**SAMUEL ABOAGYE (DATS 6401-PROGRESS REPORT 11/06/2018)**

**EXAMINING DRUG OVERDOSE DEATHS IN THE US FROM 1996 TO 2015**

**DATA CLEANING AND EXTRACTION**

I used pandas and excel for data cleaning, merging different spreadsheets to create one spreadsheet. Where it was possible, I used the calculated mean to fill in the missing values or I excluded those values if not possible. Columns that were not crucial to the project were dropped.

-**Overdose deaths by Gender**

I examined the overdose deaths for male and females and compared them with the overall overdose deaths.

**-Overdose deaths by Race**

I am using a pie chart to display the aggregate overdose deaths by race

**-Overdose deaths by Age**

I’m using a bar chart representing various age clusters to show a side by side comparison of overdose death among the various age groups. Tableau, google API and D3.js is being used for this task.

**-Overdose deaths by State**

I am using Leaflet.js and a heat map in tableau to show the spread of overdose death in each state

**-Overdose deaths by Income and Education level**

A combination of a bar graph, column charts and bubble charts to visualize the spread of overdose deaths at each income and education level.

**VIZUALIZATION**

I am using a combination of Tableau, Google API, leaflet.js, Plotly and D3.JS for visualization. Tableau is being used to create dashboards and storylines whiles leaflet.js is being used for mapping the states. Plotly and D3.js would be used to dynamically display the various demographic data.