

test.R

Noah

2022-03-07

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

```
##   as_of_date zip_code_tabulation_area local_health_jurisdiction      county
## 1 2021-01-05                92549                Riverside    Riverside
## 2 2021-01-05                92130                San Diego      San Diego
## 3 2021-01-05                92397            San Bernardino San Bernardino
## 4 2021-01-05                94563            Contra Costa    Contra Costa
## 5 2021-01-05                94519            Contra Costa    Contra Costa
## 6 2021-01-05                91042            Los Angeles    Los Angeles
##   vaccine_equity_metric_quartile      vem_source
## 1                        3 Healthy Places Index Score
## 2                        4 Healthy Places Index Score
## 3                        3 Healthy Places Index Score
## 4                        4 Healthy Places Index Score
## 5                        3 Healthy Places Index Score
## 6                        2 Healthy Places Index Score
##   age12_plus_population age5_plus_population persons_fully_vaccinated
## 1                2348.4                2461                NA
## 2                46300.3                53102                61
## 3                3695.6                4225                NA
## 4                17216.1                18896                NA
## 5                16861.2                18678                NA
## 6                23962.2                25741                NA
##   persons_partially_vaccinated percent_of_population_fully_vaccinated
## 1                        NA                NA
## 2                        27                0.001149
## 3                        NA                NA
## 4                        NA                NA
## 5                        NA                NA
## 6                        NA                NA
##   percent_of_population_partially_vaccinated
## 1                        NA
## 2                0.000508
## 3                        NA
## 4                        NA
## 5                        NA
## 6                        NA
##   percent_of_population_with_1_plus_dose booster_recip_count
## 1                        NA                NA
## 2                0.001657                NA
## 3                        NA                NA
```

```
## 4 NA NA
## 5 NA NA
## 6 NA NA
## 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?

#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?

#2022-03-01

#Q5: How many numeric columns are in this dataset?

#9 columns

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

```
## [1] 18338
```

#Q6: Note that there are "missing values" in the dataset.

#How many NA values there in the persons_fully_vaccinated column?

#18338 NAs

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

#17 percent

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

#5 days

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

#420