Project Deliverable 2

Joel: Yinzhou Lu, Zoey: Zou Yang, Cassie: Seung Hee Lee, Claire: Min Jeong Ahn

1. Introduction

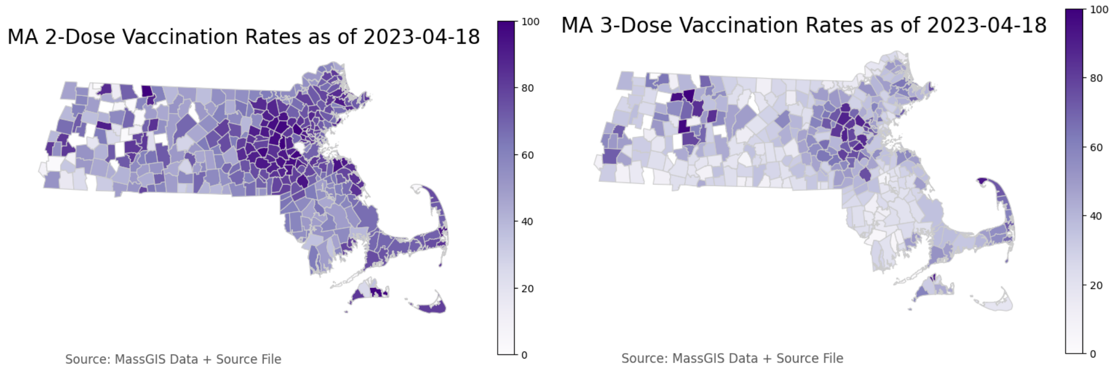

Figure 1: 2-Dose (Left) & 3-Dose (Right) COVID-19 Vaccination Coverage Across Cities and Towns in Massachusetts

In Massachusetts, vaccination disparity emerges as a formidable public health challenge, necessitating refined strategies. As depicted in Figure 1, an analysis of 2-dose and 3-dose vaccination coverage across 335 MA cities and towns reveals striking vaccination variability. Certain municipalities report full vaccination, while others, such as Towns Gosnold and Monroe, exhibit a 0% 2-dose vaccination rate. Preliminary graphical analysis of state vaccination data indicates a correlation between proximity to Boston and elevated vaccination rates. This multifaceted conundrum warrants an integrative approach that elucidates causative elements in hopes of sculpting potent interventions. Healthcare authorities in MA are concerned about vaccine uptake disparity among diverse demographic groups, especially the pediatric and geriatric cohorts.

Building upon the previous year's analysis that focused on age, race, and the impact of grassroots initiatives, we conduct an in-depth examination of a broader scope of variables including age, ethnicity, the Social Vulnerability Index (SVI), gender, economic status, employment, and population density. These factors are scrutinized to understand their correlation with vaccine hesitancy in MA, particularly in urban locales like Revere, Wellesley, and New Bedford. Our methods, encompassing census and health record data interpolation, regression, and hypothesis T-tests, aim to uncover patterns in vaccine perceptions across different demographics.

By uncovering the roots of vaccine reluctance, our overarching goal is to help foster a collaborative environment, uniting MA policymakers and community leaders to promote health inclusivity. We hope our efforts will provide actionable insights into various population strata, shaping practical policies to increase vaccine acceptance and thereby elevate the overall public health.

2. Data Collection & Project Scope

The Massachusetts Department of Public Health (MDPH) stewards a comprehensive COVID-19 Vaccination Database, indispensable for surveilling statewide immunization trajectories. This repository elucidates a granular spectrum of vaccination indices, segmented by demographic attributes such as age, gender, and race, and vaccination phases, including initial, full, and booster doses. When analyzed in conjunction with socio-economic metrics such as the Social Vulnerability Index (SVI)—which encapsulates variables of at-risk cohorts including poverty, low educational attainment, and unemployment, obtained from MDPH and US Census Bureau APIs—this nexus of data streams creates a detailed panorama of public health landscapes.

In this analysis, we examine vaccination rate data across five Massachusetts cities with disparate economic profiles (Figure 2, Left). These cities—New Bedford, Chelsea, Newton, Revere, and Wellesley—are represented in Figure 2, Right, with each city distinguished by a unique color. Our goal is to uncover the nuanced factors that influence vaccination uptake in these locations, thereby advancing our understanding of the generalizable determinants that drive immunization inequities worldwide and unraveling the intricate web of public health dynamics.

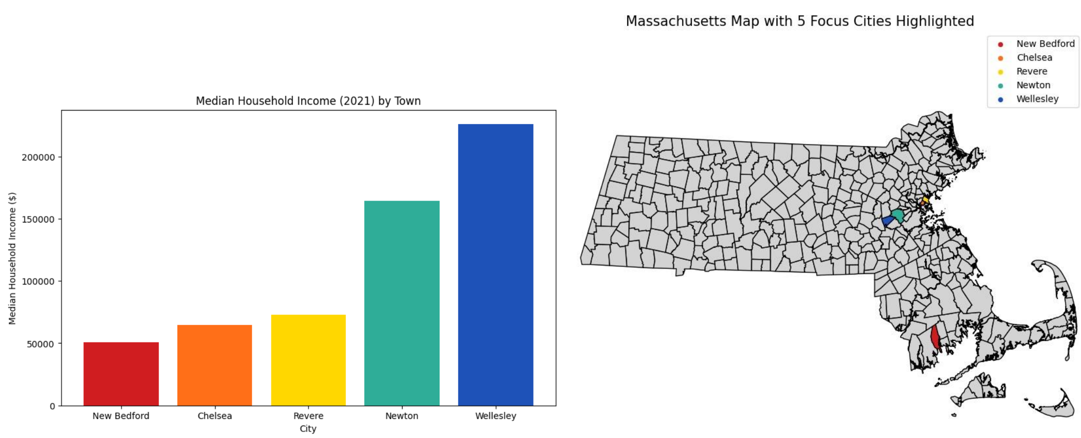

Figure 2: (Left) Median Household Incomes in 5 MA Cities (2021) and (Right) Geospatial Distribution of Five MA Cities of Interest

3. Preliminary analysis

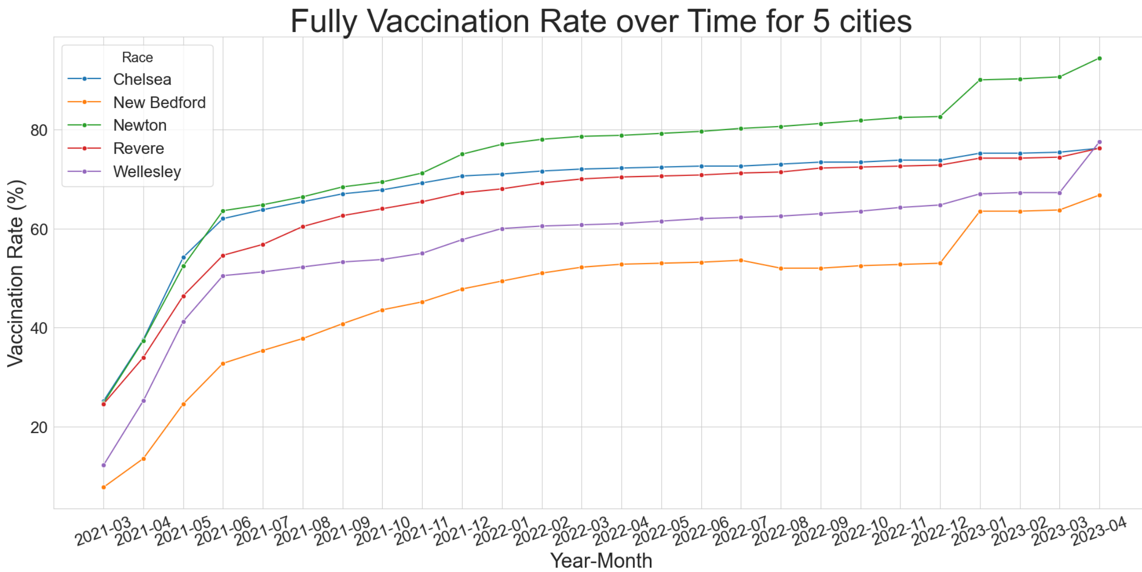

Figure 3. Fully Vaccination Rate Over Time by Five Cities.

Figure 3 delineates the progression of full vaccination rates over time in five distinct cities spanning from March 2021 to April 2023. The vaccination rates have increased across all cities, yet the gradient of increase and the final vaccination percentages point to a significant heterogeneity in public health outcomes. Newton leads with the highest vaccination rates, potentially reflective of its elderly population. Revere and Chelsea exhibit comparable rates, suggesting similar healthcare outreach or demographic patterns, while Wellesley and New Bedford follow, with New Bedford showing the lowest rates. This may indicate disparities in vaccine distribution or varying levels of vaccine hesitancy. Additionally, although Wellesley's initial vaccination rates were lower than those in Revere and Chelsea, by April 2023, they marginally surpassed the rates observed in these two cities.

Personal beliefs and attitudes can significantly influence vaccination decisions. In the elderly, these decisions are often shaped by lifelong beliefs and the potential for side effects, which may cause hesitancy towards vaccines. For children, parents' or guardians' beliefs and the social milieu shape vaccination choices, balancing perceived risks and benefits. Furthermore, Healthcare policies and public health campaigns play a pivotal role in vaccination uptake. Targeted campaigns and policies that remove financial barriers, like Medicare, can increase vaccination rates among the elderly by emphasizing the risks of non-vaccination. For children, mandatory school vaccinations and public education on community health responsibilities ensure higher vaccination rates. Finally, socioeconomic status is a key predictor due to its links with education and healthcare access. Among the elderly, factors like retirement income and social support impact vaccination decisions, with those in lower income brackets possibly prioritizing other needs due to cost concerns and facing challenges like limited transportation. For children, family socioeconomic status plays a significant role; higher status often correlates with better access to healthcare and higher vaccination rates, while lower status may lead to barriers such as costs and limited healthcare access. However, the 1-dose pediatric vaccination data revealed a curious trend that may indicate that high socioeconomic status may not necessarily predict a high vaccination rates.

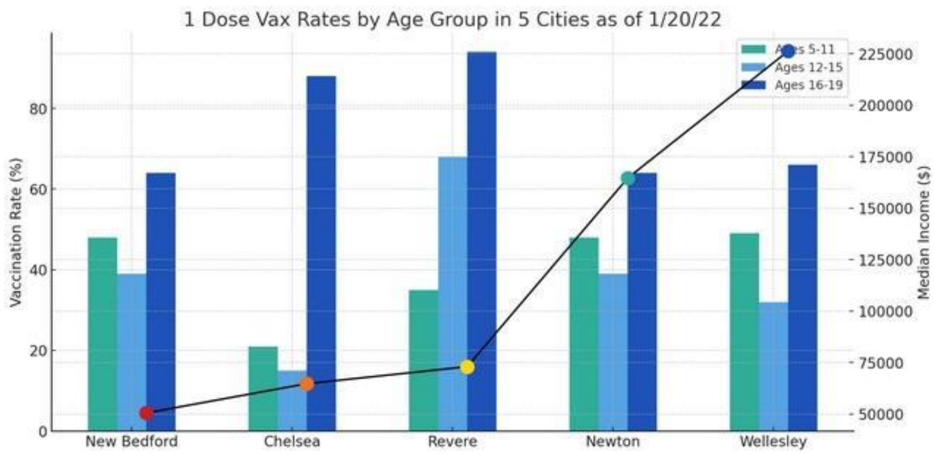

Figure 4. 1-Dose Vaccination Rates by Age Group in 5 Cities

Figure 4 depicts the first-dose vaccination rates across five cities, categorized into three pediatric age groups as of January 20, 2022. Notably, the 16-19 age group in a less affluent city such as Revere exhibits the highest vaccination rate. Meanwhile, the vaccination trends in more affluent areas like Newton and Wellesley suggest a surprising tendency for the wealthier populations to be more hesitant about COVID-19 vaccinations for their children.

Given the different vaccination rates among five cities, we are looking for the factors that may impact on the vaccination rates. As vaccination rates are highly influenced by the geographic demographic and economic indicators. In the process of exploratory data analysis, we are investigating how the income level, poverty level, employment rate, education, age, race and gender departs among five cities. Further, how these factors affect the difference of vaccination rate among five cities.

Social Vulnerability Index (SVI)

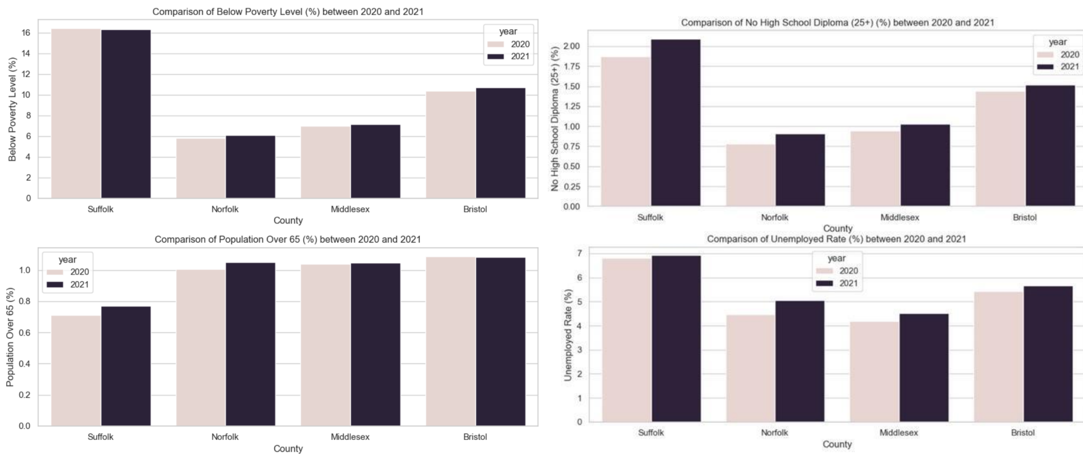

Figure 5. SVI among Five Cities

The SVI comparison utilizes bar graphs to represent four distinct socio-economic indicators: poverty level, lack of high school diploma, population over 65, and unemployment rate. Each indicator is compared between 2020 and 2021 across Suffolk, Norfolk, Middlesex, and Bristol counties.The year-over-year Poverty Level comparison suggests that Suffolk County has the highest economic vulnerability. The data on educational attainment reveals a higher rate and the highest unemployment rate in Suffolk County compared to the others. Regarding the senior population, the trend across the counties shows a consistent or slightly increasing percentage of residents over 65.

Suffolk County exhibits higher vulnerability in terms of poverty and education, which may negatively impact its resilience. The other counties, while displaying lower poverty and higher education levels, are not without their challenges. The data suggests that these counties may be better positioned to support vaccination efforts and other health initiatives due to their economic and educational advantages.Furthermore, our analysis extended to socio-economic factors, specifically income levels and unemployment rates.

Economic Factors

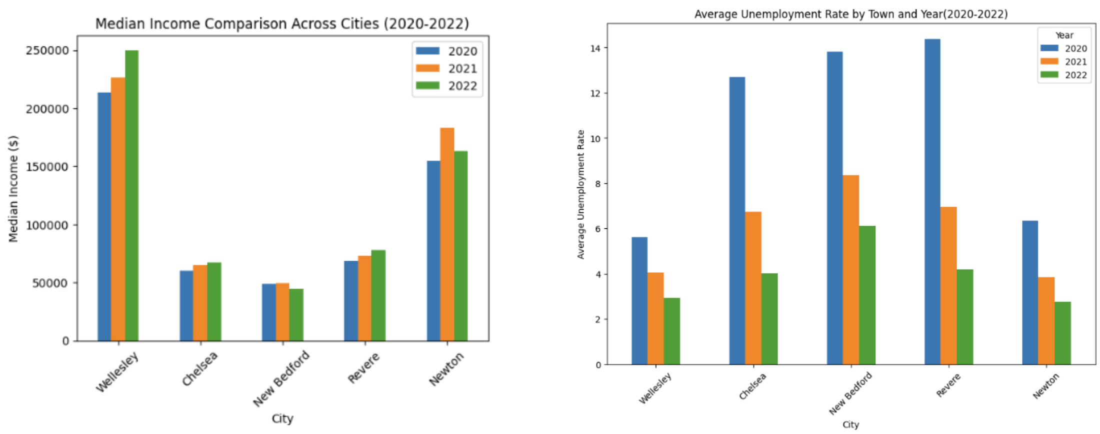

Figure 6 & 7: Income & Unemployment Rate by City and Year

Figure 6. shows that Wellesley and Newton, with the highest median incomes among the five cities studied, may experience better healthcare access and higher vaccination rates. Higher incomes typically afford residents more flexible schedules and better healthcare literacy, enabling easier access to vaccination services. The observed data suggests a correlation between economic stability and vaccination uptake, as wealthier areas like Wellesley and Newton show increased vaccine acceptance. Factors contributing to this trend may include a higher level of education regarding health and more substantial trust in healthcare systems.

Figure 7. illustrates a distinct economic pattern within the data set, showing that Chelsea, New Bedford, and Revere have experienced the highest unemployment rates over the past three years. This economic indicator is crucial as it often reflects broader socio-economic challenges within these communities, including poverty.The assumption that these three cities might have lower vaccination rates due to poverty is based on several interrelated factors associated with economic disadvantage such as health prioritization, and information and education.

Race

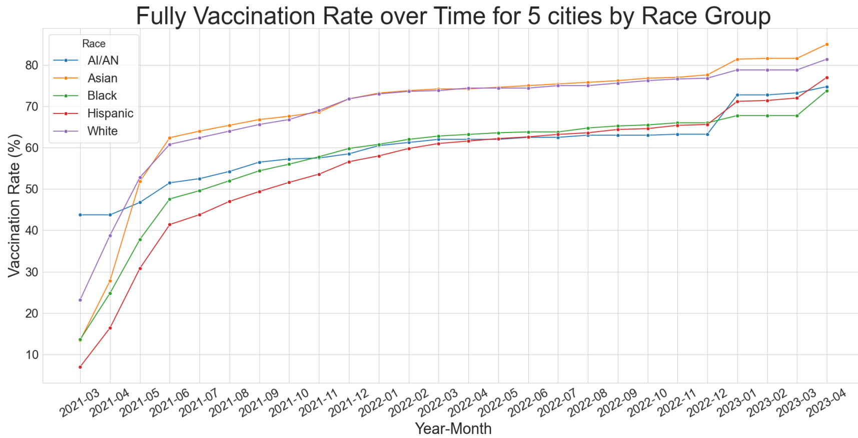

Figure 8 & 9 : 2-Dose & 3-Dose Rate over Time for 5 Cities by Race

From Figure 8, stratified by racial demographics, a clear upward trend is observed from March 2021 to April 2023 among five cities with initial rates near 20% escalating to a range between 70-80%. The data indicate a hierarchy in vaccination rates with the White population leading, followed by Asian, Black, Hispanic, and American Indian and Alaska Native (AI/AN) groups, respectively. This hierarchy suggests potential disparities that could be rooted in systemic factors such as healthcare accessibility, socioeconomic status, or vaccine hesitancy.Notably, as of the initial data point, the vaccination rate among Hispanic Americans was the lowest among the five racial groups. By the end of April 2023, however, this rate eclipsed those of both the Black and AI/AN populations.

Figure 9. reflects the booster vaccination rates across five racial groups in five cities from December 2021 to March 2023. The booster rates have risen from an average of approximately 15% to around 40% for all groups. The data indicate a hierarchy in vaccination rates with the White population leading, followed by Asian, Black, Hispanic, and AI/AN groups, respectively. It indicates that while progress has been made in administering booster shots, there is a crucial need for strategies to increase equity in vaccination efforts.

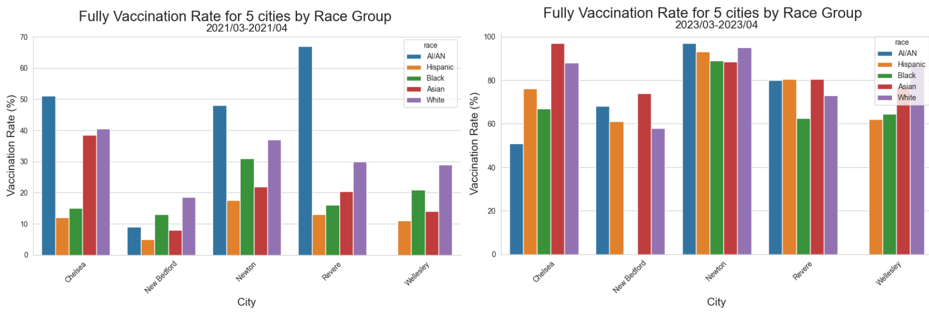

Figure 10 & 11: Histogram of Single Time Points for 5 Cities by Race

Figures 10 and 11, comprising two histograms, illuminate the variations in vaccination rates by race within each city. As of March 2021, vaccination rates for the Hispanic population were lower compared to the Black population across five cities. However, by April 2023, the Hispanic group's vaccination rates exceeded those of the Black group. Similarly, in March 2021, the Asian group's vaccination rates were lower than those of both White and Black groups. By April 2023, the vaccination rates for the Asian population outstripped those of the Black population in all five cities and surpassed the White population in Chelsea, New Bedford, Newton, and Revere.

Age

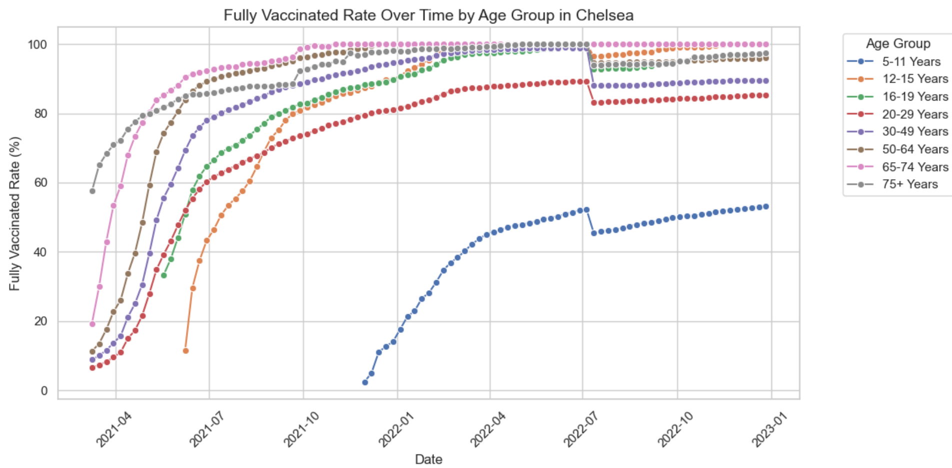

Figure 12: Fully Vaccinated Rate Over Time by Age Group in Chelsea

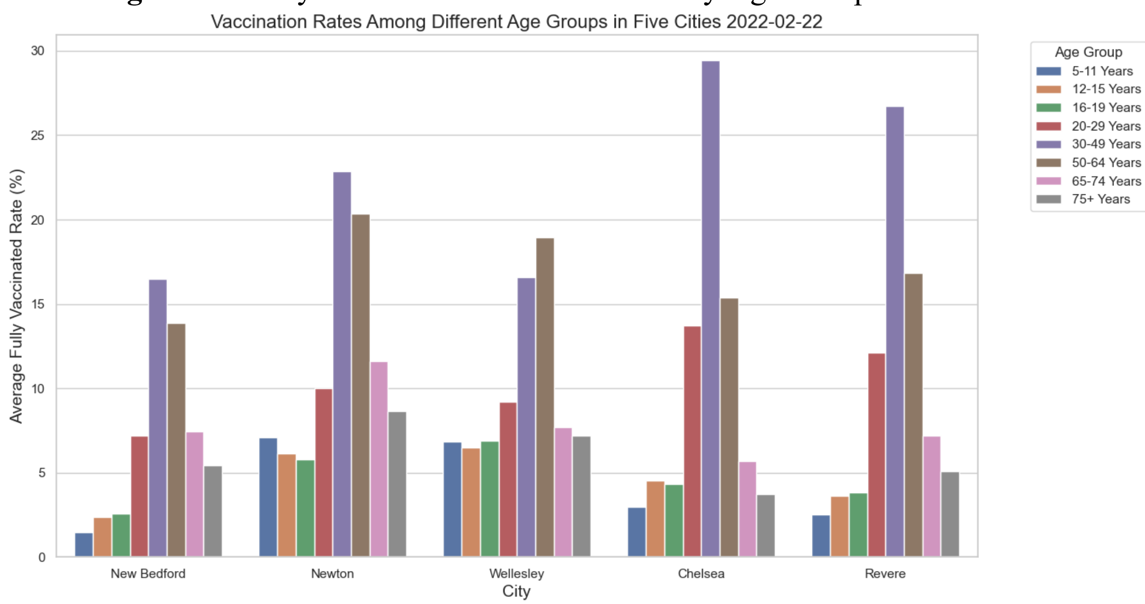

Figure 13: Vaccination Rates Among Different Age Groups in 5 Cities

Figure 12 depicts the evolution of full dose vaccination rates over time in Chelsea, while Figure 13 examines the rates across various age groups in five cities as of February 2, 2022, a date selected due to the inclusion of the 5-11 years age group in vaccination protocols. These visualizations highlight discernible patterns in age-specific vaccination uptake. Chelsea, ranking third in average vaccination rates among the age groups considered, exhibits the highest rate in the 30-49 age group, indicative of efficacious local vaccination initiatives, potentially bolstered by grassroots movements. Chelsea's vaccination rates are consistently high compared to other cities, suggesting effective outreach and education strategies that engage a broad spectrum of demographics. This observation leads to a hypothesis: Cities with higher socioeconomic status may demonstrate more uniform vaccination rates across age group

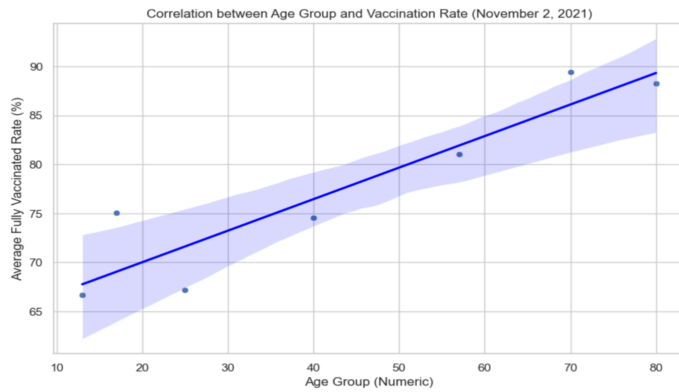

Figure 14: Pearson Correlation Between Age Group and Vaccination Rate

Figure 14 shows a scatter plot with a regression line for November 2, 2021, illustrates the correlation between age groups and vaccination rates. The Pearson correlation coefficient for this date is approximately 0.926, indicating a strong positive correlation.This high correlation coefficient suggests that as age increases, the fully vaccinated rate also tends to increase, particularly on this date. The older age groups show higher vaccination rates compared to the younger ones. This trend is visually supported by the upward slope of the regression line in the scatter plot.

Population Size

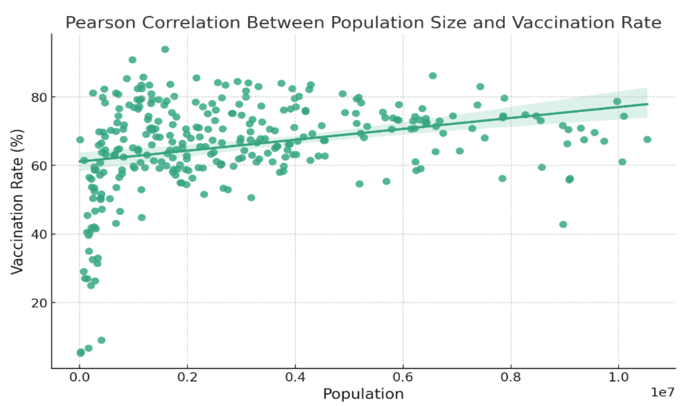

Figure 15: Pearson Correlation Between Population Size and Vaccination Rate

Figure 15 visualizes the correlation between population size and vaccination rates as of 2023-02-15 across various cities (green dots). The distribution of these points suggests that population size is not a strong linear predictor of vaccination rates, highlighting the multifaceted nature of vaccination campaigns influenced by public health policies, community outreach, and socio-economic conditions. This complexity calls for a nuanced approach to public health strategies.

Gender

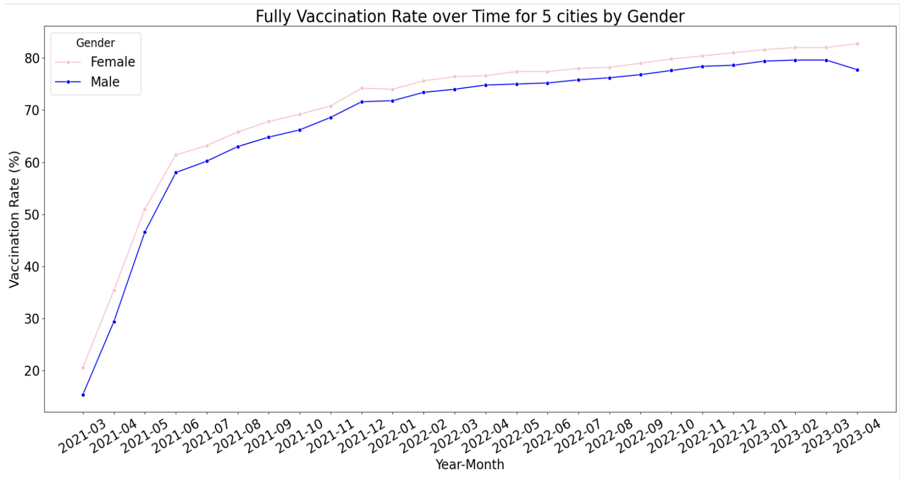

Figure 16. 2-Dose Vaccination Rates Over Time by Gender

Figure 16 delineates the evolution of vaccination rates from March 2021 to April 2023. It reveals a general ascent in vaccination uptake over the period, interspersed with minute but discernible declines. These fluctuations might be ascribed to potential lapses or inaccuracies in governmental data recording methodologies.

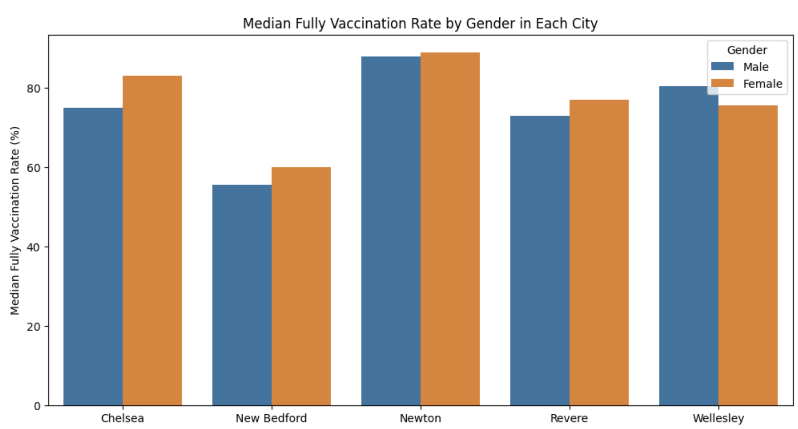

Figure 17. 2-Dose Vaccination Rates Over Time by Gender & City

Next, we refine the gender-specific vaccination data to discern disparities in distribution between men and women. Initial data from March 2021 suggested a significantly higher vaccination rate among women. Yet, the updated visual representation in Figure 17, encompassing median vaccination rates from March 2021 to April 2023, indicates a diminished gender gap. Notably, in Wellesley, vaccination rates for men now exceed those for women.

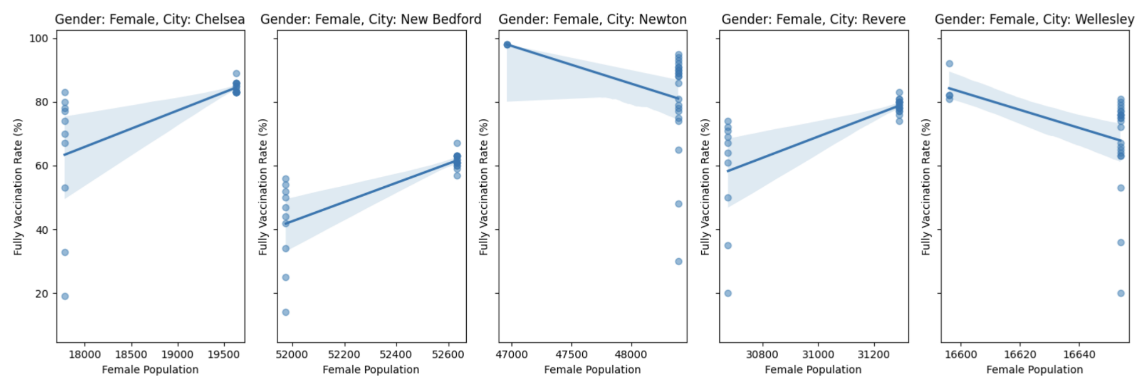

Figure 18. Comparison of Full Vaccination Rates and Female Population Proportions Across Cities

This figure delineates the relationship between the proportion of the female population and full vaccination rates across various cities. A direct correlation is generally observed, yet exceptions are notable in Wellesley and Newton. As previously indicated, Wellesley shows a higher vaccination rate in males than in females, while Newton presents comparable rates between genders. Furthermore, potential variances in data collection must be considered. The US Census data lacks regular updates on demographic shifts, challenging the presumption of a static female population over time.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:          fully_vac_rate   R-squared:                       0.300
Model:                             OLS   Adj. R-squared:                  0.286
Method:                  Least Squares   F-statistic:                     21.63
Date:                 Mon, 27 Nov 2023   Prob (F-statistic):           5.27e-18
Time:                         12:46:30   Log-Likelihood:                -1071.7
No. Observations:                  258   AIC:                             2155.
Df Residuals:                      252   BIC:                             2177.
Df Model:                            5
Covariance Type:             nonrobust
=======================================================================================
                          coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------------
Intercept              71.0236      2.370     29.968      0.000      66.356      75.691
gender[T.Female]        2.8760      1.941      1.481      0.140      -0.947       6.699
city[T.New Bedford]   -20.8077      3.058     -6.805      0.000     -26.830     -14.786
city[T.Newton]          9.8585      3.088      3.192      0.002       3.777      15.940
city[T.Revere]         -3.8654      3.058     -1.264      0.207      -9.887       2.156
city[T.Wellesley]       0.0769      3.058      0.025      0.980      -5.945       6.099
==============================================================================
Omnibus:                      114.432   Durbin-Watson:                   0.084
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              343.384
Skew:                          -2.033   Prob(JB):                     2.72e-75
Kurtosis:                       6.926   Cond. No.                         6.53
==============================================================================
```

Figure 19. Linear Regression for Gender and Vaccination Rate

Figure 19 presents a linear regression summary wherein the coefficient for gender[T.Female] is 2.8760. This indicates that being female correlates with an approximate 2.88 percentage point increase in vaccination rates relative to males. However, this association is not statistically significant, as evidenced by a p-value of 0.140.

4. Regression Analysis

Prior to executing a comprehensive regression analysis incorporating control variables, initial regressions are performed between vaccination rates and variables of interest to assess the validity of our hypotheses.

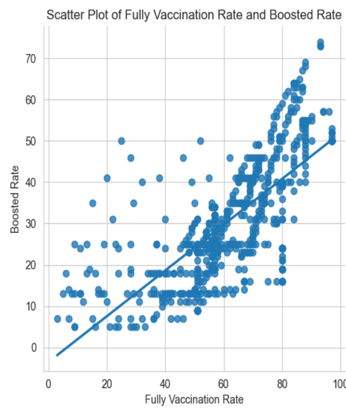

Figure 20. Scatter Plot of Full Vaccination Rate vs. Booster Vaccination Rate

Figure 20 illustrates a robust positive correlation between the rate of full vaccination and booster vaccination, corroborating our hypothesis that increased uptake of booster vaccinations is associated with higher full vaccination rates.

OLS Regression Results

Dep. Variable:	fully_vac_rate	R-squared:			0.662	
	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	38.0041	1.857	20.463	0.000	34.357	41.652
race[T.Asian]	-13.4168	1.880	-7.137	0.000	-17.109	-9.725
race[T.Black]	-10.1647	1.652	-6.154	0.000	-13.409	-6.921
race[T.Hispanic]	-9.5051	1.615	-5.886	0.000	-12.677	-6.334
race[T.White]	-18.6380	2.007	-9.284	0.000	-22.581	-14.695

Table 1: Ordinary Least Squares (OLS) Regression Analysis of Race and Vaccination Rate

The analysis reveals that all racial groups exhibit a statistically significant influence on the rate of full vaccinations. Relative to the baseline (AN/AI group), individuals in the four other racial categories demonstrate lower vaccination rates.

5. Limitations

Given the datasets we have, we may have the following limitations. 1. Correlation Does Not Imply Causation: The relationships observed between vaccination rates and various socio-demographic indicators do not necessarily imply causation. While these correlations are insightful, they should be interpreted with caution, as other unmeasured variables could also influence the outcomes. 2. Complexity of Public Health Dynamics: Vaccination uptake is influenced by a myriad of factors, including public health policies, individual beliefs, and community engagement. Our dataset captures only a fraction of these elements, potentially omitting crucial variables that play a significant role in vaccination campaigns. 3. Challenges in Measuring Community Initiatives: Without specific data from grassroots organizations, it's challenging to quantify their contribution to vaccination efforts. This gap hinders our ability to fully understand and acknowledge the role of community-based initiatives in public health outcomes.

Recommendations for Further Analysis. Investigate Other Sources of Data: To compensate for the lack of grassroots data, explore alternative sources that may provide indirect insights into the activities of these organizations, such as local news reports, community surveys, or public health bulletins. Conduct Qualitative Research: Engage in qualitative studies, including interviews and focus groups with community members and local health workers, to gather anecdotal evidence of grassroots organizations' impact on vaccination rates. Compare with Similar Cities: Analyze vaccination rates in cities with similar socio-demographic profiles but without significant grassroots activities. This comparison could offer indirect insights into the potential impact of such organizations.

6. Summary

This report on vaccine equity in MA emphasizes the critical public health challenge of vaccine hesitancy and the necessity for nuanced strategies. Building upon previous analysis, this project delves into a broader scope of variables, including age, ethnicity, gender, economic status, and the Social Vulnerability Index (SVI), to understand their correlation with vaccine hesitancy.

The preliminary analysis investigates vaccination rates among different demographics and cities, revealing significant disparities influenced by factors like socioeconomic status, education, and race. However, the lack of specific data on grassroots organizations' impact and the complexity of public health dynamics. We suggest further analysis, including exploring alternative data sources, conducting regression analysis, and expanding to other cities. In summary, the report underscores the importance of community engagement, comprehensive data collection, and collaboration among stakeholders.