Aaron J. Brown
DSC 530 - Final Project
Spring 2023

**STATISTICAL QUESTION/HYPOTHESIS**

This analysis is performed to determine if common health conditions such as: obesity, smoking, and drinking, affect the number of cases and deaths of COVID-19.Since COVID-19 is a respiratory disease, it is expected that several health conditions may increase the likelihood of contracting and dying from COVID-19, especially smoking. A positive relationship between cases/deaths by COVID and our outlined health conditions is expected to be unearthed by this analysis, relationships between the independent variables are expected to exist as well.

**OUTCOME OF EDA**

The most significant relationships among the variables were dependent variables: cases and deaths, with independent variables: obesity. In addition, the slope produced by graphing the CDF indicated that about 70% of all US counties fall within 0-50 cases. To account for outliers, we appropriated 90% of the data. Other than in the simple regression model (death*percent_adults_with_obesity), the P-values above suggest no relationship between our dependent variable (death) and our entered variables. The Simple Linear Regression Model appears to be the best fit for our data.

**WHAT DO YOU FEEL WAS MISSED DURING THE ANALYSIS?**

It is common knowledge that factors such as age and ethnicity increase the chances of mortality, not only for COVID-19, but for other conditions as well; data containing the ages or average ages of respondents could prove to be very beneficial to an analysis such as this. A lot of information regarding the relationships between COVID-19 cases and deaths and relating variables went unearthed with this analysis.

**WERE THERE ANY ASSUMPTIONS MADE YOU FELT WERE INCORRECT?**

Speculations of this analysis included: expecting that more health conditions would increase the likelihood of contracting and dying from COVID-19, especially smoking. The analysis of smoking failed to suggest a relationship between COVID-19 deaths and cases.

**WHAT CHALLENGES DID YOU FACE, WHAT DID YOU NOT FULLY UNDERSTAND?**

It proved to be rather difficult to apply code examples from the book directly to this analysis. Graphing using packages suggested by the text was especially complicated