Project title: Covid-19, Pneumonia and Influenza by Sate, Age, Gender and Race

CovidDB, Observation Date, Province / State, Country, Confirmed cases,

Team members: Michael Nosa & Sarah Mwanzi

Proposal:

To determine the number of deaths linked to Covid-19, Pneumonia and Influenza.

Question:

Is the coronavirus disease more severe than Pneumonia and Influenza?

Datasets:

Source:

<https://catalog.data.gov/dataset?q=covid&page>

<https://data.cdc.gov/api>

Extraction:

Provisional COVID-19 Death Counts by Sex, Age, and State - downloaded as csv file

Deaths involving coronavirus disease 2019 (COVID-19) by race and Hispanic origin group and age, by state - downloaded as json file

Cleaning Data:

used json module to clean and transform the data it into a more readable format

Extracted columns needed for the project

Transformation:

Transformed data from CSV and Json files into Pandas dateframes

Loading:

Created database connection

Loaded data into postgres sql.