Shao Ru Zhang
CIS 3120- Programming for Analytics
Homework #2
26th March 2021

     I chose to scrape a Wikipedia page on world happiness. These world happiness reports are published by the United Nations Sustainable Development Solutions Network. These reports are based on people's ratings and their perception of different quality of life factors from the Gallup World Poll. The rankings of happiness are based on a 0 to 10 scale.

     I chose the most recent world happiness report for 2020. It ranked 153 countries based on an average rating between the years 2017-2019. The table indicates that Finland has the highest rank and score. This is surprising since Finland is not at the top at any of the life factors except the overall score. In fact, Finland's generosity is even below the national average of 0.189. Despite not having the highest in most of the categories, I think the citizens living there are content with their lives and have a good impression and attitude towards their country.

     According to the statistical summary, each column has 153 data values/rows. The mean for each column is 5.47, 0.87, 1.16, 0.69, 0.46, 0.19, and 0.13. The standard deviation for each column is 1.11, 0.37, 0.29, 0.25, 0.14, 0.10, and 0.11. The minimum for each column is 2.57, 0, 0, 0, 0, 0, and 0. The 25% percentile for each column is 4.72, 0.58, 0.99, 0.50, 0.38, 0.12, and 0.06. The 50% for each column is 5.52, 0.92, 1.20, 0.76, 0.48, 0.18, and 0.10. The 75% percentile for each column is 6.23, 1.17, 1.39, 0.87, 0.58, 0.26, and 0.16. The maximum for each column is 7.81, 1.54, 1.55, 1.14, 0.70, 0.57, and 0.53. These columns are Score, GDP per capita, Social support, Healthy life expectancy, Freedom to make life choices, Generosity, Perceptions of corruption respectively.

     Using the pandas dataframe and the statistical summary of the numerical columns, people can utilize this information to determine the best countries to live in in terms of people's happiness. They can use this information to make decisions on where they want to live, how their country compares to other countries, and their own views compared to the averages, minimums, 25% percentiles, 50% percentiles, 75% percentiles, and the maximum ratings.

     The next step that can be taken is to analyze the data to see if there are any patterns or correlations between the numerical variables. This can help people weigh their pros and cons before moving abroad. It can also provide countries with a lot of information to keep their citizens happier.

	Score	GDP per capita	Social support	Healthy life expectancy	Freedom to make life choices	Generosity	Perceptions of corruption
count	153.000000	153.000000	153.000000	153.000000	153.000000	153.000000	153.000000
mean	5.473255	0.868778	1.155562	0.692882	0.463542	0.189366	0.130719
std	1.112288	0.372418	0.286862	0.254111	0.141139	0.100433	0.113087
min	2.567000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	4.724000	0.576000	0.987000	0.495000	0.381000	0.115000	0.056000
50%	5.515000	0.919000	1.204000	0.760000	0.483000	0.177000	0.098000
75%	6.228000	1.169000	1.387000	0.867000	0.577000	0.256000	0.163000
max	7.809000	1.537000	1.548000	1.138000	0.693000	0.570000	0.533000