After analyzing the trends of the variables among the vaccinated and unvaccinated, we designed a model that scores individuals by likeliness to be vaccinated according to the following variables:

* Race
* Sex
* Education level
* Median household income
* Trend in cost of prescription drugs related to vaccines
* Mining-dependence
* Language spoken at home
* Adherence to prescription drugs
* Region
* Segregation level
* Per-capita income
* Poverty rate
* Insurance status
* Age

Then, we divided these variables into quantitative variables and qualitative variables. For quantitative variables, we calculated the median of each variable in both vaccinated groups and non-vaccinated groups in the training file. Then we set score standards based on the medians. The higher the score, the higher probability of being vaccinated. For example, the median of census median household income in vaccinated groups is $65,000, while the median of census median household income in non-vaccinated groups is $60,000. For people whose census median household income is lower than $60,000, we gave them a score of 5. For people whose census median household income is between $60,000 and $65,000, we gave them a score of 10. And for people whose census median household income is above $65,000, we gave them a score of 15. This is because we determined that if a person reports a higher census median household income, he or she has more likelihood to get vaccinated. Other quantitative variables are evaluated and analyzed in a similar way. Different variables have different score standards since we believe that the association between each variable and vaccination status varies differently.

For qualitative variables, we divided each variable into different categories, or groups of people, and used categorical analysis to analyze the relationship between each group and group members’ vaccination status. Then we used a scoring system similar to the one we used for the quantitative variable analysis, to set score standards and to give different groups a different score.

After finishing the above steps, we scored each individual based on their quantitative and qualitative determinants. Then we ranked these individuals based on the total score they earned. A high rank represents a high score, which also shows the high likelihood of a person taking the vaccine. Correspondingly, people with low scores and low ranks are hesitant to take the vaccine.