Final Report

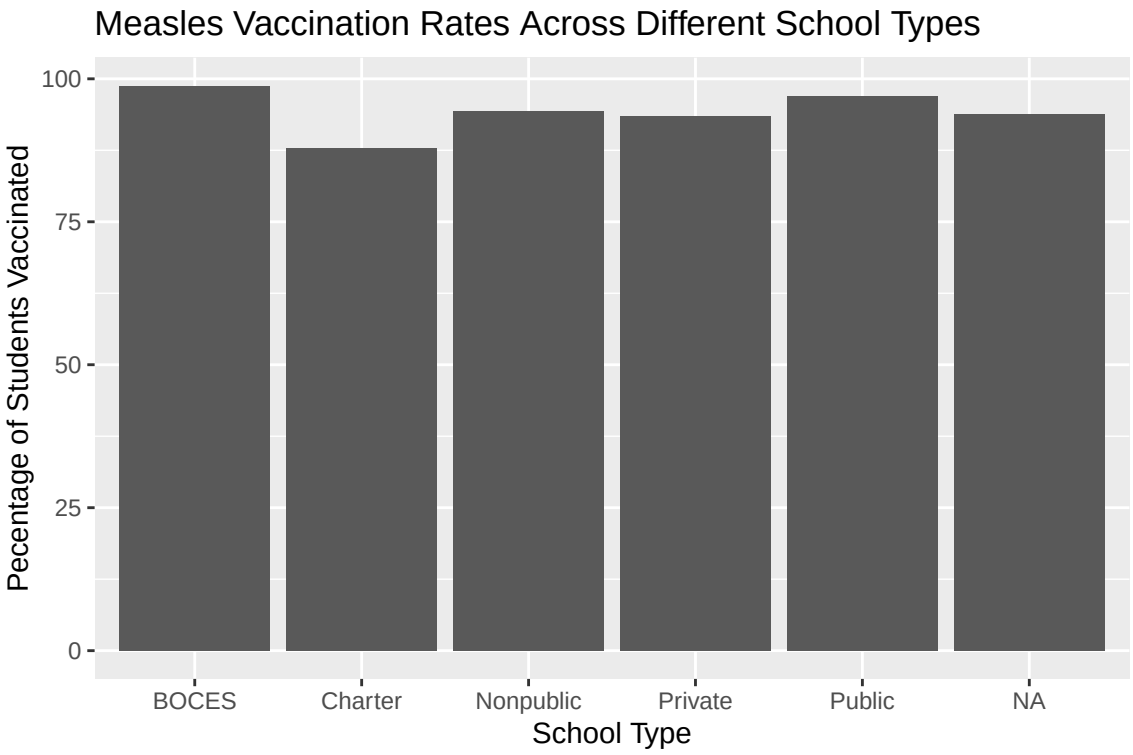
due November 16, 2021 by 11:59 PM

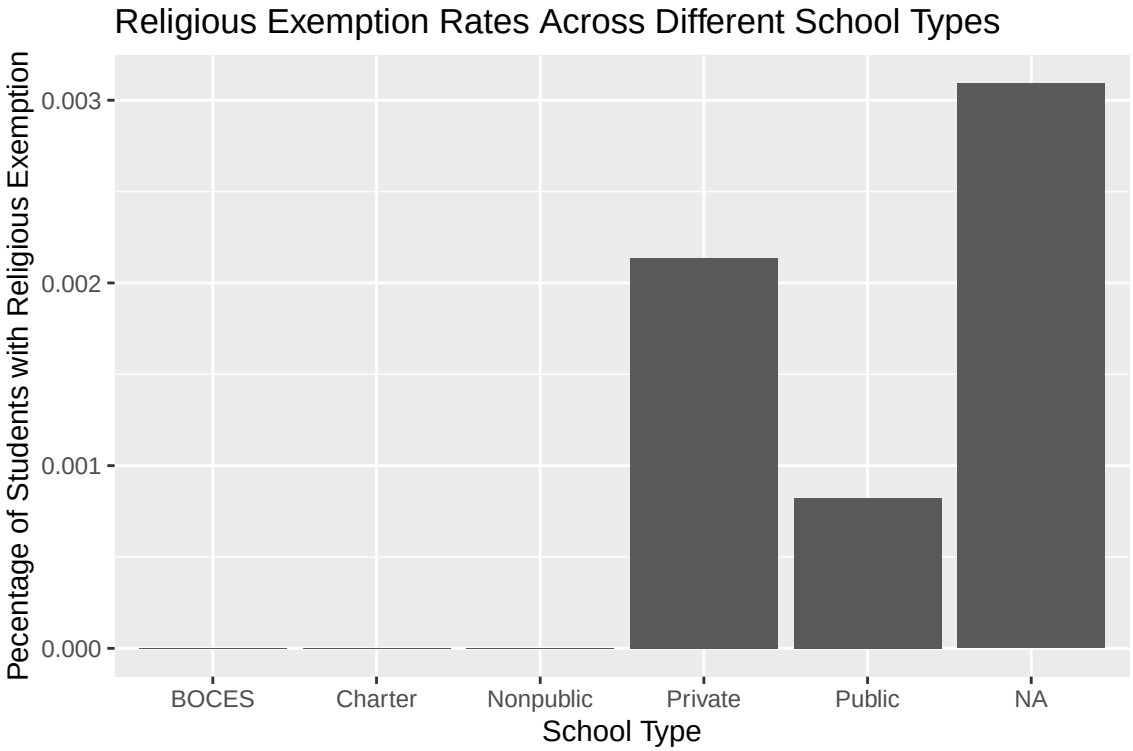
Lindsey Weyant, Ali Raich, Aden Clemente

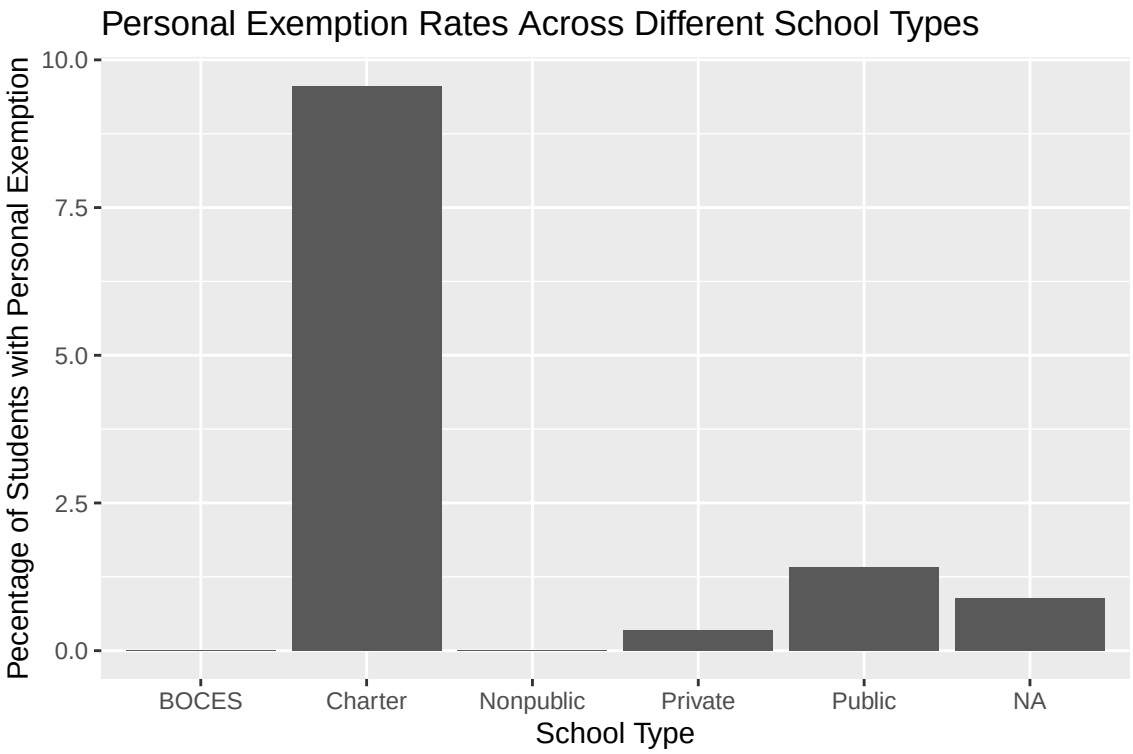
11/16/21

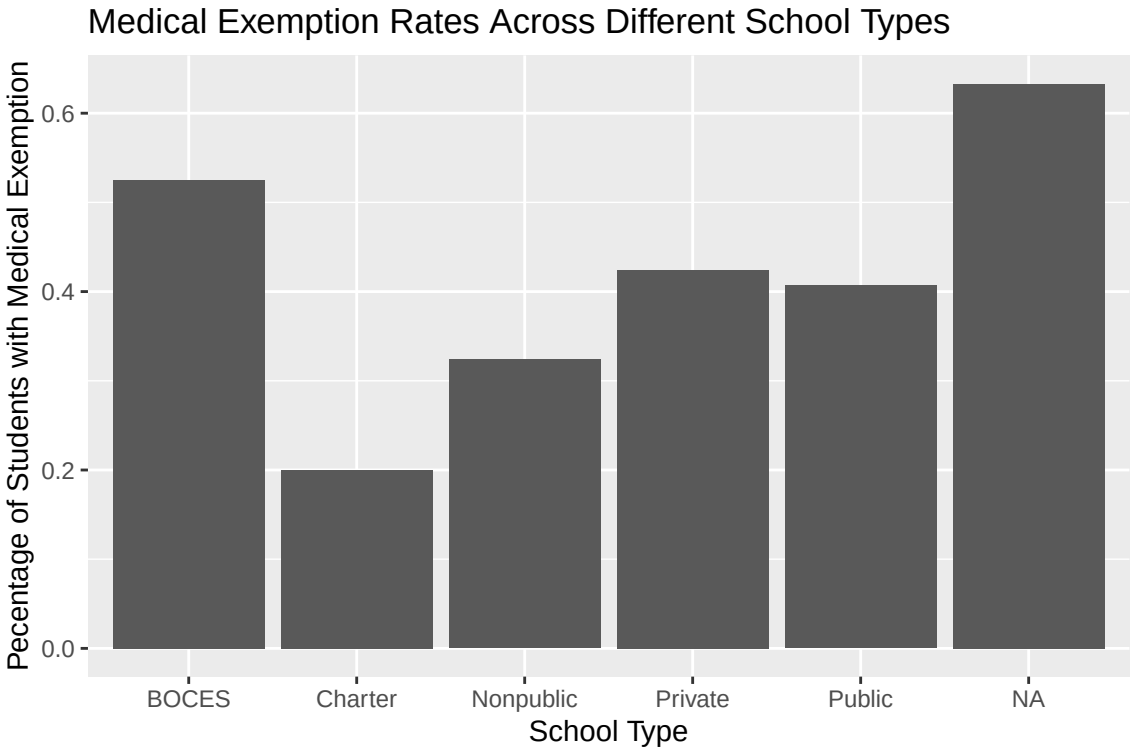
```
## [1] "abc8289fa2ba274ced76d97c7f8ee31666a2c931"
```

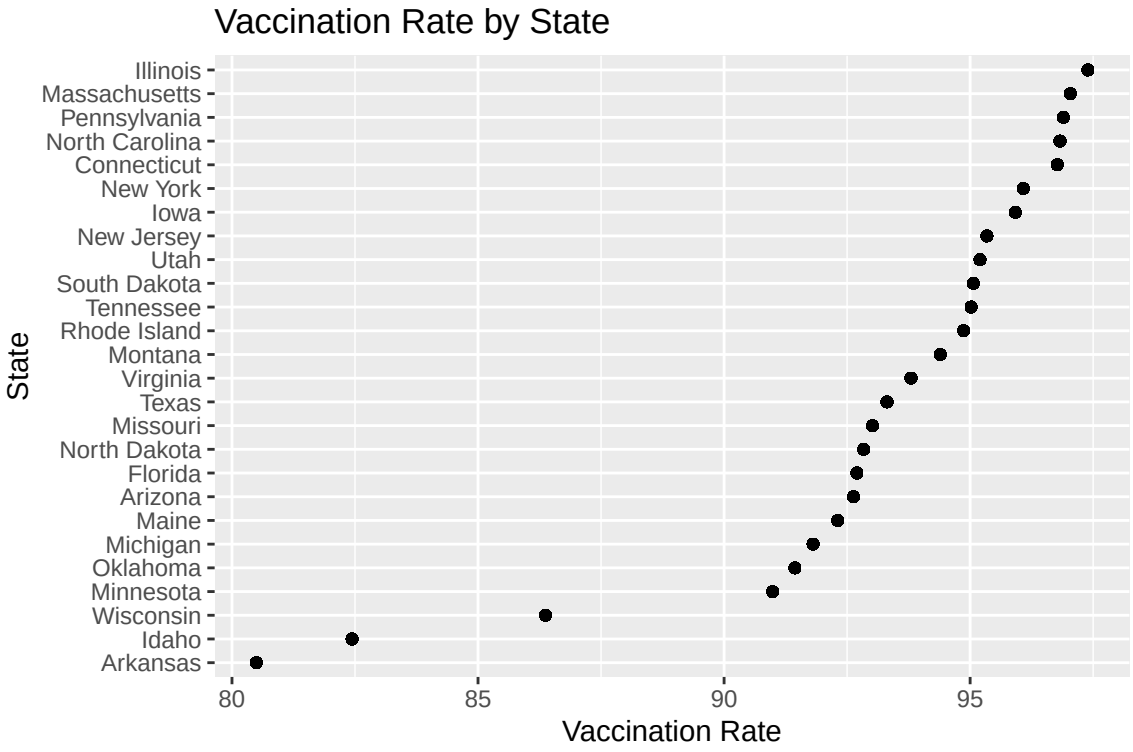
Exploratory Data Analysis

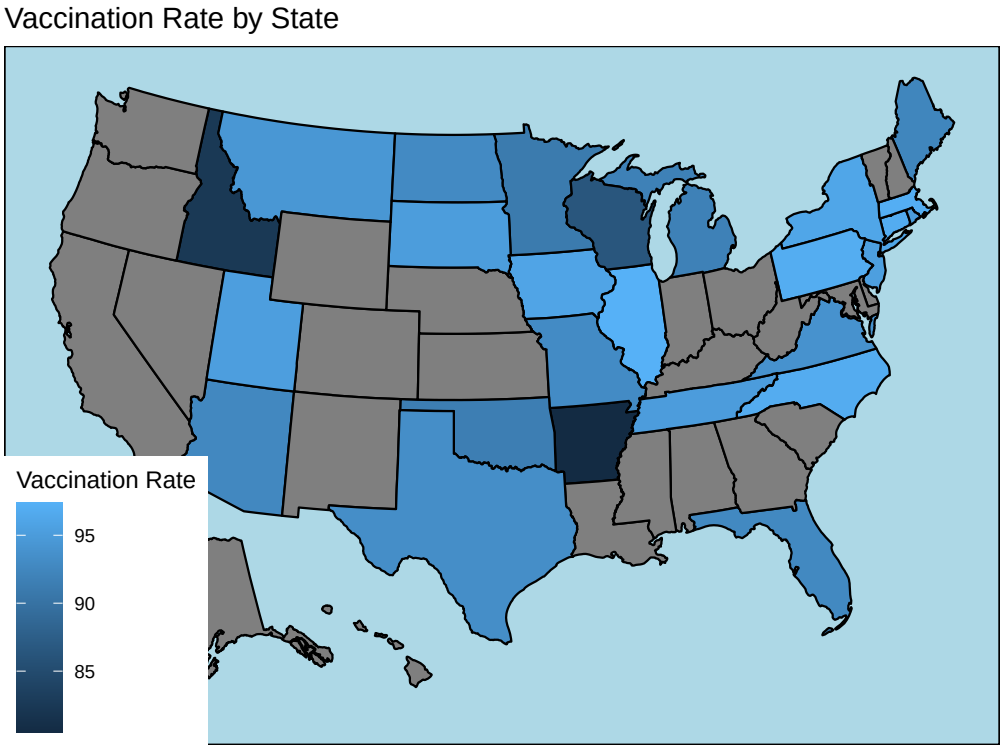


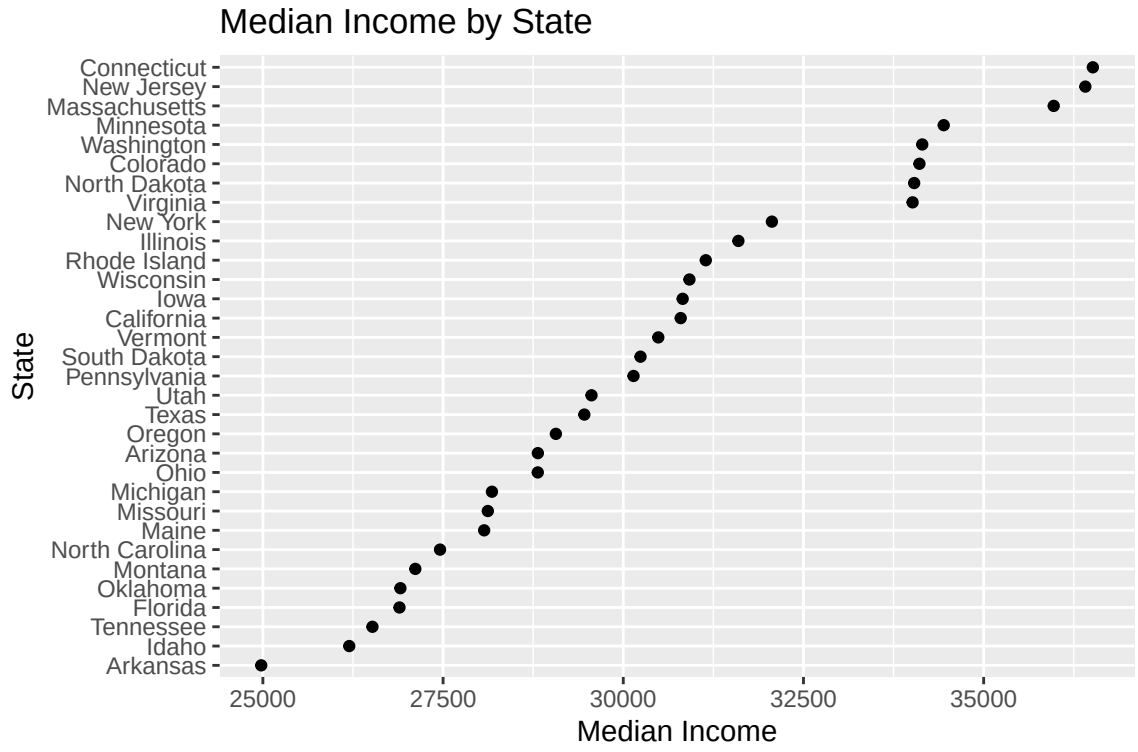



Personal Exemption Rates Across Different School Types
Pecentage of Students with Personal Exemption
10.0
7.5
5.0
2.5
0.0
BOCES
Charter
Nonpublic
Private
Public
NA
School Type


Medical Exemption Rates Across Different School Types
Pecentage of Students with Medical Exemption
0.6
0.4
0.2
0.0
BOCES
Charter
Nonpublic
Private
Public
NA
School Type


Vaccination Rate by State
Illinois
Massachusetts
Pennsylvania
North Carolina
Connecticut
New York
Iowa
New Jersey
Utah
South Dakota
Tennessee
Rhode Island
Montana
Virginia
Texas
Missouri
North Dakota
Florida
Arizona
Maine
Michigan
Oklahoma
Minnesota
Wisconsin
Idaho
Arkansas
State
80
85
90
95
Vaccination Rate


Vaccination Rate by State
Vaccination Rate
95
90
85



T-Tests and ANOVA

```
##                Df  Sum Sq Mean Sq F value Pr(>F)
## state          25  516286   20651   320.7 <2e-16 ***
## Residuals   39479 2542092      64
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
##  Welch Two Sample t-test
##
## data:  measles3$realrate by measles3$type
## t = -11.702, df = 2610.6, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group Private and group Public is not equal
## 95 percent confidence interval:
##  -4.126435 -2.941957
## sample estimates:
## mean in group Private  mean in group Public
##              93.47576              97.00995
```

```
##
##  Welch Two Sample t-test
##
## data:  measles4$realrate by measles4$type
## t = -11.423, df = 219.45, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group Charter and group Public is not equal
## 95 percent confidence interval:
##  -10.617019  -7.492476
## sample estimates:
```

```
## mean in group Charter  mean in group Public
##              87.95521              97.00995
```

```
##
##  Welch Two Sample t-test
##
## data:  measles5$realrate by measles5$type
## t = -6.5532, df = 279.42, p-value = 2.704e-10
## alternative hypothesis: true difference in means between group Charter and group Private is not equal
## 95 percent confidence interval:
##  -7.178836 -3.862267
## sample estimates:
## mean in group Charter mean in group Private
##              87.95521              93.47576
```

Regression Analysis

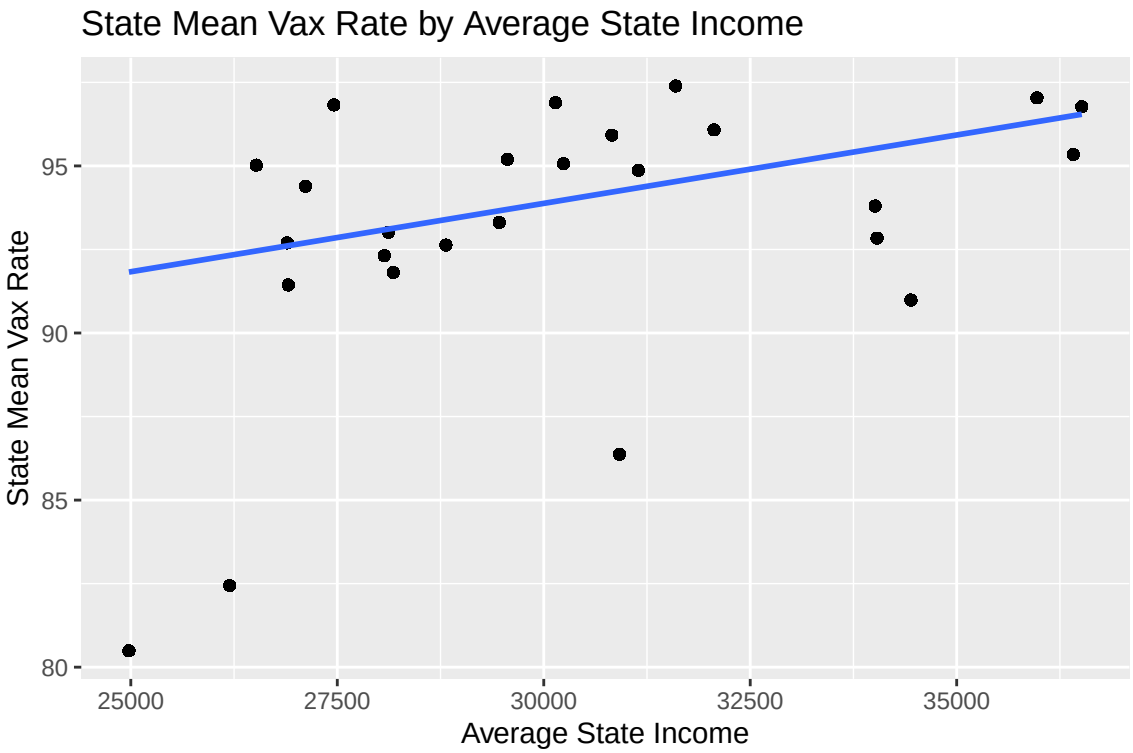
```
##
## Call:  glm(formula = cbind(numvaxx, unvaxx) ~ statefac, family = binomial,
##     data = measles)
##
## Coefficients:
##            (Intercept)         statefacArizona         statefacFlorida
##                 1.4042                  1.2150                  1.2178
##       statefacIllinois            statefacIowa           statefacMaine
##                 2.3379                  1.8029                  1.3066
##       statefacMichigan       statefacMinnesota         statefacMontana
##                 1.1680                  1.2123                  0.9553
##     statefacNew Jersey  statefacNorth Carolina    statefacNorth Dakota
##                 1.8859                  1.9320                  1.2772
##   statefacPennsylvania    statefacRhode Island    statefacSouth Dakota
##                 2.1153                  1.7268                  2.0056
##      statefacTennessee            statefacUtah        statefacVirginia
##                 1.5194                  1.6913                  1.2396
##
## Degrees of Freedom: 28126 Total (i.e. Null);  28109 Residual
##   (11347 observations deleted due to missingness)
## Null Deviance:       316100
## Residual Deviance: 177000    AIC: 255100
```

```
##
## Call:
## glm(formula = cbind(numvaxx, unvaxx) ~ statefac, family = binomial,
##     data = measles)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -154.356    -0.709     0.554     1.628    12.531
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)            1.404174   0.004625  303.61   <2e-16 ***
## statefacArizona        1.214998   0.014364   84.58   <2e-16 ***
## statefacFlorida        1.217824   0.009317  130.71   <2e-16 ***
```

```
## statefacIllinois       2.337924   0.006160  379.54   <2e-16 ***
## statefacIowa           1.802922   0.009551  188.76   <2e-16 ***
## statefacMaine          1.306578   0.026213   49.84   <2e-16 ***
## statefacMichigan       1.167997   0.011350  102.91   <2e-16 ***
## statefacMinnesota      1.212345   0.013802   87.84   <2e-16 ***
## statefacMontana        0.955268   0.011123   85.88   <2e-16 ***
## statefacNew Jersey     1.885880   0.017059  110.55   <2e-16 ***
## statefacNorth Carolina 1.931960   0.015660  123.37   <2e-16 ***
## statefacNorth Dakota   1.277233   0.033663   37.94   <2e-16 ***
## statefacPennsylvania   2.115323   0.016716  126.54   <2e-16 ***
## statefacRhode Island   1.726839   0.048012   35.97   <2e-16 ***
## statefacSouth Dakota   2.005628   0.050716   39.55   <2e-16 ***
## statefacTennessee      1.519426   0.016118   94.27   <2e-16 ***
## statefacUtah           1.691345   0.010156  166.54   <2e-16 ***
## statefacVirginia       1.239632   0.013847   89.53   <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: 316104  on 28126  degrees of freedom
## Residual deviance: 177040  on 28109  degrees of freedom
##   (11347 observations deleted due to missingness)
## AIC: 255099
##
## Number of Fisher Scoring iterations: 5
```

```
## parsnip model object
##
## Fit time:  1.8s
##
## Call:
## stats::lm(formula = statemean ~ estimate, data = data)
##
## Coefficients:
## (Intercept)     estimate
##   8.160e+01    4.092e-04
```

```
## `geom_smooth()` using formula 'y ~ x'
```



Equation for predicting state mean vax rate: $\hat{y} = 81.60 + 0.0004092 * x_i$