TidyTuesday: Vaccination in schools

Kyla McConnell

2/25/2020

Vaccination rates in schools

Prepare dataset

The dataset has a lot of NAs and -1s inn both the 'overall' column and the 'mmr' column. First, I create a dataset that takes the 'overall' vaccination rate, unless this is either NA or -1.. in this case, I take the 'mmr' ("mumps, measles, rubella", a common childhood vaccine) vaccination rate (as an approximation).

```
measles <- measles %>%
  mutate(vac =
    ifelse(
    (!is.na(measles$overall)& measles$overall != "-1"),
    measles$overall,
    measles$mmr
    )
         )%>%
  filter(vac != -1)%>%
  mutate(non_vac = 100 - vac)%>%
  filter(lng < 1)#remove problematic/erroneous longitude value
```

Next, I create a dataset that separates schools with less than 50% vaccination rate, and create a column for the amount of UNvaccinated students

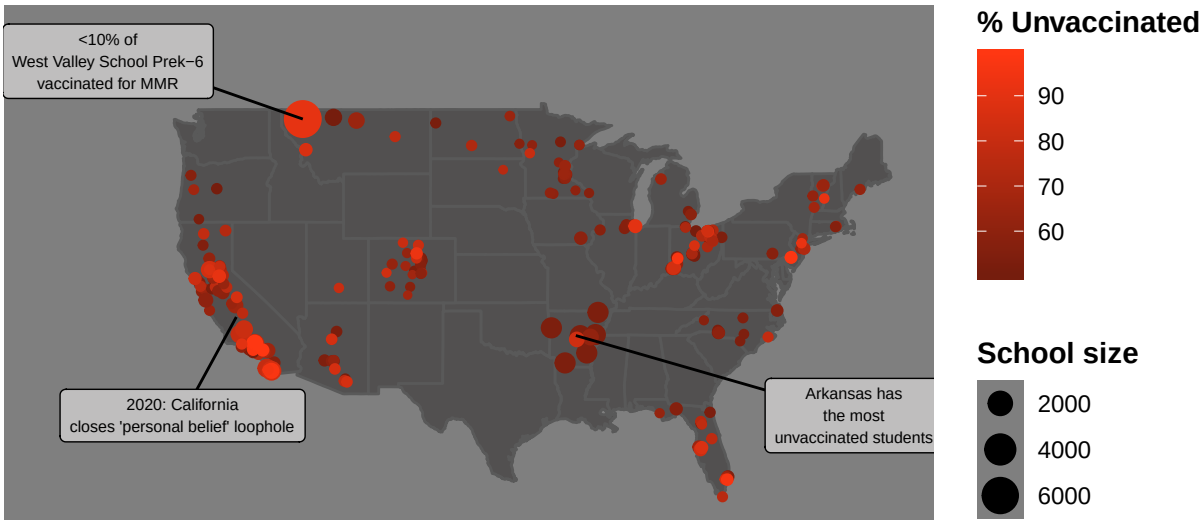
```
measles_low <- measles %>%
  filter(vac < 50)
```

Plot

```
g <- ggplot()+
  geom_sf(data=states, fill="#525050")+
  geom_point(data=measles_low, aes(x=lng, y=lat, size=enroll, color=non_vac))+
  theme_dark()+
  scale_color_gradient(low="#751D0D", high="#FF3712", name="% Unvaccinated")+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank(),
        axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank(),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        plot.title = element_text(face="bold"), #family = "Garamond",
        legend.title = element_text(face="bold")
```

```
        )+
  scale_size(name="School size")+
  ggtitle("Hotspots for outbreak?\n% of unvaccinated students in America's schools")+
  geom_segment(aes(x= -70, xend = -92.37, y= 30, yend = 34.969704))+
  annotate(geom="label", fill="#BFBEBE", x= -70, y= 30, label="Arkansas has\nthe most\nunvaccinated stu
  geom_segment(aes(x= -135, xend = -114.33, y= 54, yend = 48.41))+
  annotate(geom="label", fill="#BFBEBE", x= -130, y=52, label="<10% of\nWest Valley School Prek-6\nvacc
  geom_segment(aes(x= -124, xend = -119.68, y= 30, yend = 36.12))+
  annotate(geom="label", fill="#BFBEBE", x= -124, y=30, label="2020: California\ncloses 'personal belie

g
```



For labels: Check for state with highest percent unvaccinated

```
measles %>%
  group_by(state) %>%
  summarize(mean_non_vac = mean(non_vac))%>%
  arrange(desc(mean_non_vac))
```

```
## # A tibble: 31 x 2
##    state      mean_non_vac
##    <chr>             <dbl>
##  1 Arkansas          19.5
##  2 Idaho             17.6
##  3 Washington        15.4
##  4 Wisconsin         13.6
##  5 Colorado          12.1
##  6 Ohio              10.9
##  7 Vermont            9.27
##  8 Minnesota          8.96
##  9 Oklahoma           8.56
## 10 Oregon             8.49
## # ... with 21 more rows
```