Project Proposal

County-Level Analysis of Covid-19 Vaccination rate in the United States.

Team – **Group 3** Alekhya Kotha Mohammed Zain Badal Pallavilatha Vadlamudi Haricharan Mukthineni Pratyusha Gadasanthala

# INTRODUCTION

The Covid-19 pandemic has claimed over a million lives in the United States to date. Safe and effective vaccines have been available since late 2020, providing a path out of the pandemic. However, over a year into the vaccination campaign, rates continue to lag behind the target vaccination rate in many parts of the country. As of February 2023, only around 65% of the US population is fully vaccinated. Achieving high coverage across all communities is critical to controlling the spread of Covid-19. Personal characteristics such as age have been demonstrated to influence vaccination rates, complicating matters. In this complex system, some counties stand out as outliers due to their behaviors, which differ significantly from those of their neighbors. This project seeks to uncover the factors driving the differences in Covid-19 vaccination rates at the county level across the US. It will compile and analyze vaccination data alongside other demographic, socioeconomic, health, and political factors.

# PROBLEM STATEMENT

Understanding and mitigating regional disparities in COVID-19 vaccination remains crucial. We need to identify key factors like economic status, education, and state policies impacting vaccination rates. Analyzing outliers – counties with vastly different outcomes – might reveal crucial insights. Predicting when 80% of the population will be vaccinated across all regions to achieve herd immunity requires considering into account complex demographics and behavior. Exploring the correlation between vaccination rates and infectious numbers at the county level, while accounting for other variables, will assess the effectiveness of vaccination campaigns. Investigating potential data inaccuracies and violations is crucial for reliable analyses. Finally, a comprehensive analysis of COVID-19 and county-specific data will inform data-driven recommendations to improve vaccination rates and reduce infections, ultimately optimizing public health interventions.

# OBJECTIVE

The primary objectives of this project are as follows:

* Analyzing vaccination data at the county level in the United States to identify variations in vaccination rates among communities
* Examining the role of vaccine hesitancy in shaping vaccination behavior
* Identifying counties with significantly higher or lower vaccination rates compared to neighboring areas and investigate underlying reasons
* Determining how different age groups influence the vaccination rates at county level

# LITERATURE REVIEW

Jingling Gao, yuqi Guo, and Lilian Ademu (2022) conducted a longitudinal analysis to examine the association between public fear of COVID-19 (PFC) and county-level COVID-19 vaccination rates in the United States. The paper "Associations between Public Fear of COVID-19 and Number of COVID-19 Vaccinations: A County-Level Longitudinal Analysis" used county-level data from 2856 counties across 50 states, excluding Texas, from February 1 to July 1, 2021. The PFC was measured by mining Twitter data using an emotion detection tool called Test2emotion. Random-effects models were adopted to analyze the associations while controlling for county demographics, vaccine policies, and other variables. The results showed PFC has a positive association with county vaccination numbers. However, the impact of PFC on vaccinations varies based on county income level, unemployment rate, and vaccine incentive policies. Higher PFC links to higher vaccinations for lower-middle income counties but not higher income counties. PFC has diverging associations with vaccinations above and below 6% unemployment rates. Also, higher PFC associates with increasing vaccinations under lottery incentives but decreasing vaccinations under bonus incentives. The study concludes that considering public emotions like fear and socioeconomic factors is critical when making policies aimed at increasing vaccination rates.

# METHODOLYGY

In our analysis of COVID-19 vaccination rates at the county level in the United States, we will employ multiple regression analysis techniques. Specifically, we will use ordinary least squares (OLS) regression to measure the effect of various factors on vaccination rates. OLS regression is a widely-used method that estimates the relationships between a dependent variable (vaccination rates) and multiple independent variables (such as age, geographic location, vaccine hesitancy and communities) by minimizing the sum of the squared differences between the observed and predicted values. This technique allows us to quantify the impact of each independent variable on vaccination rates while controlling for the effects of other variables.

Furthermore, we may utilize logistic regression analysis to examine the probability of a county achieving certain vaccination thresholds based on its characteristics. Logistic regression is suitable when the dependent variable is binary or categorical, such as whether a county has achieved a specific vaccination rate benchmark. By fitting a logistic regression model, we can assess the likelihood of a county reaching a certain vaccination rate based on its demographic, economic, and policy-related attributes. This technique enables us to understand the factors that contribute to the likelihood of successful vaccination campaigns at the county level, providing insights for targeted interventions and policy recommendations.

# REFERENCES

Gao, Jingjing & Guo, Yuqi & Ademu, Lilian. (2022). Associations between Public Fear of COVID-19 and Number of COVID-19 Vaccinations: A County-Level Longitudinal Analysis. Vaccines. 10. 1422. 10.3390/vaccines10091422.

[https://www.researchgate.net/publication/363142328\_Associations\_between\_Public\_Fea](https://www.researchgate.net/publication/363142328_Associations_between_Public_Fear_of_COVID-19_and_Number_of_COVID-19_Vaccinations_A_County-Level_Longitudinal_Analysis) [r\_of\_COVID-19\_and\_Number\_of\_COVID-19\_Vaccinations\_A\_County-](https://www.researchgate.net/publication/363142328_Associations_between_Public_Fear_of_COVID-19_and_Number_of_COVID-19_Vaccinations_A_County-Level_Longitudinal_Analysis) [Level\_Longitudinal\_Analysis](https://www.researchgate.net/publication/363142328_Associations_between_Public_Fear_of_COVID-19_and_Number_of_COVID-19_Vaccinations_A_County-Level_Longitudinal_Analysis)

Topmiller, Michael & Carrozza, Mark & McCann, Jessica & Rankin, Jennifer. (2022). Identifying health centers in areas with low COVID-19 vaccination rates & high rates of vaccine hesitancy. Annals of family medicine. 2734. 10.1370/afm.20.s1.2734. [https://www.researchgate.net/publication/359840526\_Identifying\_health\_centers\_in\_area](https://www.researchgate.net/publication/359840526_Identifying_health_centers_in_areas_with_low_COVID-19_vaccination_rates_high_rates_of_vaccine_hesitancy) [s\_with\_low\_COVID-19\_vaccination\_rates\_high\_rates\_of\_vaccine\_hesitancy](https://www.researchgate.net/publication/359840526_Identifying_health_centers_in_areas_with_low_COVID-19_vaccination_rates_high_rates_of_vaccine_hesitancy)