COVID-19 Vaccination Rates

Duc Nguyen

Getting Started

Import vaccination data

```
vax <- read.csv("covid19vaccinesbyzipcode_test.csv")

head(vax)
```

```
  as_of_date zip_code_tabulation_area local_health_jurisdiction         county
1 2021-01-05                   93562            San Bernardino  San Bernardino
2 2021-01-05                   93437             Santa Barbara   Santa Barbara
3 2021-01-05                   93445          San Luis Obispo San Luis Obispo
4 2021-01-05                   93442          San Luis Obispo San Luis Obispo
5 2021-01-05                   93444          San Luis Obispo San Luis Obispo
6 2021-01-05                   93453          San Luis Obispo San Luis Obispo
  vaccine_equity_metric_quartile                vem_source
1                              1 Healthy Places Index Score
2                             NA            No VEM Assigned
3                              2 Healthy Places Index Score
4                              3 Healthy Places Index Score
5                              3 Healthy Places Index Score
6                              3 Healthy Places Index Score
  age12_plus_population age5_plus_population tot_population
1                1469.5                 1668           1771
2                2494.5                 2871           3387
3                6116.7                 6762           7106
4               10005.2                10615          10917
5               18951.8                20522          21331
6                2373.6                 2499           2578
  persons_fully_vaccinated persons_partially_vaccinated
```

```
1                        NA                             NA
2                        NA                             NA
3                        NA                             NA
4                        NA                             NA
5                        NA                             NA
6                        NA                             NA
  percent_of_population_fully_vaccinated
1                                     NA
2                                     NA
3                                     NA
4                                     NA
5                                     NA
6                                     NA
  percent_of_population_partially_vaccinated
1                                         NA
2                                         NA
3                                         NA
4                                         NA
5                                         NA
6                                         NA
  percent_of_population_with_1_plus_dose booster_recip_count
1                                     NA                  NA
2                                     NA                  NA
3                                     NA                  NA
4                                     NA                  NA
5                                     NA                  NA
6                                     NA                  NA
  bivalent_dose_recip_count eligible_recipient_count
1                        NA                        0
2                        NA                        1
3                        NA                        0
4                        NA                        1
5                        NA                        0
6                        NA                        0
                                                                redacted
1 Information redacted in accordance with CA state privacy requirements
2 Information redacted in accordance with CA state privacy requirements
3 Information redacted in accordance with CA state privacy requirements
4 Information redacted in accordance with CA state privacy requirements
5 Information redacted in accordance with CA state privacy requirements
6 Information redacted in accordance with CA state privacy requirements
```

Q1. What column details the total number of people fully vaccinated?

```
names(vax)
```

```
 [1] "as_of_date"
 [2] "zip_code_tabulation_area"
 [3] "local_health_jurisdiction"
 [4] "county"
 [5] "vaccine_equity_metric_quartile"
 [6] "vem_source"
 [7] "age12_plus_population"
 [8] "age5_plus_population"
 [9] "tot_population"
[10] "persons_fully_vaccinated"
[11] "persons_partially_vaccinated"
[12] "percent_of_population_fully_vaccinated"
[13] "percent_of_population_partially_vaccinated"
[14] "percent_of_population_with_1_plus_dose"
[15] "booster_recip_count"
[16] "bivalent_dose_recip_count"
[17] "eligible_recipient_count"
[18] "redacted"
```

A1: persons_fully_vaccinated

Q2. What column details the Zip code tabulation area?

A2: zip_code_tabulation_area

Q3. What is the earliest date in this dataset?

```
head(vax$as_of_date)
```

```
[1] "2021-01-05" "2021-01-05" "2021-01-05" "2021-01-05" "2021-01-05"
[6] "2021-01-05"
```

A3: 2021-01-05

Q4. What is the latest date in this dataset?

```
tail(vax$as_of_date)
```

```
[1] "2022-11-15" "2022-11-15" "2022-11-15" "2022-11-15" "2022-11-15"
[6] "2022-11-15"
```

A4: 2022-11-15

Let's call the 'skim()' function from the skimr package to get a quick overview of this dataset

```
skimr::skim(vax)
```

Table 1: Data summary

Name	vax
Number of rows	172872
Number of columns	18
_____	
Column type frequency:	
character	5
numeric	13
_____	
Group variables	None

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
as_of_date	0	1	10	10	0	98	0
local_health_jurisdiction	0	1	0	15	490	62	0
county	0	1	0	15	490	59	0
vem_source	0	1	15	26	0	3	0
redacted	0	1	2	69	0	2	0

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
zip_code_tabulation_area	0	1.00	93665.11	1817.39	90001	92257.75	93658.50	95380.50	97635.0	
vaccine_equity_metric_quartile	8526	0.95	2.44	1.11	1	1.00	2.00	3.00	4.0	
age12_plus_population	0	1.00	18895.04	18993.88	0	1346.95	13685.13	31756.12	88556.7	
age5_plus_population	0	1.00	20875.22	21105.98	0	1460.50	15364.00	34877.00	101902.0	
tot_population	8428	0.95	23372.72	22628.51	12	2126.00	18714.00	38168.00	111165.0	
persons_fully_vaccinated	15440	0.91	13309.15	14740.07	11	859.00	7687.00	22253.00	87305.0	
persons_partially_vaccinated	15440	0.91	1679.13	1993.86	11	157.00	1158.00	2483.00	39201.0	
percent_of_population_fully_vaccinated	18986	0.89	0.54	0.26	0	0.36	0.58	0.73	1.0	
percent_of_population_partially_vaccinated	18986	0.89	0.08	0.09	0	0.05	0.06	0.08	1.0	
percent_of_population_with_1_plus_dose	19822	0.89	0.60	0.26	0	0.42	0.64	0.79	1.0	

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
booster_recip_count	70642	0.59	5701.06	6972.68	11	276.00	2546.00	9513.00	58301.0	
bivalent_dose_recip_count	156937	0.09	1512.94	1994.71	11	101.00	662.00	2236.00	16790.0	
eligible_recipient_count	0	1.00	12114.80	14551.97	0	438.00	5520.00	20714.00	86817.0	

Q5. How many numeric columns are in this dataset?

A5: 13

Q6. Note that there are "missing values" in the dataset. How many NA values there in the persons_fully_vaccinated column?

```
na_persons_fully_vaccinated <- sum(is.na(vax$persons_fully_vaccinated))
na_persons_fully_vaccinated
```

```
[1] 15440
```

A6: 15440

Q7. What percent of persons_fully_vaccinated values are missing (to 2 significant figures)?

```
round((na_persons_fully_vaccinated/nrow(vax))*100, 2)
```

```
[1] 8.93
```

A7: 8.93%

Q8. [Optional]: Why might this data be missing?

A8: I think that this data be missing might due to some people did not get the vaccine.

Working with dates

Install the "lubridate" package in R Console by using the 'install.packages("lubridate")' command.

```
library(lubridate)
```

```
Warning: package 'lubridate' was built under R version 4.2.2
```

```
Warning: package 'timechange' was built under R version 4.2.2
```

What is today's date

```
today()
```

```
[1] "2022-11-22"
```

Convert our date data in the as_of_date column into a lubridate format (year-month-day)

```
vax$as_of_date <- ymd(vax$as_of_date)
```

How many days have passed since the first vaccination reported in this dataset?

```
today()
```

```
[1] "2022-11-22"
```

```
vax$as_of_date[1]
```

```
[1] "2021-01-05"
```

```
today() - vax$as_of_date[1]
```

```
Time difference of 686 days
```

Using the last and the first date value we can now determine how many days the dataset span?

```
vax$as_of_date[nrow(vax)]
```

```
[1] "2022-11-15"
```

```
  vax$as_of_date[1]
```

```
[1] "2021-01-05"
```

```
  vax$as_of_date[nrow(vax)] - vax$as_of_date[1]
```

```
Time difference of 679 days
```

Q9. How many days have passed since the last update of the dataset?

```
  today()
```

```
[1] "2022-11-22"
```

```
  vax$as_of_date[nrow(vax)]
```

```
[1] "2022-11-15"
```

```
  today() - vax$as_of_date[nrow(vax)]
```

```
Time difference of 7 days
```

A9: 7 days (as of today is 2022-11-22)

Q10. How many unique dates are in the dataset (i.e. how many different dates are detailed)?

```
  length(unique(vax$as_of_date))
```

```
[1] 98
```

A10: 98 unique dates

Working with ZIP codes

Install the "zipcodeR" package in R Console to work with postal ZIP codes by using the 'install.packages("zipcodeR")' command.

```
library(zipcodeR)
```

```
geocode_zip('92037')
```

```
# A tibble: 1 x 3
  zipcode   lat   lng
  <chr>   <dbl> <dbl>
1 92037    32.8 -117.
```

Calculate the distance between the centroids of any two ZIP codes in miles, e.g.

```
zip_distance('92037','92109')
```

```
  zipcode_a zipcode_b distance
1     92037     92109     2.33
```

Pull census data about ZIP code areas

```
reverse_zipcode(c('92037', "92109"))
```

```
# A tibble: 2 x 24
  zipcode zipcode_~1 major~2 post_~3 common_c~4 county state   lat   lng timez~5
  <chr>   <chr>      <chr>   <chr>       <blob> <chr>  <chr> <dbl> <dbl> <chr>
1 92037   Standard   La Jol~ La Jol~ <raw 20 B> San D~ CA     32.8 -117. Pacific
2 92109   Standard   San Di~ San Di~ <raw 21 B> San D~ CA     32.8 -117. Pacific
# ... with 14 more variables: radius_in_miles <dbl>, area_code_list <blob>,
#   population <int>, population_density <dbl>, land_area_in_sqmi <dbl>,
#   water_area_in_sqmi <dbl>, housing_units <int>,
#   occupied_housing_units <int>, median_home_value <int>,
#   median_household_income <int>, bounds_west <dbl>, bounds_east <dbl>,
#   bounds_north <dbl>, bounds_south <dbl>, and abbreviated variable names
#   1: zipcode_type, 2: major_city, 3: post_office_city, ...
```

Pull data for all ZIP codes in the dataset

```
zipdata <- reverse_zipcode(vax$zip_code_tabulation_area)
```

Take a look at the zipdata

```
head(zipdata)
```

```
# A tibble: 6 x 24
  zipcode zipcode_~1 major~2 post_~3 common_c~4 county state   lat   lng timez~5
  <chr>   <chr>      <chr>   <chr>       <blob> <chr>  <chr> <dbl> <dbl> <chr>
1 90001   Standard   Los An~ Los An~ <raw 44 B> Los A~ CA     34.0 -118. Pacific
2 90002   Standard   Los An~ Los An~ <raw 47 B> Los A~ CA     34.0 -118. Pacific
3 90003   Standard   Los An~ Los An~ <raw 23 B> Los A~ CA     34.0 -118. Pacific
4 90004   Standard   Los An~ Los An~ <raw 34 B> Los A~ CA     34.1 -118. Pacific
5 90005   Standard   Los An~ Los An~ <raw 34 B> Los A~ CA     34.1 -118. Pacific
6 90006   Standard   Los An~ Los An~ <raw 23 B> Los A~ CA     34.0 -118. Pacific
# ... with 14 more variables: radius_in_miles <dbl>, area_code_list <blob>,
#   population <int>, population_density <dbl>, land_area_in_sqmi <dbl>,
#   water_area_in_sqmi <dbl>, housing_units <int>,
#   occupied_housing_units <int>, median_home_value <int>,
#   median_household_income <int>, bounds_west <dbl>, bounds_east <dbl>,
#   bounds_north <dbl>, bounds_south <dbl>, and abbreviated variable names
#   1: zipcode_type, 2: major_city, 3: post_office_city, ...
```

Focus on the San Diego area

Subset to San Diego county only areas

Install the "dplyr" package in R Console to work with postal ZIP codes by using the 'install.packages("dplyr")' command.

```
library(dplyr)
```

```
sd <- filter(vax, county == "San Diego")
```

```
nrow(sd)
```

```
[1] 10486
```

Take a look at the sd

```
head(sd)
```

```
  as_of_date zip_code_tabulation_area local_health_jurisdiction    county
1 2021-01-05                   92083                San Diego San Diego
2 2021-01-05                   92066                San Diego San Diego
3 2021-01-05                   92091                San Diego San Diego
4 2021-01-05                   92103                San Diego San Diego
5 2021-01-05                   92113                San Diego San Diego
6 2021-01-05                   92116                San Diego San Diego
  vaccine_equity_metric_quartile                vem_source
1                              2 Healthy Places Index Score
2                              1    CDPH-Derived ZCTA Score
3                              4    CDPH-Derived ZCTA Score
4                              4 Healthy Places Index Score
5                              1 Healthy Places Index Score
6                              3 Healthy Places Index Score
  age12_plus_population age5_plus_population tot_population
1               32246.5                36283          39509
2                 589.5                  685            693
3                1238.3                 1303           1313
4               32146.4                33213          34700
5               47799.7                53883          58408
6               30255.7                31673          33408
  persons_fully_vaccinated persons_partially_vaccinated
1                       11                          462
2                       NA                           NA
3                       NA                           NA
4                       37                         1386
5                       11                          381
6                       29                          910
  percent_of_population_fully_vaccinated
1                               0.000278
2                                     NA
3                                     NA
4                               0.001066
5                               0.000188
6                               0.000868
  percent_of_population_partially_vaccinated
1                                   0.011694
2                                         NA
3                                         NA
4                                   0.039942
5                                   0.006523
6                                   0.027239
  percent_of_population_with_1_plus_dose booster_recip_count
```

```
1                                0.011972                  NA
2                                      NA                  NA
3                                      NA                  NA
4                                0.041008                  NA
5                                0.006711                  NA
6                                0.028107                  NA
  bivalent_dose_recip_count eligible_recipient_count
1                        NA                       11
2                        NA                        0
3                        NA                        0
4                        NA                       37
5                        NA                       11
6                        NA                       29
                                                                                redacted
1 Information redacted in accordance with CA state privacy requirements
2 Information redacted in accordance with CA state privacy requirements
3 Information redacted in accordance with CA state privacy requirements
4 Information redacted in accordance with CA state privacy requirements
5 Information redacted in accordance with CA state privacy requirements
6 Information redacted in accordance with CA state privacy requirements
```

Subset of all San Diego county areas with a population of over 10,000.

```
sd.10 <- filter(vax, county == "San Diego" &
                  age5_plus_population > 10000)
```

Take a look at the sd.10

```
head(sd.10)
```

```
  as_of_date zip_code_tabulation_area local_health_jurisdiction    county
1 2021-01-05                   92083                 San Diego San Diego
2 2021-01-05                   92103                 San Diego San Diego
3 2021-01-05                   92113                 San Diego San Diego
4 2021-01-05                   92116                 San Diego San Diego
5 2021-01-05                   92118                 San Diego San Diego
6 2021-01-05                   92075                 San Diego San Diego
  vaccine_equity_metric_quartile                vem_source
1                              2 Healthy Places Index Score
2                              4 Healthy Places Index Score
3                              1 Healthy Places Index Score
```

```
4                            3 Healthy Places Index Score
5                            3 Healthy Places Index Score
6                            4 Healthy Places Index Score
  age12_plus_population age5_plus_population tot_population
1               32246.5                36283          39509
2               32146.4                33213          34700
3               47799.7                53883          58408
4               30255.7                31673          33408
5               19835.0                21470          22548
6               11136.3                12177          12752
  persons_fully_vaccinated persons_partially_vaccinated
1                       11                          462
2                       37                         1386
3                       11                          381
4                       29                          910
5                       NA                           NA
6                       NA                           NA
  percent_of_population_fully_vaccinated
1                               0.000278
2                               0.001066
3                               0.000188
4                               0.000868
5                                     NA
6                                     NA
  percent_of_population_partially_vaccinated
1                                   0.011694
2                                   0.039942
3                                   0.006523
4                                   0.027239
5                                         NA
6                                         NA
  percent_of_population_with_1_plus_dose booster_recip_count
1                               0.011972                  NA
2                               0.041008                  NA
3                               0.006711                  NA
4                               0.028107                  NA
5                                     NA                  NA
6                                     NA                  NA
  bivalent_dose_recip_count eligible_recipient_count
1                        NA                       11
2                        NA                       37
3                        NA                       11
4                        NA                       29
```

```
5                        NA                        10
6                        NA                         8
                                                                     redacted
1 Information redacted in accordance with CA state privacy requirements
2 Information redacted in accordance with CA state privacy requirements
3 Information redacted in accordance with CA state privacy requirements
4 Information redacted in accordance with CA state privacy requirements
5 Information redacted in accordance with CA state privacy requirements
6 Information redacted in accordance with CA state privacy requirements
```

Q11. How many distinct zip codes are listed for San Diego County?

```r
length(unique(sd$zip_code_tabulation_area))
```

```
[1] 107
```

A11: 107 distinct zip codes

Q12. What San Diego County Zip code area has the largest 12 + Population in this dataset?

```r
max_12_pop <- which.max(sd$age12_plus_population)
sd[max_12_pop, "zip_code_tabulation_area"]
```

```
[1] 92154
```

A12: 92154

Q13. What is the overall average "Percent of Population Fully Vaccinated" value for all San Diego "County" as of "2022-11-15"?

```r
date_2022.11.15 <- filter(vax, county == "San Diego" &
                            as_of_date == "2022-11-15")

head(date_2022.11.15)
```

```
  as_of_date zip_code_tabulation_area local_health_jurisdiction    county
1 2022-11-15                    92127                 San Diego San Diego
2 2022-11-15                    92111                 San Diego San Diego
3 2022-11-15                    92122                 San Diego San Diego
4 2022-11-15                    92129                 San Diego San Diego
```

```
5 2022-11-15                    92119                San Diego San Diego
6 2022-11-15                    92024                San Diego San Diego
  vaccine_equity_metric_quartile                vem_source
1                              4 Healthy Places Index Score
2                              3 Healthy Places Index Score
3                              4 Healthy Places Index Score
4                              4 Healthy Places Index Score
5                              4 Healthy Places Index Score
6                              4 Healthy Places Index Score
  age12_plus_population age5_plus_population tot_population
1               38942.3                46080          49935
2               44075.0                48160          50693
3               44091.1                45951          48071
4               46449.1                51493          54762
5               21444.8                23472          24831
6               44405.4                48477          51381
  persons_fully_vaccinated persons_partially_vaccinated
1                    41150                         3362
2                    34968                         3524
3                    37846                         5167
4                    43573                         3337
5                    18505                         1362
6                    39175                         3487
  percent_of_population_fully_vaccinated
1                               0.824071
2                               0.689799
3                               0.787294
4                               0.795679
5                               0.745238
6                               0.762441
  percent_of_population_partially_vaccinated
1                                   0.067328
2                                   0.069517
3                                   0.107487
4                                   0.060936
5                                   0.054851
6                                   0.067866
  percent_of_population_with_1_plus_dose booster_recip_count
1                               0.891399               27742
2                               0.759316               21623
3                               0.894781               26519
4                               0.856615               29989
5                               0.800089               12406
```

```
6                                0.830307                25755
  bivalent_dose_recip_count eligible_recipient_count redacted
1                      7809                    40694       No
2                      5715                    34693       No
3                      8646                    37523       No
4                      8783                    43106       No
5                      3950                    18279       No
6                      8722                    38922       No
```

```
overall_average <- function(x) {
x[is.na(x)] <- 0
mean(x)
}
```

```
round(overall_average(date_2022.11.15$percent_of_population_fully_vaccinated)*100, 2)
```

```
[1] 68.3
```

```
mean(date_2022.11.15$percent_of_population_fully_vaccinated, na.rm = TRUE)
```

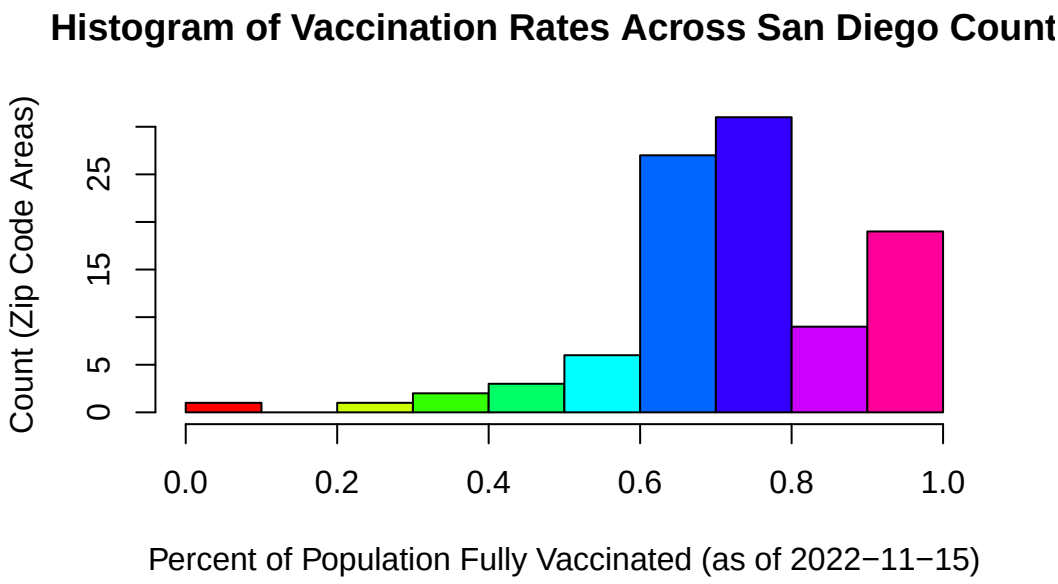
```
[1] 0.7381765
```

A13: 68.30%

Q14. Using either ggplot or base R graphics make a summary figure that shows the distribution of Percent of Population Fully Vaccinated values as of "2022-11-15"?

A14:

```
hist(date_2022.11.15$percent_of_population_fully_vaccinated,
     main = "Histogram of Vaccination Rates Across San Diego County",
     xlab = "Percent of Population Fully Vaccinated (as of 2022-11-15)",
     ylab = "Count (Zip Code Areas)",
     col = rainbow(10))
```



Focus on UCSD/La Jolla

UC San Diego resides in the 92037 ZIP code area and is listed with an age 5+ population size of 36,144.

```
ucsd <- filter(sd, zip_code_tabulation_area == "92037")
ucsd[1, ]$age5_plus_population
```

```
[1] 36144
```

Q15. Using ggplot make a graph of the vaccination rate time course for the 92037 ZIP code area

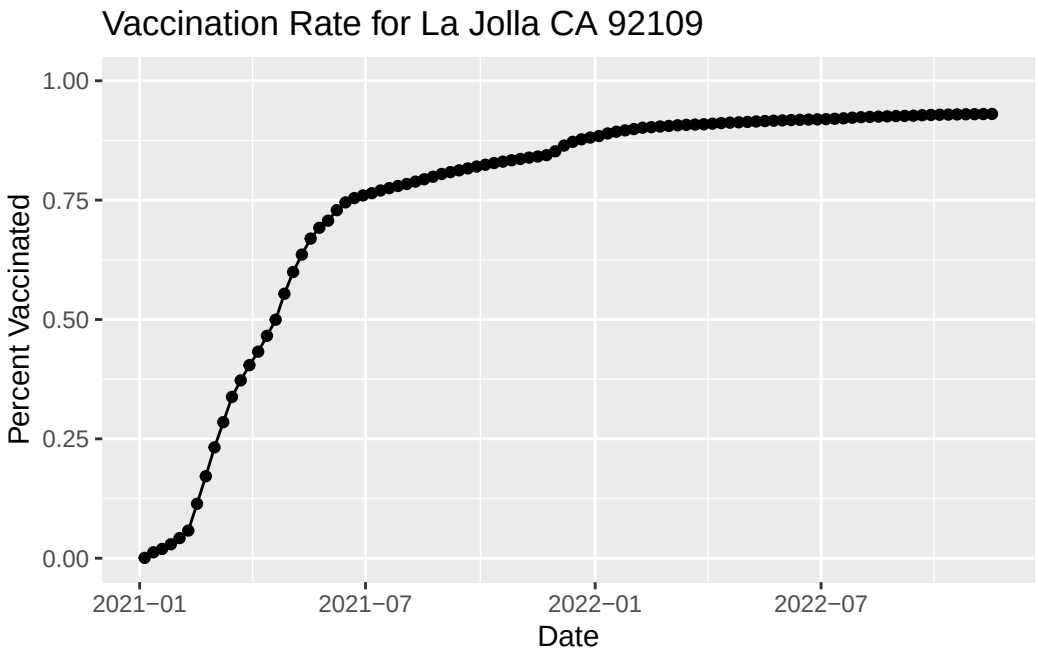
A15:

```
library(ggplot2)
```

```
ucsd_plot <- ggplot(ucsd) +
  aes(x = ucsd$as_of_date,
      y = ucsd$percent_of_population_fully_vaccinated) +
  geom_point() +
```

```
  geom_line(group = 1) +
  ylim(c(0,1)) +
  labs(title = "Vaccination Rate for La Jolla CA 92109",
       x = "Date", y = "Percent Vaccinated")

print(ucsd_plot)
```



Comparing to similar sized areas

Look across every zip code area with a population at least as large as that of the zip code 92037 (which has an age 5+ population size of 36,144) on as_of_date "2022-02-22".

```
# Subset to all CA areas with a population as large as 92037
vax.36 <- filter(vax, age5_plus_population > 36144 &
                   as_of_date == "2022-11-15")

head(vax.36)
```

```
  as_of_date zip_code_tabulation_area local_health_jurisdiction       county
1 2022-11-15                   92127                 San Diego    San Diego
2 2022-11-15                   92201                 Riverside    Riverside
```

```
3 2022-11-15                    92111                San Diego    San Diego
4 2022-11-15                    92122                San Diego    San Diego
5 2022-11-15                    92129                San Diego    San Diego
6 2022-11-15                    94561              Contra Costa Contra Costa
  vaccine_equity_metric_quartile                 vem_source
1                              4 Healthy Places Index Score
2                              1 Healthy Places Index Score
3                              3 Healthy Places Index Score
4                              4 Healthy Places Index Score
5                              4 Healthy Places Index Score
6                              3 Healthy Places Index Score
  age12_plus_population age5_plus_population tot_population
1               38942.3                46080          49935
2               55960.9                61733          65726
3               44075.0                48160          50693
4               44091.1                45951          48071
5               46449.1                51493          54762
6               34548.9                39272          42473
  persons_fully_vaccinated persons_partially_vaccinated
1                    41150                         3362
2                    44078                         7311
3                    34968                         3524
4                    37846                         5167
5                    43573                         3337
6                    32347                         1804
  percent_of_population_fully_vaccinated
1                               0.824071
2                               0.670633
3                               0.689799
4                               0.787294
5                               0.795679
6                               0.761590
  percent_of_population_partially_vaccinated
1                                   0.067328
2                                   0.111235
3                                   0.069517
4                                   0.107487
5                                   0.060936
6                                   0.042474
  percent_of_population_with_1_plus_dose booster_recip_count
1                               0.891399               27742
2                               0.781868               21043
3                               0.759316               21623
```

```
4                              0.894781               26519
5                              0.856615               29989
6                              0.804064               18909
  bivalent_dose_recip_count eligible_recipient_count redacted
1                      7809                    40694       No
2                      3965                    43955       No
3                      5715                    34693       No
4                      8646                    37523       No
5                      8783                    43106       No
6                      4323                    32167       No
```

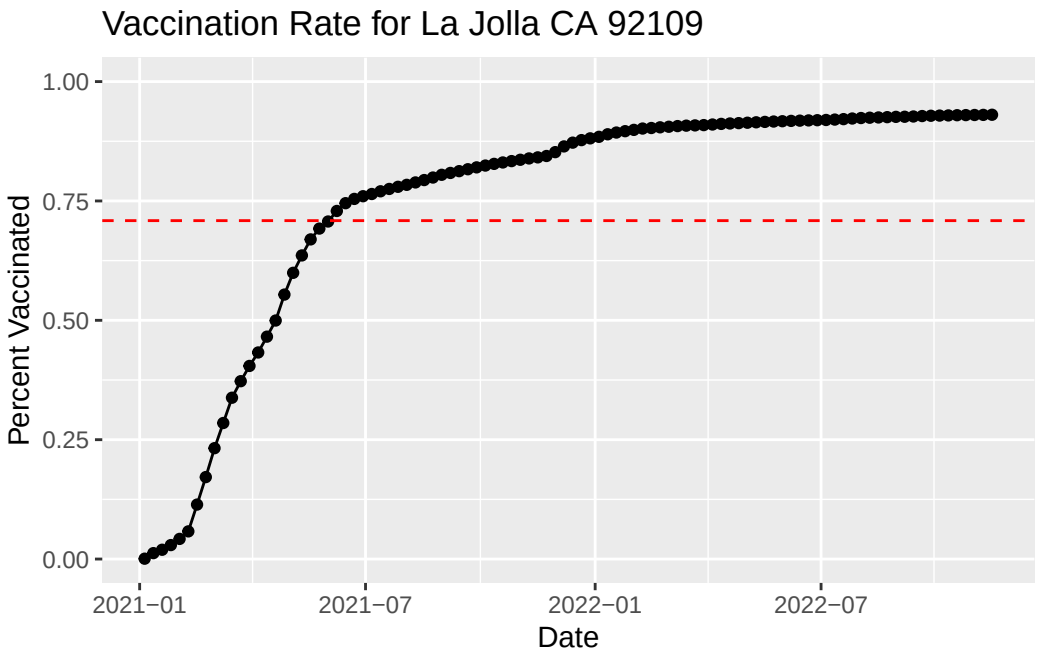
Q16. Calculate the mean "Percent of Population Fully Vaccinated" for ZIP code areas with a population as large as 92037 (La Jolla) as_of_date "2022-11-15". Add this as a straight horizontal line to your plot from above with the geom_hline() function?

A16:

```
mean_1 <- mean(vax.36$percent_of_population_fully_vaccinated)
mean_1
```

```
[1] 0.7088141
```

```
ucsd_plot +
  geom_hline(aes(yintercept = mean_1), color = "red", linetype = 2)
```



Q17. What is the 6 number summary (Min, 1st Qu., Median, Mean, 3rd Qu., and Max) of the "Percent of Population Fully Vaccinated" values for ZIP code areas with a population as large as 92037 (La Jolla) as_of_date "2022-11-15"?

A17:

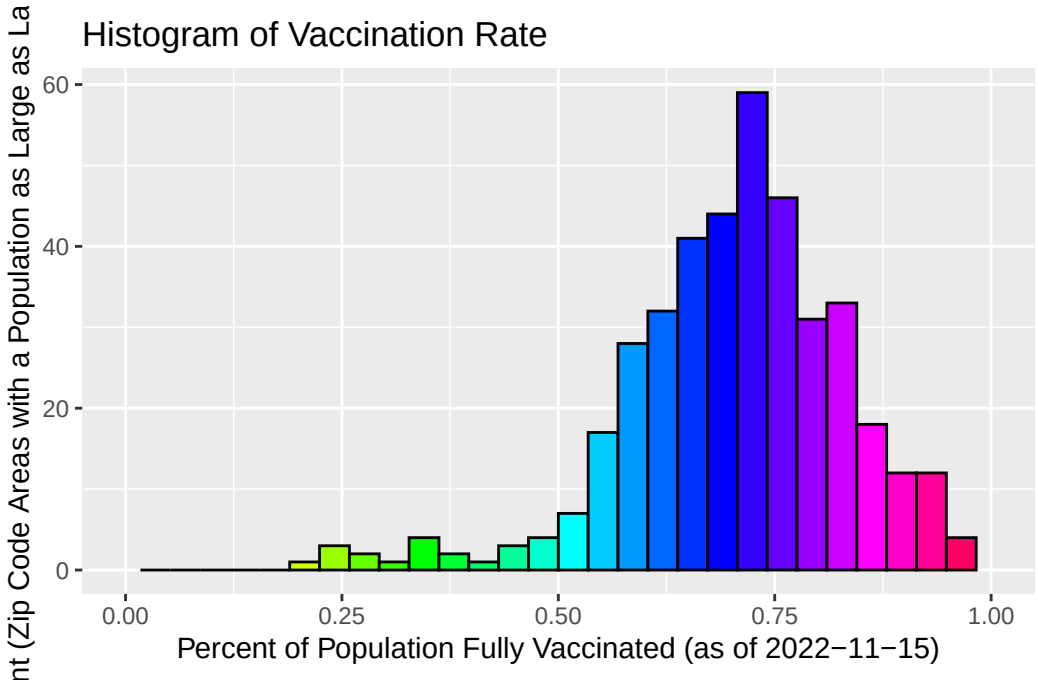
```
summary(vax.36$percent_of_population_fully_vaccinated)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
 0.1986  0.6338  0.7162  0.7088  0.7893  1.0000 
```

Q18. Using ggplot generate a histogram of this data.

A18:

```
ggplot(vax.36) +
  aes(percent_of_population_fully_vaccinated) +
  geom_histogram(bins = 30, color = "black", fill = rainbow(30)) +
  xlim(0.00, 1.00) +
  labs(title = "Histogram of Vaccination Rate",
       x = "Percent of Population Fully Vaccinated (as of 2022-11-15)",
       y = "Count (Zip Code Areas with a Population as Large as La Jolla)")
```



Q19. Is the 92109 and 92040 ZIP code areas above or below the average value you calculated for all these above?

The average value:

```
mean_1
```

```
[1] 0.7088141
```

For the 92109 zip code area:

```
vax %>% filter(as_of_date == "2022-11-15") %>%
  filter(zip_code_tabulation_area=="92109") %>%
  select(percent_of_population_fully_vaccinated)
```

```
  percent_of_population_fully_vaccinated
1                               0.695676
```

For the 92040 zip code area:

```
vax %>% filter(as_of_date == "2022-11-15") %>%
  filter(zip_code_tabulation_area=="92040") %>%
  select(percent_of_population_fully_vaccinated)
```

```
  percent_of_population_fully_vaccinated
1                               0.547251
```

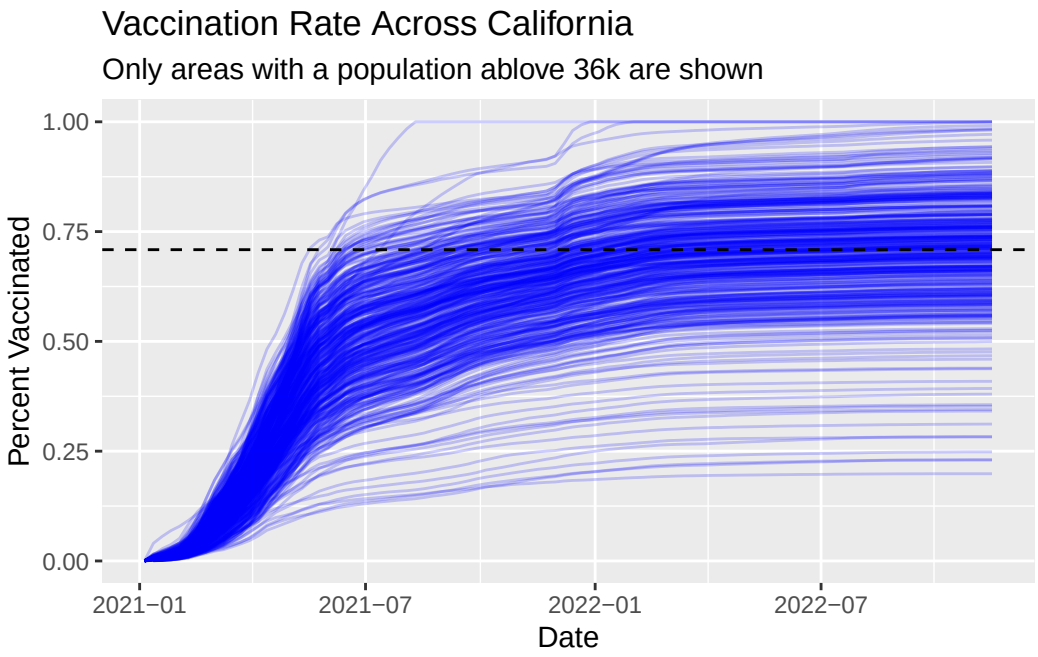
A19: The 92109 and 92040zip code areas is below the average value.

Q20. Finally make a time course plot of vaccination progress for all areas in the full dataset with a age5_plus_population > 36144.

A20:

```
vax.36.all <- filter(vax, age5_plus_population > 36144)
```

```
ggplot(vax.36.all) +
  aes(as_of_date,
      percent_of_population_fully_vaccinated,
      group = zip_code_tabulation_area) +
  geom_line(alpha = 0.2, color = "blue") +
  ylim(0.00, 1.00) +
  labs(x = "Date",
       y = "Percent Vaccinated",
       title = "Vaccination Rate Across California",
       subtitle = "Only areas with a population ablove 36k are shown") +
  geom_hline(yintercept = mean_1, linetype = 2)
```



Q21. How do you feel about traveling for Thanksgiving Break and meeting for in-person class afterwards?

A21: I feel enthusiastic about the upcoming travel trip with my wife and my little girl during Thanksgiving Break, and I also feel excited about the in-person class afterward to learn new things.