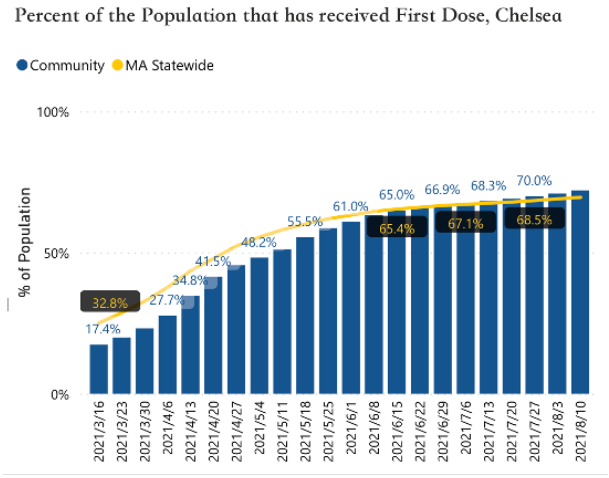
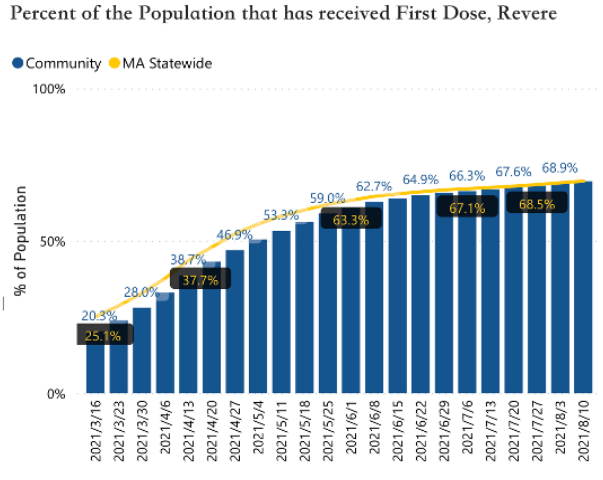
Deliverable 4 - Vaccine Equity Team 2 Report

Introduction

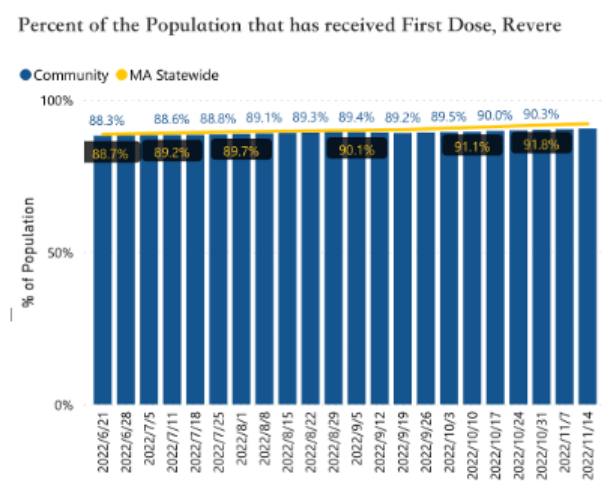
During the pandemic of COVID-19, vaccination became the most effective way to prevent people getting infected. In Massachusetts, La Colaborativa, a grassroots organization, made a great effort to help to vaccinate Latine and Black communities in Chelsea. With the help of volunteers from La Colaborativa, Chelsea became one of the communities which have the highest vaccination rates. As we can see statistics and graphs from the website of the Massachusetts government. The rates of receiving the first dose increased sharply from 17.4% which is obviously lower than the average percentage of MA Statewide, to 70% in about 4 months (As we can see from picture 1-a). Moreover, it only took three months for Chelsea to catch up with the average percentage. In comparison, Revere had a higher rate than Chelsea at the beginning, which is about 20% (As we can see from picture 1-b). However, Revere still did not reach the average rates till now, while the rate of Chelsea is almost 100% (As we can see from picture 1-c and 1-d). In this article, we would illustrate how La Colaborativa helped the Chelsea community detailly by putting emphasis on the age of Vaccination rates that includes both one dose and fully vaccinated.



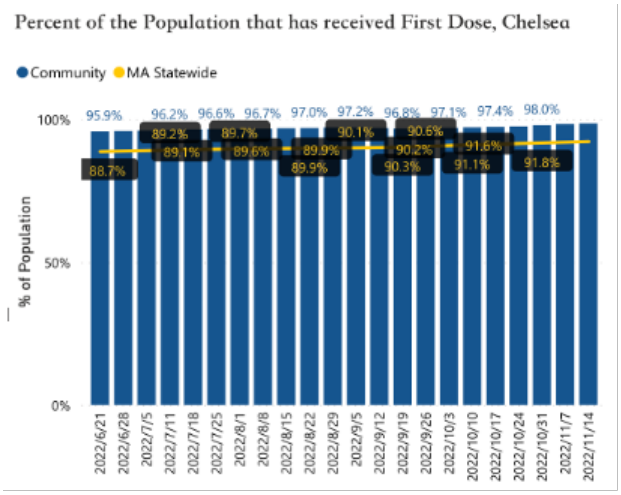
1-a



1-b



1-c



1-d

Algorithm and formula:

Indeed, for those tree circumstances, since we need to verify whether La Colaborativa has significant influence on vaccination, the null hypothesis and alternative hypothesis will be the same:
H0: La Colaborativa doesn't have a significant influence on (adult/pediatric/booster) vaccination in Chelsea. (Which means $\mu_C = \mu_{othercity}$)
Ha: La Colaborativa has a significant influence on (adult/pediatric/booster) vaccination in Chelsea. (Which means $\mu_C > \mu_{othercity}$)
So here we need to run a one-tailed hypothesis t-test, and the output, which is p-value, once is less than 0.05, then we can believe that it is credible to reject the null hypothesis and accept our alternative hypothesis–the grassroot organizations have significant influence on vaccination. The reason that p-value should be less than 0.05 for us to reject the null hypothesis is that the exceed in mean may be caused by accident or coincidence, and our test works to prevent our result from accidents. If p-value is greater than 0.05, it means that the possibility that our conclusion is caused by accidents is greater than 5%, which violates the rigorous rule of statistics to think the evidence is credible enough to reject the null hypothesis.

To run the hypothesis t-test, we need to compute the mean, standard deviation of each group of data, and also a t table to get p-value. The formation we use is below.

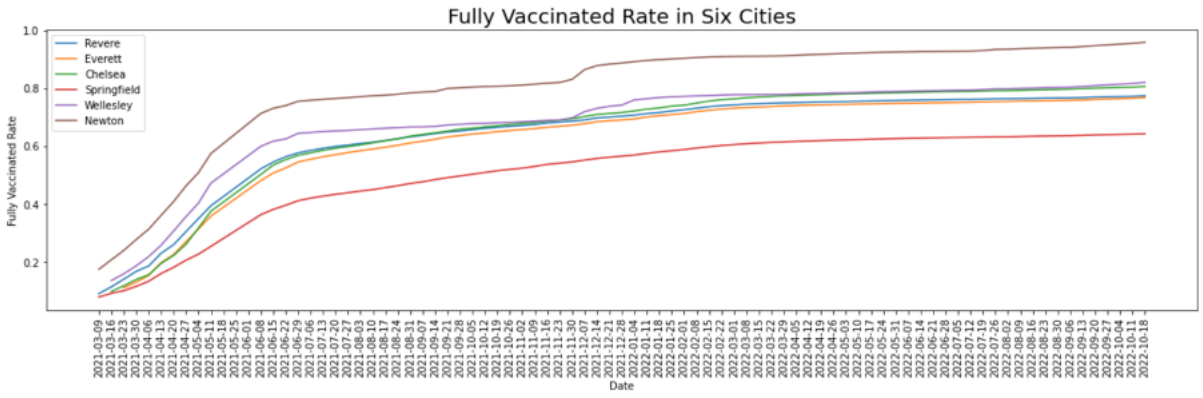
$$t = \frac{\overline{x_1} + \overline{x_2}}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where s means standard deviation, x-bar means mean and n means population.

By the t value we compute from the formula, we can use a t table to compute the p-value with 95% confidence interval. There is also an easier way to compute the p-value–by applying the stats package in python, we can use a function to compute p-value with two cities' data.

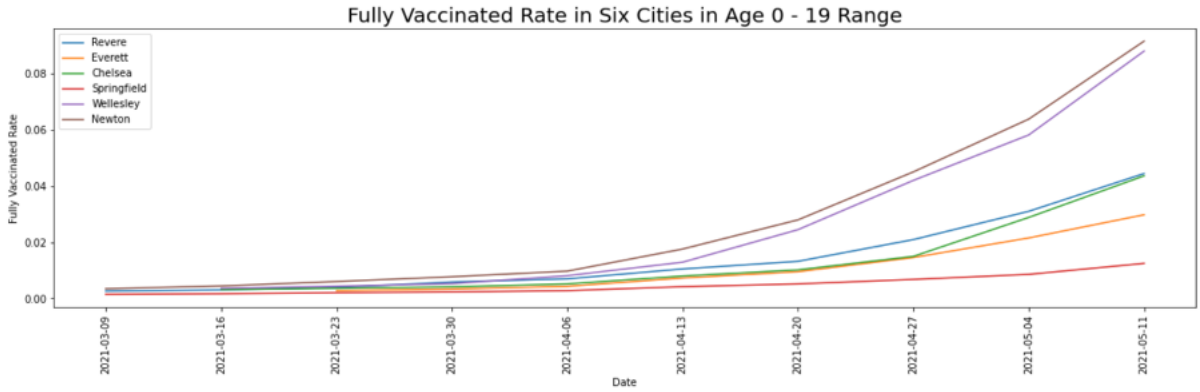
Analysis of the Fully Vaccinated Rate at Different Ages

We graphed the following fully vaccinated rate graphs for six age ranges (age 0-19, age 20-29, age 30-49, age 50-65, age 60-75, age 75+) and each line in the graphs corresponds to a city. The first graph includes the vaccination rate in all age ranges. The purpose is to figure out the characteristics and discrepancies of vaccination rate between different cities in each age range.



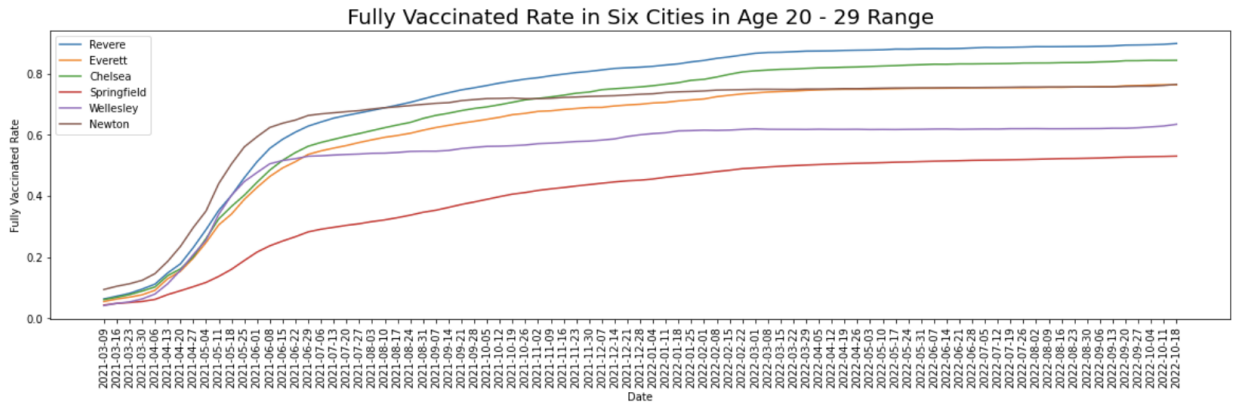
2-a

Plot 2-a presents the fully vaccinated proportion in six cities (Revere, Chelsea, Springfield, Everett, Newton, and Wellesley). Based on this plot, June 2021 is a turning point, the date from which the growth rate of fully vaccinated starts to slow down, especially for Revere, Everett, Chelsea, Wellesley, and Newton. Overall, Newton had the highest fully vaccinated rate and Springfield had the lowest fully vaccinated rate. Chelsea's fully vaccinated growth rate was slightly higher than Everett, Wellesley, and Revere's.



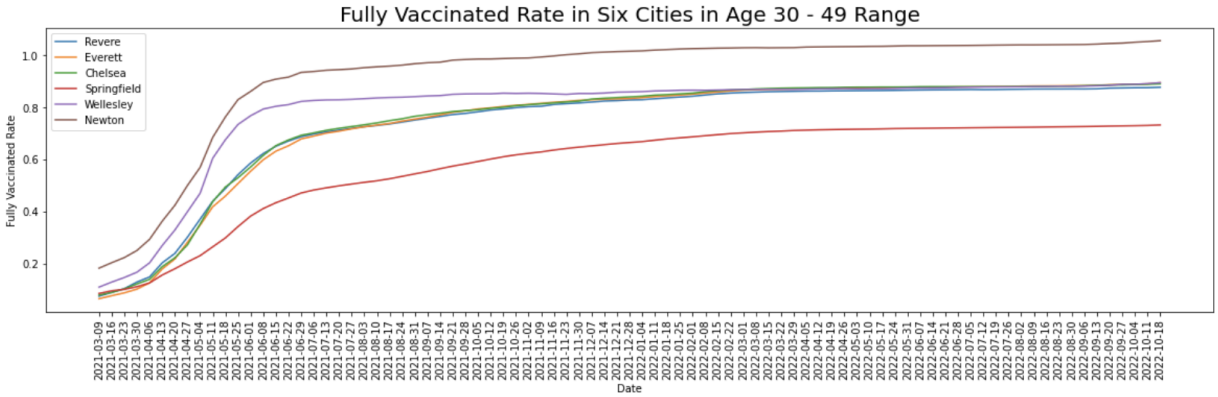
2-b

Plot 2-b shows the change in the proportion of fully vaccinated children and adolescents (aged 0-19) in six cities. From this graph, we can see that around April 10, 2021, the proportion of people aged 0-19 in Wellesley and Newton who are fully vaccinated suddenly increased. Around April 25, 2021, Revere, Chelsea and Everett had a significant increase in the proportion of fully vaccinated people aged 0-19. But in general, the proportion of people aged 0-19 who were fully vaccinated by May 2021 is less than 1%, which is very small.



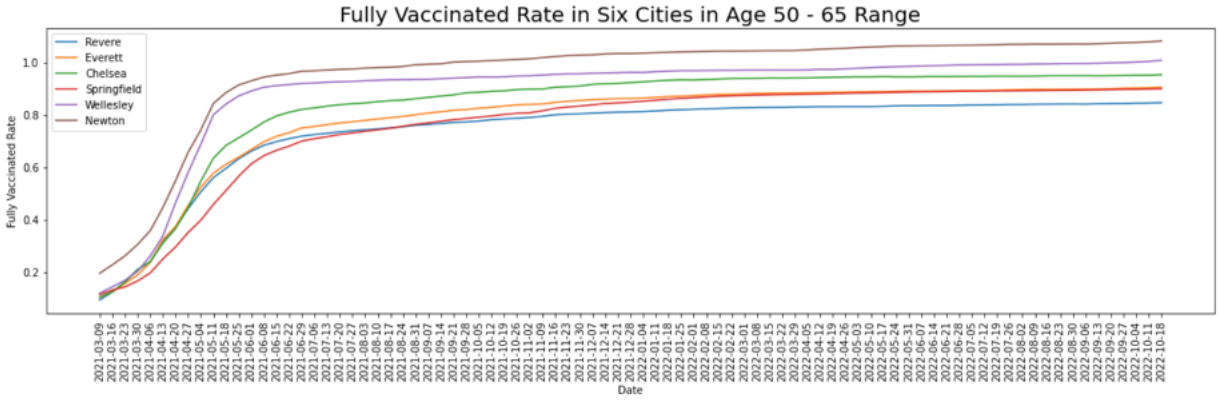
2-c

Plot 2-c presents the change in the proportion of fully vaccinated adults aged 20-29 in six cities. During April 2021 to June 2021, except for Springfield, the fully vaccinated proportion in other five cities has increased significantly. Percentage of fully vaccinated people in Newton and Wellesley remained basically the same after that, and did not continue to increase. After June 2021, the proportion of fully vaccinated in Chelsea, and Revere had a significant increase. They became cities with the two highest fully vaccinated rates.



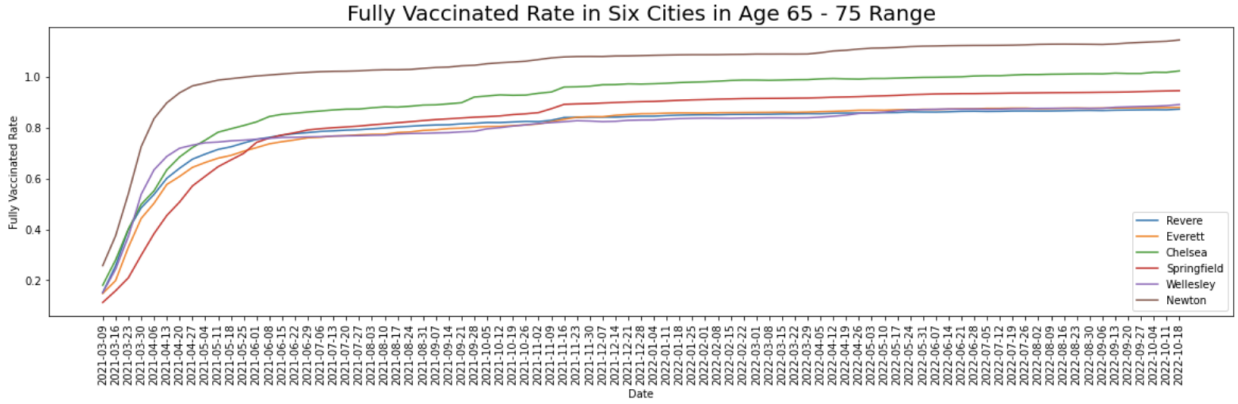
2-d

Plot 2-d presents the change in the proportion of fully vaccinated adults aged 30-49 in six cities. During April 2021 to June 2021, the fully vaccinated proportion has increased significantly. After that, the increase rate in all cities did not have a significant increase. Chelsea, Everett, Wellesley and Revere had the same fully vaccinated rate which was 80%.



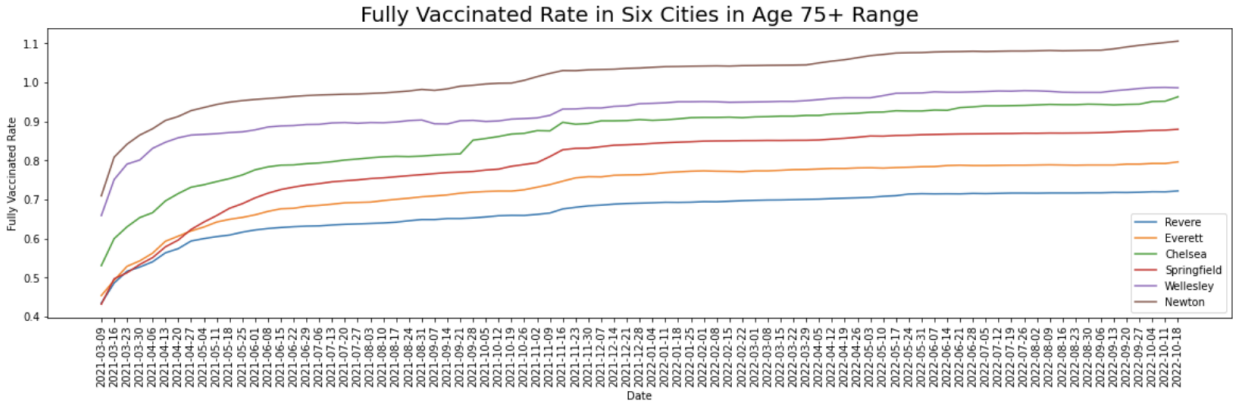
2-e

From plot 2-e, we can see that from April 2021 to May 2021, the proportion of fully vaccinated in these six cities has increased significantly. Springfield has a higher growth rate than the other five cities after June 2021.



2-f

Plot 2-f shows the change in the proportion of fully vaccinated people aged 65-75 in six cities. From March 2021 to May 2021, the proportion of fully vaccinated in all cities has increased significantly. Compared with other cities, Chelsea had a higher growth rate and had a relatively high vaccination rate after April 2021.

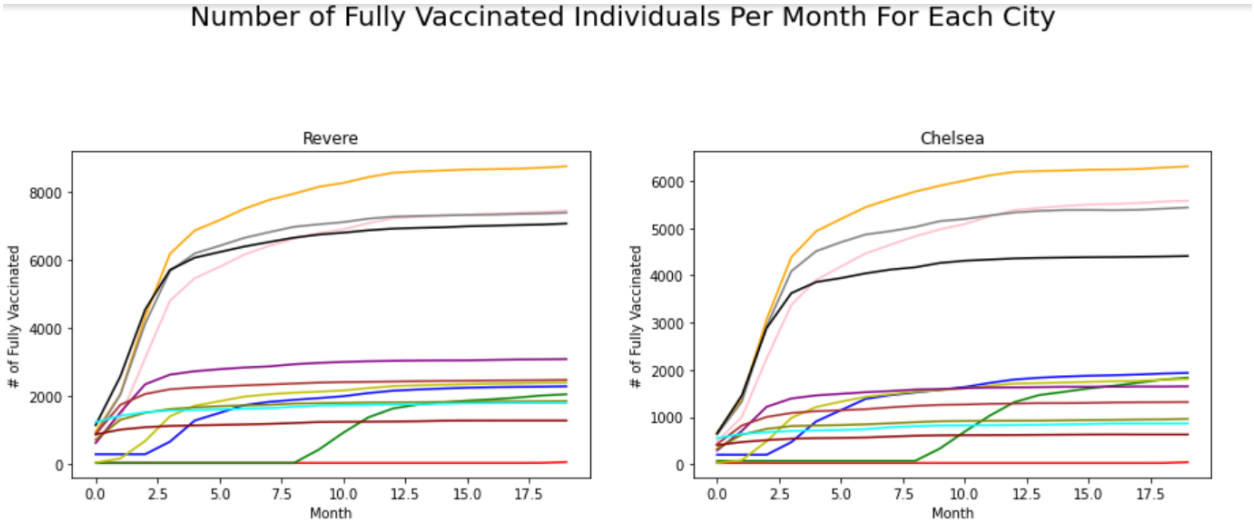


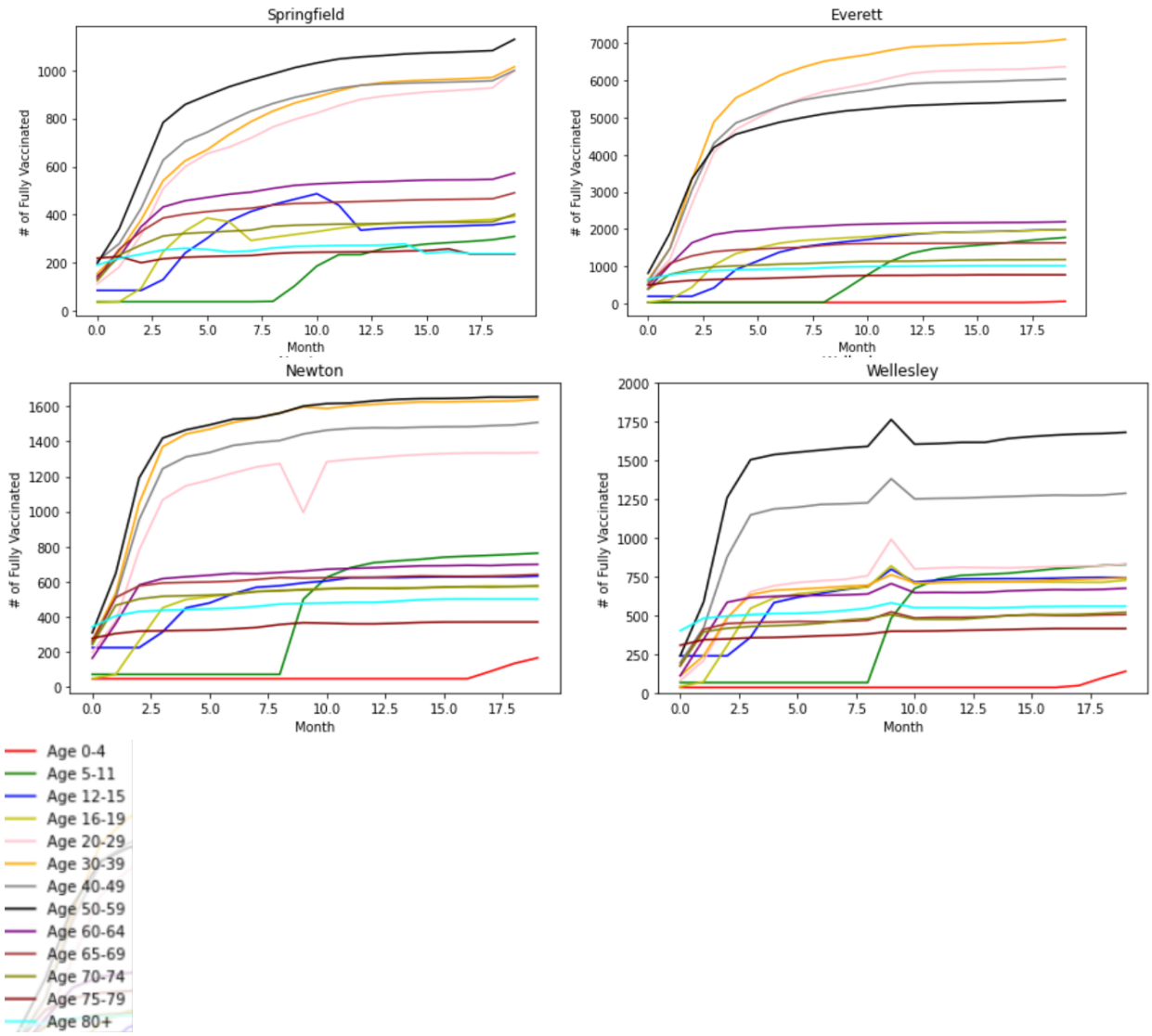
2-g

Plot 2-g shows the change in the proportion of fully vaccinated people aged 75+ in six cities. From March 2021 to April 2021, the proportion of fully vaccinated in all cities has increased significantly. Springfield and Chelsea had higher growth rates than the other four cities after April 2021. Compared with other cities, Chelsea had a relatively high vaccination rate.

By separately analyzing the growth and changes in the proportion of complete vaccinations in the six age groups, we can clearly find that people in the 65-75 and 75+ age groups received a large number of vaccinations from March 2021 to April 2021, and people in the 50-65 age group received a large number of vaccinations from April 2021 to May 2021. In the May and June of 2021, a large number of people in the 20-29 and 30-49 age groups were vaccinated in mid-June 2021. Across the age groups, Newton had the greatest percentage of complete vaccinations among the six cities, and Springfield had the lowest percentage of complete vaccinations. However, the vaccination rate of 50-75+ older people in Springfield is the highest among the six cities. Chelsea and Revere have similar fully vaccinated rates and growth rates, but in the 65-75+ age group, Chelsea have higher rates of complete vaccination and growth rates than Revere. Based on the average vaccination rate of each age group, Chelsea had a relatively high vaccination rate in the 20-29, 65-75 and 75+ age groups.

Analysis on Fully-vaccinated people at different ages in different cities





We can look into each city for different ages. Focus on 0-4 ages, we can find that for Revere, Chelsea, and Everett, none of 0-4 age people had vaccinated, and Springfield's data were lacking–we have no data about 0-4 ages' data. However, in Newton and Wellesley, after 16 months, about 200 people at that age had fully vaccinated when the data ended. Based on the data we have collected, small children (0-4 years old) are not paid attention to for vaccination, and that may be caused by the shortage of vaccination at the beginning of the pandemic, or the reason may be the beginning vaccination are not tested that can be taken by infants or small children.

Focus on age 5-11, after 8 months, children at that age started to be vaccinated. This situation may be caused by the vaccines not allowed to be taken on children at that age.

For people at the age of 12-15, all cities had a small number of children who were fully vaccinated at the beginning, and after two months, the number started to increase. However,

there may be outliers to make the data in Springfield abnormal–over months 10 to 12, the number of fully vaccinated children decreases about 150. Since other cities' data does not have this problem, this may be caused by children's death or fake data or other reasons.

For people at the age of 16-19, similar to that of 12-15, increased after the first month for fully vaccinating. Moreover, a similar problem also occurs. From 5 to 7 months, numbers of children also decreased, and this may also be caused by children's death.

For other ages, data is similar and all of them increase all the time (the first five months grows faster). However, something goes wrong in Wellesley and Newton. In Newton, there is an abnormal concave (in month 9) in the graph of 20-29, and this may be caused by wrong data since except the concave, the rest of the data is normal and month 8&10's data is similar to each other. Similarly, in Wellesley, for all graphs about age over 20, there are convexes in month 9, and in month 8 and 10 are similar. That may also be caused by wrong data.

Conclusion

Based on the graphs provided, we can draw many conclusions as to how different age groups were affected by vaccine rate. We can see that in general much younger age groups (i.e 0-4 years old) tended to have much lower rates of vaccination when compared to other much older age groups. This stark disproportionality can be attributed to less data provided for lower age groups as compared to older age groups. This result also correlates with common perception that those of older age groups are the only ones vulnerable to disease and thus the only ones who require a vaccination, when in fact the data does not show much the corresponding discrimination that would suggest that conclusion. Currently, our model consists of basic statistical inferences as to fill in missing data with values that match the general positive proportional relationship between age group and vaccination rate. As such, additional considerations such as normalization, implementation of additional machine learning algorithms, such as clustering or neural networks, will be heavily considered to eliminate bias and tighten the data so that more accurate conclusions may be discerned.