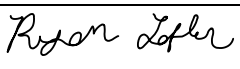
An Analysis of COVID-19 in the U.S.

Measuring Cases, Deaths, Vaccination Counts, Demographic Profiles and Americans' Responses to the Pandemic

Ryan Lafler, Aradhya Agrawal, Aishwarya Bhangarshettra, Bhagyashri Patil

Database Tables:

(1) Vaccinations Time Series table: *https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh*
(2) COVID-19 Case Count and Deaths Time Series Table: https://github.com/nytimes/covid-19-data
(3) Five-Year American Community Survey Demographics Table (2019 data): https://www.census.gov/programs-surveys/acs
(4) Twitter COVID-19 Table: https://drive.google.com/file/d/1mcCC-N07TK0YJkLGmbMDPrplrnXvcA_L/view?usp=sharing

Group Name	COVID-19 U.S. Response	**Date**	9/27/2021
Member Name	Ryan Lafler	**Signature**	
Member Name	Aradhya Agrawal	**Signature**	
Member Name	Aishwarya Bhangarshettra	**Signature**	
Member Name	Bhagyashri Patil	**Signature**	
Project Statement (200 words)	The vaccination dataset is a time series showing the number of vaccines administered, by U.S. County, since January 2021. It provides detailed descriptions of vaccination numbers by age group and frequencies for those partially and fully vaccinated. The COVID-19 case count and death table shows the number of cases and deaths in all U.S. counties beginning on January 21, 2020; and updating to the current date. The data were published to GitHub by the New York Times. The demographics table was collected from the five-year American Community Survey published by the Census Bureau from 2014-2019 (their most recent publication). Units of interest are U.S. Counties using variables measuring median age, median household income, educational attainment and racial composition.		

	These three tables are linked together by a primary key identifying a county's FIPS code, permitting relationships and analysis to be conducted between tables inside our COVID-19 database. Analysis will be centered on how COVID-19 spread through counties and what characteristics were shared among worst-impacted counties. Furthermore, what characteristics are shared among counties with high- and low-vaccination rates? The final table contains text scraped from Twitter, hosted on Kaggle, that analyzes people's views about the vaccines from a global perspective. Capturing the human response to a suddenly changed world, this table provides a unique qualitative perspective in understanding the spread of disinformation, relief and the "return to normalcy". It contains approximately 300,000 tweets collected over time from August 2020.
Bullet points (objective and/or steps toward project completion)	1) Understand and clean the data tables, understand relationships connecting them to theme of project 2) Analyze demographic characteristics of U.S. counties 3) Measure Covid cases and deaths reported in U.S. counties as a spatial time series (how do cases and deaths change on the map over time?) 4) Measure Covid vaccine counts by U.S. counties: which counties boast high vaccine-counts relative to population? Low-vaccine counts? 5) Assess whether trends exist between county demographics and confirmed Covid cases/vaccine count 6) Conduct basic text analysis on tweets concerning Covid-19 topics beginning in August 2020
Project Weekly schedules and individual assignments	Week 1 (9/27-10/1): Project Kickoff • All members: Examine datasets, develop fundamental understanding of the data being collected. Discuss methods for implementing data and the types of analysis we can conduct in the project.
	Week 2 (10/4-10/8): Explore and clean dataset 1 • Ryan Lafler: • Aradhya Agrawal: • Aishwarya Bhangarshettra: • Bhagyashri Patil:

	Week 3 (10/11-10/15): Explore and clean dataset 2 • Ryan Lafler: • Aradhya Agrawal: • Aishwarya Bhangarshettra: • Bhagyashri Patil:
	Week 4 (10/18-10/22): Exploratory Data Analysis 1 • Ryan Lafler: • Aradhya Agrawal: • Aishwarya Bhangarshettra: • Bhagyashri Patil:
	Week 5 (10/25-10/29): Exploratory Data Analysis 2 • Ryan Lafler: • Aradhya Agrawal: • Aishwarya Bhangarshettra: • Bhagyashri Patil:
	Week 6 (11/1-11/5): Exploratory Data Analysis 3 • Ryan Lafler: • Aradhya Agrawal: • Aishwarya Bhangarshettra: • Bhagyashri Patil:
	Week 7 (11/8-11/12): Start webpage and start analyzing findings • Ryan Lafler: • Aradhya Agrawal: • Aishwarya Bhangarshettra: • Bhagyashri Patil:
	Week 8 (11/15-11/19): Update webpage and explain findings • Ryan Lafler: • Aradhya Agrawal: • Aishwarya Bhangarshettra: • Bhagyashri Patil:
	Week 9 (11/22-11/26): Update webpage and explain findings 2 • Ryan Lafler: • Aradhya Agrawal: • Aishwarya Bhangarshettra: • Bhagyashri Patil:
	Week 10 (11/29-12/3): Finalize Webpage • Ryan Lafler: • Aradhya Agrawal: • Aishwarya Bhangarshettra: • Bhagyashri Patil:
	Week 11 (12/6-12/8): Finalize Project • All members: Practice presentation and make any final adjustments