

# class17

Daniel Xu

```
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 |                          |          |

```
tail(vax)
```

|        | as_of_date | zip_code_tabulation_area | local_health_jurisdiction | county     |
|--------|------------|--------------------------|---------------------------|------------|
| 172867 | 2022-11-15 | 93720                    | Fresno                    | Fresno     |
| 172868 | 2022-11-15 | 95062                    | Santa Cruz                | Santa Cruz |

|        |  |                              |                            |        |
|--------|--|------------------------------|----------------------------|--------|
| 172869 | 2022-11-15                                 | 93711                        | Fresno                     | Fresno |
| 172870 | 2022-11-15                                 | 95713                        | Placer                     | Placer |
| 172871 | 2022-11-15                                 | 93673                        | Tulare                     | Tulare |
| 172872 | 2022-11-15                                 | 95746                        | Placer                     | Placer |
|        | vaccine_equity_metric_quartile             |                              | vem_source                 |        |
| 172867 |  | 3                            | Healthy Places Index Score |        |
| 172868 |  | 3                            | Healthy Places Index Score |        |
| 172869 |  | 3                            | Healthy Places Index Score |        |
| 172870 |  | 3                            | Healthy Places Index Score |        |
| 172871 |  | 1                            | CDPH-Derived ZCTA Score    |        |
| 172872 |  | 4                            | Healthy Places Index Score |        |
|        | age12_plus_population                      | age5_plus_population         | tot_population             |        |
| 172867 | 40357.3                                    | 44412                        | 47081                      |        |
| 172868 | 32943.4                                    | 35251                        | 36737                      |        |
| 172869 | 33662.9                                    | 36191                        | 38007                      |        |
| 172870 | 9386.5                                     | 10084                        | 10472                      |        |
| 172871 | 615.7                                      | 705                          | 754                        |        |
| 172872 | 20588.8                                    | 22923                        | 23934                      |        |
|        | persons_fully_vaccinated                   | persons_partially_vaccinated |                            |        |
| 172867 | 33480                                      | 2986                         |                            |        |
| 172868 | 28801                                      | 2161                         |                            |        |
| 172869 | 26086                                      | 2334                         |                            |        |
| 172870 | 4141                                       | 346                          |                            |        |
| 172871 | 443  | 54                           |                            |        |
| 172872 | 16979                                      | 1108                         |                            |        |
|        | percent_of_population_fully_vaccinated     |                              |                            |        |
| 172867 | 0.711115                                   |                              |                            |        |
| 172868 | 0.783978                                   |                              |                            |        |
| 172869 | 0.686347                                   |                              |                            |        |
| 172870 | 0.395435                                   |                              |                            |        |
| 172871 | 0.587533                                   |                              |                            |        |
| 172872 | 0.709409                                   |                              |                            |        |
|        | percent_of_population_partially_vaccinated |                              |                            |        |
| 172867 | 0.063423                                   |                              |                            |        |
| 172868 | 0.058824                                   |                              |                            |        |
| 172869 | 0.061410                                   |                              |                            |        |
| 172870 | 0.033040                                   |                              |                            |        |
| 172871 | 0.071618                                   |                              |                            |        |
| 172872 | 0.046294                                   |                              |                            |        |
|        | percent_of_population_with_1_plus_dose     | booster_recip_count          |                            |        |
| 172867 | 0.774538                                   | 20576                        |                            |        |
| 172868 | 0.842802                                   | 20075                        |                            |        |
| 172869 | 0.747757                                   | 15928                        |                            |        |

|                           |                          |          |
|---------------------------|--------------------------|----------|
| 172870                    | 0.428475                 | 2384     |
| 172871                    | 0.659151                 | 199      |
| 172872                    | 0.755703                 | 11492    |
| bivalent_dose_recip_count | eligible_recipient_count | redacted |
| 172867                    | 5265                     | 33361 No |
| 172868                    | 7234                     | 28554 No |
| 172869                    | 4569                     | 26006 No |
| 172870                    | 740                      | 4127 No  |
| 172871                    | 43                       | 442 No   |
| 172872                    | 3809                     | 16877 No |

Q1 persons\_fully\_vaccinated Q2) zip\_code\_tabulation\_area Q3)2021-01-05 Q4) 2022-11-15

```
library(skimr)
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 | mean     | sd      | p0   | p25       | p50      | p75      | p100    | hist |
|--|-----------|----------|----------|---------|------|-----------|----------|----------|---------|------|
| zip_code_tabulation_area                   | 0         | 1.00     | 93665.11 | 1817.39 | 0000 | 192257.75 | 3658.50  | 5380.50  | 7635.0  |      |
| vaccine_equity_metric_052tile              | 0         | 0.95     | 2.44     | 1.11    | 1    | 1.00      | 2.00     | 3.00     | 4.0     |      |
| age12_plus_population                      | 0         | 1.00     | 18895.01 | 8993.88 | 0    | 1346.95   | 13685.13 | 1756.18  | 8556.7  |      |
| age5_plus_population                       | 0         | 1.00     | 20875.24 | 1105.98 | 0    | 1460.50   | 15364.00 | 14877.00 | 1902.0  |      |
| tot_population                             | 8428      | 0.95     | 23372.72 | 2628.51 | 2    | 2126.00   | 18714.00 | 18168.00 | 11165.0 |      |
| persons_fully_vaccinated                   | 15440     | 0.91     | 13309.15 | 4740.07 | 1    | 859.00    | 7687.00  | 22253.00 | 87305.0 |      |
| persons_partially_vaccinated               | 15440     | 0.91     | 1679.13  | 1993.86 | 1    | 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_1_plus_dose          | 18986     | 0.89     | 0.60     | 0.26    | 0    | 0.42      | 0.64     | 0.79     | 1.0     |      |
| booster_recip_count                        | 70642     | 0.59     | 5701.06  | 972.68  | 1    | 276.00    | 2546.00  | 9513.00  | 58301.0 |      |
| bivalent_dose_recip_count                  | 156937    | 0.09     | 1512.94  | 1994.71 | 1    | 101.00    | 662.00   | 2236.00  | 16790.0 |      |
| eligible_recipient_count                   | 0         | 1.00     | 12114.80 | 4551.97 | 0    | 438.00    | 5520.00  | 20714.00 | 86817.0 |      |

Q5) 13 Q6) 15440 Q7)  $15440/172872 = 0.089 = 8.9\%$  Q8) information redacted by California privacy

```
library(lubridate)
```

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

Loading required package: timechange

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

Attaching package: 'lubridate'

The following objects are masked from 'package:base':

date, intersect, setdiff, union

```
today()
```

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

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

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

Time difference of 7 days

Q9) 7 days

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

```
[1] "2021-01-05" "2021-01-12" "2021-01-19" "2021-01-26" "2021-02-02"
[6] "2021-02-09" "2021-02-16" "2021-02-23" "2021-03-02" "2021-03-09"
[11] "2021-03-16" "2021-03-23" "2021-03-30" "2021-04-06" "2021-04-13"
[16] "2021-04-20" "2021-04-27" "2021-05-04" "2021-05-11" "2021-05-18"
[21] "2021-05-25" "2021-06-01" "2021-06-08" "2021-06-15" "2021-06-22"
[26] "2021-06-29" "2021-07-06" "2021-07-13" "2021-07-20" "2021-07-27"
[31] "2021-08-03" "2021-08-10" "2021-08-17" "2021-08-24" "2021-08-31"
[36] "2021-09-07" "2021-09-14" "2021-09-21" "2021-09-28" "2021-10-05"
[41] "2021-10-12" "2021-10-19" "2021-10-26" "2021-11-02" "2021-11-09"
[46] "2021-11-16" "2021-11-23" "2021-11-30" "2021-12-07" "2021-12-14"
[51] "2021-12-21" "2021-12-28" "2022-01-04" "2022-01-11" "2022-01-18"
[56] "2022-01-25" "2022-02-01" "2022-02-08" "2022-02-15" "2022-02-22"
[61] "2022-03-01" "2022-03-08" "2022-03-15" "2022-03-22" "2022-03-29"
[66] "2022-04-05" "2022-04-12" "2022-04-19" "2022-04-26" "2022-05-03"
[71] "2022-05-10" "2022-05-17" "2022-05-24" "2022-05-31" "2022-06-07"
[76] "2022-06-14" "2022-06-21" "2022-06-28" "2022-07-05" "2022-07-12"
[81] "2022-07-19" "2022-07-26" "2022-08-02" "2022-08-09" "2022-08-16"
[86] "2022-08-23" "2022-08-30" "2022-09-06" "2022-09-13" "2022-09-20"
[91] "2022-09-27" "2022-10-04" "2022-10-11" "2022-10-18" "2022-10-25"
[96] "2022-11-01" "2022-11-08" "2022-11-15"
```

Q10) 98

```
library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

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

```
nrow(sd)
```

```
[1] 10486
```

```
unique(sd$zip_code_tabulation_area)
```

```
[1] 92083 92066 92091 92103 92113 92116 92118 92060 92075 92084 92104 91911
[13] 91941 91963 91980 91977 92040 92054 92056 92129 92127 92004 92007 92037
[25] 92020 92025 92027 92065 91915 92003 92021 92059 92024 92028 92070 92071
[37] 92101 92110 92111 92119 92008 92036 92106 92108 92109 92121 92128 92105
[49] 92107 91901 91902 92067 92078 92117 92123 92139 92173 92081 92132 92134
[61] 92135 92140 92145 92147 92155 92058 92019 92010 92011 92055 91945 91917
[73] 91931 92069 92064 92086 92126 92130 91905 91910 92114 92102 91934 91914
[85] 92082 91916 92061 91932 92122 92154 92120 92115 92131 92009 92029 91935
[97] 92057 91950 92014 91906 91913 91978 91942 91948 92026 91962 92124
```

Q11) 107

```
which.max(sd$age12_plus_population)
```

```
[1] 90
```

```
sd$zip_code_tabulation_area[90]
```

```
[1] 92154
```

Q12)92154

```
sd.filt <- filter(vax, county == "San Diego", as_of_date == "2022-11-15")
mean(sd.filt$percent_of_population_fully_vaccinated, na.rm=TRUE)
```

```
[1] 0.7381765
```

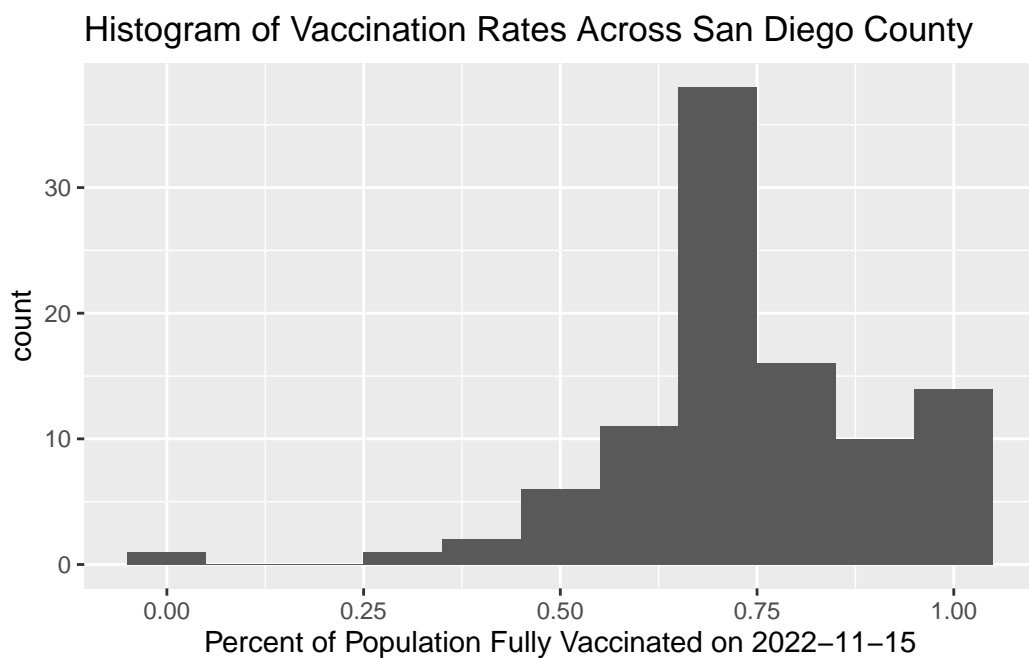
Q13) 0.74

```
library(ggplot2)
```

Q14)

```
ggplot(sd.filt,aes(percent_of_population_fully_vaccinated))+geom_histogram(binwidth=0.10)+
```

Warning: Removed 8 rows containing non-finite values (stat\_bin).



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

```
[1] 36144
```



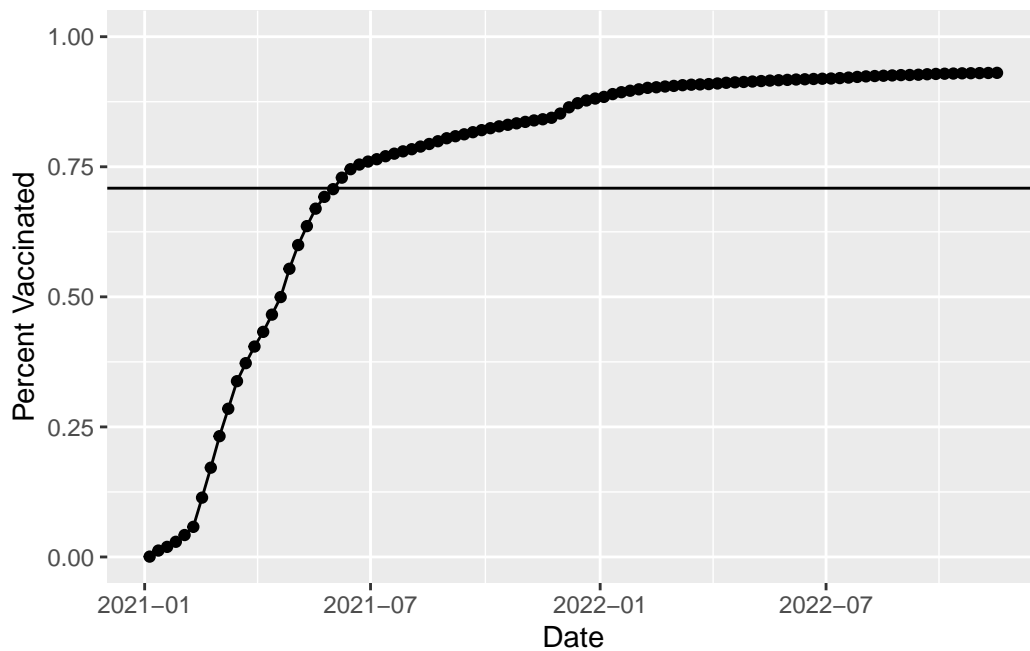
Q15)

```
plot<-ggplot(ucsd) +  
  aes(as_of_date,  
      percent_of_population_fully_vaccinated) +  
  geom_point() +  
  geom_line(group=1) +  
  ylim(c(0,1)) +  
  labs(x="Date", y="Percent Vaccinated")
```

Q16)

```
vax.36 <- filter(vax, age5_plus_population > 36144 &  
                 as_of_date == "2022-11-15")  
mean<-mean(vax.36$percent_of_population_fully_vaccinated)
```

```
plot+geom_hline(yintercept=0.7088141)
```



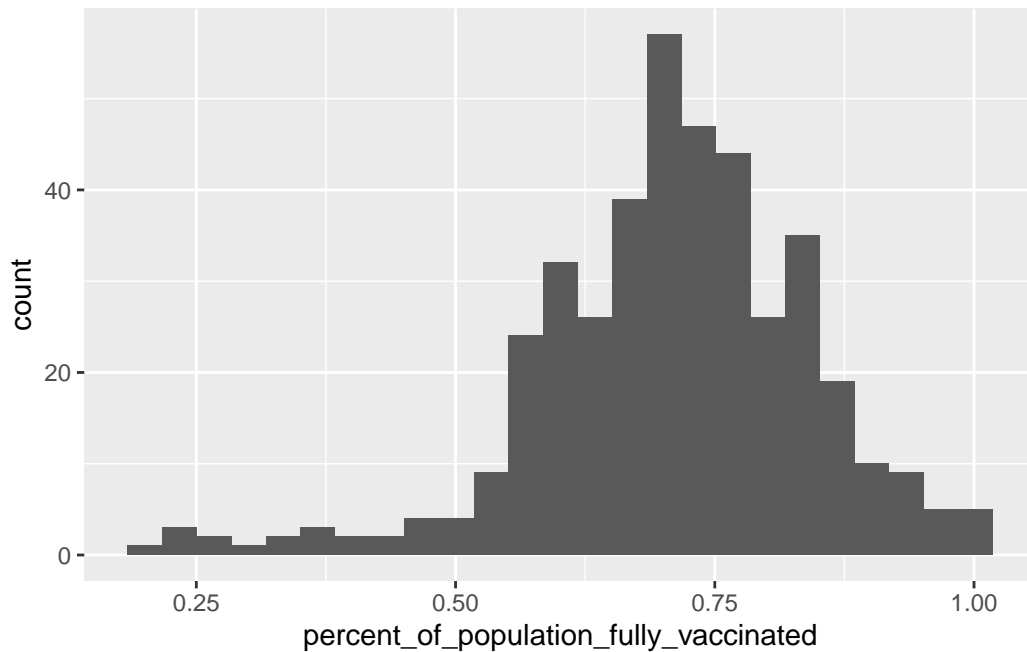
Q17)

```
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 |

Q17) Min. 1st Qu. Median Mean 3rd Qu. Max. 0.1986 0.6338 0.7162 0.7088 0.7893 1.0000  
 Q18)

```
ggplot(vax.36,aes(percent_of_population_fully_vaccinated))+geom_histogram(bins=25)
```



Q19)

```
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
```

```
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
```

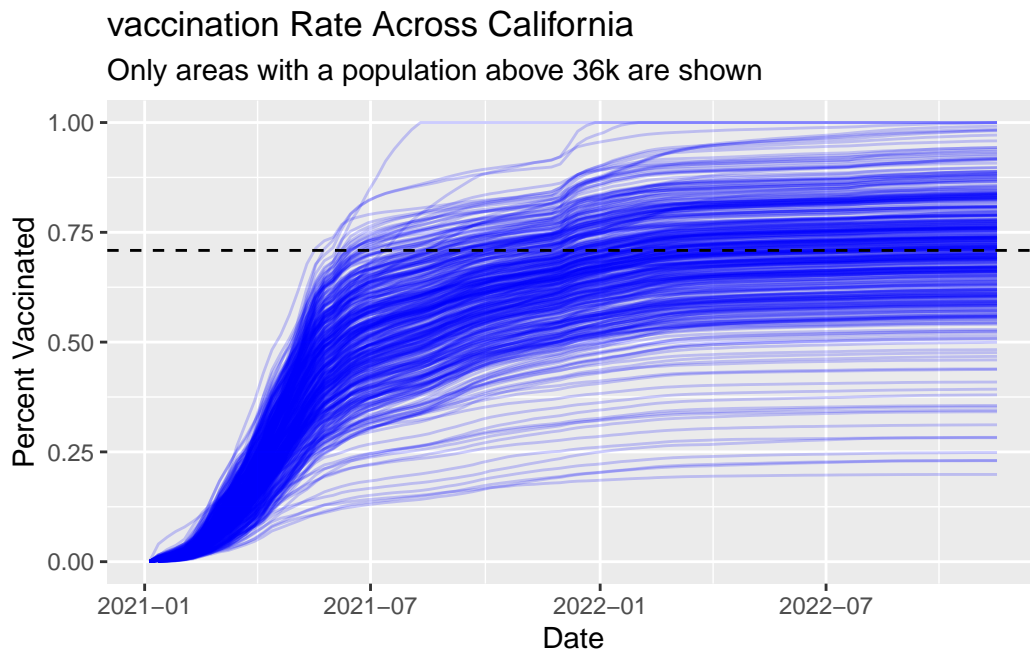
Both these zip code values are lower than the mean value of 0.7088

Q20

```
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(c(0,1.00)) +
  labs(x="Date", y="Percent Vaccinated",
       title="vaccination Rate Across California",
       subtitle="Only areas with a population above 36k are shown") +
  geom_hline(yintercept = 0.7088141, linetype="dashed")
```

Warning: Removed 183 row(s) containing missing values (geom\_path).



Q21 I am more wary about the risks