**ISRAEL ADEBOWALE ADEDEJI**

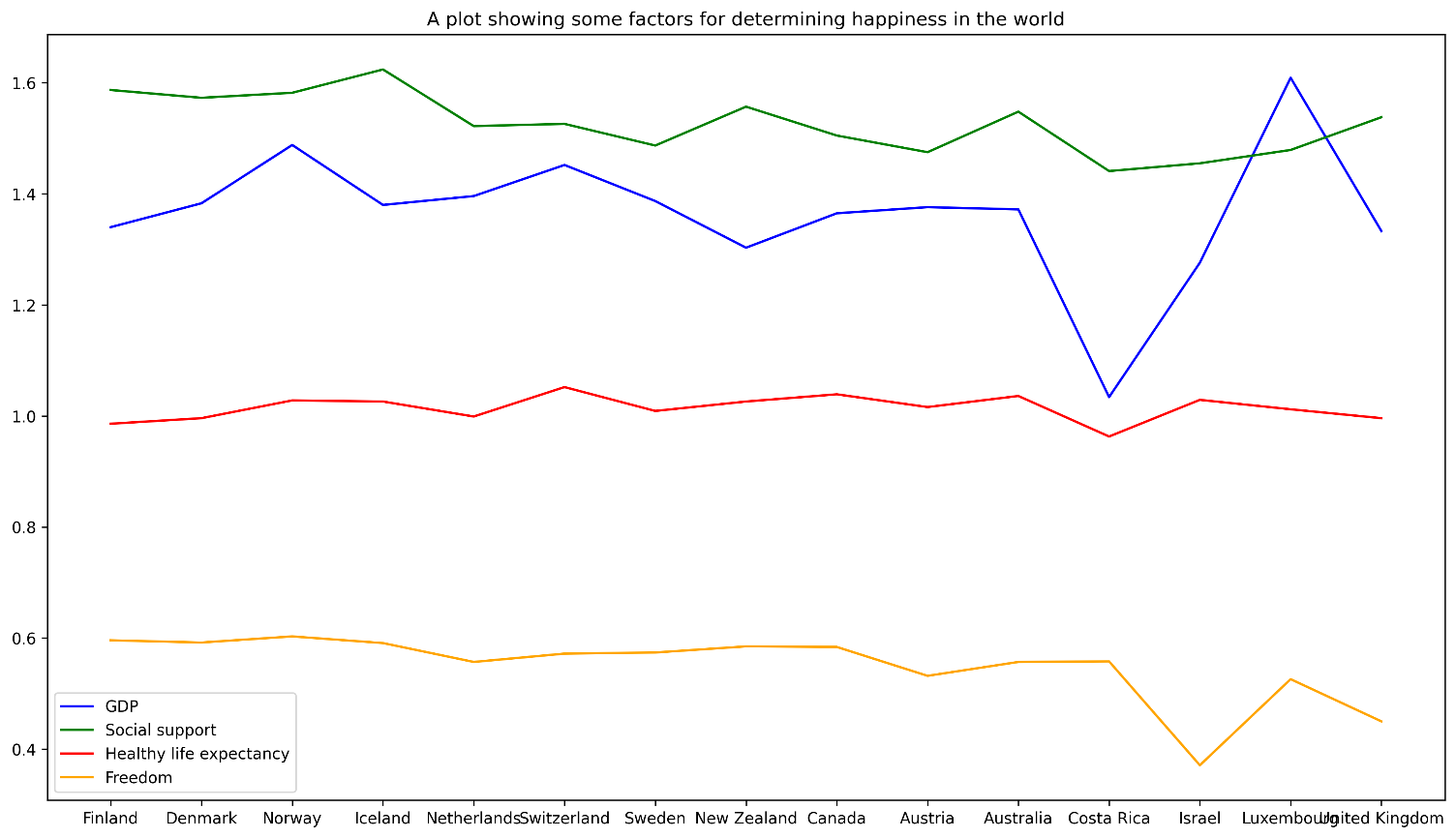
**ASSIGNMENT 1: VISUALIZATION**

**APPLIED DATA SCIENCE 1**

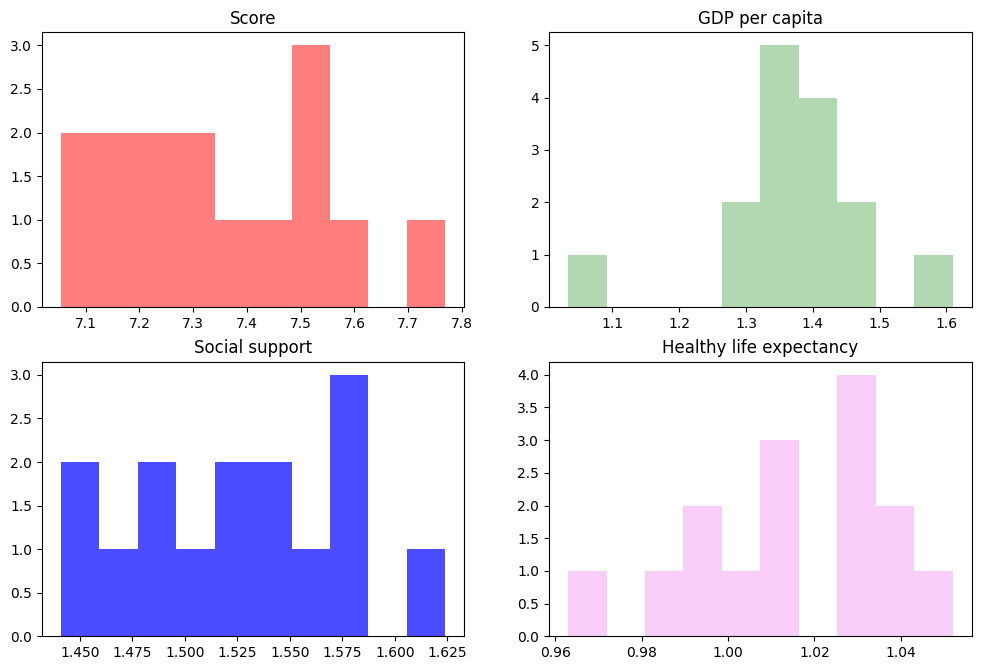
The dataset used is a zip file and it can be found at <https://storage.googleapis.com/kaggle-data-sets/894/813759/bundle/archive.zip?X-Goog-Algorithm=GOOG4-RSA-SHA256&X-Goog-Credential=gcp-kaggle-com%40kaggle-161607.iam.gserviceaccount.com%2F20221110%2Fauto%2Fstorage%2Fgoog4_request&X-Goog-Date=20221110T233311Z&X-Goog-Expires=259200&X-Goog-SignedHeaders=host&X-Goog-Signature=> but due to the Kaggle’s API and domain structure, the link might have expired by the time you click on the link. Fortunately, the page source of the dataset has been provided and it is states as: <https://www.kaggle.com/datasets/unsdsn/world-happiness?select=2019.csv>, and the reference to the extracted CSV file name is “2019.csv”.

The main clause of the dataset is a world happiness report and the underlying factors that influence the state of global happiness are: Score, GDP per capita, Social support, Healthy life expectancy, Freedom to make life choices, Generosity, and Perceptions to corruption. The aforementioned factors are described based on the rank of the countries over the world. The dataset contains 156 rows of countries according to their rank and 9 columns of factors which have been stated earlier in the second line of the second paragraph. The factors that influence the report are what give the countries their respective scores.

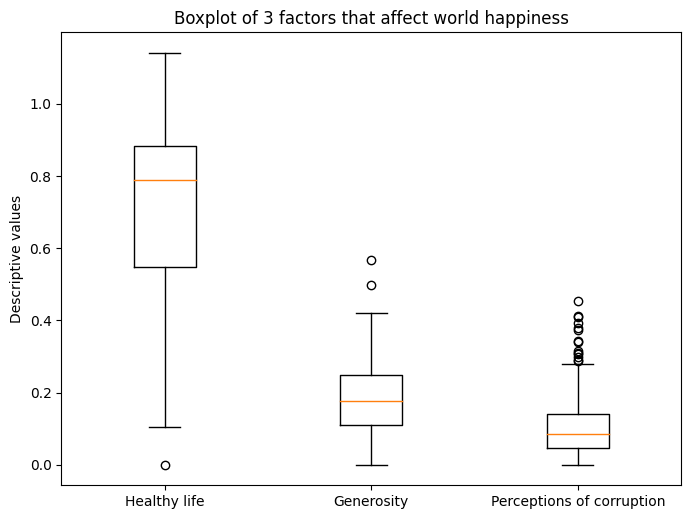
The visualizations of the plots are shown below:



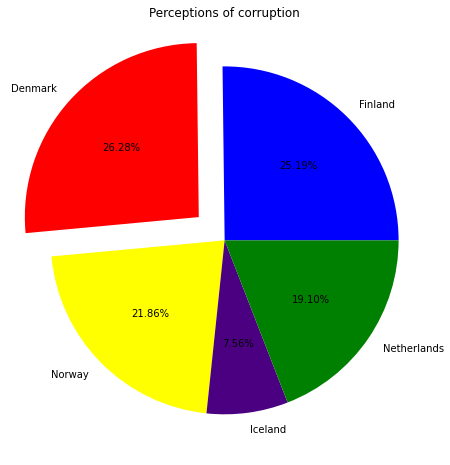
The first visualization I stated is the multiple line plots and it states the degree of the factors that affect happiness. I plot the lines for the first 15 countries and made use of GDP per capita, Social support, Healthy life expectancy, and Freedom to make life choices as the dependent variables. In terms of GDP per capita, Luxemborg showed the highest while Costa Rica was the least in the trend, also, Iceland peaks in terms of social support while Costa Rica had the lowest value in that section. Additionally, Norways slightly supercedes some countries when it comes to Healthy life expectancy while Israel was the lowest in this factor. All countries showed little difference between each other in regards to Healthy life expectancy.



The above plot is a subplot of histogram which states the distribution of the factors on each plot. The values around the GDP per capita are skewed towards the centre, while for social support and Score it shows left-skewness. Healthy life expectancy shows a left skewness.



The boxplot shows the descriptive statistics of Healthy life expectancy, Generosity and Perceptions of corruption. Perceptions of corruption has the most outlier in the plot. There is a significant difference between the maximum and minimum values of the Healthy life expectancy of the countries. All the factors have a minimum value of 0.



In the pie chart above, Perceptions of corruption was used to determine the distribution of values around top 5 countries in the world in terms of happiness. Denmark has the highest perception about corruption, although 3 other countries weren’t far off in terms of this factor. Iceland has the least perception of corruption.