

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import pickle
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import r_squared, variance_inflation_factor
from sklearn.metrics import accuracy_score, confusion_matrix, roc_curve, roc_auc_score
from sklearn.metrics import auc
from sklearn.preprocessing import StandardScaler
import warnings
warnings.filterwarnings('ignore')
from sklearn.preprocessing import PowerTransformer

# Notebook is building the font cache; this may take a moment.

df = pd.read_csv('https://raw.githubusercontent.com/darcyclint/SDData/master/happiness_data_dataset.csv')

df

Out[2]:
   Country      Region  Happiness Rank  Happiness Score  Standard Error  Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  Dystopia Residual
0  Switzerland  Western Europe      1      7.587      0.0411      13961  134951      0.94143  0.66537      0.41978  0.29678  2.51738
1  Iceland      Western Europe      2      7.561      0.04884      130232  140223      0.94784  0.62877      0.14145  0.48300  2.70201
2  Denmark      Western Europe      3      7.527      0.03328      132548  136058      0.87464  0.64938      0.48357  0.34139  2.49304
3  Norway      Western Europe      4      7.522      0.03880      145900  133095      0.88521  0.66973      0.36503  0.34699  2.46531
4  Canada      North America      5      7.427      0.05553      132629  132261      0.90563  0.63297      0.32657  0.45811  2.45176
...
153 Rwanda      Sub-Saharan Africa  154      3.465      0.05464      622268  67770      0.42864  0.50201      0.55191  0.22628  0.67042
154 Benin      Sub-Saharan Africa  155      3.340      0.05056      62865  35386      0.31910  0.48450      0.09010  0.18200  1.63328
155 Syria      Middle East and Northern Africa  156      3.006      0.05015      666320  47489      0.72193  0.15684      0.18906  0.47179  0.32858
156 Burundi      Sub-Saharan Africa  157      2.905      0.06858      601530  411587      0.22396  0.11850      0.10062  0.19727  1.83302
157 Togo      Sub-Saharan Africa  158      2.839      0.06727      620668  413995      0.28443  0.16453      0.10731  0.16681  1.56726

158 rows x 12 columns

df.columns

Out[3]:
Index(['Country', 'Region', 'Happiness Rank', 'Happiness Score',
       'Standard Error', 'Economy (GDP per Capita)', 'Family',
       'Health (Life Expectancy)', 'Freedom', 'Trust (Government Corruption)',
       'Generosity', 'Dystopia Residual'],
      dtype='object')

df.isnull().sum()

Out[4]:
Country      0
Region      0
Happiness Rank      0
Happiness Score      0
Standard Error      0
Economy (GDP per Capita)      0
Family      0
Health (Life Expectancy)      0
Freedom      0
Trust (Government Corruption)      0
Generosity      0
Dystopia Residual      0
dtypes: int64 0

df.describe()

Out[5]:
   Happiness Rank  Happiness Score  Standard Error  Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  Dystopia Residual
count  158.000000      158.000000      158.000000      158.000000      158.000000      158.000000      158.000000      158.000000      158.000000      158.000000
mean      79.49271      5.75754      0.047895      0.946137      0.919046      0.630259      0.428615      0.164342      0.237766      2.06977
std      45.75483      1.145010      0.017146      0.403121      0.272369      0.247078      0.150869      0.140034      0.126485      0.53550
min      1.000000      2.839000      0.018480      0.000000      0.000000      0.000000      0.000000      0.000000      0.000000      0.338350
25%      40.250000      4.526000      0.037268      0.545803      0.851623      0.419185      0.328130      0.061679      0.155553      1.759410
50%      79.500000      5.232500      0.043940      0.910245      1.029510      0.696705      0.435515      0.107220      0.216130      2.095415
75%      118.750000      6.243750      0.052300      1.158448      1.214405      0.811013      0.549002      0.180255      0.308683      2.462415
max      158.000000      7.587000      0.136930      1.690420      1.402230      1.025230      0.669730      0.551910      0.795800      3.602140

df['New Happiness Score'] = df['Economy (GDP per Capita)'] + df['Family'] + df['Health (Life Expectancy)'] + df['Freedom'] + df['Trust (Government Corruption)'] + df['Generosity']

df

Out[6]:
   Country      Region  Happiness Rank  Happiness Score  Standard Error  Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  Dystopia Residual  New Happiness Score
0  Switzerland  Western Europe      1      7.587      0.0411      13961  134951      0.94143  0.66537      0.41978  0.29678  2.51738  5.09958
1  Iceland      Western Europe      2      7.561      0.04884      130232  140223      0.94784  0.62877      0.14145  0.48300  2.70201  4.83691
2  Denmark      Western Europe      3      7.527      0.03328      132548  136058      0.87464  0.64938      0.48357  0.34139  2.49304  5.05564
3  Norway      Western Europe      4      7.522      0.03880      145900  133095      0.88521  0.66973      0.36503  0.34699  2.46531  5.05601
4  Canada      North America      5      7.427      0.05553      132629  132261      0.90563  0.63297      0.32657  0.45811  2.45176  4.97518
...
153 Rwanda      Sub-Saharan Africa  154      3.465      0.05464      622268  67770      0.42864  0.50201      0.55191  0.22628  0.67042  2.79462
154 Benin      Sub-Saharan Africa  155      3.340      0.05056      62865  35386      0.31910  0.48450      0.09010  0.18200  1.63328  1.70681
155 Syria      Middle East and Northern Africa  156      3.006      0.05015      666320  47489      0.72193  0.15684      0.18906  0.47179  0.32858  2.67771
156 Burundi      Sub-Saharan Africa  157      2.905      0.06858      601530  411587      0.22396  0.11850      0.10062  0.19727  1.83302  1.07152
157 Togo      Sub-Saharan Africa  158      2.839      0.06727      620668  413995      0.28443  0.16453      0.10731  0.16681  1.56726  1.27171

158 rows x 13 columns

df.drop(['Country', 'Region', 'Happiness Rank', 'Happiness Score', 'Standard Error', 'Dystopia Residual'], inplace=True, axis=1)
df

Out[11]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

158 rows x 7 columns

df_corr = df.corr()

df_corr

Out[13]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
Economy (GDP per Capita)  1.000000  0.646239      0.816478  0.370300      0.307885  0.010465      0.897034
Family                    0.646239  1.000000      0.521106  0.441518      0.255605  0.007519      0.767976
Health (Life Expectancy)  0.816478  0.521106      1.000000  0.360477      0.246815  0.008151      0.641187
Freedom                   0.370300  0.441518      0.360477  1.000000      0.485524  0.379166      0.631547
Trust (Government Corruption)  0.307885  0.255605      0.246815  0.493534      1.000000  0.276123      0.484350
Generosity                0.010465  0.007513      0.008335  0.373916      0.276123  1.000000      0.270083
New Happiness Score       0.897034  0.767976      0.641187  0.631547      0.484350  0.270083      1.000000

sns.heatmap(df_corr, annot=True)

plt.show()

Economy (GDP per Capita)  1      0.65  0.82  0.37  0.31  0.01  0.9
Family                    0.65  1      0.53  0.44  0.21  0.048  0.76
Health (Life Expectancy)  0.82  0.53  1      0.36  0.25  0.11  0.84
Freedom                   0.37  0.44  0.36  1      0.49  0.37  0.63
Trust (Government Corruption)  0.31  0.21  0.25  0.49  1      0.28  0.48
Generosity                0.01  0.048  0.11  0.37  0.28  1      0.27
New Happiness Score       0.9  0.76  0.84  0.63  0.48  0.27  1

Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0  4078140  Economy (GDP per Capita)
1  190183  Family
3  3157655  Health (Life Expectancy)
3  1740443  Freedom
4  1419084  Trust (Government Corruption)
5  1288610  Generosity

df = df[['Economy (GDP per Capita)', 'Family', 'Health (Life Expectancy)', 'Freedom', 'Trust (Government Corruption)', 'Generosity']]

df

Out[15]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity
0      13961  134951      0.94143  0.66537      0.41978  0.29678
1      130232  140223      0.94784  0.62877      0.14145  0.48300
2      132548  136058      0.87464  0.64938      0.48357  0.34139
3      145900  133095      0.88521  0.66973      0.36503  0.34699
4      132629  132261      0.90563  0.63297      0.32657  0.45811
...
152     622268  67770      0.42864  0.50201      0.55191  0.22628
153     62865  35386      0.31910  0.48450      0.09010  0.18200
154     666320  47489      0.72193  0.15684      0.18906  0.47179
155     601530  411587      0.22396  0.11850      0.10062  0.19727
156     601530  411587      0.22396  0.11850      0.10062  0.19727
157     620668  413995      0.28443  0.16453      0.10731  0.16681
158     620668  413995      0.28443  0.16453      0.10731  0.16681

df

Out[16]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

df

Out[17]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

df

Out[18]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

df

Out[19]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

df

Out[20]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

df

Out[21]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

df

Out[22]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

df

Out[23]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

df

Out[24]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.43054      0.06977  0.33128  1.48459
157     60153  411587      0.28443  0.16453      0.10731  0.16681  1.27171
156     60153  411587      0.22396  0.11850      0.10062  0.19727  1.07152
147     60780  400000      0.06699  0.48879      0.06289  0.23835  0.95552

df

Out[25]:
   Economy (GDP per Capita)  Family  Health (Life Expectancy)  Freedom  Trust (Government Corruption)  Generosity  New Happiness Score
0      13961  134951      0.94143  0.66537      0.41978  0.29678  5.09958
8      145900  133095      0.88521  0.66973      0.36503  0.34699  5.05601
27     149042  107960      0.78913  0.64040      0.52208  0.32573  5.05456
28     132548  136058      0.87464  0.64938      0.48357  0.34139  5.03504
8      125018  131967      0.90837  0.63938      0.42862  0.47501  5.02183
...
152     61982  430285      0.30335  0.23414      0.09719  0.36510  1.62245
130     61604  441134      0.22562  0.
```