

Group 1

SSC 442: Project Outline

18 February 2020

### Chasing the Coronavirus

#### Data Collection Plan:

Current data regarding 2019 novel coronavirus (2019-nCov cases) was made available by John Hopkins University on [an online dashboard](#) that is frequently updated with new cases. Data from that dashboard was downloaded in .csv format from the following [linked website](#).

#### Analysis Plan:

We plan to analyze the data above by pursuing three specific topics of interest: main locations of case confirmations, the amount of recoveries and deaths when compared to confirmations, and the amount of confirmed cases over time (in other words, the probability of spread).

1. Locations of case confirmations
  - a. When analyzing the location of the virus confirmations, we can use gapminder or use the map library to show the different locations where the virus is prominent.
2. The amount of recoveries and deaths when compared to confirmations
  - a. For this data visualization, we can use the map library to geographically display the amount of deaths, recoveries, and latent cases per location (country and province) through circles of varying sizes.
3. The amount of confirmed cases over time
  - a. To analyze the amount of confirmed cases over time, we're going to create a slider tool that will begin on the date of the first confirmed case and end with the most recent update to our data.
  - b. As well, predict the probability of an epidemic in the United States by performing compartmental models (a combination of if/then and loop statements).

- i. The compartmental models will include four groups (untouched, latent, dead, and recovered) and be linked by foreign travel patterns, domestic/foreign interactions, and policy established by the United States (specifically, Customs & Border Protection) to prevent the spread of the coronavirus.