Exploratory Data Analysis of World Happiness Report 2021

Importing necessary libaries

In [ ]:
```
import numpy as np
import pandas as pd
import seaborn as sb
import matplotlib.pyplot as plt
%matplotlib inline
```

setting the style of plot

In [ ]:
```
sb.set_style('darkgrid')
plt.rcParams['font.size'] = 15
plt.rcParams['figure.figsize'] = (10, 7)
plt.rcParams['figure.facecolor'] = '#FFE5B4'
```

In [ ]:
```
data = pd.read_csv('DFF.csv')
data.head()
```

In [ ]:
```
data_col = ['Country name', 'Regional indicator', 'Ladder score', 'Logged GDP per capita', 'Social support', 'Healthy life expectancy', 'Freedom to make life
```

In [ ]:
```
data = data[data_col].copy()
data.head()
```

In [ ]:
```
happy_df = data.rename( columns = {'Country name':'Country_name',
                        'Regional indicator':'Regional_indicator',
                        'Ladder score':'Happiness_score',
                        'Logged GDP per capita':'Logged_GDP_per_capita',
                        'Social support':'Social_support',
                        'Healthy life expectancy':'Healthy_life_expectancy',
                        'Freedom to make life choices':'Freedom_to_make_life_choices',
                        'Perceptions of corruption':'Perceptions_of_corruption'} )


happy_df.head()
```

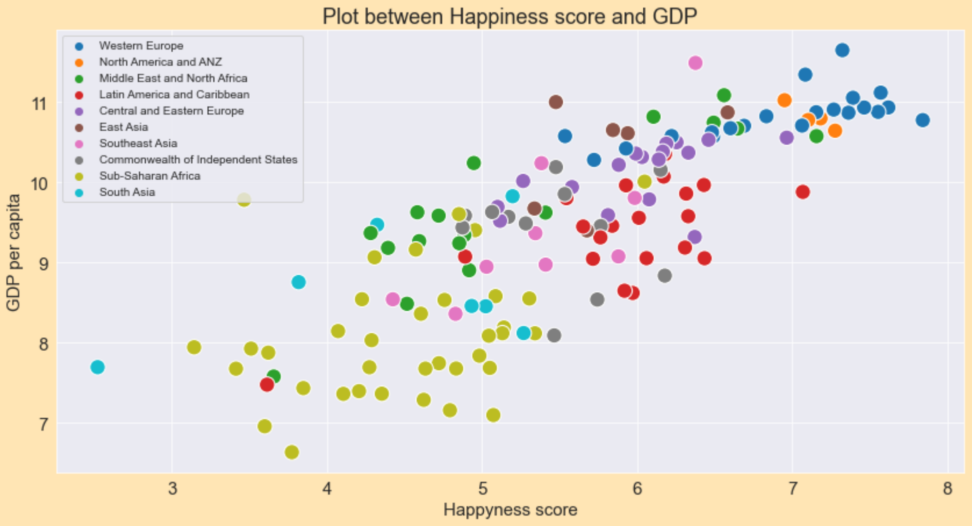
In [ ]:
```
happy_df.isnull().sum()
```

Scatter plot between Happiness score and GDP

In [89]:
```
plt.rcParams['figure.figsize'] = (14, 7)
plt.title('Plot between Happiness score and GDP')
sb.scatterplot(x = happy_df.Happiness_score, y = happy_df.Logged_GDP_per_capita, hue = happy_df.Regional_indicator, s = 180)

plt.legend(loc = 'upper left', fontsize = 10)
plt.xlabel('Happyness score')
plt.ylabel('GDP per capita')
```

Out[89]:
```
Text(0, 0.5, 'GDP per capita')
```



Logged GDP per capita of countries Region wise

In [90]:
```
gdp_region = happy_df.groupby('Regional_indicator')['Logged_GDP_per_capita'].sum()
print(gdp_region)
```

```
Regional_indicator
Central and Eastern Europe            171.854
Commonwealth of Independent States    112.822
East Asia                              62.206
Latin America and Caribbean           187.400
Middle East and North Africa          164.324
North America and ANZ                  43.238
South Asia                             60.778
Southeast Asia                         84.793
Sub-Saharan Africa                    290.707
Western Europe                        227.277
Name: Logged_GDP_per_capita, dtype: float64
```

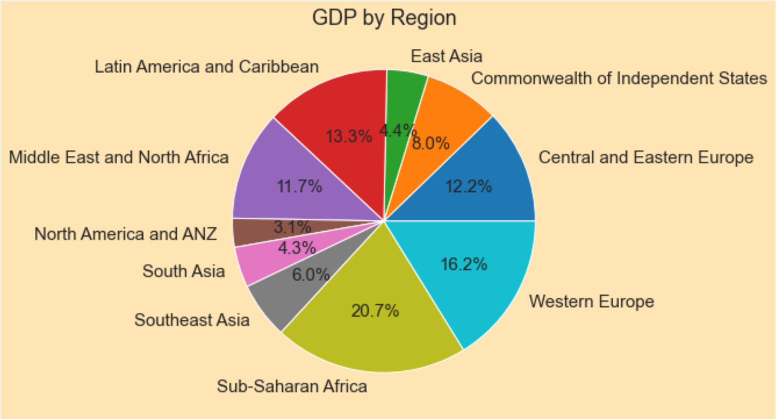
Pie chart to show GDP by Region wise

In [104…

```
gdp_region.plot.pie(autopct = '%1.1f%%')
plt.title('GDP by Region')
plt.ylabel('')
```

Out[104…

```
Text(0, 0.5, '')
```



Total number of countries in each region

In [105…

```
total_country = happy_df.groupby('Regional_indicator')[['Country_name']].count()
print(total_country)
```

```
                                    Country_name
Regional_indicator                              
Central and Eastern Europe                    17
Commonwealth of Independent States            12
East Asia                                      6
Latin America and Caribbean                   20
Middle East and North Africa                  17
North America and ANZ                          4
South Asia                                     7
Southeast Asia                                 9
Sub-Saharan Africa                            36
Western Europe                                21
```

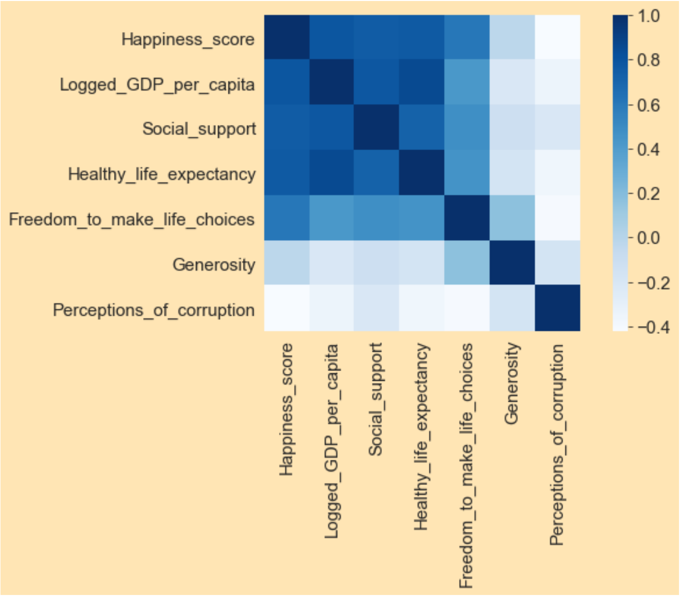
Correlation Map or matrix

In [93]:

```
cor = happy_df.corr(method = "pearson")
f, ax = plt.subplots(figsize = (10, 5))
sb.heatmap(cor, mask = np.zeros_like(cor, dtype=np), cmap="Blues", square=True, ax=ax)
```

Out[93]:

```
<AxesSubplot:>
```



Corruption in different Regions

In [94]:

```
corruption = happy_df.groupby('Regional_indicator')[['Perceptions_of_corruption']].mean()
print(corruption)
```

```
                                    Perceptions_of_corruption
Regional_indicator                                           
Central and Eastern Europe                           0.850529
Commonwealth of Independent States                   0.725083
East Asia                                            0.683333
Latin America and Caribbean                          0.792600
Middle East and North Africa                         0.762235
North America and ANZ                                0.449250
South Asia                                           0.797429
Southeast Asia                                       0.709111
Sub-Saharan Africa                                   0.765944
Western Europe                                       0.523095
```

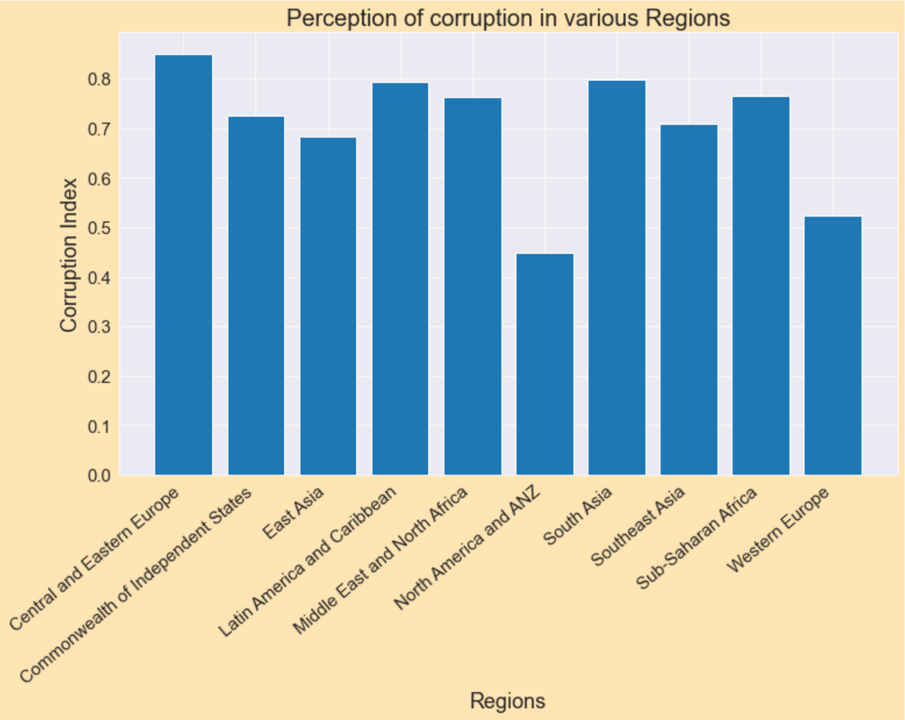
Bar graph to show the perception of corruption in various Regions

In [95]:

```
plt.rcParams['figure.figsize'] = (12, 7)
plt.title('Perception of corruption in various Regions', fontsize = 20)
plt.xlabel('Regions', fontsize = 18)
plt.ylabel('Corruption Index', fontsize = 18)
plt.xticks(rotation = 40, ha = 'right')
plt.bar(corruption.index, corruption.Perceptions_of_corruption)
```

Out[95]:

```
<BarContainer object of 10 artists>
```



Top 10 Happiest Countries and Bottom 10 least Happiest Counrties

In [96]:

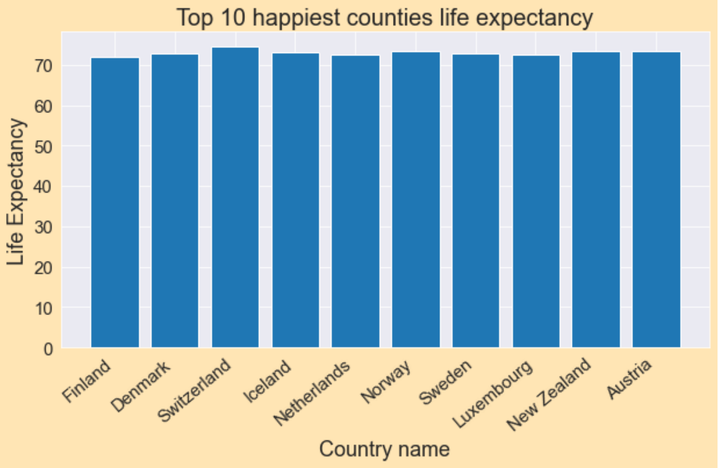
```
top_10 = happy_df.head(10)
bottom_10 = happy_df.tail(10)
```

In [146…

```
plt.rcParams['figure.figsize'] = (10, 5)
plt.title('Top 10 happiest counties life expectancy', fontsize = 20)
plt.xlabel('Country name', fontsize = 18)
plt.ylabel('Life Expectancy', fontsize = 18)
plt.xticks(rotation = 40, ha = 'right')
plt.bar(top_10.Country_name, top_10.Healthy_life_expectancy)
```

Out[146…

```
<BarContainer object of 10 artists>
```

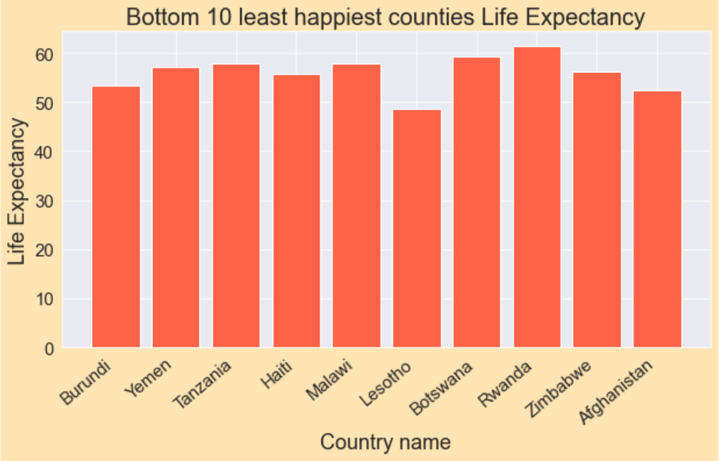


In [145…

```
plt.rcParams['figure.figsize'] = (10, 5)
plt.title('Bottom 10 least happiest counties Life Expectancy', fontsize = 20)
plt.xlabel('Country name', fontsize = 18)
plt.ylabel('Life Expectancy', fontsize = 18)
plt.xticks(rotation = 40, ha = 'right')
plt.bar(bottom_10.Country_name, bottom_10.Healthy_life_expectancy, color='tomato')
```

Out[145…

```
<BarContainer object of 10 artists>
```



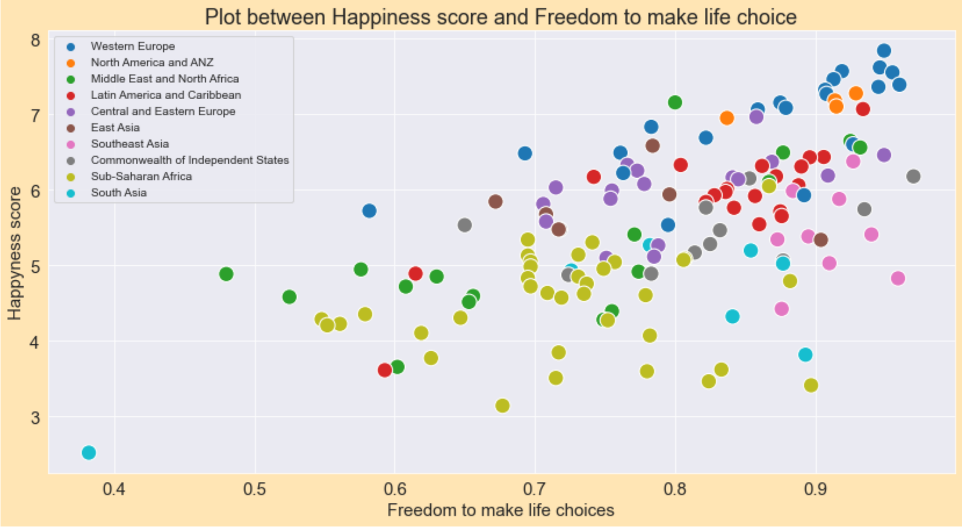
Scatter plot between Happiness score and Freedom to make life choice

In [100…

```
plt.rcParams['figure.figsize'] = (14, 7)
plt.title('Plot between Happiness score and Freedom to make life choice')
sb.scatterplot(x = happy_df.Freedom_to_make_life_choices, y = happy_df.Happiness_score, hue = happy_df.Regional_indicator, s = 180)
plt.legend(loc = 'upper left', fontsize = 10)
plt.xlabel('Freedom to make life choices')
plt.ylabel('Happyness score')
```

Out[100…

```
Text(0, 0.5, 'Happyness score')
```



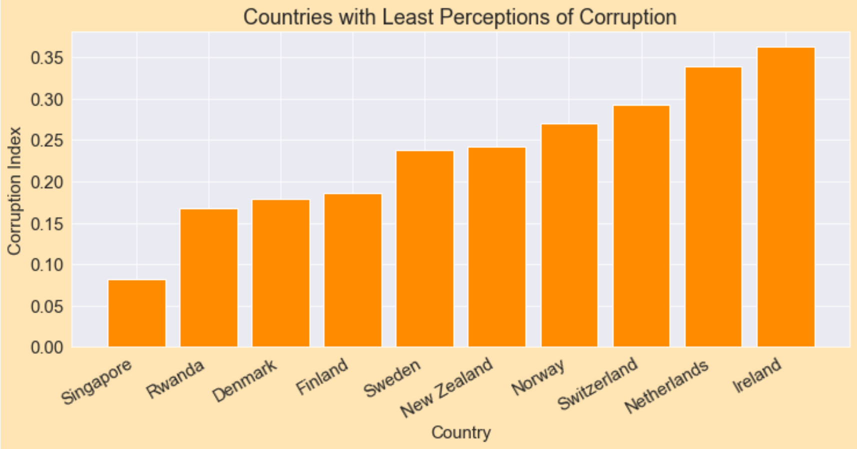
Bar plot to show the Countries with Least Perceptions of Corruption

In [136…

```
country = happy_df.sort_values(by='Perceptions_of_corruption').head(10)
plt.rcParams['figure.figsize'] = (12,5)
plt.title('Countries with Least Perceptions of Corruption')
plt.xlabel('Country', fontsize = 15)
plt.ylabel('Corruption Index', fontsize = 15)
plt.xticks(rotation = 30, ha='right')
plt.bar(country.Country_name, country.Perceptions_of_corruption, color='darkorange')
```

Out[136…

```
<BarContainer object of 10 artists>
```

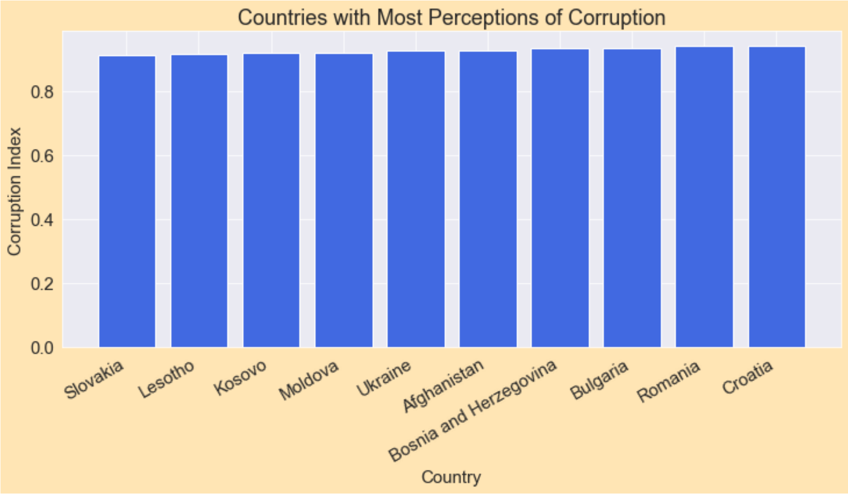


Bar plot to show the Countries with Most Perceptions of Corruption

In [134…

```
country = happy_df.sort_values(by='Perceptions_of_corruption').tail(10)
plt.rcParams['figure.figsize'] = (12,5)
plt.title('Countries with Most Perceptions of Corruption')
plt.xlabel('Country', fontsize = 15)
plt.ylabel('Corruption Index', fontsize = 15)
plt.xticks(rotation = 30, ha='right')
plt.bar(country.Country_name, country.Perceptions_of_corruption, color='royalblue')
```

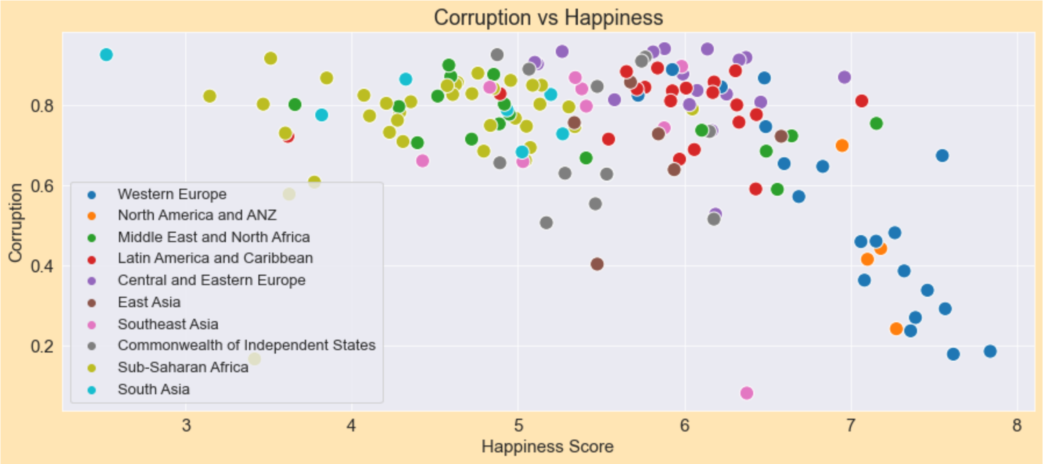
```
Out[134… <BarContainer object of 10 artists>
```



Scatter Plot between Corruption and Happiness

```
In [103…
plt.rcParams['figure.figsize'] = (15, 6)
plt.title('Corruption vs Happiness')
sb.scatterplot(x = happy_df.Happiness_score, y = happy_df.Perceptions_of_corruption, hue=happy_df.Regional_indicator, s=150)
plt.legend(loc = 'lower left', fontsize = '13')
plt.xlabel('Happiness Score')
plt.ylabel('Corruption')
```

```
Out[103… Text(0, 0.5, 'Corruption')
```


Corruption vs Happiness
Corruption
Happiness Score
● Western Europe
● North America and ANZ
● Middle East and North Africa
● Latin America and Caribbean
● Central and Eastern Europe
● East Asia
● Southeast Asia
● Commonwealth of Independent States
● Sub-Saharan Africa
● South Asia