

# **Hasnah Said**

## **A5 - Extension Plan**

### **November 10, 2021**

#### **Motivation/Problem Statement**

To slow the spread of COVID-19, many places in the world implemented different policies to slow the spread of the virus. These policies included lockdowns, canceling events and concerts, moving schools and work to a virtual setting, and restricting travel and public transportation. The main analysis I am interested in conducting for the remainder of this project revolves around the question: how visits to places, such as work and parks, have changed during the pandemic in Fairfield County, Connecticut?

I will also extend the visualization I created in A4 to include a timeline of COVID19 related deaths & hospitalizations, positive tests, and rate of vaccination. I will use some of the factors to see if they had an impact on the rate of infection and the number of confirmed cases. I will briefly look at COVID19's economic impact in my county and specifically look at how the rate of unemployment changed since the beginning of the pandemic.

#### **Research questions and/or hypotheses**

The timeline I will be looking at for this analysis is the same as A4: February 1, 2020 through October 15, 2021:

1. How did COVID19 change the movement of people in 6 categories of places?
2. Is there a correlation between movements trends and the number of positive COVID cases?
3. Did vaccination and masking have an impact on slowing the spread of COVID19?
4. What was COVID19's impact on unemployment in Fairfield County?

#### **Data used**

The datasets I will use for the analysis extension are:

[Google's Community Report 2020 and 2021:](#)

This data is released by Google and it is aimed to help officials understand how movements changed during the pandemic. The Community Mobility Report shows movement trends by region, across different categories of places. The US dataset has county-level data that I will be using for this analysis.

The data was collected from users who opted-in Location History setting (default is off) on their Google accounts, so this data represents a sample of the users and not the entire population. The user's information was anonymized before publishing and there might be empty fields where the data did not meet Google's privacy and quality thresholds.

**License:** I'll use the following citing from [Google](#) to attribute the data:  
*Google LLC "Google COVID-19 Community Mobility Reports".*  
*<https://www.google.com/covid19/mobility/> Accessed: <date>.*

#### [John Hopkins University's COVID data on Kaggle:](#)

This dataset provides daily information on COVID19 confirmed cases and related deaths

**License:** This dataset uses Creative Commons attribution, I'll also include attribution from JHU Github Repo:

- [Attribution 4.0 International \(CC BY 4.0\)](#)
- *COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University" or "JHU CSSE COVID-19 Data" for short, and the url:*  
<https://github.com/CSSEGISandData/COVID-19>

#### [CDC Daily Vaccination data. County:](#)

CDC's daily vaccination data shows the overall number of vaccines administered in the US at the county level

**License:** The data on the CDC is powered and licensed by [Socrata](#). Which uses Creative Commons, Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0). I'll use the CDCs URL for attribution.

The following datasets are from Connecticut Open Data Portal. They have data on COVID19 hospitalization, test results, and deaths. I will be using them to create visualizations of COVID19 impact on Fairfield county as well as analyzing the correlation between movement and the number of confirmed cases.

The third dataset is on the unemployment claims in Connecticut and I will use it to conduct an analysis on COVID's impact on unemployment and how it changed from pre-pandemic days. This data is provided on the town level so I will be using the last dataset to extract data on my county.

- [COVID-19 Test Results by Date of Specimen Collection \(By County\)](#)
- [COVID-19 Cases, Hospitalizations, and Deaths \(By County\)](#)
- [Initial Claims for Unemployment Benefits by Town](#)
- [Connecticut Towns - Crosswalk with Tax Codes and FIPS Codes](#)

#### **License:**

The data in CT Open Data Portal is licensed by [Socrata](#). Which uses Creative Commons, Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0). I'll use CT Open Portal URLs for attributions.

### **Unknowns and dependencies**

The biggest unknown I anticipate is whether or not I have all the data I need to answer the questions and create the visualizations of COVID19 in my county. The unemployment claims dataset is given by town so I will need an additional dataset to extract towns in my county.

## **Methodology**

I will be mainly creating visualizations to show trends and show the results of my analysis.

The first step is to clean and process the datasets. I will remove any unnecessary data for this analysis and then combine the datasets based on my county's name or FIPS. The main analytical method I will be is Descriptive Statistics, which includes calculating ratios and percentages. After that, I will plot the results over the timeline of the analysis

## **Timeline to completion**

There are about 4 weeks till the next deliverable is due. I plan on working on this project during the weekends and weekdays. The milestones I set for myself are as follows:

1. Milestone 1: Complete the initial data processing by Nov 16, 2021
2. Milestone 2: Conduct the analysis and have the results by Dec 1, 2021
3. Milestone 3:
  - a. Create the slides for A6 by Dec 6, 2021
  - b. Began writing the final report
4. Milestone 4: Submit the slides on Dec 7, 2021