Project Ideas

Covid-19 – Nytimes US Covid

* Covid State Count Dataset
* Covid County Dataset

Explanatory Data Analysis - Micah

* Plot of cumulative cases by state
* Visualziing spatial distribution of cases through time

Analysis

* Auto Correlation between Cases and Deaths
* Linear Correlation between States Cases/Deaths and Demographics
  + Population
  + Density
* Looking at case rate correlation with key holidays/events
* Web Scraping of Twitter

**Project Outline**

**Name –** Exploring Temporal Trends in NYTimes US State and County Covid 19 Data

**Introduction**

* **C**ovid 19 cases was a novel respitory virus that quickly spread across US in 2020
* Cases exponentially grew in different states at different times
  + Early spring peak for NY
  + Summer Rise for South (Florida, Texas, Arizona)
  + Fall Rise for West (Ca)
* New Death rate lagged behind cases by a few weeks
  + Incubation time for virus
  + Takes time for people to get sick, go to hospital, die
* ***Can we develop a model to predict future deaths based on current case numbers?***

**Methodology**

* **Exploratory Data Analysis - Micah**
  + Visualizing timeseries for average daily cases and deaths by state
  + Spatial trends
    - Box plots grouped by states
  + Visuals and statistics for monthly trends in cases by states
    - Box plot of average daily cases per month by state
    - Box plot of average daily deaths per month
  + Animated Map of cases across states
    - Shows when different regions were at their peak
* **Clutering analysis for heriartchy of states**
* **Timeseries lag anayslis**
  + Developing Lag-1 Autogressive model to predict daily deaths as a function of cases
  + What is the lag between the deaths and cases?
  + Is it consistent between states?

**Results**

* EDA
* Clustering Analysis
* Timeseries Analysis
  + How well does AR-1 Model Perform?

**Discussion**

* **Are there relationship between states? - Clustering**
  + What do they have common?
    - Location
    - Poltiics
    - Demographcs
* **Temporal Trends**
  + Does the lag between cases and deaths make sense
  + Can we predict with cases and/or deaths

**Conclusions**

* What are the implications for public policy?

**Appendix –**

* Figures, Tables, Codes