Final Report
due November 16, 2021 by 11:59 PM

Lindsey Weyant, Ali Raich, Aden Clemente

11/16/21

#Load Data

```
census_api_key("abc8289fa2ba274ced76d97c7f8ee31666a2c931", overwrite = TRUE, install = TRUE)
```

Your original .Renviron will be backed up and stored in your R HOME directory if needed.

Your API key has been stored in your .Renviron and can be accessed by Sys.getenv("CENSUS_API_KEY").
To use now, restart R or run `readRenviron("~/.Renviron")`

[1] "abc8289fa2ba274ced76d97c7f8ee31666a2c931"

```
if (FALSE) {
census_api_key("abc8289fa2ba274ced76d97c7f8ee31666a2c931", install = TRUE)
# First time, reload your environment so you can use the key without restarting R.
readRenviron("~/.Renviron")
# You can check it with:
Sys.getenv("CENSUS_API_KEY")
}
```

#Research Question:

How do measles vaccination rates vary across the country and demographics in schools?

overall vaccination status vs. state, overall vaccination status vs. type of school, each type of exemption (personal, religious, and medical) vs. state exemption vs. type of school. To analyze vaccination and exemption rates by states, we will use spatial data to show the change in these rates across the country. Then, we can use two-sample t-tests to test for significance of vaccination and exemption rates between different types of schools. If there are significantly lower vaccination rates in private schools vs. other types of schools, this will support our main hypothesis.

#Variable Manipulation

```
measles <- measles %>%
  filter(overall != (-1)) %>%
  mutate(numvaxx = round(enroll*overall*.01)) %>%
  mutate(unvaxx = enroll - numvaxx)%>%

  mutate(numxmed = round(enroll*.01*xmed)) %>%
  mutate(numxmed = ifelse(is.na(numxmed), 0, numxmed)) %>%

  mutate(numxrel = round(enroll*.01*xrel)) %>%
  mutate(numxrel = ifelse(is.na(numxrel), 0, numxrel)) %>%

  mutate(numxper = round(enroll*.01*xper)) %>%
  mutate(numxper = ifelse(is.na(numxper), 0, numxper)) %>%
```

```
  mutate(numxother = unvaxx - numxper - numxmed - numxrel) %>%
  mutate(numxother = ifelse(is.na(numxother), 0, numxother)) %>%

  mutate(statefac = as.factor(state))
```

Logistic Regression

```
measlereg <- glm(cbind(numvaxx, unvaxx) ~ statefac + type, data=measles, family = binomial)
measlereg
```
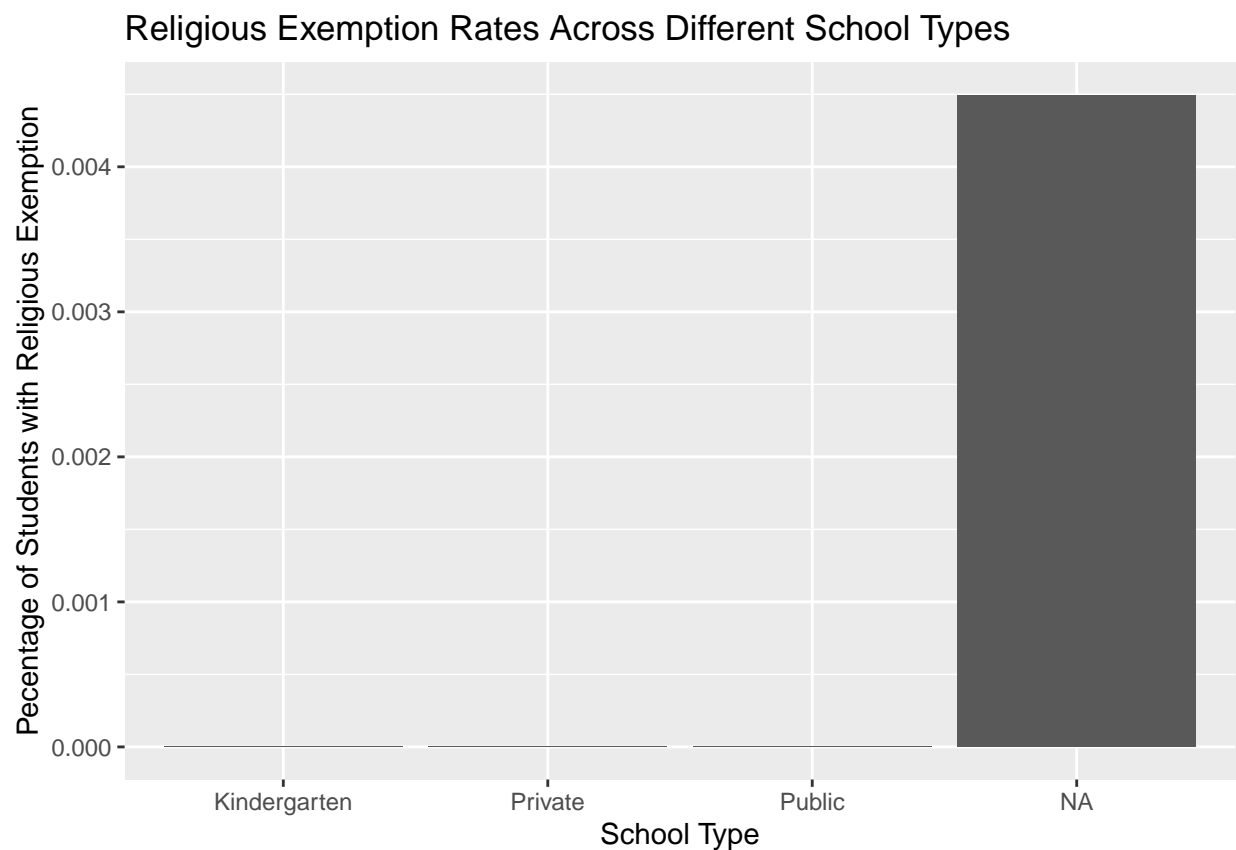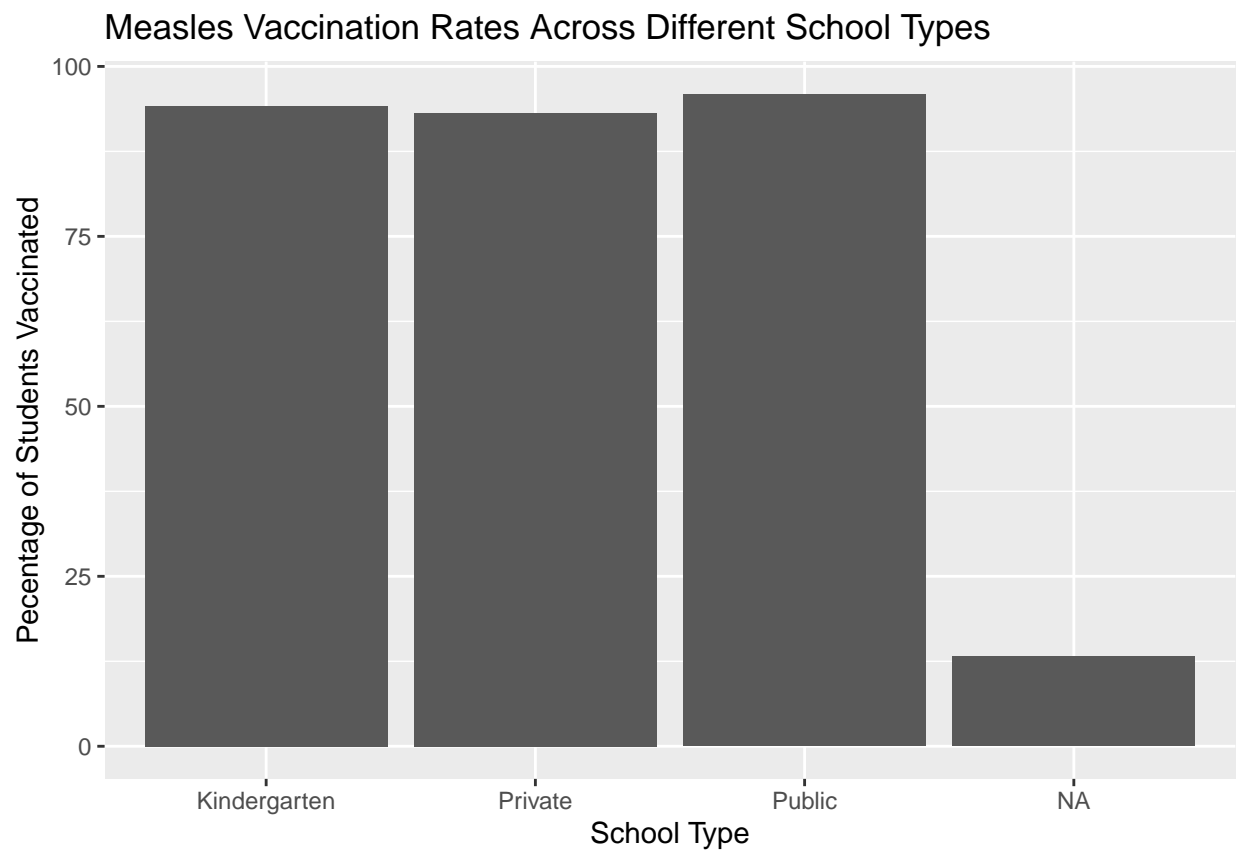
```
##
## Call:  glm(formula = cbind(numvaxx, unvaxx) ~ statefac + type, family = binomial,
##     data = measles)
##
## Coefficients:
##      (Intercept)  statefacColorado       statefacOhio        typePrivate
##           3.7220           -1.6471            -0.8457            -0.8424
##       typePublic
##          -0.6406
##
## Degrees of Freedom: 18645 Total (i.e. Null);  18641 Residual
##    (20243 observations deleted due to missingness)
## Null Deviance:        114500
## Residual Deviance: 101300     AIC: 153400
```

```
summary(measlereg)
```

```
##
## Call:
## glm(formula = cbind(numvaxx, unvaxx) ~ statefac + type, family = binomial,
##     data = measles)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -59.896   -0.292    0.737    1.357   17.086
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)       3.722031   0.049580   75.07   <2e-16 ***
## statefacColorado -1.647126   0.048295  -34.11   <2e-16 ***
## statefacOhio     -0.845687   0.009185  -92.07   <2e-16 ***
## typePrivate      -0.842418   0.047546  -17.72   <2e-16 ***
## typePublic       -0.640616   0.049617  -12.91   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 114508  on 18645  degrees of freedom
## Residual deviance: 101255  on 18641  degrees of freedom
##    (20243 observations deleted due to missingness)
## AIC: 153354
##
```
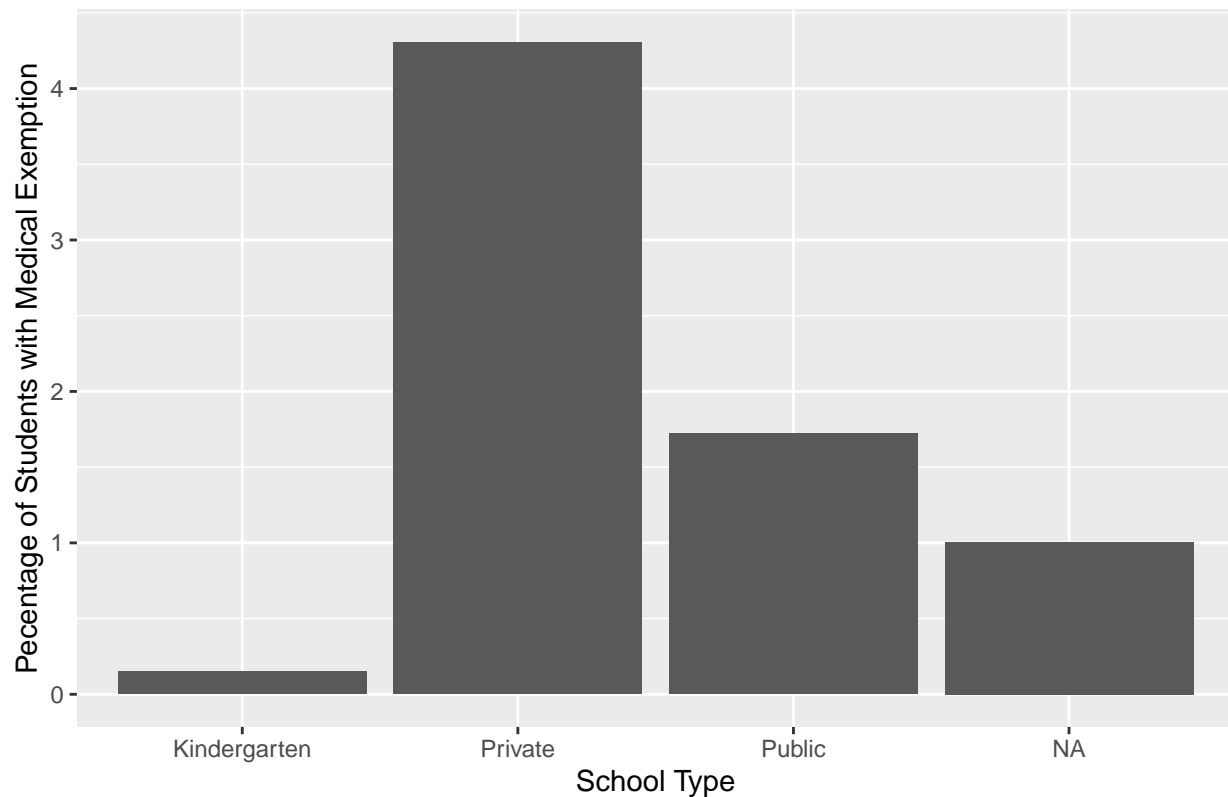
```
## Number of Fisher Scoring iterations: 5
```

Preliminary Bar Graphs

Measles Vaccination Rates Across Different School Types

Religious Exemption Rates Across Different School Types

Personal Exemption Rates Across Different School Types
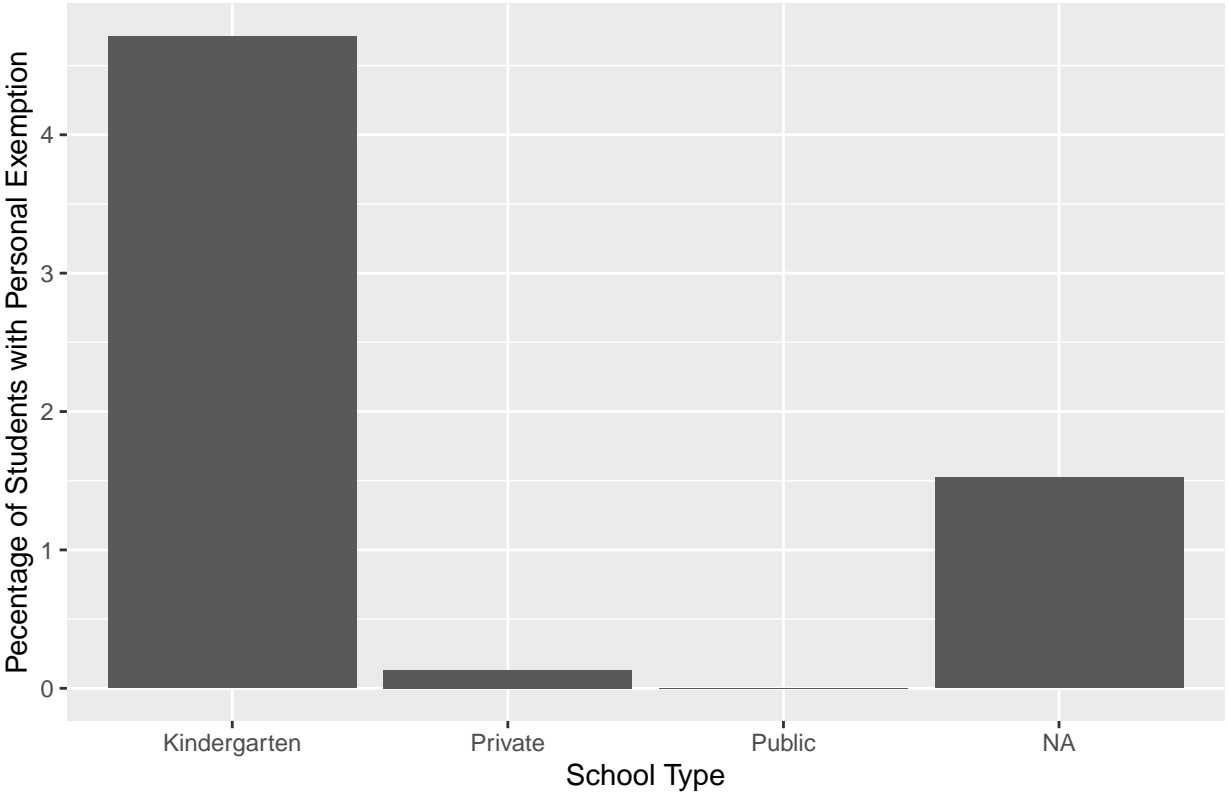
Medical Exemption Rates Across Different School Types
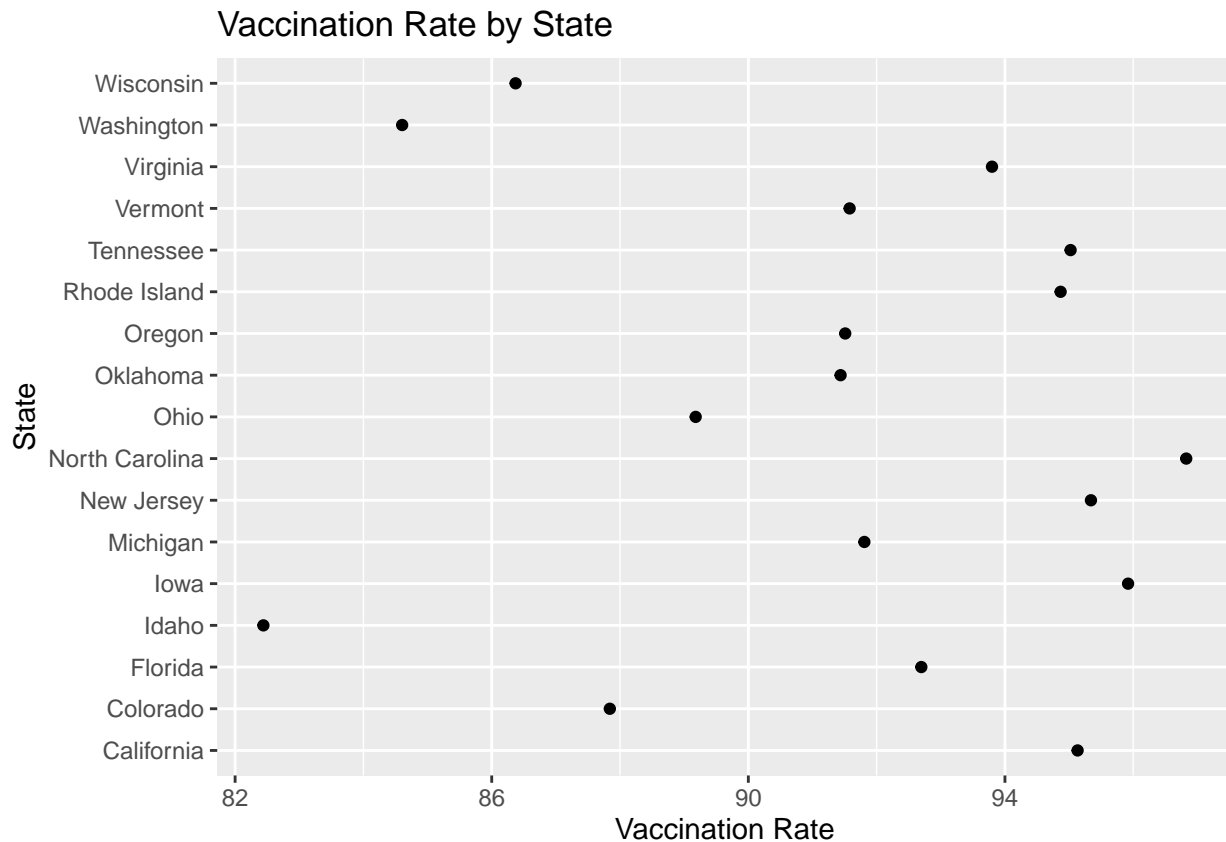

vaccination by state graph

```
measles %>%
  count(state) %>%
  group_by(state)
```

```
## # A tibble: 17 x 2
## # Groups:   state [17]
##    state              n
##    <chr>          <int>
##  1 California     14224
##  2 Colorado        1505
##  3 Florida         2672
##  4 Idaho            467
##  5 Iowa            1163
##  6 Michigan        2351
##  7 New Jersey      2044
##  8 North Carolina  2084
##  9 Ohio            2917
## 10 Oklahoma        1052
## 11 Oregon           806
## 12 Rhode Island     215
## 13 Tennessee       1152
## 14 Vermont          338
## 15 Virginia        1413
```

```
## 16 Washington      1978
## 17 Wisconsin       2508
```

```
measles %>%
  filter(overall != (-1)) %>%
  group_by(state) %>%
  summarise(statemean = mean(overall)) %>%
  ggplot(aes(x = statemean, y = state)) +
  geom_point() +
  labs(x = "Vaccination Rate", y = "State", title = "Vaccination Rate by State")
```



```
plotdata <- measles %>%
filter(overall != (-1)) %>%
group_by(state) %>%
summarise(statemean = mean(overall))
plot_usmap(data=plotdata, values = "statemean") +
  labs(title = "Vaccination Rate by State", fill = "Vaccination Rate") +
  theme(panel.background = element_rect(color = "black", fill = "lightblue"))
```

Vaccination Rate by State
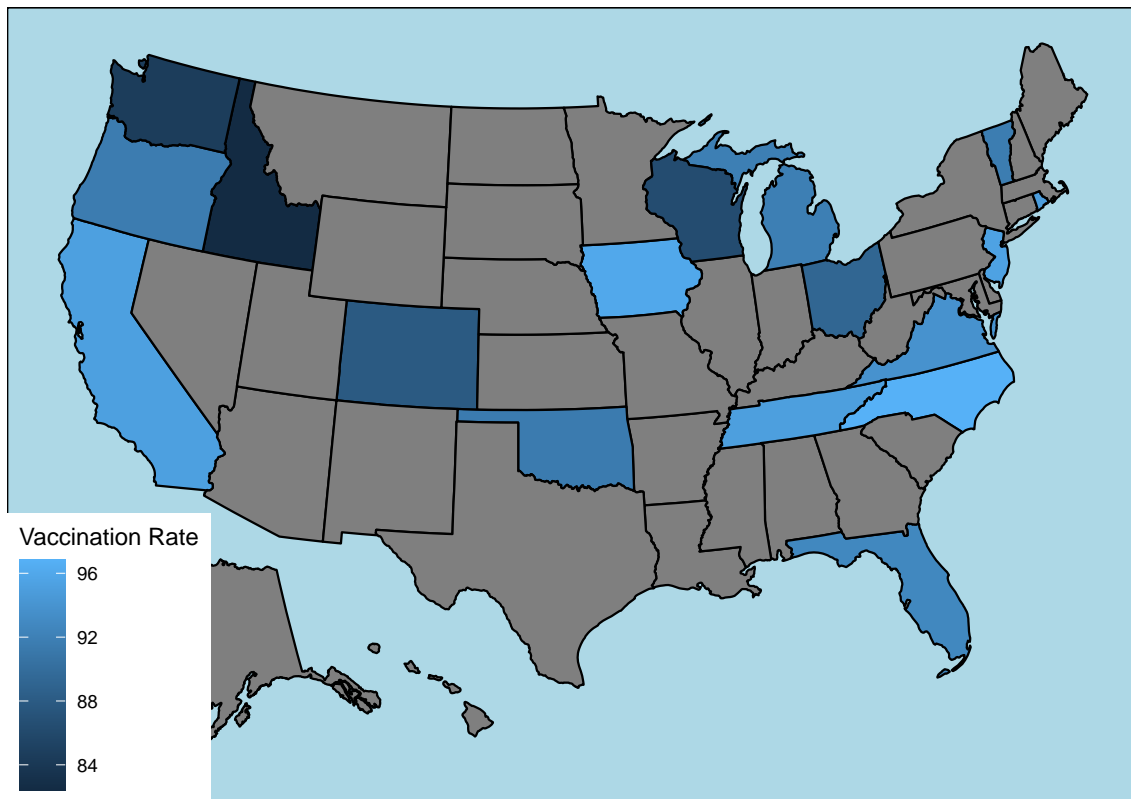

```
# measles %>%
#   pivot_longer(cols = numxmed:numxother, names_to = "group", values_to = "count") %>%
#   ggplot(aes(x = type, y = count, (fill = group))) +
#           geom_bar(position = "stack")
```