In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [3]:

```
data=pd.read_csv(r"C:\Users\user\Downloads\2015 - 2015.csv")
data
```

Out[3]:

	Country	Region	Happiness Rank	Happiness Score	Standard Error	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Fr
0	Switzerland	Western Europe	1	7.587	0.03411	1.39651	1.34951	0.94143	0
1	Iceland	Western Europe	2	7.561	0.04884	1.30232	1.40223	0.94784	0
2	Denmark	Western Europe	3	7.527	0.03328	1.32548	1.36058	0.87464	0
3	Norway	Western Europe	4	7.522	0.03880	1.45900	1.33095	0.88521	0
4	Canada	North America	5	7.427	0.03553	1.32629	1.32261	0.90563	0
...	...	...	...	...	...	...	...	...	
153	Rwanda	Sub-Saharan Africa	154	3.465	0.03464	0.22208	0.77370	0.42864	0
154	Benin	Sub-Saharan Africa	155	3.340	0.03656	0.28665	0.35386	0.31910	0
155	Syria	Middle East and Northern Africa	156	3.006	0.05015	0.66320	0.47489	0.72193	0
156	Burundi	Sub-Saharan Africa	157	2.905	0.08658	0.01530	0.41587	0.22396	0
157	Togo	Sub-Saharan Africa	158	2.839	0.06727	0.20868	0.13995	0.28443	0

158 rows × 12 columns

In [4]:
```
data.head()
```

Out[4]:

	Country	Region	Happiness Rank	Happiness Score	Standard Error	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Freec
0	Switzerland	Western Europe	1	7.587	0.03411	1.39651	1.34951	0.94143	0.66
1	Iceland	Western Europe	2	7.561	0.04884	1.30232	1.40223	0.94784	0.62
2	Denmark	Western Europe	3	7.527	0.03328	1.32548	1.36058	0.87464	0.64
3	Norway	Western Europe	4	7.522	0.03880	1.45900	1.33095	0.88521	0.66
4	Canada	North America	5	7.427	0.03553	1.32629	1.32261	0.90563	0.63

In [5]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 158 entries, 0 to 157
Data columns (total 12 columns):
 #   Column                         Non-Null Count  Dtype  
---  ------                         --------------  -----  
 0   Country                        158 non-null    object 
 1   Region                         158 non-null    object 
 2   Happiness Rank                 158 non-null    int64  
 3   Happiness Score                158 non-null    float64
 4   Standard Error                 158 non-null    float64
 5   Economy (GDP per Capita)       158 non-null    float64
 6   Family                         158 non-null    float64
 7   Health (Life Expectancy)       158 non-null    float64
 8   Freedom                        158 non-null    float64
 9   Trust (Government Corruption)  158 non-null    float64
 10  Generosity                     158 non-null    float64
 11  Dystopia Residual              158 non-null    float64
dtypes: float64(9), int64(1), object(2)
memory usage: 14.9+ KB
```

In [6]:
```
data.describe()
```

Out[6]:

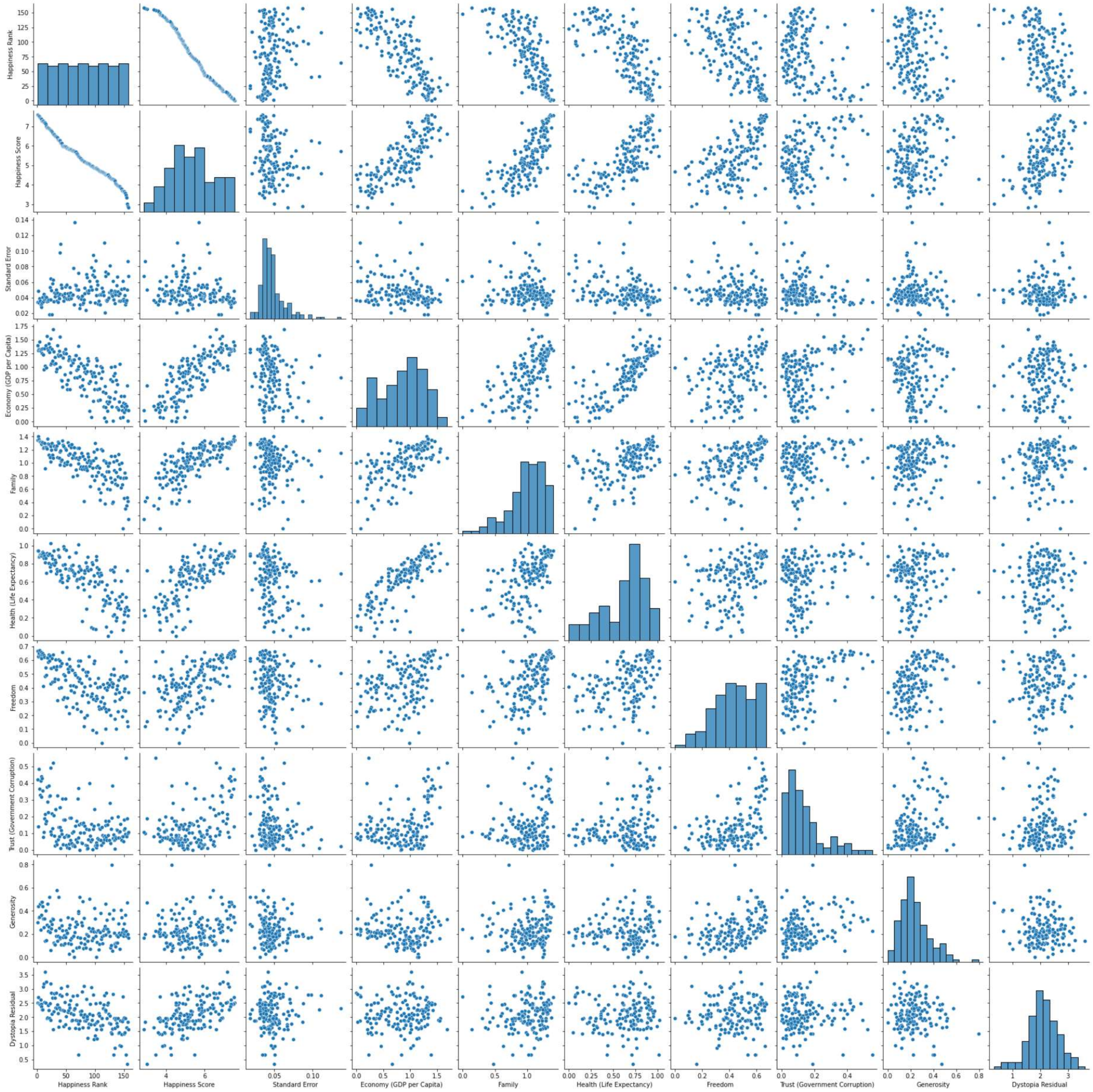
	Happiness Rank	Happiness Score	Standard Error	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Freedom	(G (
count	158.000000	158.000000	158.000000	158.000000	158.000000	158.000000	158.000000	
mean	79.493671	5.375734	0.047885	0.846137	0.991046	0.630259	0.428615	
std	45.754363	1.145010	0.017146	0.403121	0.272369	0.247078	0.150693	
min	1.000000	2.839000	0.018480	0.000000	0.000000	0.000000	0.000000	
25%	40.250000	4.526000	0.037268	0.545808	0.856823	0.439185	0.328330	
50%	79.500000	5.232500	0.043940	0.910245	1.029510	0.696705	0.435515	
75%	118.750000	6.243750	0.052300	1.158448	1.214405	0.811013	0.549092	
max	158.000000	7.587000	0.136930	1.690420	1.402230	1.025250	0.669730	

In [7]:
```
data.columns
```

Out[7]:
```
Index(['Country', 'Region', 'Happiness Rank', 'Happiness Score',
       'Standard Error', 'Economy (GDP per Capita)', 'Family',
       'Health (Life Expectancy)', 'Freedom', 'Trust (Government Corruptio
n)',
       'Generosity', 'Dystopia Residual'],
      dtype='object')
```

In [8]:
```
sns.pairplot(data)
```

Out[8]:
```
<seaborn.axisgrid.PairGrid at 0x1b035c47280>
```



In [13]:

```
da=data[[ 'Happiness Rank', 'Happiness Score',
        'Standard Error', 'Family',
        'Health (Life Expectancy)', 'Freedom', 'Trust (Government Corruption)',
        'Generosity', 'Dystopia Residual']]
da
```

Out[13]:

	Happiness Rank	Happiness Score	Standard Error	Family	Health (Life Expectancy)	Freedom	Trust (Government Corruption)	Generosity
0	1	7.587	0.03411	1.34951	0.94143	0.66557	0.41978	0.29678
1	2	7.561	0.04884	1.40223	0.94784	0.62877	0.14145	0.43630
2	3	7.527	0.03328	1.36058	0.87464	0.64938	0.48357	0.34139
3	4	7.522	0.03880	1.33095	0.88521	0.66973	0.36503	0.34699
4	5	7.427	0.03553	1.32261	0.90563	0.63297	0.32957	0.45811
...	...	...	...	...	...	...	...	...
153	154	3.465	0.03464	0.77370	0.42864	0.59201	0.55191	0.22628
154	155	3.340	0.03656	0.35386	0.31910	0.48450	0.08010	0.18260
155	156	3.006	0.05015	0.47489	0.72193	0.15684	0.18906	0.47179
156	157	2.905	0.08658	0.41587	0.22396	0.11850	0.10062	0.19727
157	158	2.839	0.06727	0.13995	0.28443	0.36453	0.10731	0.16681

158 rows × 9 columns

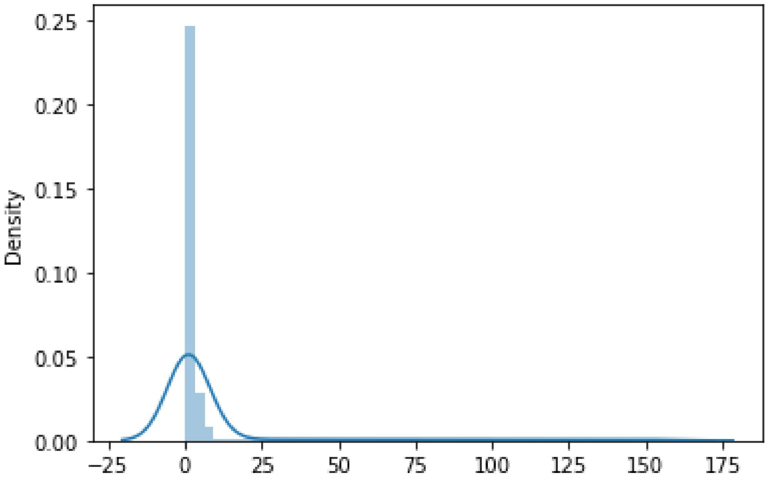
In [14]:

```
sns.distplot(da)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: Fut
ureWarning: `distplot` is a deprecated function and will be removed in a futu
re version. Please adapt your code to use either `displot` (a figure-level fu
nction with similar flexibility) or `histplot` (an axes-level function for hi
stograms).
  warnings.warn(msg, FutureWarning)
```

Out[14]:

```
<AxesSubplot:ylabel='Density'>
```

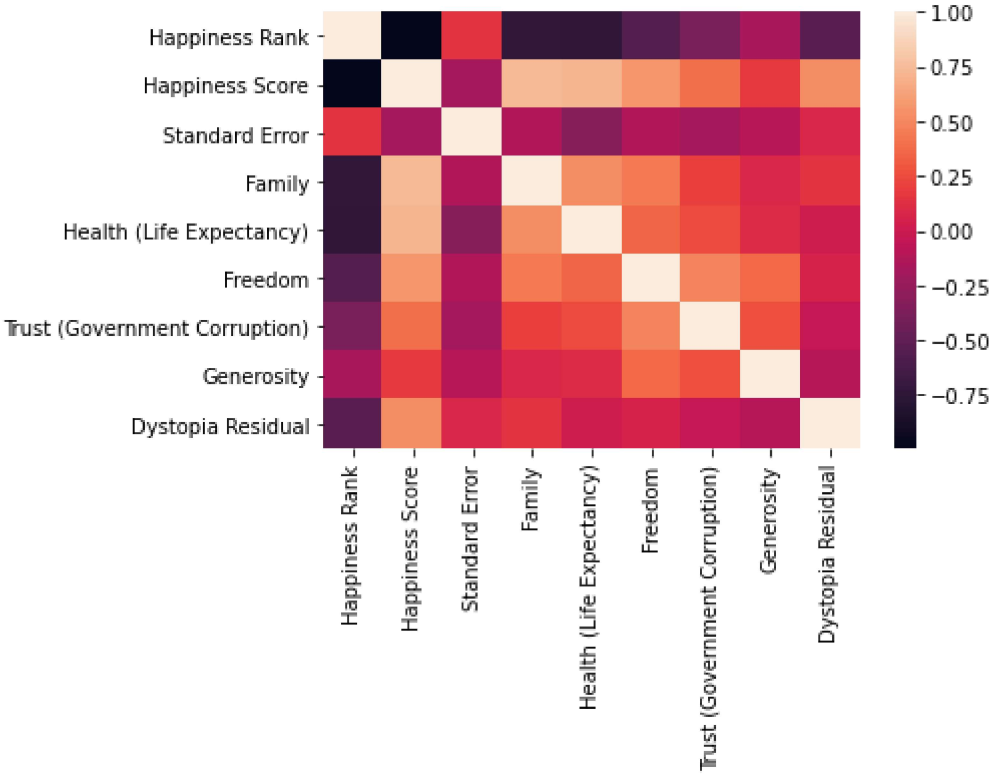


In [15]:

```
sns.heatmap(da.corr())
```

Out[15]:

```
<AxesSubplot:>
```



In [16]:
```
x=da[['Happiness Rank', 'Happiness Score',
       'Standard Error', 'Family',
       'Health (Life Expectancy)', 'Trust (Government Corruption)',
       'Generosity', 'Dystopia Residual']]
y=da['Freedom']
```

In [17]:
```
from sklearn.model_selection import train_test_split

x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3)
```

In [18]:
```
from sklearn.linear_model import LinearRegression

lr=LinearRegression()
lr.fit(x_train,y_train)
```

Out[18]:
```
LinearRegression()
```

In [19]:
```
print(lr.intercept_)
```

```
-0.05717906284556812
```

In [20]:
```
coeff = pd.DataFrame(lr.coef_,x.columns,columns=['Co-efficient'])
coeff
```

Out[20]:

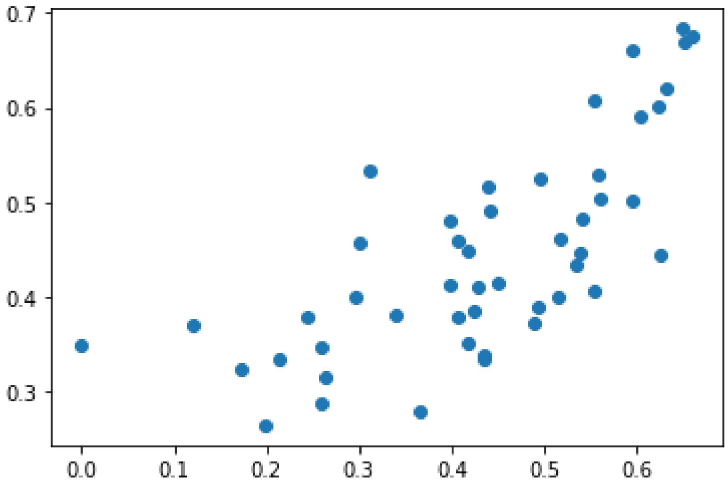
	Co-efficient
Happiness Rank	0.000579
Happiness Score	0.259168
Standard Error	-0.234183
Family	-0.249744
Health (Life Expectancy)	-0.410238
Trust (Government Corruption)	0.006356
Generosity	0.094358
Dystopia Residual	-0.217003

In [21]:

```
prediction=lr.predict(x_test)
plt.scatter(y_test,prediction)
```

Out[21]:

```
<matplotlib.collections.PathCollection at 0x1b0406e8880>
```



In [23]:

```
print(lr.score(x_test,y_test))
```

```
0.5164860718462851
```

In [ ]: