Ron was in charge of collecting the data. We started with data from Kaggle, but it was found out that it is not working. While Our project requires to use label data, the hashtags in Kaggle’s data were not significantly towards positive or negative directions. That is why we decided to scrape our one data, with hashtags we can label. In order to choose the right hashtags, we took the existing data set and viewed hashtags from the kaggle set. Then, we picked ones that were being used in relation to the vaccine that were clearly pro or against. We used **tweepy** for the scraping, together with Twitter developer account credentials. We faced scraping limitations made by Twitter, which made the process harder. Another obstacle we faced was that we can scrape data from tweepy only from the previous week.