Milestone #2

Virginia Chan, Patrick Traynor, Antoinette Stetzenmeyer

9/29/2021

```
#load libraries
library(tidyverse)
```

```
## Warning in system("timedatectl", intern = TRUE): running command 'timedatectl'
## had status 1
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.2     v dplyr   1.0.6
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(lubridate)
```

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

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

Description of dataset What is the data source? (1-2 sentences on where the data is coming from, dates included, etc.) How does the dataset relate to the group problem statement and question? Problem: Do counties with younger median age values less likely to have higher counts of vaccinated persons?

```
library(readr)
cov_vax_admin <- read_csv("cov_vax_admin.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   X1 = col_double(),
##   as_of_date = col_character(),
##   zip_code_tabulation_area = col_double(),
##   local_health_jurisdiction = col_character(),
##   county = col_character(),
##   vaccine_equity_metric_quartile = col_double(),
##   vem_source = col_character(),
##   age12_plus_population = col_double(),
##   persons_fully_vaccinated = col_double(),
##   persons_partially_vaccinated = col_double(),
##   redacted = col_character()
```

```
## )
```

```
view(cov_vax_admin)
```

```
#ca_county_demographics <- read_csv("ca_county_demographics.csv")
#View(ca_county_demographics)
```