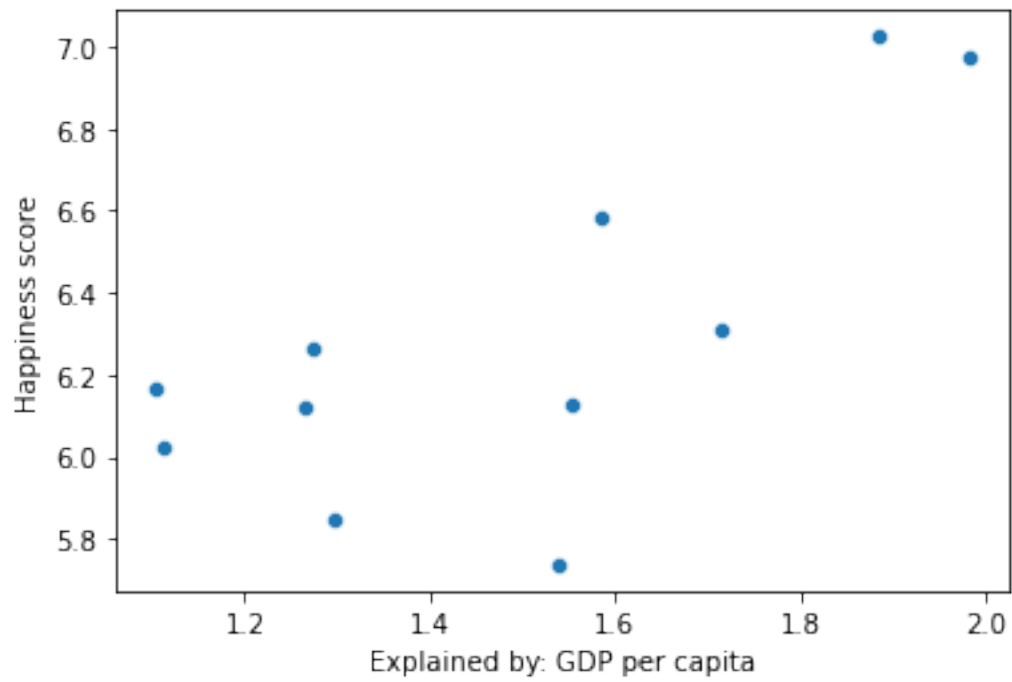
```
[115]: sns.scatterplot(x=NAdat['Explained by: Freedom to make life_  

↪choices'],y=NAdat['Happiness score'])  

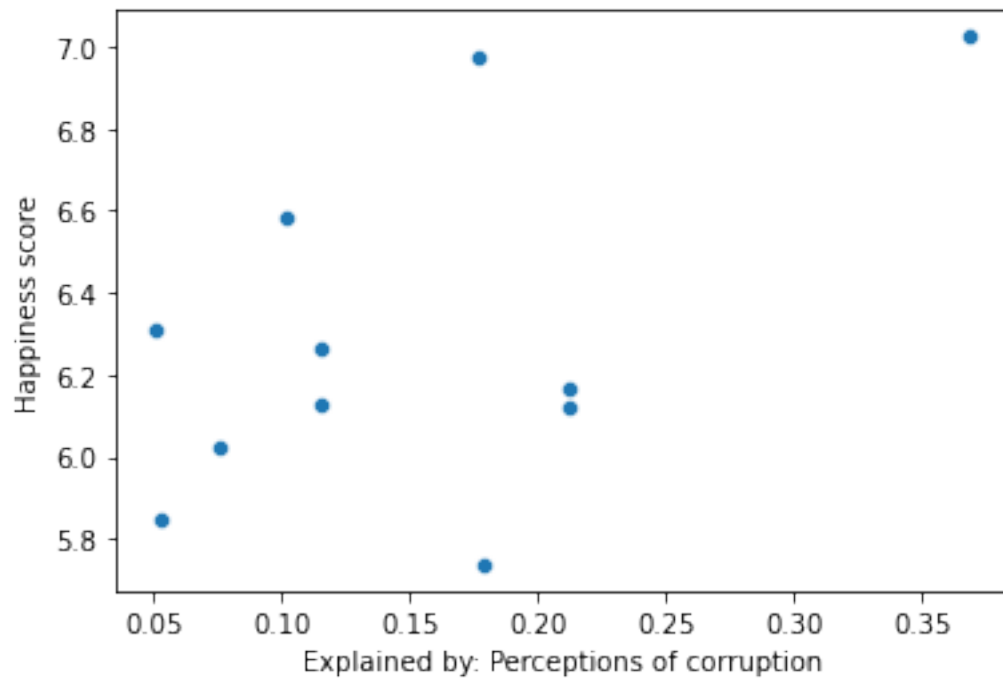
plt.show()
```



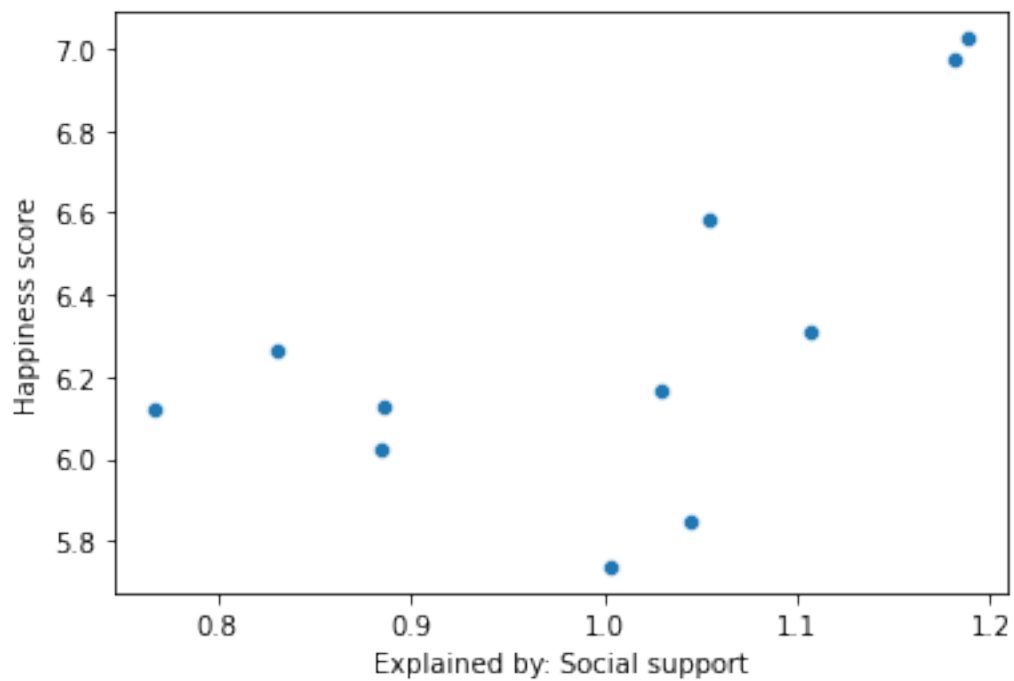
```
[116]: sns.scatterplot(x=NAdata['Explained by: GDP per capita'],y=NAdata['Happiness_↵score'])  
plt.show()
```



```
[117]: sns.scatterplot(x=NAdata['Explained by: Perceptions of_↵corruption'],y=NAdata['Happiness score'])  
plt.show()
```



```
[118]: sns.scatterplot(x=NAdata['Explained by: Social support'],y=NAdata['Happiness_↪score'])
plt.show()
```



## 2.2 South America

```
[121]: SAdata=data.filter(["RANK","Country"],axis=1)
SAdata = data[data["Continent"] == "SA"]
SAdata.head()
```

```
[121]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low	\
29	30	Uruguay	6.474	6.562	6.386	
37	38	Brazil	6.293	6.384	6.202	
43	44	Chile	6.172	6.262	6.082	
56	57	Argentina	5.967	6.090	5.844	
65	66	Colombia	5.781	5.903	5.659	

	Dystopia (1.83) + residual	Explained by: GDP per capita	\
29	1.974	1.615	
37	2.361	1.462	
43	2.040	1.651	
56	1.891	1.592	
65	1.971	1.452	

	Explained by: Social support	Explained by: Healthy life expectancy	\
29	1.180	0.672	
37	1.044	0.615	
43	1.080	0.748	
56	1.102	0.662	
65	0.929	0.720	

	Explained by: Freedom to make life choices	Explained by: Generosity	\
29	0.665	0.103	
37	0.546	0.131	
43	0.460	0.124	
56	0.555	0.081	
65	0.545	0.087	

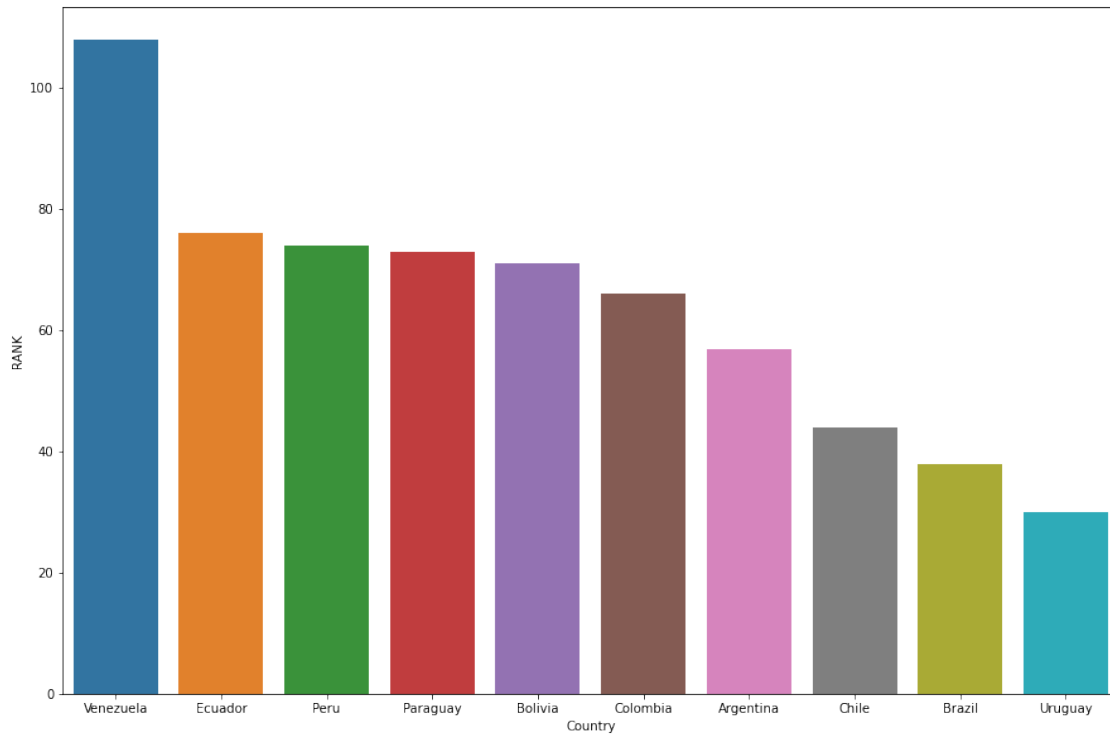
	Explained by: Perceptions of corruption	Country code	Continent
29	0.265	UY	SA
37	0.134	BR	SA
43	0.069	CL	SA
56	0.085	AR	SA
65	0.077	CO	SA

```
[122]: SAdata.shape
```

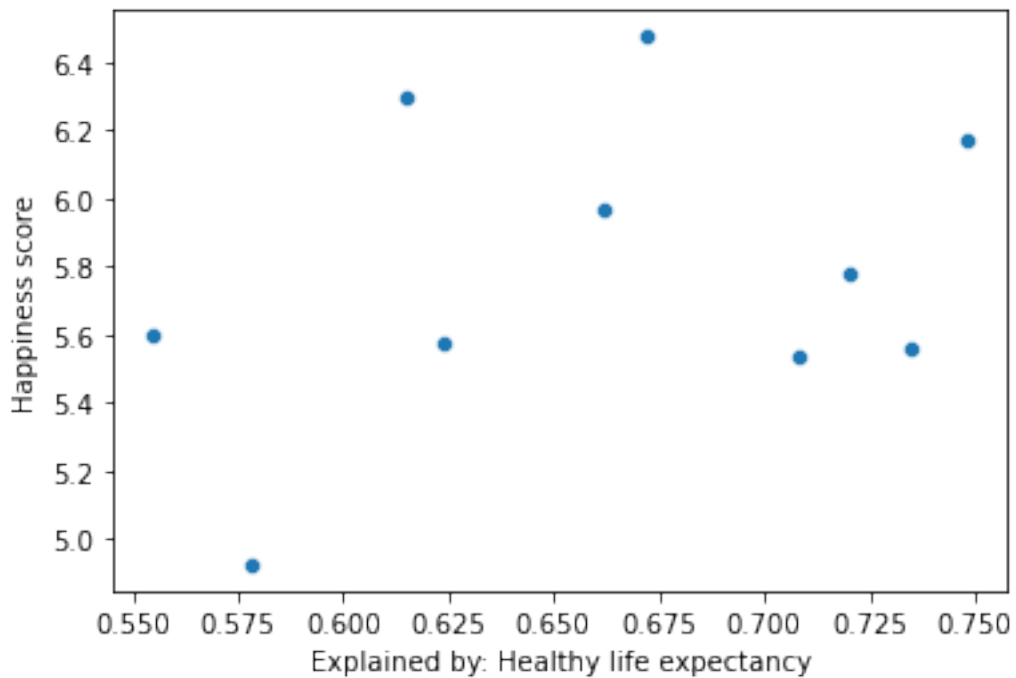
```
[122]: (10, 14)
```

```
[123]: plt.figure(figsize=(15,10))
sns.barplot(x='Country',
            y='RANK',
            data=SAdata.nlargest(10, 'RANK'))
```

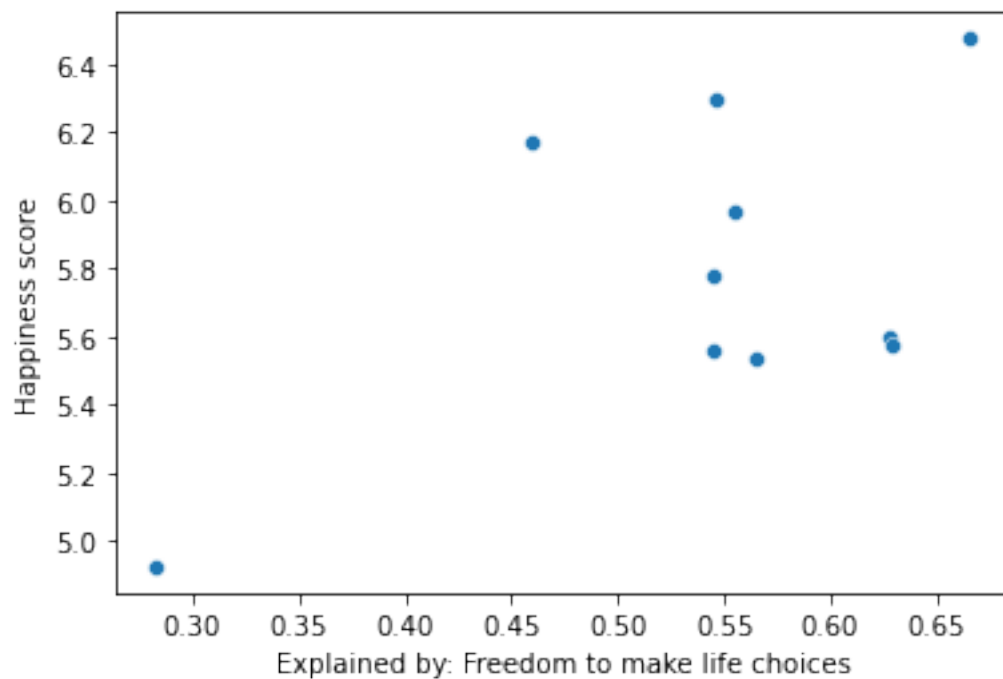
```
[123]: <AxesSubplot:xlabel='Country', ylabel='RANK'>
```



```
[124]: sns.scatterplot(x=SAdata['Explained by: Healthy life_
↪expectancy'], y=SAdata['Happiness score'])
plt.show()
```

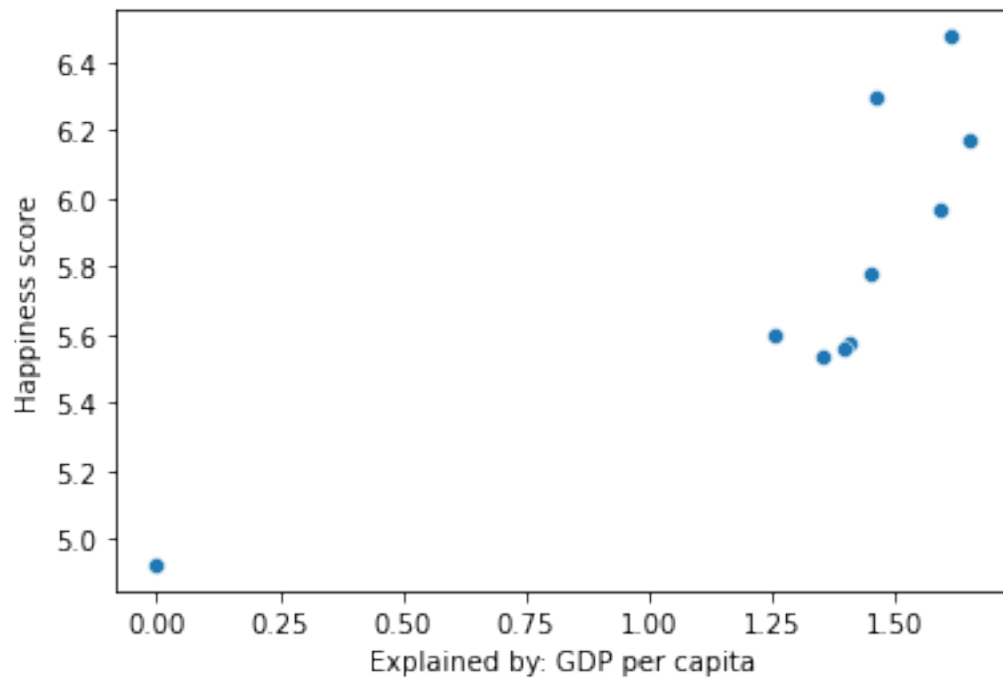


```
[125]: sns.scatterplot(x=SAdata['Explained by: Freedom to make life_
↳choices'],y=SAdata['Happiness score'])
plt.show()
```

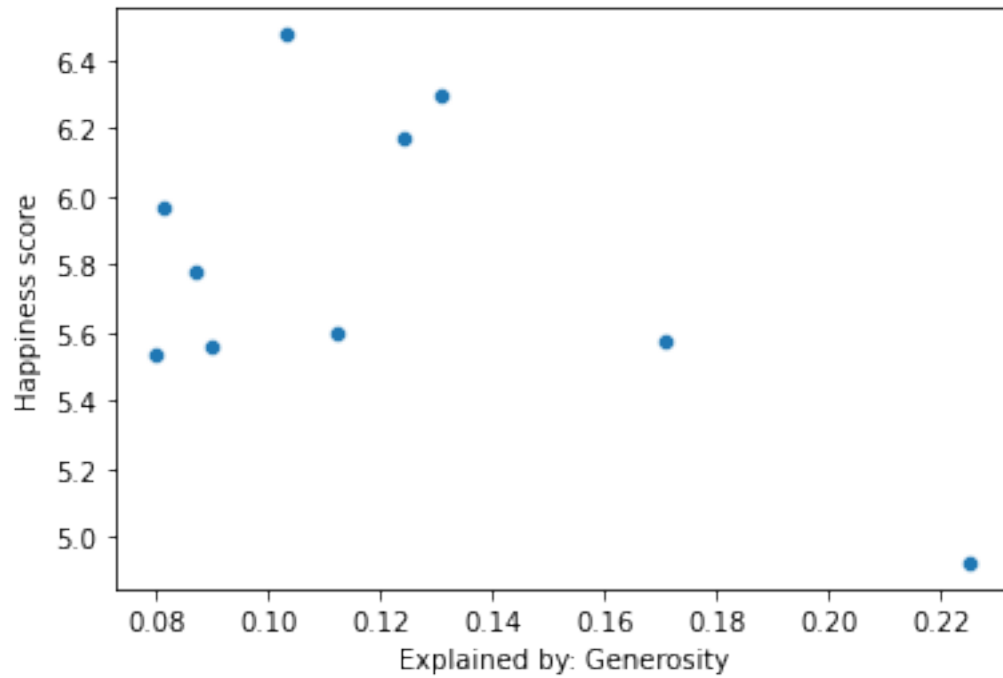




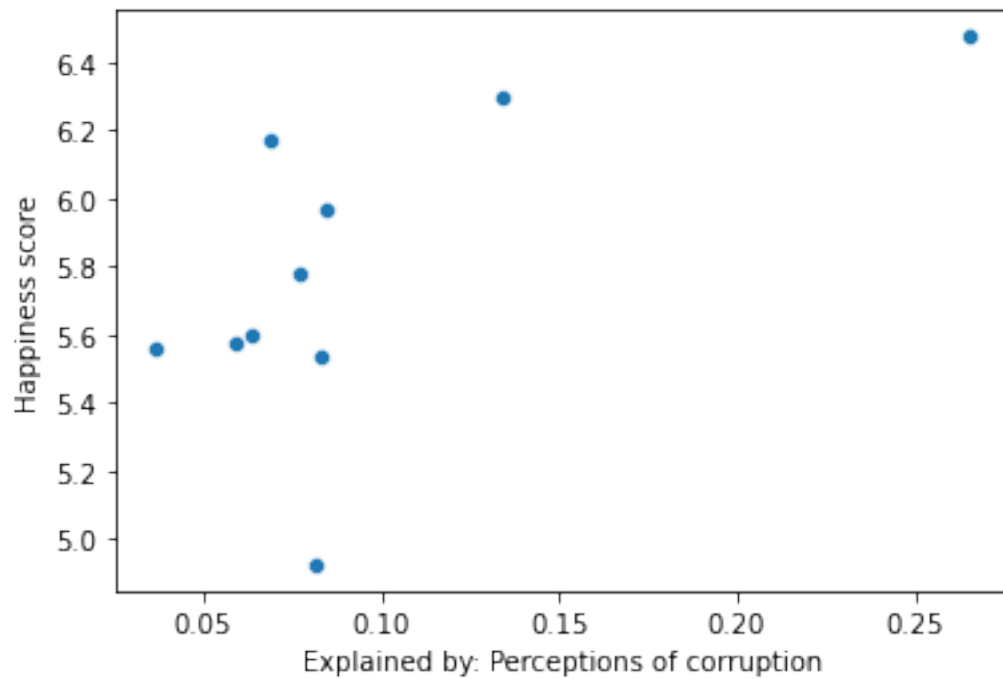
```
[126]: sns.scatterplot(x=SAdata['Explained by: GDP per capita'],y=SAdata['Happiness_↵score'])  
plt.show()
```



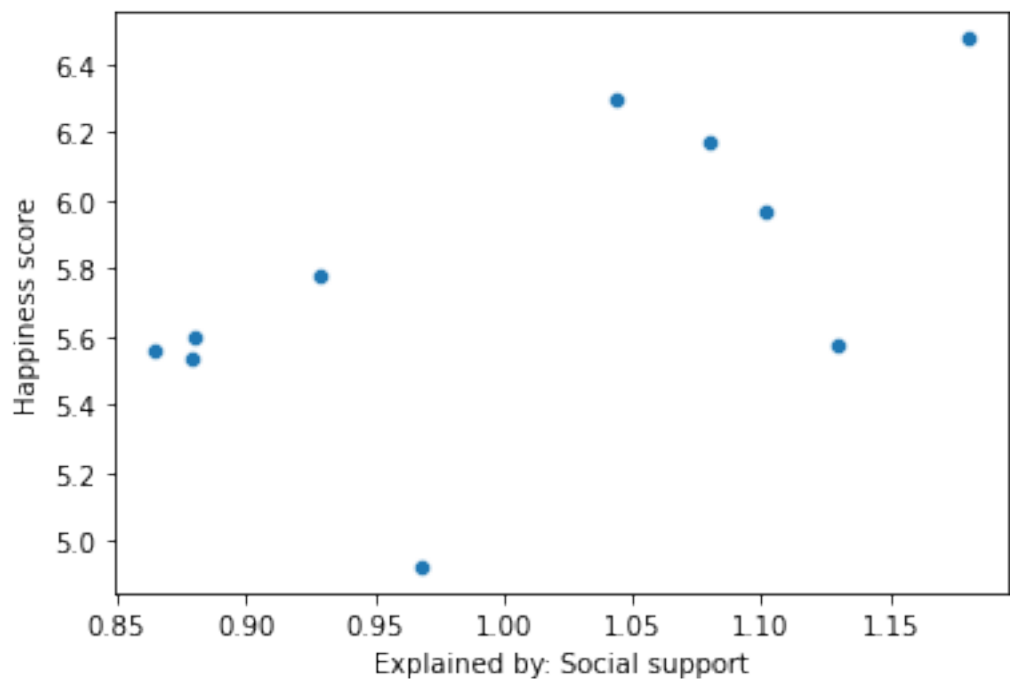
```
[127]: sns.scatterplot(x=SAdata['Explained by: Generosity'],y=SAdata['Happiness_↵score'])  
plt.show()
```



```
[128]: sns.scatterplot(x=SAdata['Explained by: Perceptions of_
↳corruption'],y=SAdata['Happiness score'])
plt.show()
```



```
[129]: sns.scatterplot(x=SAdata['Explained by: Social support'],y=SAdata['Happiness_
↪score'])
plt.show()
```



## 2.3 Africa

```
[131]: AFdata=data.filter(["RANK","Country"],axis=1)
AFdata = data[data["Continent"] == "AF"]
AFdata.head()
```

```
[131]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low	\
51	52	Mauritius	6.071	6.184	5.958	
85	86	Libya	5.330	5.543	5.118	
87	88	Ivory Coast	5.235	5.384	5.085	
90	91	South Africa	5.194	5.315	5.073	
92	93	Gambia	5.164	5.409	4.918	

	Dystopia (1.83) + residual	Explained by: GDP per capita	\
51	1.969	1.591	
85	1.544	1.476	
87	2.652	1.094	
90	1.742	1.425	
92	2.531	0.785	

	Explained by: Social support	Explained by: Healthy life expectancy \
51	1.116	0.568
85	0.943	0.606
87	0.442	0.322
90	1.088	0.361
92	0.621	0.369

	Explained by: Freedom to make life choices	Explained by: Generosity \
51	0.589	0.131
85	0.477	0.106
87	0.451	0.149
90	0.442	0.089
92	0.367	0.388

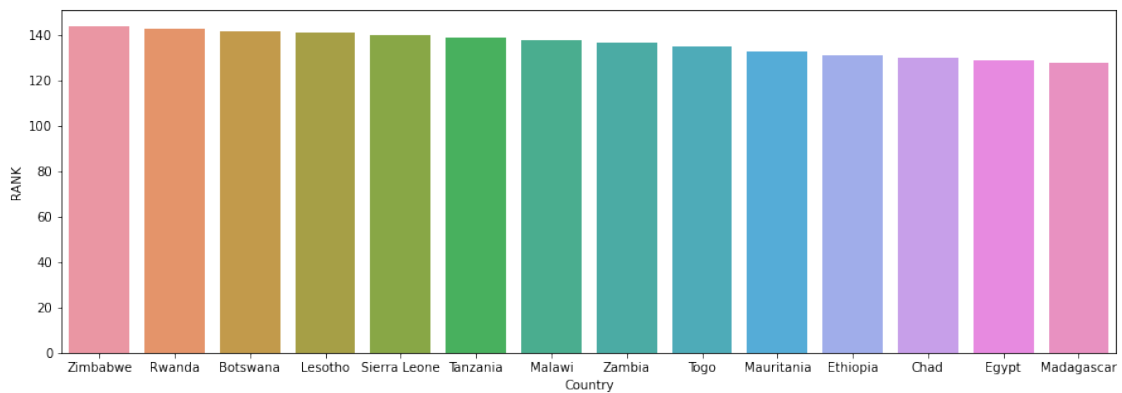
	Explained by: Perceptions of corruption	Country code	Continent
51	0.107	MU	AF
85	0.179	LY	AF
87	0.124	CI	AF
90	0.046	ZA	AF
92	0.103	GM	AF

```
[132]: AFdata.shape
```

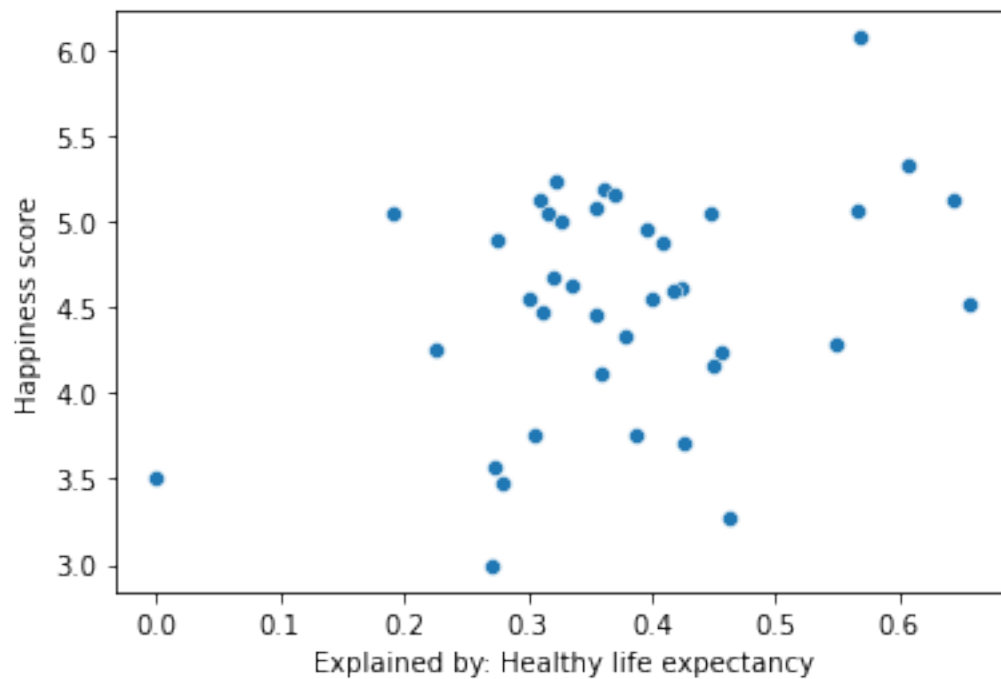
```
[132]: (39, 14)
```

```
[134]: plt.figure(figsize=(15,5))
sns.barplot(x='Country',
            y='RANK',
            data=AFdata.nlargest(14, 'RANK'))
```

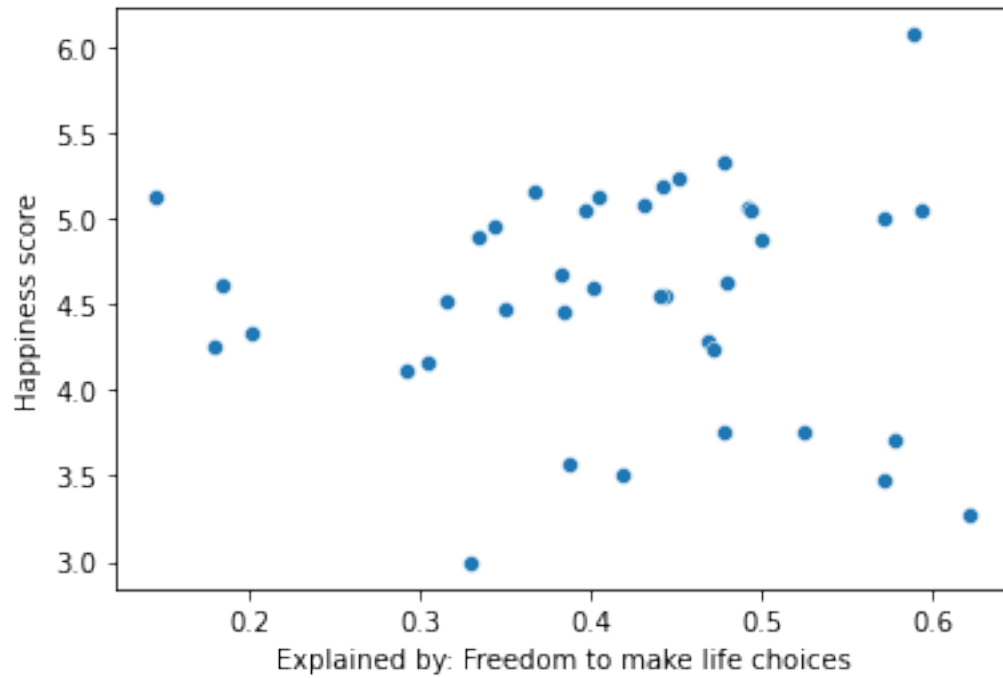
```
[134]: <AxesSubplot:xlabel='Country', ylabel='RANK'>
```



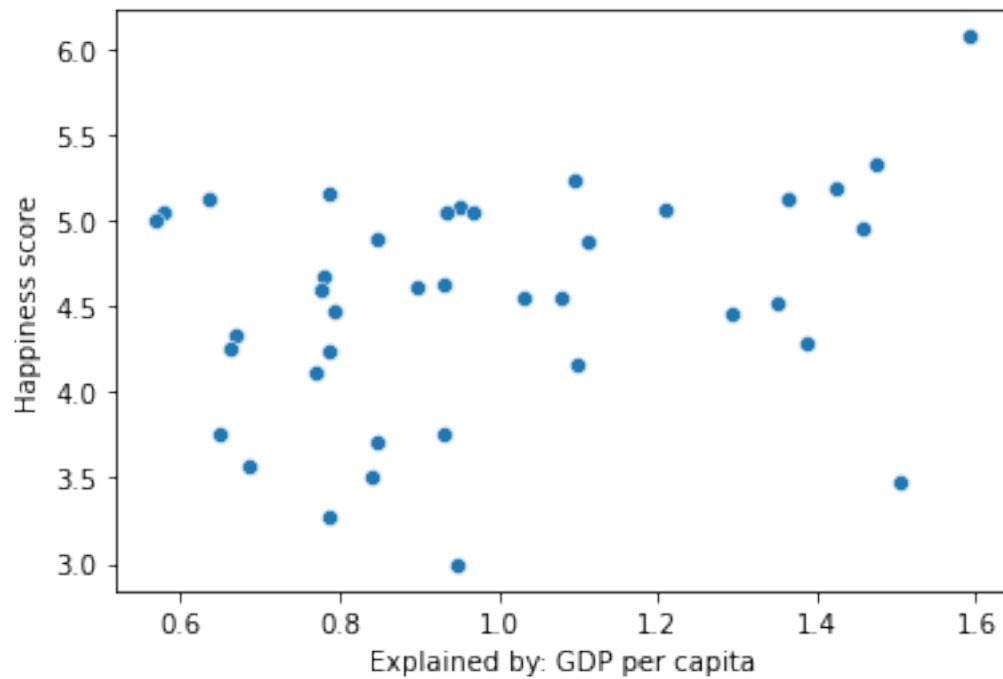
```
[135]: sns.scatterplot(x=AFdata['Explained by: Healthy life_↵  
↵expectancy'],y=AFdata['Happiness score'])  
plt.show()
```



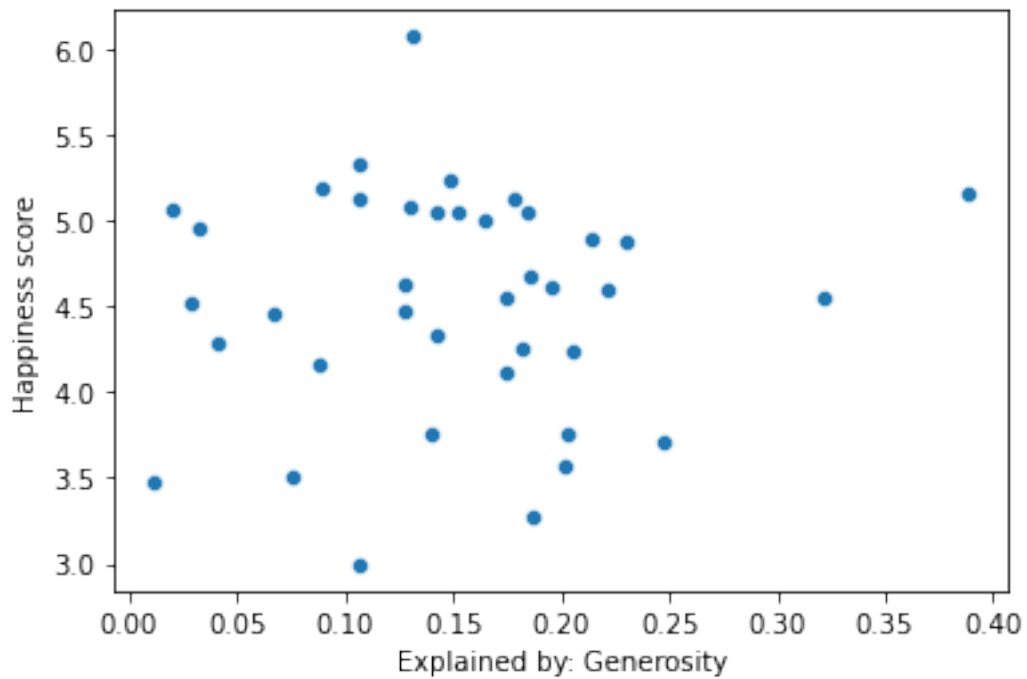
```
[136]: sns.scatterplot(x=AFdata['Explained by: Freedom to make life_↵  
↵choices'],y=AFdata['Happiness score'])  
plt.show()
```



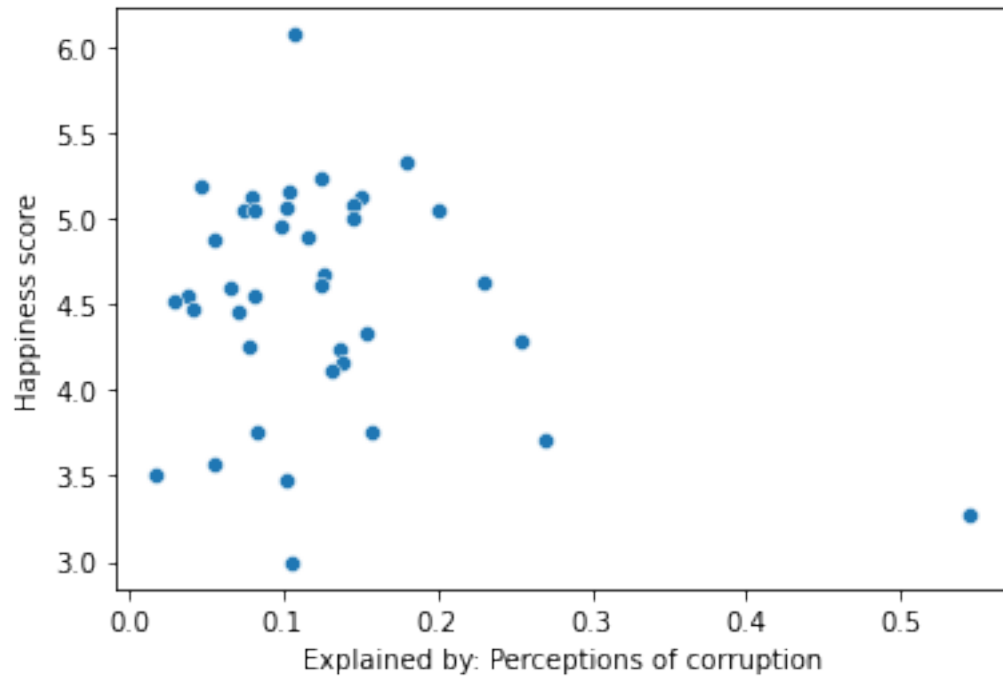
```
[137]: sns.scatterplot(x=AFdata['Explained by: GDP per capita'],y=AFdata['Happiness_
↪score'])
plt.show()
```



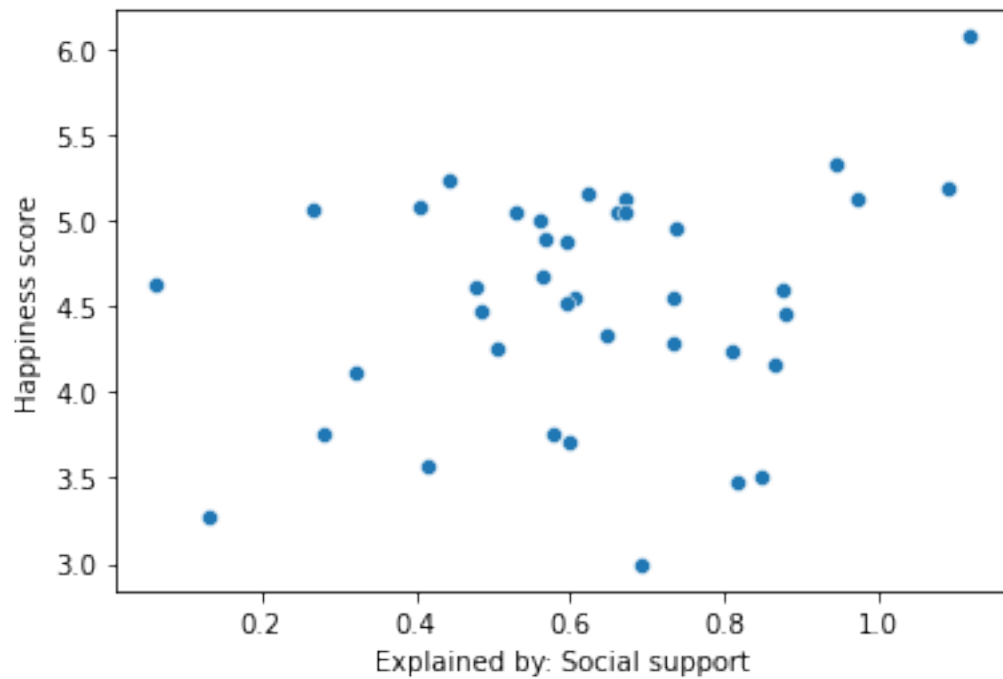
```
[138]: sns.scatterplot(x=AFdata['Explained by: Generosity'],y=AFdata['Happiness_↵score'])  
plt.show()
```



```
[139]: sns.scatterplot(x=AFdata['Explained by: Perceptions of_↵corruption'],y=AFdata['Happiness score'])  
plt.show()
```



```
[140]: sns.scatterplot(x=AFdata['Explained by: Social support'],y=AFdata['Happiness_
↪score'])
plt.show()
```





## 2.4 Europe

```
[141]: EUdata=data.filter(["RANK","Country"],axis=1)
EUdata = data[data["Continent"] == "EU"]
EUdata.head()
```

```
[141]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low	\
0	1	Finland	7.821	7.886	7.756	
1	2	Denmark	7.636	7.710	7.563	
2	3	Iceland	7.557	7.651	7.464	
3	4	Switzerland	7.512	7.586	7.437	
4	5	Netherlands	7.415	7.471	7.359	

	Dystopia (1.83) + residual	Explained by: GDP per capita	\
0	2.518	1.892	
1	2.226	1.953	
2	2.320	1.936	
3	2.153	2.026	
4	2.137	1.945	

	Explained by: Social support	Explained by: Healthy life expectancy	\
0	1.258	0.775	
1	1.243	0.777	
2	1.320	0.803	
3	1.226	0.822	
4	1.206	0.787	

	Explained by: Freedom to make life choices	Explained by: Generosity	\
0	0.736	0.109	
1	0.719	0.188	
2	0.718	0.270	
3	0.677	0.147	
4	0.651	0.271	

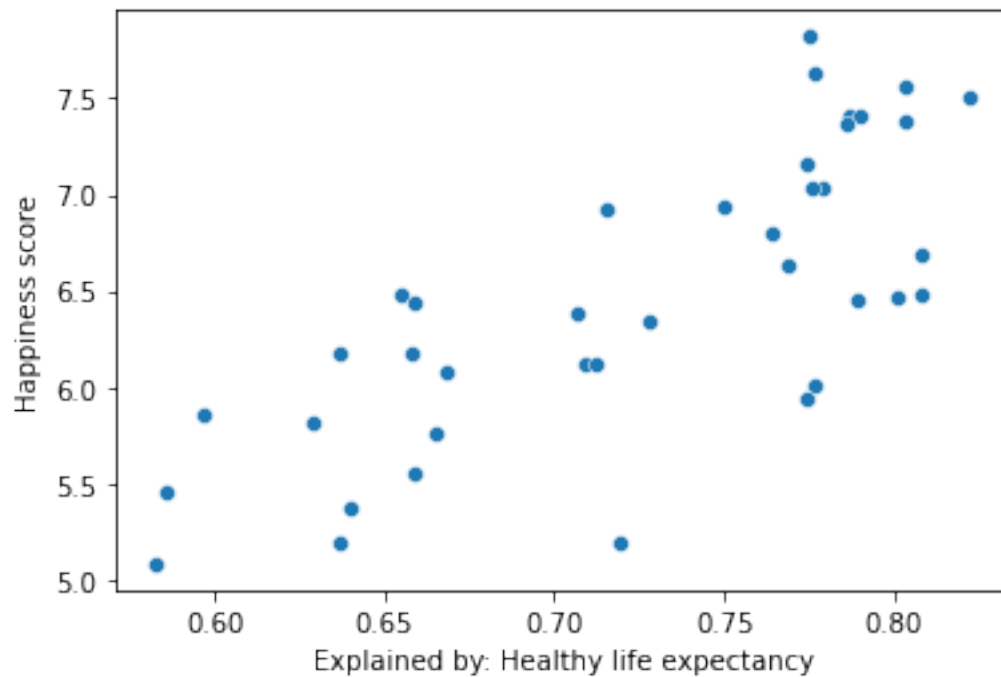
  

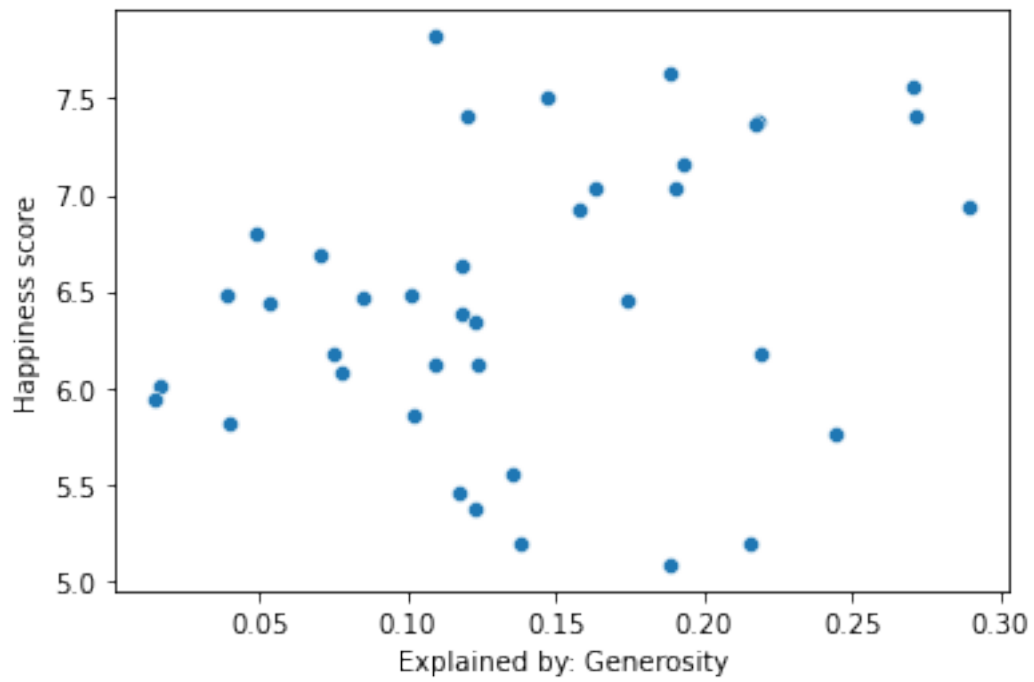
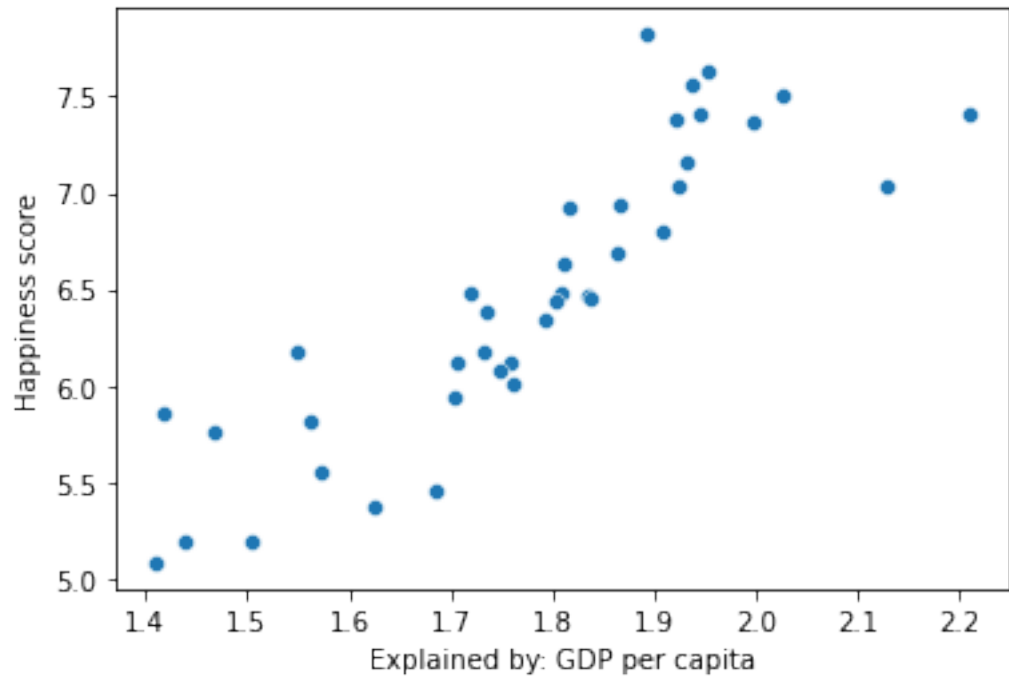
	Explained by: Perceptions of corruption	Country code	Continent
0	0.534	FI	EU
1	0.532	DK	EU
2	0.191	IS	EU
3	0.461	CH	EU
4	0.419	NL	EU

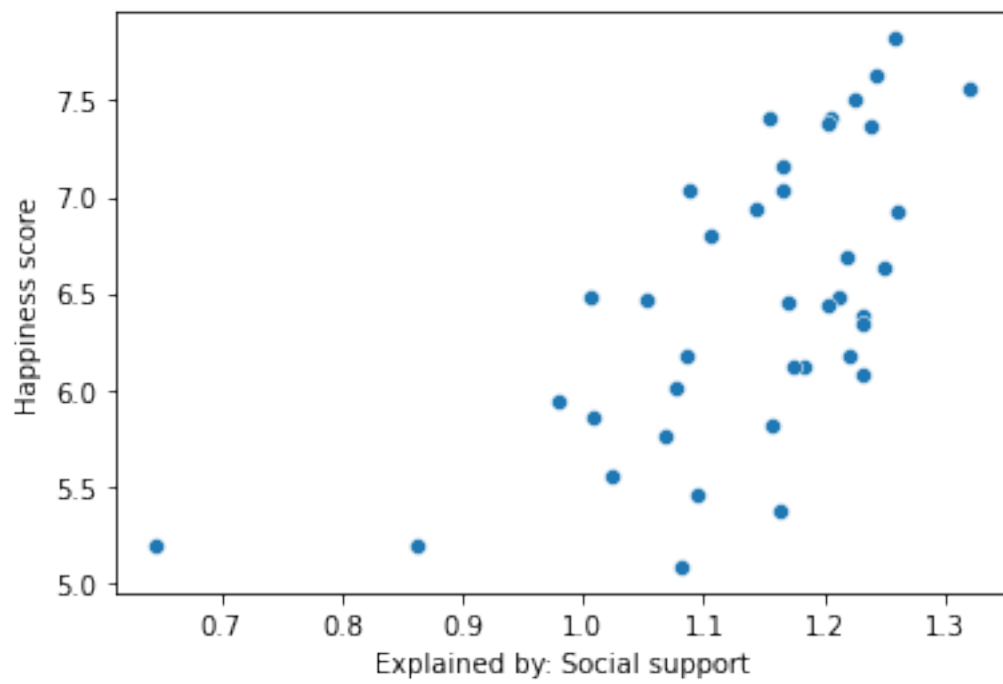
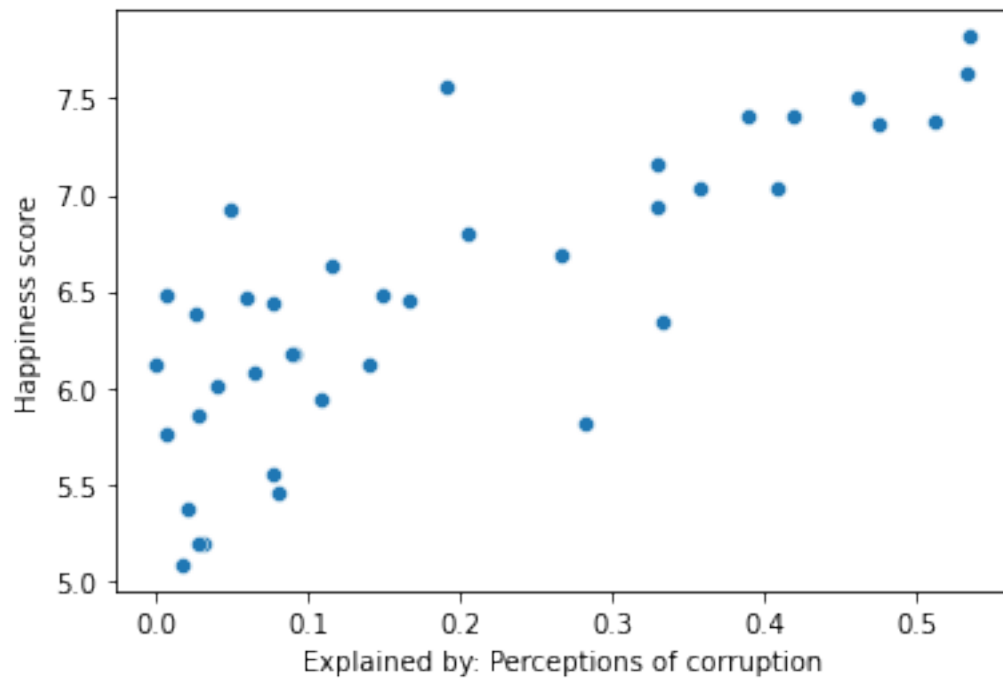
```
[142]: EUdata.shape
```

```
[142]: (39, 14)
```

```
[150]: sns.scatterplot(x=EUdata['Explained by: Healthy life_
↳expectancy'],y=EUdata['Happiness score'])
plt.show()
sns.scatterplot(x=EUdata['Explained by: GDP per capita'],y=EUdata['Happiness_
↳score'])
plt.show()
sns.scatterplot(x=EUdata['Explained by: Generosity'],y=EUdata['Happiness_
↳score'])
plt.show()
sns.scatterplot(x=EUdata['Explained by: Perceptions of_
↳corruption'],y=EUdata['Happiness score'])
plt.show()
sns.scatterplot(x=EUdata['Explained by: Social support'],y=EUdata['Happiness_
↳score'])
plt.show()
```







## 2.5 Other Continent

```
[148]: OCdata=data.filter(["RANK","Country"],axis=1)
OCdata = data[data["Continent"] == "OC"]
OCdata.head()
```

```
[148]:
```

	RANK	Country	Happiness score	Whisker-high	Whisker-low	\
9	10	New Zealand	7.200	7.279	7.120	
11	12	Australia	7.162	7.244	7.081	

	Dystopia (1.83) + residual	Explained by: GDP per capita	\
9	1.954	1.852	
11	2.011	1.900	

	Explained by: Social support	Explained by: Healthy life expectancy	\
9	1.235	0.752	
11	1.203	0.772	

	Explained by: Freedom to make life choices	Explained by: Generosity	\
9	0.680	0.245	
11	0.676	0.258	

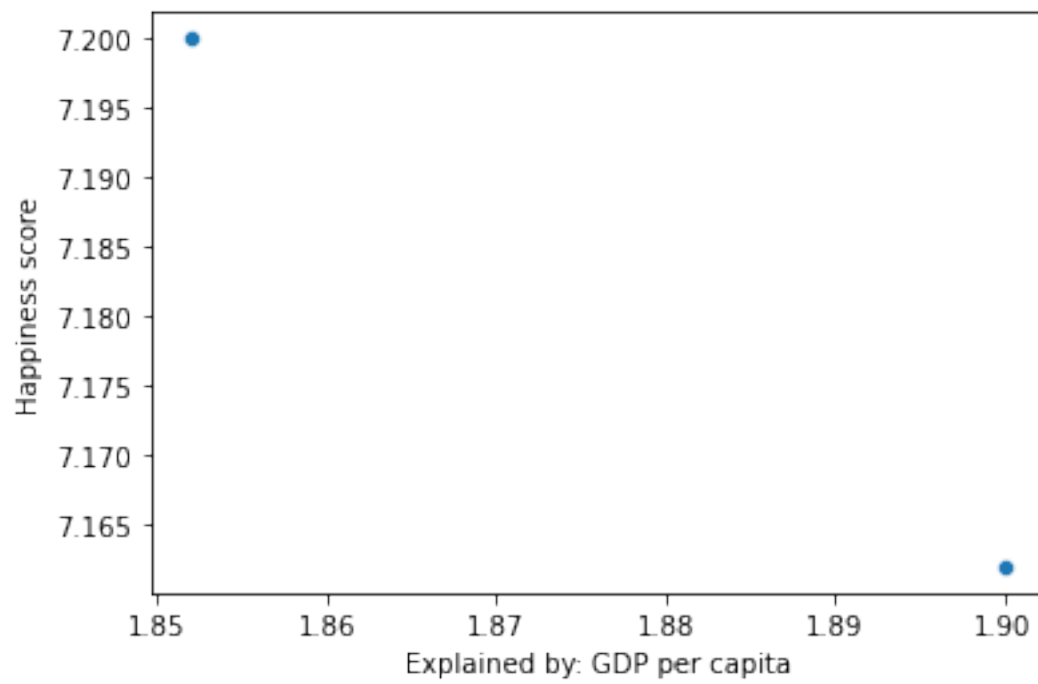
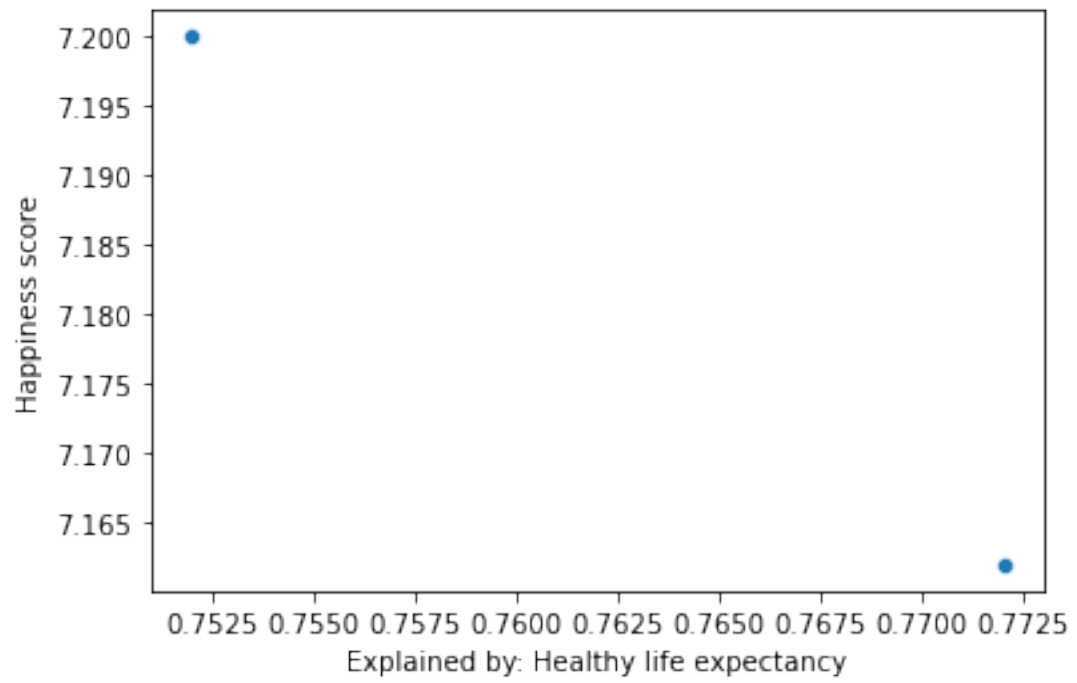
  

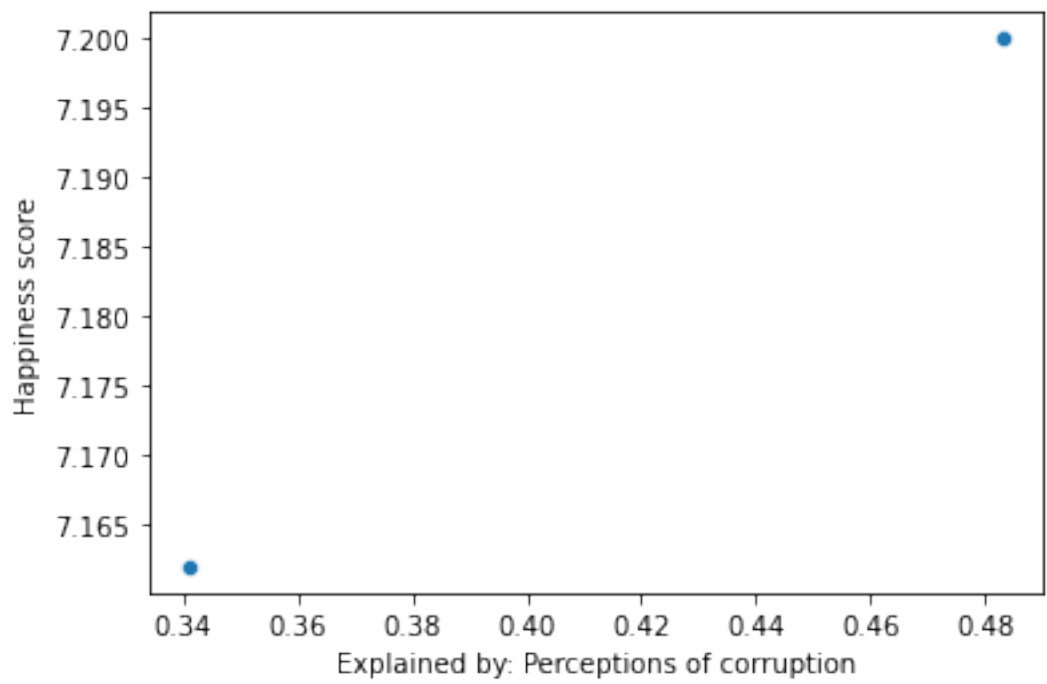
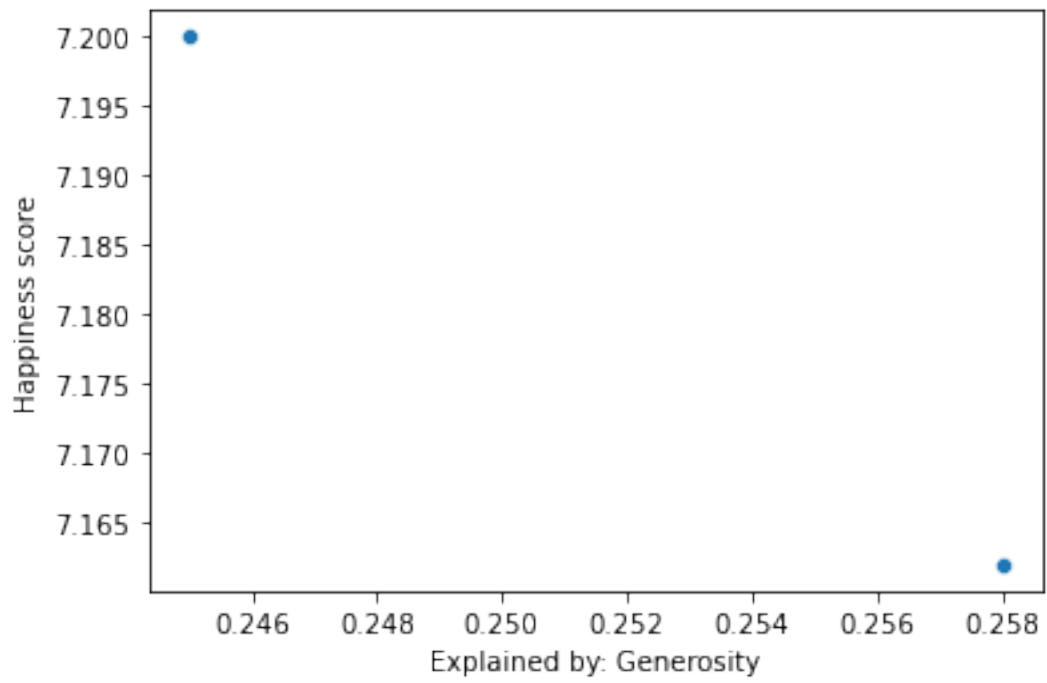
	Explained by: Perceptions of corruption	Country code	Continent
9	0.483	NZ	OC
11	0.341	AU	OC

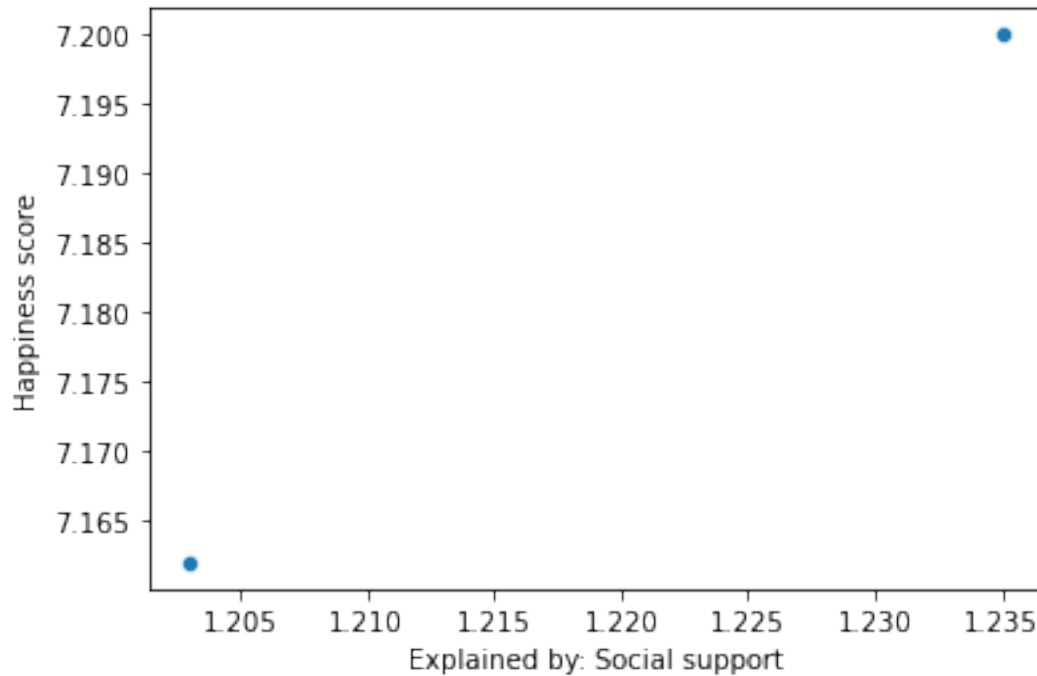
```
[149]: OCdata.shape
```

```
[149]: (2, 14)
```

```
[151]: sns.scatterplot(x=OCdata['Explained by: Healthy life_
↳expectancy'],y=OCdata['Happiness score'])
plt.show()
sns.scatterplot(x=OCdata['Explained by: GDP per capita'],y=OCdata['Happiness_
↳score'])
plt.show()
sns.scatterplot(x=OCdata['Explained by: Generosity'],y=OCdata['Happiness_
↳score'])
plt.show()
sns.scatterplot(x=OCdata['Explained by: Perceptions of_
↳corruption'],y=OCdata['Happiness score'])
plt.show()
sns.scatterplot(x=OCdata['Explained by: Social support'],y=OCdata['Happiness_
↳score'])
plt.show()
```



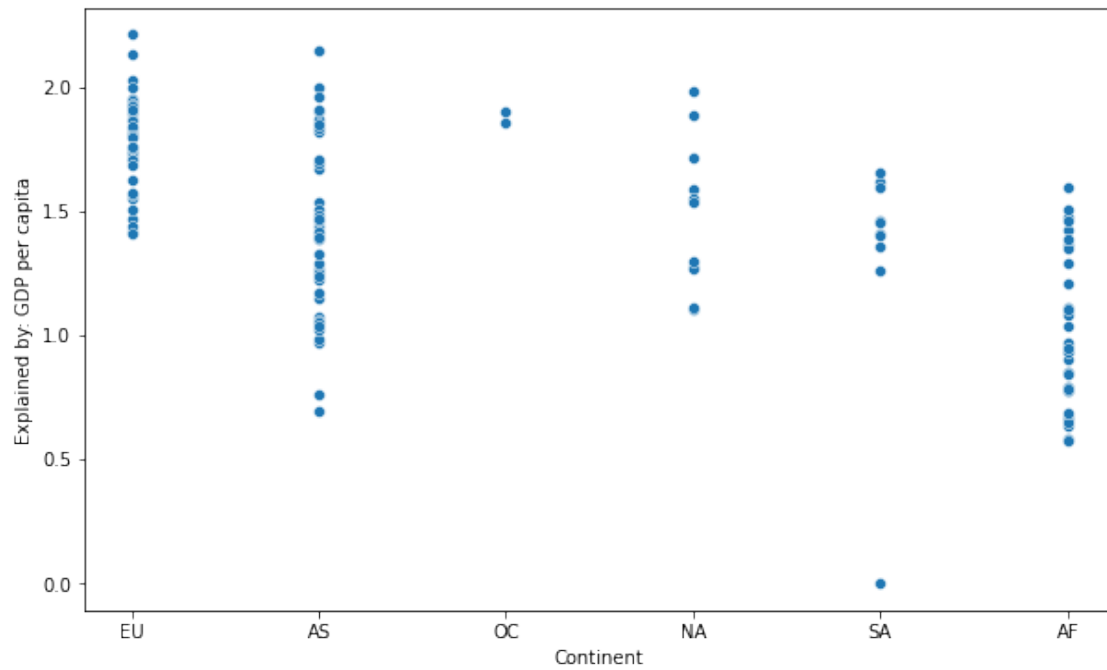




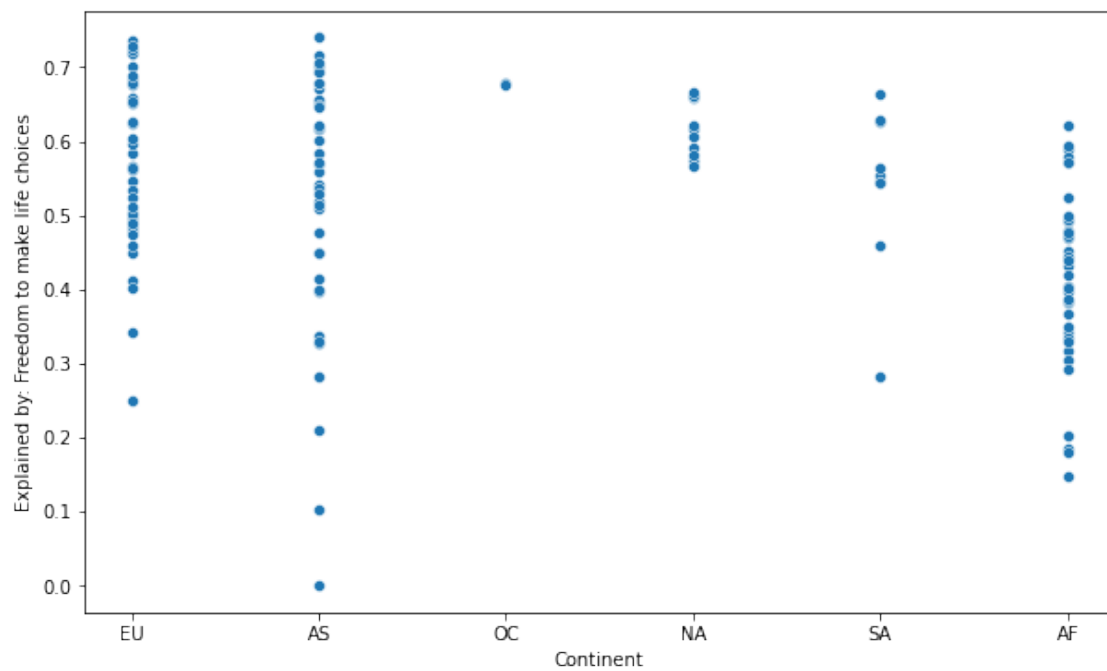
```
[154]: EUdata.to_csv(path_or_buf="C:\\Users\\MSI\\Desktop\\VIT DS\\Python for Data_
        ↪Science\\EUdata.csv")
        AFdata.to_csv(path_or_buf="C:\\Users\\MSI\\Desktop\\VIT DS\\Python for Data_
        ↪Science\\AFdata.csv")
        NAdata.to_csv(path_or_buf="C:\\Users\\MSI\\Desktop\\VIT DS\\Python for Data_
        ↪Science\\NAdata.csv")
        SAdata.to_csv(path_or_buf="C:\\Users\\MSI\\Desktop\\VIT DS\\Python for Data_
        ↪Science\\SAdata.csv")
        OCdata.to_csv(path_or_buf="C:\\Users\\MSI\\Desktop\\VIT DS\\Python for Data_
        ↪Science\\OCdata.csv")
```

```
[157]: plt.figure(figsize=(10,6))
        sns.scatterplot(x=data['Continent'],y=data['Explained by: GDP per capita'])
        plt.show()
```

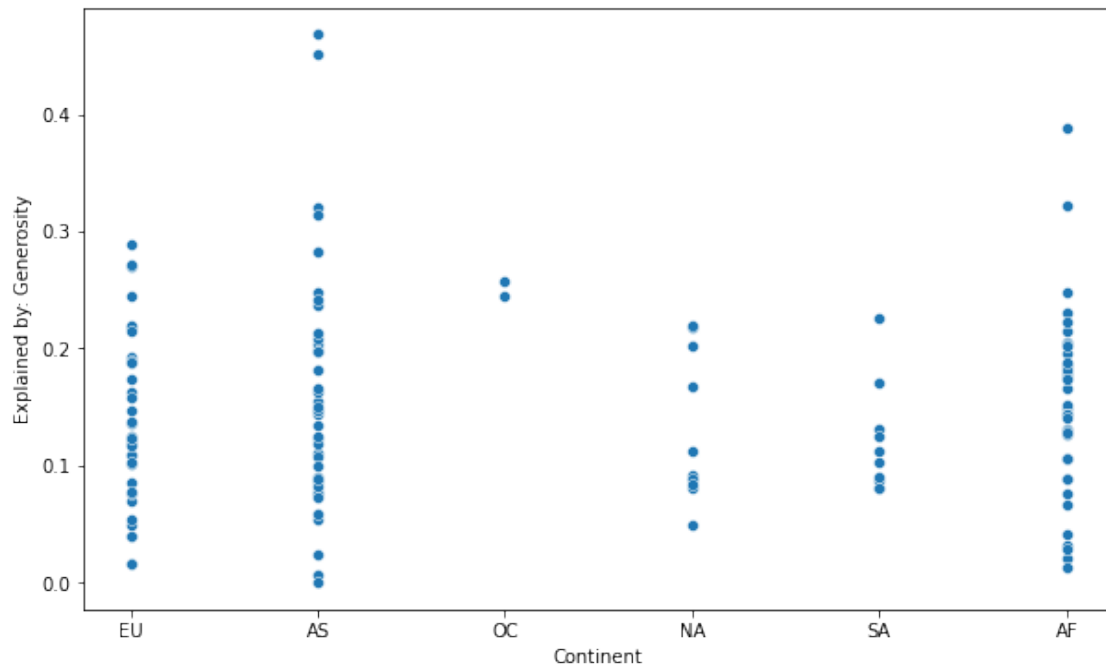




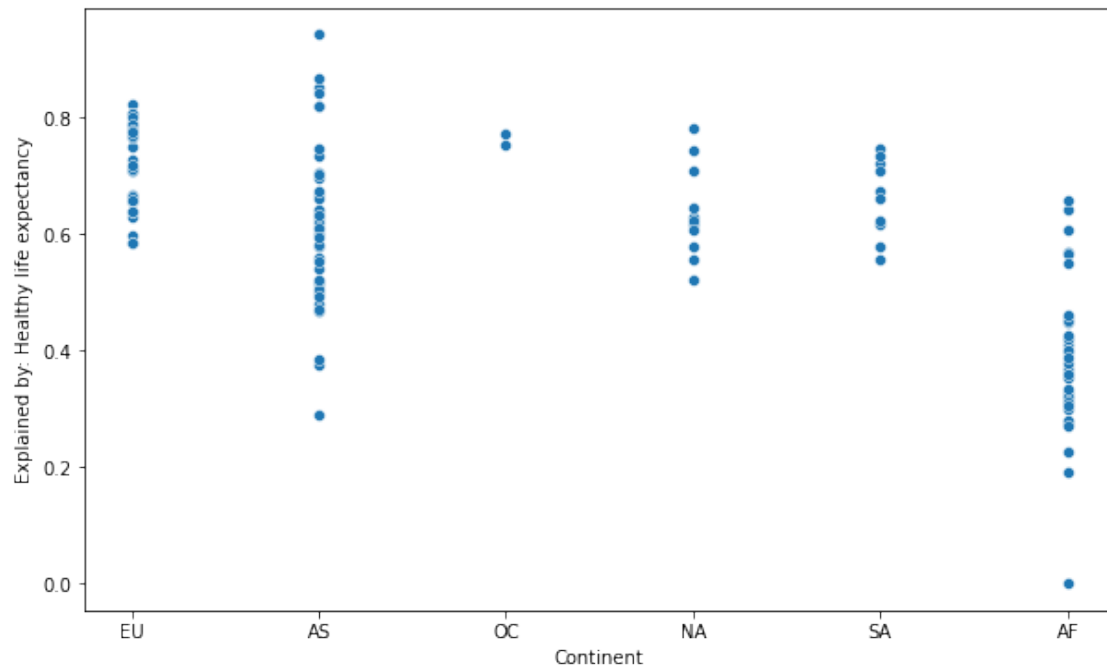
```
[158]: plt.figure(figsize=(10,6))
sns.scatterplot(x=data['Continent'],y=data['Explained by: Freedom to make life_
↵choices'])
plt.show()
```



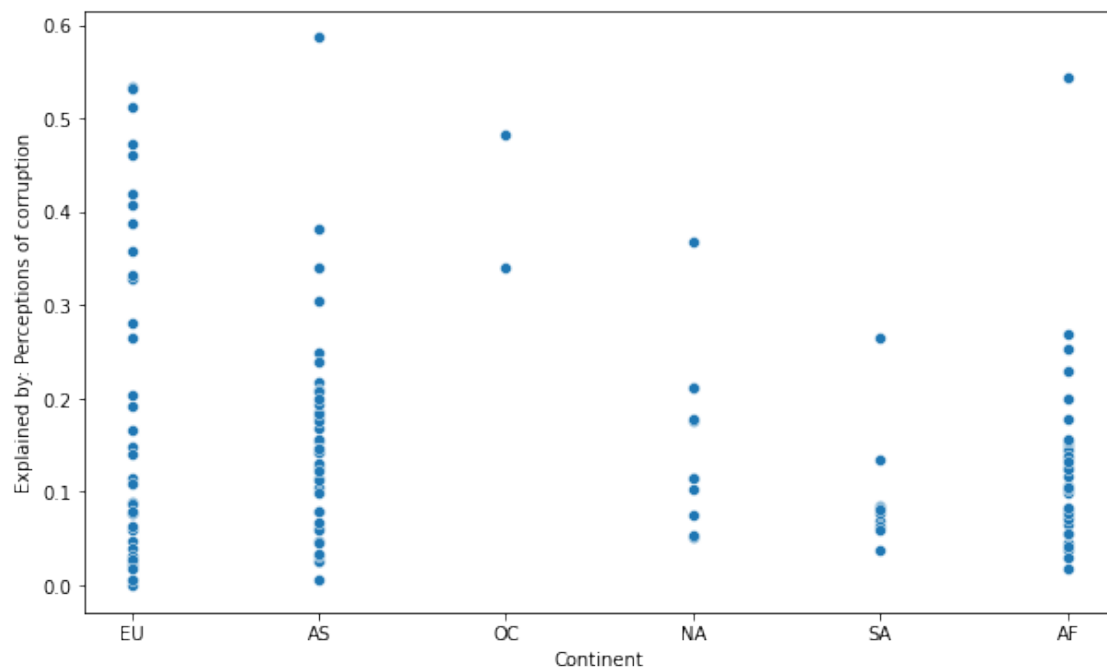
```
[159]: plt.figure(figsize=(10,6))
sns.scatterplot(x=data['Continent'],y=data['Explained by: Generosity'])
plt.show()
```



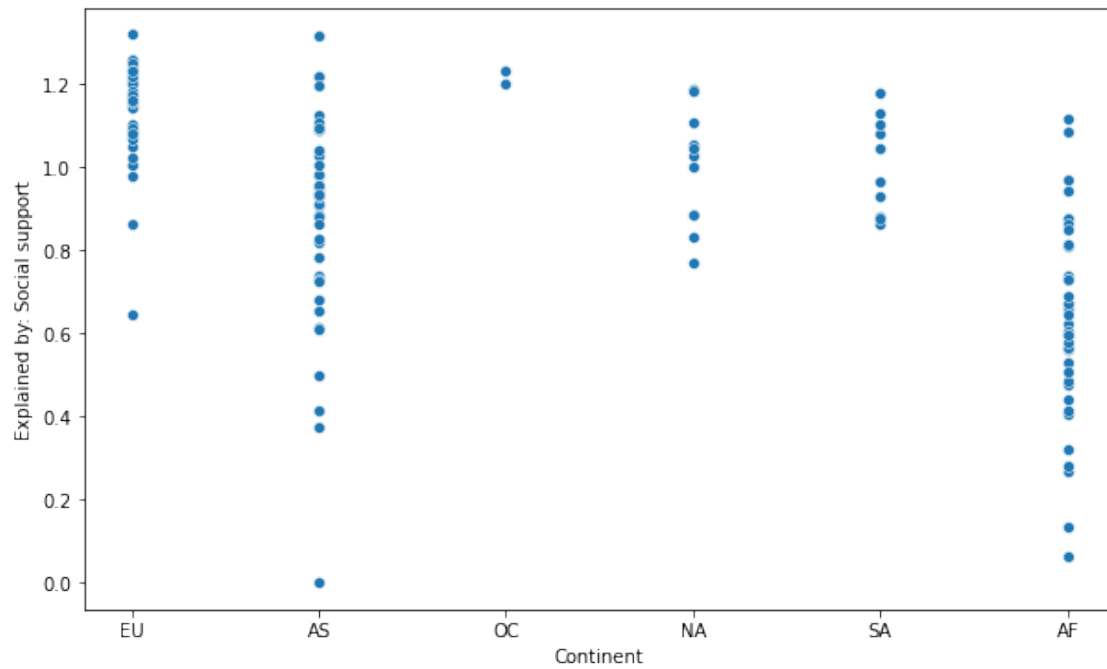
```
[160]: plt.figure(figsize=(10,6))
sns.scatterplot(x=data['Continent'],y=data['Explained by: Healthy life_
↵expectancy'])
plt.show()
```



```
[161]: plt.figure(figsize=(10,6))
sns.scatterplot(x=data['Continent'],y=data['Explained by: Perceptions of_
↪corruption'])
plt.show()
```



```
[162]: plt.figure(figsize=(10,6))
sns.scatterplot(x=data['Continent'],y=data['Explained by: Social support'])
plt.show()
```



```
[ ]:
```