We are analyzing the Exploratory Data Analysis on kaggle data set :World Happiness Report up to 2020.

We will explain the happiness indexes of different different country of the world and explain the data in a statistics method

Data Source : https://www.kaggle.com/mathurinache/world-happiness-report

In [1]:

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

In [2]:

```
data = pd.read_csv("2015.csv")
data
```

Out[2]:

	Country	Region	Happiness Rank	Happiness Score	Standard Error	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Freedom	Trust (Government Corruption)	Generosity	Dys Res
0	Switzerland	Western Europe	1	7.587	0.03411	1.39651	1.34951	0.94143	0.66557	0.41978	0.29678	2.5
1	Iceland	Western Europe	2	7.561	0.04884	1.30232	1.40223	0.94784	0.62877	0.14145	0.43630	2.7
2	Denmark	Western Europe	3	7.527	0.03328	1.32548	1.36058	0.87464	0.64938	0.48357	0.34139	2.4
3	Norway	Western Europe	4	7.522	0.03880	1.45900	1.33095	0.88521	0.66973	0.36503	0.34699	2.4
4	Canada	North America	5	7.427	0.03553	1.32629	1.32261	0.90563	0.63297	0.32957	0.45811	2.4
...	...	...	...	...	...	...	...	...	...	...	...	
153	Rwanda	Sub-Saharan Africa	154	3.465	0.03464	0.22208	0.77370	0.42864	0.59201	0.55191	0.22628	0.6
154	Benin	Sub-Saharan Africa	155	3.340	0.03656	0.28665	0.35386	0.31910	0.48450	0.08010	0.18260	1.6
155	Syria	Middle East and Northern Africa	156	3.006	0.05015	0.66320	0.47489	0.72193	0.15684	0.18906	0.47179	0.3
156	Burundi	Sub-Saharan Africa	157	2.905	0.08658	0.01530	0.41587	0.22396	0.11850	0.10062	0.19727	1.8
157	Togo	Sub-Saharan Africa	158	2.839	0.06727	0.20868	0.13995	0.28443	0.36453	0.10731	0.16681	1.5

158 rows × 12 columns

In [3]:

```
print(data.shape)
```

```
(158, 12)
```

We are having 158 rows which indicate the country and 12 column which indicates the essential details of the country

In [4]:

```
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
```

We need to check null values in dataset

In [5]:

```
print("Is there any null value in dataset:",data.isnull().values.any())
```

```
Is there any null value in dataset: False
```

In [6]:

```
print("We need to check the null in detail:",data.isnull().sum())
```

```
We need to check the null in detail: 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
dtype: int64
```

We got confirmation there is no null values

We are have to create the bar chart in order to happiness score

In [7]:

```
import plotly.express as px
fig = px.bar(data,x ='Country',y='Happiness Score',height = 800)
fig.update_layout(title = 'Arrangment of countries in descending order of happiness',titlefont_size
=20)
fig.show()
```

We found the better arrangement of country in order to happiness score

We are have to create the bar chart in order to happiness rank

In [8]:

```
fig = px.bar(data,x ='Country', y = 'Happiness Rank',height=800)
fig.update_layout(title = "Arrangment of Country in descending order of Happiness Rank")
fig.show()
```

We can identitfy the happiness rank and happiness score both are following the opposite direction. If the happiness rank is low that's means happiness score is having high value

In [9]:

```
fig = px.pie(data,values = 'Economy (GDP per Capita)',names = 'Region',title = 'Region Wise economy
',height=650)
fig.show()
```

We found the continent of Europe has a good score of GDP per Capita compared to others. Australian has least score GPD per Capita

In [10]:

```
import plotly.graph_objects as go
fig = go.Figure()
x0 = data['Economy (GDP per Capita)'].values
x1 = data['Health (Life Expectancy)'].values
x2 = data['Trust (Government Corruption)'].values
x3 = data['Dystopia Residual'].values

fig.add_trace(go.Histogram(x=x0))
fig.add_trace(go.Histogram(x=x1))
fig.add_trace(go.Histogram(x=x2))
fig.add_trace(go.Histogram(x=x3))

#overlay both histogram
fig.update_layout(barmode='overlay')
fig.update_traces(opacity=1)
fig.show()
```

In [11]:

```
fig = px.bar(data,x='Country',y = 'Family',color = 'Health (Life Expectancy)',height=800)
fig.update_layout(title = 'Family Vs Heath Factors',titlefont_size=20)
fig.show()
```

We found the highest life Expectency "Iceland Country" and lowest life expectency "Togo"

In [12]:

```
fig = px.bar(data,x = 'Country',y = 'Freedom',color = 'Health (Life Expectancy)',height=800)
fig.update_layout(title = 'Country Vs Freedom',titlefont_size=20)
fig.show()
```

We found the highest freedom for Norway Country where life Expectancy 0.88521 and least freedom for Greece Country where Life Expectancy 0.88213

In [13]:

```
fig = px.line(data,x = 'Freedom',y = 'Health (Life Expectancy)',color = 'Region')
fig.show()
```

Freedom of people regarding health concerns is maximum in Southeastern Asia

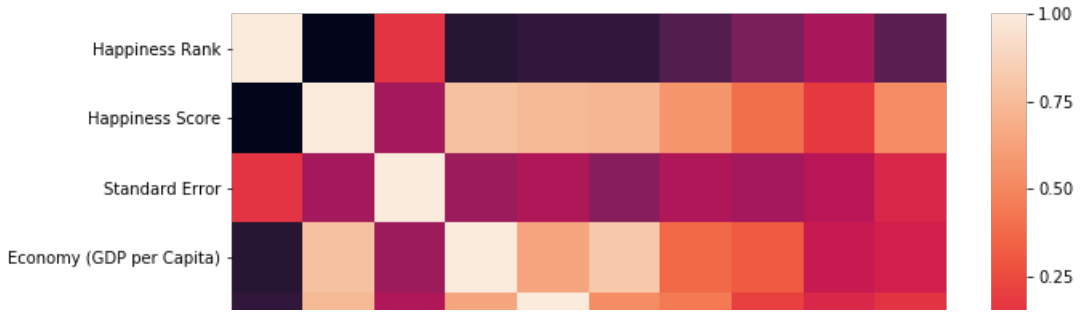
Correlation matrix denotes a score of correlation among various parameters of the dataset

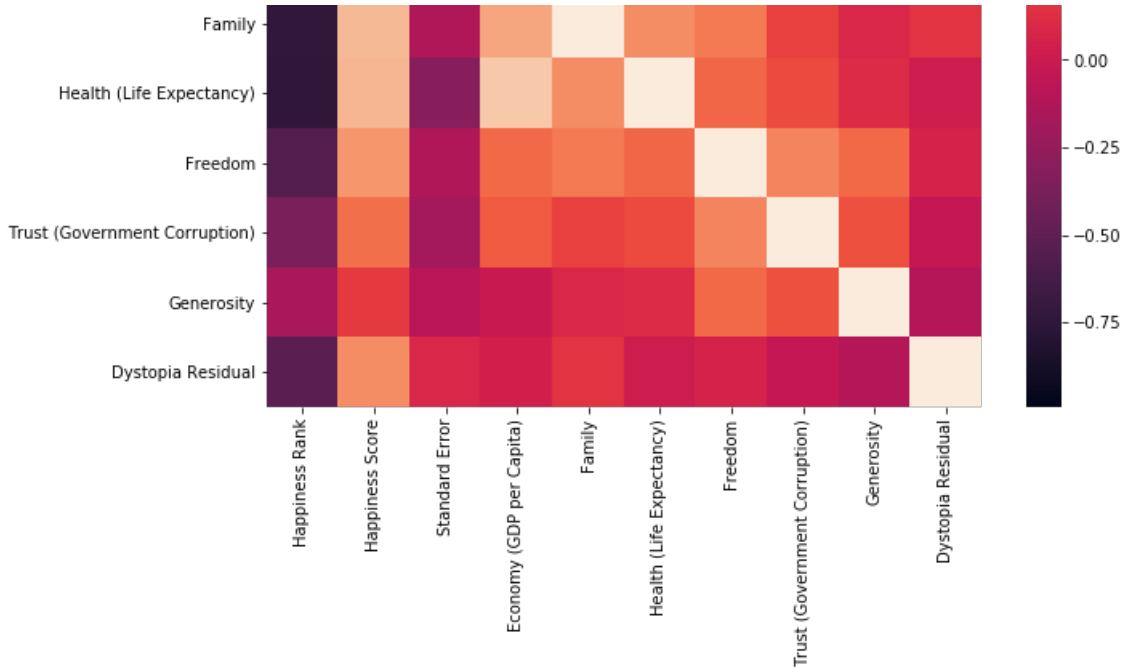
In [14]:

```
import seaborn as sns
plt.figure(figsize=(10,8))
sns.heatmap(data.corr())
```

Out[14]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x212ef3eac88>
```





Conclusion:

Hence from this project we can see that family,happiness and health depends a lot on GDP and government trust.
Happiness depends on Health, Money and Trust.So the countries like Switzerland, Iceland and Singapore have a greater
happiness count than most other countries. African countries like Burundi and Togo have a lower happiness count.