I decided on a dataset from the World Health Organization on Suicide Statistics from Kaggle. Growing up and in college there was always discussions and anecdotal evidence that males are more likely to commit suicide than females. There were also comparisons between developed countries (i.e. United States) to developing countries (i.e. Argentina), in that the developing countries have less suicides. Also, that suicide is more prevalent today than it was 40 years ago.

The dataset has six variables:

1. Country
2. Year, 1979 to 2016
3. Sex, male and female
4. Age, age group
5. Suicide No, Number of Suicides
6. Population, all living people within the country

To answer the questions that I have, I will be focusing on all the variables.

Questions that I will be asking to confirm if the above anecdotal evidence is true:

1. Do middle age males commit suicide more than females?
2. Do suicides happen more in developed countries than less developed countries?
3. Were there more recent suicides committed than there were 10, 20 and even 30 years ago?
4. What trends, if any, are among the demographic?
5. Does the United States lead in suicides?

During my exploratory data analysis, I did find that United States did in fact lead the world in suicides. It was followed by Romania, Sweden, Iceland, United Kingdom and Mongolia. I do believe that Sweden practices in assisted/passive suicides. Which has developed countries with more suicides that are documented.

The age groups 14 to 24 and 34 to 54 do commit suicide with middle age people committing suicide more.

It looked like the mid-1990s and early 2000s were the years that were hit by suicide the most.

I believe it would have been easier if I would have recoded sex male = 1 and female =2 to get a better analysis of the sex that committed the most suicides.

The challenges that I faced was Python, I see that I am going to have to practice more and become more confident. Also, time management.

Reference:

<https://www.kaggle.com/szamil/who-suicide-statistics#who_suicide_statistics.csv>