## CSE 2050 Honors Project

Marco Nardone Guerra

## Project Outline

- The purpose of this project is to extract as many significant insights as possible from a dataset.
- Before this, the dataset must be cleaned up and analyzed to plan out what to do.
- Afterwards, the student should creatively visualize the data in the most optimal way possible for the specific data points.
- Finally, use statmodels or similar package to make predictions and visualize them.

## Goals/Steps

- Import datasets
- Clean-up datasets (delete unnecessary columns, change NaN values, etc.)
- Create Visualizations:
  - Use time series datasets to create a global map and mark magnitude of coronavirus deaths/infections depending on location (i believe plotly is best for this)
  - Create a function that uses these time series and graphs specific ones and specific countries
    (i.e., an input of "country" and recovery/death/infections); Make equal one for US states
  - o Create a US covid dashboard with Plotly Dash-
    - Should have map mentioned prior, with bar chart per state, pie chart of % infected dead, recovered, etc
- Finally, create a forecast visualization with the <u>statsmodels package</u>