Sarah Lufkin - Annotated File Showing Data Analytics

For my **live source**, I spoke to Associate Professor of Anesthesiology Joseph Caiati at Brown University, and I conducted an interview over the phone. His phone number is 401-243-3729. I was connected to Associate Professor Caiati through a family connection, I expressed my need to find a live source in the medical field to my family and a family member gave me his contact information.

The Cancer Mortality Rate Data set:
https://docs.google.com/spreadsheets/d/1pvjWFviFmpzcqbaCEVB0YYQ36AAixstHSd-VVgOBRoE/edit?usp=sharing

The first spreadsheet, The Cancer Mortality Rate Data set, used two different imported datasheets. One of which is the population of each state in the United States from Wikipedia's "List of U.S. states by population" and the second one is "Rate of Cancers in the United States, 2018" from the CDC website. I imported the state population data onto one tab and deleted all unnecessary columns, leaving a column for the states and the second column for all of their populations. I added a second tab for the cancer mortality rate data and deleted all unnecessary columns beside the states and the number of deaths there were. I added a third tab and put the state and population columns from the first tab into the third tab and then went back to the second tab. In the second tab, I rearranged the data to line up with the order of the first tab to make sure that all information was in the right place with the right state. I calculated the percentage of mortality due to cancer by dividing the case count by the population. The last tab is a column with the states and the calculation from the previous tab, which I then turn into a percentage. I put this final (4th) tab/spreadsheet into Datawrapper for my data visualization.

BMI Percentage Data Set:
https://docs.google.com/spreadsheets/d/1udcTJIDoUHBYGSy0beMirDtcdmSF6NLtMojvR7qnuW0/edit?usp=sharing

For the second data set, I imported one form of data called "Percent of obese adults (Body Mass Index of 30+)" from the State of Childhood Obesity and then put this data into Datawrapper for my data visualization.

Current Cigarette Use Among Adults (Behavior Risk Factor Surveillance System) 2018:
https://docs.google.com/spreadsheets/d/14QWUmVKV8WYHITrhTQNOaKKOHrD2sr4L9r31Dm8-oH4/edit?usp=sharing

For the third data set, I imported a data set "Current Cigarette Use Among Adults (Behavior Risk Factor Surveillance System) 2018" from the CDC and put that into Datawrapper for my data visualization.