

Individual Assignment 1

Simran Singh

10/2/2021

Task 0: Loading the Data

The following code loads the in 3 pre-filtered vectors for cases, county names, and vaccination rates. The vectors are defined as follows: *case_count* - Provided by NC department of public health, shows the rate of COVID infections per 100k residents by county *vaccination_rate* - Provided by the CDC, shows the % of residents that are fully vaccinated by county *nc_county_list* - Shows the county name for each entry in the case count and vaccination vectors.

For example, the first entry in *nc_county_list* is 'Transylvania' meaning the first value of the *case_count* and *vaccination_rate* vectors correspond to case count and vaccination rate in Transylvania County.

```
# Load the case_count, vaccination_rate, and county vectors
load("C:/Users/Simran/Downloads/NC_Covid_Data_Sep-4-2021.RData")
```

Task 1

```
##Get average cases?
nc_average_vax = mean(case_count)
nc_average_vax
```

```
## [1] 508.07
```

```
#average cases is 508.07
##which county has the highest vaccination rate
highest_rate = which.max(vaccination_rate)
highest_vacc_rate = nc_county_list[highest_rate]
highest_vacc_rate
```

```
## [1] "Martin"
```

```
#The county with the highest vaccination rate is Martin
```

TASK 2 # Filter function # ? Found this on Stack overflow... No idea how it works!

```
filter_func = function(input_vector, county_names, filt_counties){
  filt = county_names %in% filt_counties
  output_vector = input_vector[filt]
  return(output_vector)
}
```

```
#We define a variable named filter_func. This variable takes 3 arguments: input_vector, county_names
#, filt_counties. These 3 arguments are vectors that are being created by the function. This function
#checks to see if each element in county_names is present in filt_counties. The '%in%' portion returns
#a true or false depending on if the county matches in filt_counties and county_names. Then the
#elements listed as true are put into the output_vector.
```

```
durham_counties = c("Wake", "Granville", "Person", "Orange", "Chatham", "Durham")
```

```
# ? what are the sizes of the resulting vectors? Are they the same size?
```

```
filtered_cases = filter_func(input_vector = case_count,
                             county_names = nc_county_list,
                             filt_counties = durham_counties)
```

```
filtered_vax = filter_func(input_vector = vaccination_rate,
                           county_names = nc_county_list,
                           filt_counties = durham_counties)
```

```
filtered_cases
```

```
## [1] 296 309 393 295 374 274
```

```
filtered_vax
```

```
## [1] 44.3 46.8 56.6 47.3 58.0 62.5
```

```
#resulting vectors are both size [1:6], so they are the same size.
```

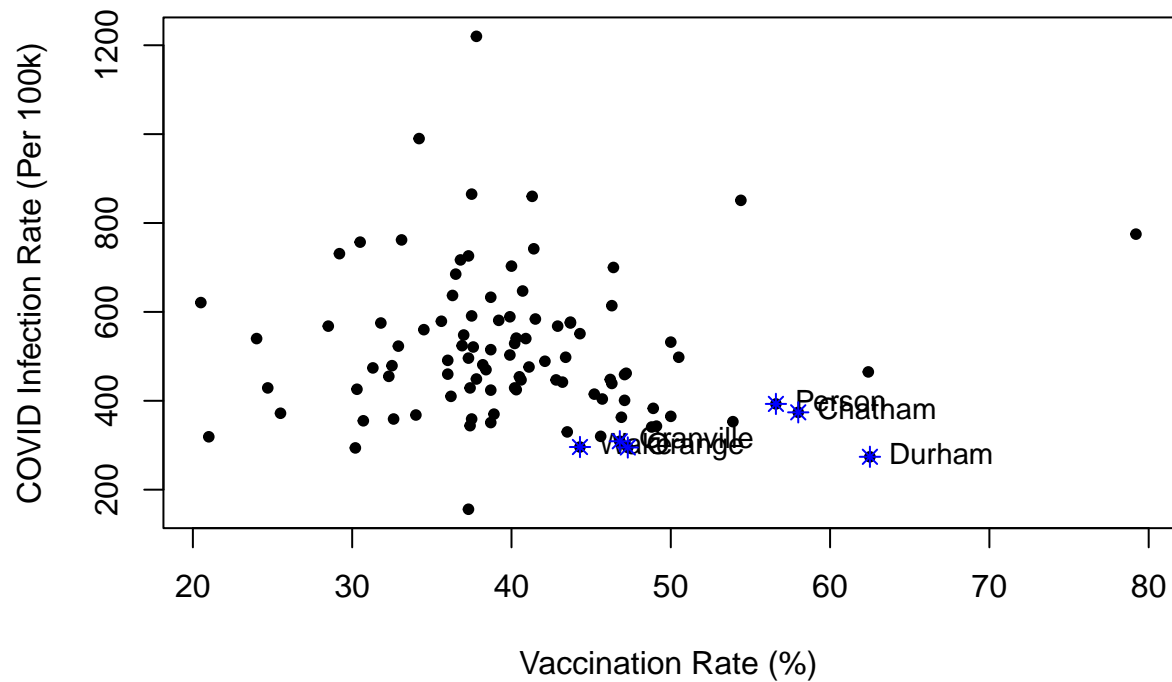
Task 3

```
# Plots
```

```
plot(x = vaccination_rate,
     y = case_count,
     pch = 20,
     main = "Vaccination Rates vs COVID Infection Rates",
     xlab = 'Vaccination Rate (%)',
     ylab = 'COVID Infection Rate (Per 100k)'
     )
```

```
for (i in 1:length(durham_counties)){
  points(x = filtered_vax[i], y = filtered_cases[i], pch = 8, col = 'blue')
  text(x = filtered_vax[i], y = filtered_cases[i], labels = durham_counties[i],
       cex=.9, font=1.5, pos=4)
}
```

Vaccination Rates vs COVID Infection Rates



? what does this mean?

*#the first line of code generates a scatter plot to visualize the data. This scatter plot compares
#Covid Infection rate with vaccination rate. The for loop creates a loop of sequences that goes
#from values 1 to thelength of durham_counties. This loop then creates labeled points in the scatter
#plot for all the counties in Durham. The x and y coordinates for the points are found by using the
#data specific for the counties, where x = filtered_vax and y = filtered_cases*