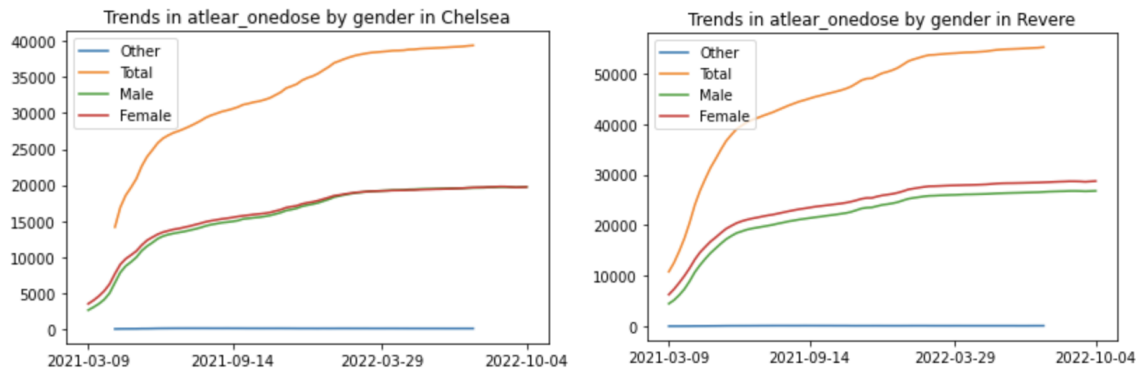


We grouped all data by gender, age, and race for Chelsea and Revere. We have done some analysis for these batches of data. For example, we can clearly see the trends in at least one dose by gender in both cities. Number of at least one dose in both cities has a great increase before March 29, 2022.



The data presented to us has significant gaps in the locations of interests, Revere and Chelsea. Specifically, for several locations (zip codes), there is no data available for vaccination status. One key limitation of our data is that a significant portion of individuals did not specify their race when they received their vaccinations. Thus we will be forced to remove this potentially useful data as it may contribute to inaccurate results. Sources which can provide us data on the vaccination status of individuals in Black and Latino communities, provide us with little data. This can pose risks in not only achieving our project goal but also ensuring that our data is representative of its respective populations. For our project goal, we are hoping to shed light on the significance of vaccine inequity among Black and Latino communities by revealing the significantly lower numbers of individuals in Black and Latino communities being vaccinated as compared to other more affluent groups. Based on preliminary analysis of the data, the numbers reveal significantly lower numbers of individuals in Black and Latino groups being vaccinated as compared to other groups. However, the fact that the numbers are insignificant when compared to their respective total populations may be due to unreported data. Thus more data, especially data on unreported vaccination, may prove to be beneficial in bolstering the claims of not only our analysis, but future analysis on this subject.

By simply analyzing the percentage of least-one-dose and fully vaccinated on population, it is easy to find out that Chelsea's percentage is much higher than Revere's, both at least-one and fully (from the graph below, we can see that almost 20 percent higher for one dose, and 10 percent for fully vaccinated). It is also important to analyze different situations between gender, age groups, and races. As there is a lack of second booster's data (since most people haven't received second booster), we may use models like KNN to train former data like first booster, fully vaccinated and partially vaccinated percentage to predict the data of second booster for different races, ages, and genders.

