Lab17

Karen Guerrero

```
# Import vaccination data
vax <- read.csv("Statewide.csv")
```

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

#Q2. What column details the Zip code tabulation area? zip_code_tabulation_area

#Q3. What is the earliest date in this dataset? 2021-01-05

#Q4. What is the latest date in this dataset? 2022-11-22

```
head(vax)
```

	as_of_date	zip_code_tabulation_area	local_health_jurisdiction	county
1	2021-01-05	92240	Riverside	Riverside
2	2021-01-05	91302	Los Angeles	Los Angeles
3	2021-01-05	93420	San Luis Obispo	San Luis Obispo
4	2021-01-05	91901	San Diego	San Diego
5	2021-01-05	94110	San Francisco	San Francisco
6	2021-01-05	91902	San Diego	San Diego

	vaccine_equity_metric_quartile	vem_source
1	1	Healthy Places Index Score
2	4	Healthy Places Index Score
3	3	Healthy Places Index Score
4	3	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	29270.5	33093	35278
2	23163.9	25899	26712
3	26694.9	29253	30740

```
4              15549.8                 16905           18162
5              64350.7                 68320           72380
6              16620.7                 18026           18896
  persons_fully_vaccinated persons_partially_vaccinated
1                       NA                           NA
2                       15                          614
3                       NA                           NA
4                       NA                           NA
5                       17                         1268
6                       15                          397
  percent_of_population_fully_vaccinated
1                                     NA
2                               0.000562
3                                     NA
4                                     NA
5                               0.000235
6                               0.000794
  percent_of_population_partially_vaccinated
1                                         NA
2                                   0.022986
3                                         NA
4                                         NA
5                                   0.017519
6                                   0.021010
  percent_of_population_with_1_plus_dose booster_recip_count
1                                     NA                  NA
2                               0.023548                  NA
3                                     NA                  NA
4                                     NA                  NA
5                               0.017754                  NA
6                               0.021804                  NA
  bivalent_dose_recip_count eligible_recipient_count
1                        NA                        2
2                        NA                       15
3                        NA                        4
4                        NA                        8
5                        NA                       17
6                        NA                       15
                                                    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

```r
tail(vax)
```

```
       as_of_date zip_code_tabulation_area local_health_jurisdiction
174631 2022-11-22                    94066                  San Mateo
174632 2022-11-22                    92254                   Riverside
174633 2022-11-22                    94065                  San Mateo
174634 2022-11-22                    92280             San Bernardino
174635 2022-11-22                    94929                      Marin
174636 2022-11-22                    92313             San Bernardino
               county vaccine_equity_metric_quartile                    vem_source
174631      San Mateo                              4 Healthy Places Index Score
174632      Riverside                              1 Healthy Places Index Score
174633      San Mateo                              4 Healthy Places Index Score
174634 San Bernardino                             NA           No VEM Assigned
174635          Marin                              4    CDPH-Derived ZCTA Score
174636 San Bernardino                              2 Healthy Places Index Score
       age12_plus_population age5_plus_population tot_population
174631             37730.3                40903          43101
174632              7882.3                 8985           9779
174633             10465.5                11778          12461
174634                 0.0                    0             NA
174635               174.2                  218            254
174636             10842.9                11847          12547
       persons_fully_vaccinated persons_partially_vaccinated
174631                    38105                         2889
174632                     9456                         1688
174633                    11238                          889
174634                       NA                           NA
174635                       NA                           NA
174636                     7948                          600
       percent_of_population_fully_vaccinated
174631                               0.884086
174632                               0.966970
174633                               0.901854
174634                                     NA
174635                                     NA
174636                               0.633458
       percent_of_population_partially_vaccinated
```

```
174631                                        0.067029
174632                                        0.172615
174633                                        0.071343
174634                                              NA
174635                                              NA
174636                                        0.047820
       percent_of_population_with_1_plus_dose booster_recip_count
174631                               0.951115               27085
174632                               1.000000                3840
174633                               0.973197                8701
174634                                     NA                  NA
174635                                     NA                  NA
174636                               0.681278                4522
       bivalent_dose_recip_count eligible_recipient_count
174631                      9127                    37620
174632                       372                     9430
174633                      3456                    11021
174634                        NA                       14
174635                        NA                      159
174636                      1085                     7921
                                                                  redacted
174631                                                                  No
174632                                                                  No
174633                                                                  No
174634 Information redacted in accordance with CA state privacy requirements
174635 Information redacted in accordance with CA state privacy requirements
174636                                                                  No
```

#Q5. How many numeric columns are in this dataset? 13

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

Table 1: Data summary

Name	vax
Number of rows	174636
Number of columns	18
Column type frequency:	
character	5
numeric	13

Table 1: Data summary

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	99	0
local_health_jurisdiction	0	1	0	15	495	62	0
county	0	1	0	15	495	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	8613	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	31756.13	88556.7	
age5_plus_population	0	1.00	20875.24	21105.98	0	1460.50	15364.00	34877.00	101902.0	
tot_population	8514	0.95	23372.77	22628.51	12	2126.00	18714.00	38168.00	111165.0	
persons_fully_vaccinated	14921	0.91	13466.31	14722.46	11	883.00	8024.00	22529.00	87186.0	
persons_partially_vaccinated	14921	0.91	1707.50	1998.80	11	167.00	1194.00	2547.00	39204.0	
percent_of_population_fully_vaccinated	18065	0.89	0.55	0.25	0	0.39	0.59	0.73	1.0	
percent_of_population_partially_vaccinated	18065	0.89	0.08	0.09	0	0.05	0.06	0.08	1.0	
percent_of_population_with_1_plus_dose	19562	0.89	0.61	0.25	0	0.46	0.65	0.79	1.0	
booster_recip_count	70421	0.60	5655.17	6867.49	11	280.00	2575.00	9421.00	58304.0	
bivalent_dose_recip_count	156958	0.10	1646.02	2161.84	11	109.00	719.00	2443.00	18109.0	
eligible_recipient_count	0	1.00	12309.19	14555.83	0	466.00	5810.00	21140.00	86696.0	

```
sum(is.na(vax$persons_fully_vaccinated))
```

```
[1] 14921
```

```
library(lubridate)
```

```
Loading required package: timechange
```

```
Attaching package: 'lubridate'
```

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

```
    date, intersect, setdiff, union
```

```
    today()
```

```
[1] "2022-12-02"
```

```
    # This will give an Error!
    ##today() - vax$as_of_date[1]
```

```
    # Specify that we are using the year-month-day format
    vax$as_of_date <- ymd(vax$as_of_date)
```

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

```
Time difference of 696 days
```

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

```
Time difference of 686 days
```

```
    library(zipcodeR)
```

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

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