TITLE OF PROJECT	
Contact	Name: Runsheng Wang, Jinqi Lu Email:runsheng@bu.edu, jinqilu@bu.edu Cell Phone: N/A
Organization	N/A
Organization Description	N/A
Project Type	Data Science
Project Description	We seek to unveil COVID-19 insights through the compilation and analysis of a multitude of diverse datasets. We hope to explore, predict, and infer patterns that COVID-19 exhibits in order to highlight the impact of COVID-19 on different groups, as well as offering guidelines and advice to better protect our community from this current infectious disease and future ones.
Data Sets	● Multiple COVID-19 datasets from Kaggle as well as JHU and CDS will be combined into a large dataset that contains COVID-19 information that we need (vaccine data included) ● APIs from OpenStreetMap will be used to import spatial data and create layered maps ● APIs from Google map and NOAA will be used to obtain heatmaps. Uber data and taxi data will be used alongside with heatmaps to identify traffic hotspots ● We will download socioeconomic data (e.g. income, development index, urbanization level, insurance) from government websites as well as providers ● We will crawl twitter and reddit to obtain sentiment data to measure the attitude of the general public towards COVID-19 and mask-wearing ● We will use a dataset from Kaggle of face mask images to potentially build a image recognition system that identifies whether one is wearing a mask or not
Suggested Steps	We will first download and scrape the required dataset. We will then perform cleaning and combine everything into one big dataset. We will use the following eight techniques to analyze our input datasets ● clustering ● classification ● regression ● network analysis ● sentiment analysis (news, reddit, twitter, analysis with emojis) ● forecast/prediction ● recommender system ● spatial system

Questions to be answered in Analysis	#Very specific questions that the clients wants answered Travel advisory based on current pandemic trends and other natural disasters. Relationship between COVID-19 trend and Income/Temperature/supply/urbanization/insurance COVID-19 vs past infectious diseases Stay at home condition based on heatmap(traffic/toll/taxi) Spread pattern
Additional Information	Other relevant information