Data Visualization Final Project Proposal

Project Title: Effects of COVID-19 on the working population in the U.S.

Group Members:
Shauna Han (shh2145@columbia.edu),
Toby Law (kl3343@columbia.edu),
Wenyi Li (wl2770@columbia.edu),
Yu Wang(yw3760@columbia.edu)

GitHub repository: https://github.com/QMSS-G5063-2022/Group_A_COVID-19_insights

Abstract: We want to investigate how the Covid-19 pandemic affects the Labor Market of the United States. For this project, we plan to focus on several indicators in employment population, for example, unemployment rate by states/industries, population movement activity track, workers paid hourly rates and sentiment analysis for unemployment. We will focus on the United States at both the national and state levels.
We aim to generalize five visualization graphs using techniques such as ggplot2, leaflet, tmap, web scraping, text analysis, and more.

Visualizations:
1. **Unemployment rate**
   - Figure 1: A line chart to capture changes overtime, on national average, perhaps can contrast between different industries (static plot, with annotations indicate COVID-19 waves or lockdown periods)
   - Figure 2: An interactive map showing all 50 states, with a timeline toggle, plus individual values for each state. Data rage from Jan 2019 - Dec 2021, either monthly or quarterly values
   - National data: https://data.bls.gov/timeseries/LNS14000000?years_option=all_years
   - States data: https://www.bls.gov/lau/

2. **Average weekly wage - how wage and working hours have been impacted by the COVID-19 pandemic**
   - Datasets: https://data.bls.gov/cew/apps/data_views/data_views.htm#tab=Tables - data finder for weekly wage by industry https://www.bls.gov/web/empsit/ceseeb4a.htm - data table for weekly wage and hours, segregated by industry
   - Main figure: average weekly wage changes across time by industries (before and at the height of the COVID-19 pandemic)
   - Categorized by industry
   - Data range from 2017 - 2021, yearly
   - Line chart

- Auxiliary figure: changes in weekly working hours across time
- (timeframe: 2018 - 2021) Dataset on world Employment and Social Outlook
- https://www.ilo.org/wesodata/?chart=Z2VuZGVyPVsiVG90YWwiXSZ1bml0PSJOdW1iZXIiJnNlY3Rvcj1bXSZ5ZWFyRnJvbT0xOTkxJmluY29tZT1bXSZpbmRpY2F0b3I9WyJ1bmVtcGxveW1lbnQiLCJ0b3RhbExhYm91ckZvcmNlIiwid29ya2luZ0hvdXJzIl0mc3RhdHVzPVtdJnJlZ2lvbj1bIldvcmxkIl0mY291bnRyeT1bIlVuaXRlZCBTdGF0ZXMiXSZ3b3JraW5nUG92ZXJ0eT1bXSZ5ZWFyVG89MjAyMyZ2aWV3Rm9ybWF0PSJDaGFydCImYWdlPVsiQWdlMTVwbHVzIl0mbGFuZ3VhZ2U9ImVuIg%3D%3D

3. **Movement of population**
   - Map-based visualization: since cities have gone into lockdown during the pandemic, people have switched to working from home and offices have closed down, their day-to-day movement ranges and patterns have drastically changed. We aim to visualize differences in people's movement patterns across time before and during the pandemic.
   - Data for movement ranges of Facebook users consenting to share their location: https://dataforgood.facebook.com/dfg/tools/movement-range-maps
   - Geospatial data for the US, for mapping purposes: https://gadm.org/download_country.html

4. **Sentiment analysis for unemployment**
   - Mine the information posted on twitter site about unemployment using hashtags
   - Perform text mining to generate word cloud
   - Time range: Jan 2020 to Dec 2021
   - Useful resource → https://ieeexplore.ieee.org/document/7456920; https://towardsdatascience.com/how-to-navigate-analytics-job-search-during-covid-19-b6d2159fac8e

Any significant hurdles/doubts that we can foresee: We each need a better grasp on data visualization techniques as well as text mining techniques at the time of writing. There could be unforeseen challenges that we encounter during data cleaning, dealing with missing or inaccurate data.

Initial Brainstorming Memo:
- How the COVID-19 pandemic has affected work - unemployment rate, work hours change, income change - trends across time
- Geospatial data for the US on state level https://gadm.org/download_country.html
- Map-based visualization: since people can work from home and offices have closed down, people's workplace can be flexible and some choose to move home or relocate to

other places to work. We can visualize movement of people across time before and during the pandemic https://dataforgood.facebook.com/dfg/tools/movement-range-maps