Logan’s script:

**Intro:**

Introduce group members and what our project entailed. The main goals of our project were to find the best way, in our opinion, to determine who has a better overall healthcare system between the United States and Canada. We chose Canada because it is our neighbor and they are well known by most to have a centralized healthcare system that is boasted as being better than ours. Another goal was to determine areas in which each country reigns supreme for their healthcare systems. Finally we wanted to try to find a way that we could theorize how each country could adopt the aspects that they are weaker in from the winning country. **(transition to “the data” slide then into the next)** The source we used was from the OECD which is an organization that works to build better policies for better lives. They also had some of the largest datasets we could find and they were very thorough and well sorted. Information was from a few years ago but all information gathered is up to date for modern times. **(next slide)** Once we found the source we wanted to use we had to then determine what data we wanted to use because the json files containing the data were absolutely massive and not all variables matter for what our project is focused on. It was then just a task of filtering what we wanted as far as data and removing the rest.

**Dataframes:**

To begin the analysis of the data it first had to be sorted and organized into something that was, well, usable. Jerry and Peter both took care of their own datasets before pushing the csv files for me to handle the code for the project. From there it was all about manipulating the data to create dataframes that are easy to work with and easy to read from a visual aspect. Dataframes for the main data, winners of each variable/indicator, weight numbers, and additional math were all created in the process of parsing the data to be easier to work with for the python code.

**Code explanation 1:**

The python code that I wrote consisted of two main parts. I did write more code than what is going to be shown but its mostly just claiming variables that were used as well as creating the different dataframes with pandas. The first main piece of python code was used to pull data in order to basically just give a huge list that corresponded to which country won which variable/indicator. The output of this code gave what I then turned into a different data frame for the second bit of code

**Code explanation 2:**

For the second bit of important python code I used the same method as before but on a newly created dataframe that I ran some math through to give an output of the scores for each country after iterating through every row of the dataframe. This total score was just the addition of each respective weight score corresponding to the row in which each country won. This code is what spit out the final scores that will be presented to you next.

**Overall scores:**

After all the data parsing and python code was ran, the winner was determined to be Canada based on our chosen parameters and scoring system. You can see the scores were not even close between the two and the further analysis given by Jerry and Peter with their findings should help to clarify visually the scoring discrepancy.