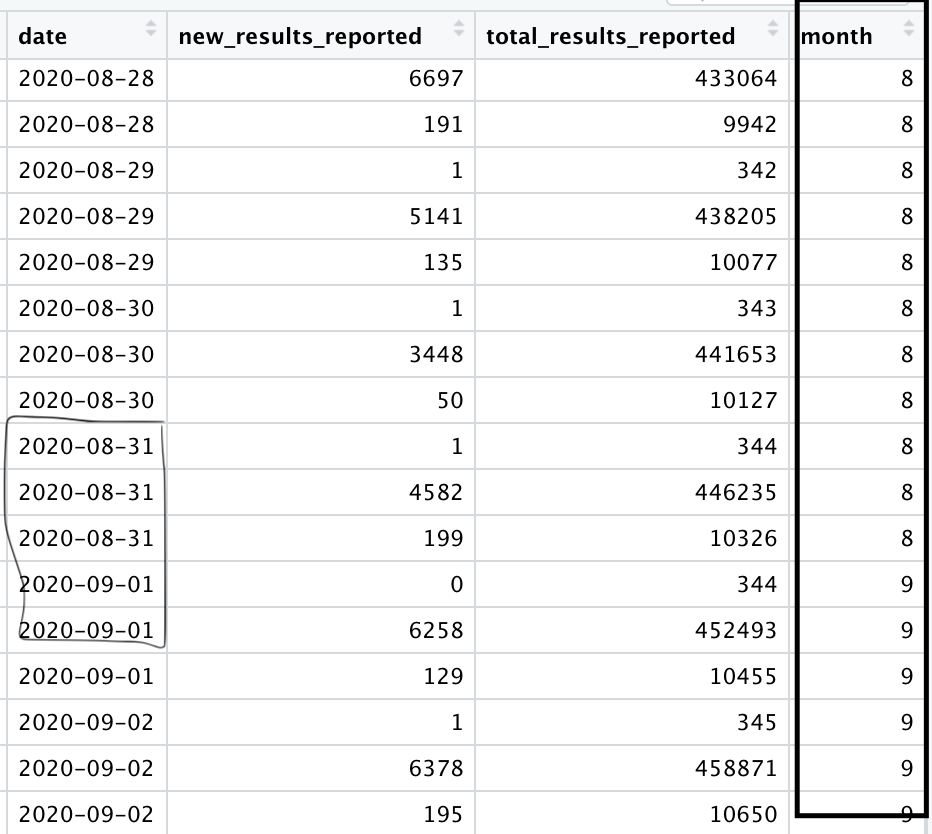
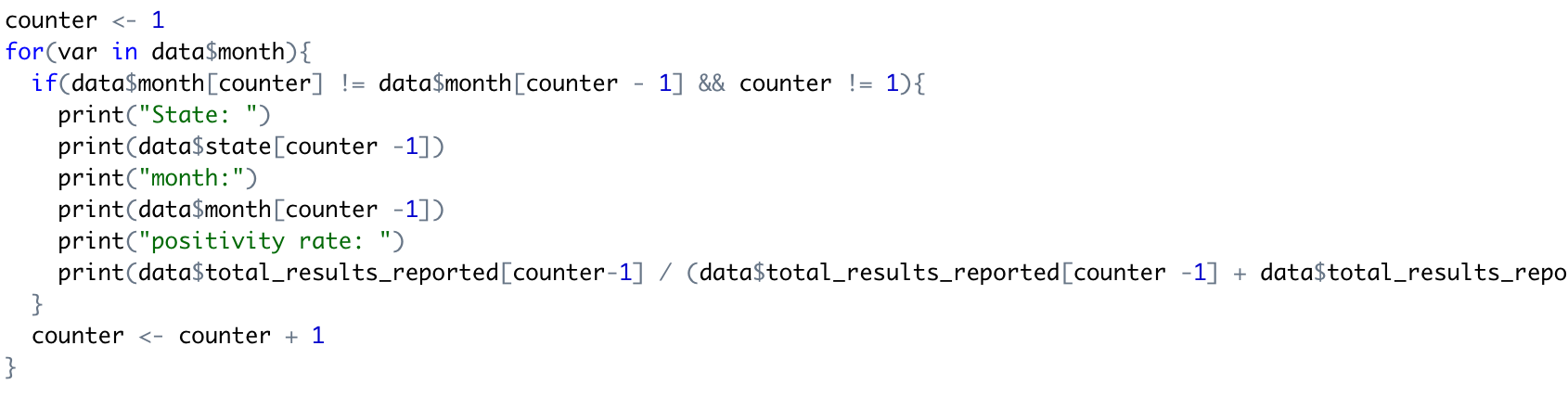
Nikhil Gopal

For this problem set, I unfortunately was not able to correctly compute the desired data. I decided to attack the problem by generating a column for the specific months (not including the full date) within R:

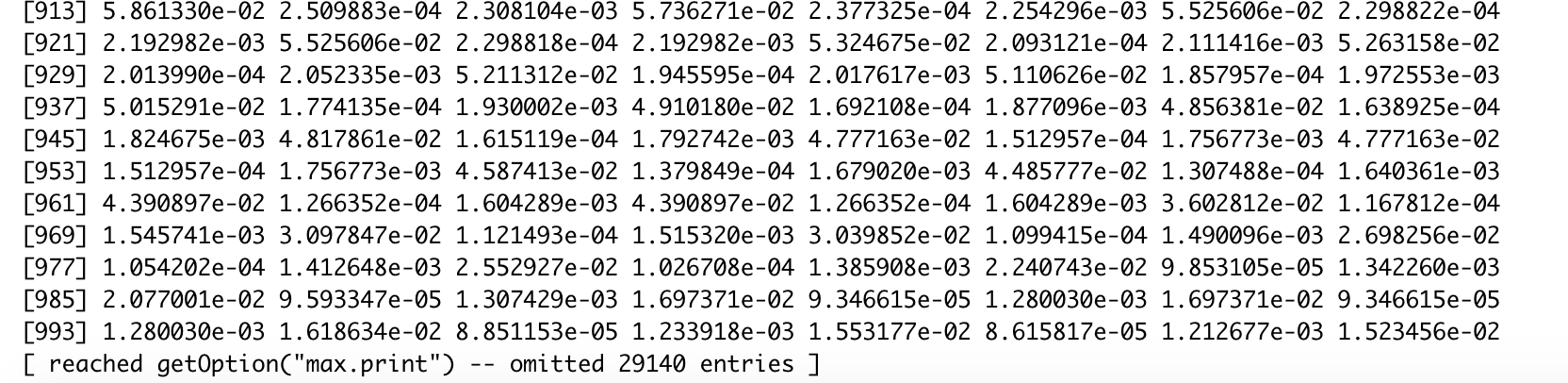


Covid case positivity rates can be calculated by dividing the total positive cases by the sum of the positive, negative and inconclusive results. My plan was to iterate through the dataset’s date column, and then detect when the month changes, and then use indexing to calculate Covid positivity rates by capturing the three indexes before the date change, and using those total case # values to calculate the Covid positivity rate for a given state in a given month:



The values that got cut off are just data$total\_results\_reported[counter-2] and data$total\_results\_reported[counter-3], which represent the negative and inconclusive values respectively.

Puzzlingly, when I ran my code, the last print statement would end up printing multiple values, when it should just be printing one double which is the COVID positivity rate:



When I ask the console to print out those values with specific indexes specified, it prints out one value only as intended, which is how I intended the program to run. Ideally, the program should give me a printed list of each state, the month and the positivity rate:

> data$total\_results\_reported[32]/(data$total\_results\_reported[32] + data$total\_results\_reported[31] + data$total\_results\_reported[30])

[1] 0.9769628

I cannot seem to figure out why I am getting multiple values for the covid positivity rate when it should be just one. I have made an honest effort to figure out the problem and unfortunately cannot. I plan on discussing this with the TA in office hours, since I know that I am close to the solution, and there must be some small technical error keeping my program from running as I intended.