**Introduction**

Depression is something most people deal with behind closed doors. As time has gone on, we have become more open about our thoughts or feelings which allows for a better understanding of how others could potentially feel. Due to the widespread effect depression can have, numerous studies depict the differing variables that may be a central cause for common depression. Within this dataset, [here](https://catalog.data.gov/dataset/adult-depression-lghc-indicator), you can see the differing basic variables (such as ethnicity, age, income, etc.) that may have a basic cause on depression. However, [here](https://www.nih.gov/news-events/nih-research-matters/factors-affect-depression-risk), you can see the external factors that can lead to depression and that is what I’d like to analyze.

**Research Questions**

Question: Do the choices we make have a direct effect on our mood?

Explanation: Being that anxiety and depression have become more “popular”, what factors play a larger role in creating this feeling. When it comes down to it, what can the people who experience this do to become better each day? My goal is to be able to analyze the data and come up with a few different methods or ideas that may be a catalyst for change.

**Data to be Used**

I will be studying and analyzing data from data.gov which will analyze adult depression in California only. As well as [this](https://www.kaggle.com/twinkle0705/mental-health-and-suicide-rates) to analyze depression and suicide rates in pretty much every country. I cannot find data that includes depression and external factors such as drugs, low income, poor quality of life, etc. However, I’m hoping I will be able to draw conclusions based on what is common in the US compared to other countries. I have uploaded these important data sets onto Github.

**Approach**

To keep track of this information, I plan on keeping the data within excel as a base layout of the data. From there, I need to ensure that the data is accurate and useful for the conclusion I seek. After I ensure the data is correct, I want to begin to code the data and create graphs that allow me to see the important entries. I feel like there is some use that will come from me having a clean layout of not only the data itself but the visuals it provides. Also, in comparison to other charts I program, I plan on comparing the two to give me more of an edge when figuring out the reasons behind increased depression in one country oppose to increased cases in other. Like I mentioned earlier, finding the exact data I wanted was not possible. However, I believe I will be able to analyze this data and use deductive reasoning to decipher the conclusion I need. This will also provide more insight to the countries in question as well as the methods behind them. For example, if Bhutan is one the happiest countries in the world and is much simpler than America, maybe it should tell us something about ourselves. Having these visuals such will allow for a more coordinated understanding of the data at hand.

I plan on using different charts such as regression models to see if we can predict depression based on differing variables. Within my research, I found a ton of data regarding depression prediction based on tweets, tone, etc. I may also choose to utilize this data as a means of backing up my findings. I am uncertain (as of now) on how to perform cross-validation, but I plan on figuring out the concept and utilizing within my analysis. The main conclusion I want out of this is to determine the controllable factors or reasons as to one may feel a certain way.