https://www.kaggle.com/unsdsn/world-happiness?select=2019.csv

Section 1:

Can happiness be related to whole countries?

We got the data from Kaggle. The data was collected from the United Nations at an event celebrating International Day of Happiness on March 20th.

Cases:

-Country (156)

Variables:

-Social support

-Overall rank

-Score

-capita

-Healthy life expectancy

-Freedom to make life choices

-Generosity

-Perceptions of corruption

Section 2:

(we need help)

Section 3:

- The outcome would be overall happiness and the predictor would be based on the individual country
- The comparison groups we will use is the countries against one another, in order to find patterns in countries and their location based on happiness of the population
- Very preliminary exploratory data analysis, including some summary statistics and visualizations, along with some explanation on how they help you learn more about your data. (You can add to these later as you work on your project.)
- The method(s) that you believe will be useful in answering your question(s). (You can update these later as you work on your project.)
- What results from these specific statistical methods are needed to support your hypothesized answer?

1. **Mean**

The mean will allow us to determine if there is an overall trend between happiness and the country a person lives in. This statistical method is simple to calculate and it gives up a good insight on the data.

2. **Standard Deviation**

This statistical method will provide a number used to tell how spread out the data is from the mean. If the number is closer to the mean it means that the numbers are close together. Therefore the data is concise and there are little to no outliers.

3. **Regression**

A regression model will explain the relationship between the dependent variable and independent variable. This analysis will give us an idea whether the relationship between out two variables are either strong or weak.

4. **Sample Size Determination**

We will pick a sample from the large data we collected. There are 156 countries we could analyze but we will determine the right size of the sample to come up with accurate results and data. We will calculate the standard deviation for each of the variables and utilize the set of data where it is more consistent or where the data is statistically significant.