In this analysis the dataset that I picked is the world happiness report of 2019 found on Kaggle. I picked this set because I wanted to see how people perceived happiness before covid became a huge impact for the whole world. First, I imported the necessary libraries to plot the visuals and then imported the dataset on Google Colab. The first plot I decided to do was a line plot between happiness score and GDP per capita. I wanted to see if GDP per capita will have any correlations with happiness score. From the line plot, I can see that there is a clear positive correlation between happiness score and GDP per capita. Next, I made a line plot between happiness score and social support (which is defined as having someone to count on when times are hard) to see if there is a correlation between these two factors. From this line plot I can see that there is a positive correlation between these two factors as well. After that, I made a line plot between happiness score and healthy life expectancy to see if there is a correlation between these two factors. From the line plot I can tell that there is a positive correlation between these two factors. Then, I made a line plot between happiness score and freedom to make life choices to see if there is a correlation between these two factors. From this plot I can see that there is a positive correlation between these two factors. Then, I made a line plot between happiness score and generosity (which is defined as donations to charities per GDP per capita). Surprisingly, there is no clear correlation between happiness score and generosity. The last line plot I made was between happiness score and perceptions of corruption. From this line plot I can see that there is no clear correlation between these two factors, which is surprising to see because one would think that the less corrupt the people think the government is, the happier they will be.

After the line plots, I made a box plot of the 6 factors that are recorded in the world happiness score and see the spread between all the countries. There are no outliers in GDP per capita; but there are outliers in the other 5 categories. The category with the most outliers is perceptions of corruption. To see the distribution clearer, I made a violin plot overlapped with swarm plot to see the distribution of the data in each of the 6 categories. From the combined plot of violin and swarm, I can see that GDP per capita has the highest spread of values while perceptions of corruption have the spread of values.

After looking at the distribution, I made a bar chart of the top 5 and bottom 5 countries and found that the top 5 happiest countries are all in Europe and the 5 saddest countries are in Africa and middle east region. From the data perceptions of corruption and freedom to make life choices are people's perceptions, I decided not to look at joint plots of these categories. Making joint plots between GDP per capita and healthy life expectancy, GDP per capita and social support, and healthy life expectancy and social support, I found that they all have a positive correlation between each other. This is in line with my thoughts; the more money a country has, the more likely they will live longer and have more people they can depend on. The surprising thing is that there is no correlation between GDP per capita and generosity. This is surprising because you would think that the more GDP a country has, the more likely they will donate to charities. But reality shows that this is not true.

These are all the things I have learned from using visuals on the world happiness report.