Again, I'm using the plasma color palette from the viridis R package to show hot and cold spots intuitively, and again the color scale for the number of cases is shown on log scale. One nice thing about this color scale (at least as of the time of writing) is that the changes in color correspond pretty nicely to each order of magnitude on the log scale. As with before, this image is set to update daily, so this post should be current throughout the coronavirus pandemic.

One design choice in this animation is different than in the Arkansas visualization. As discussed in the previous post, I elected to use the median county size (rounded to the nearest 5,000) for the per capita calculations. A commenter mentioned that powers of 10 are customary in public health reporting. While I completely agree that's customary, I chose the median value of 20,000 to use for per capita calculations as providing a better intuitive feel for the actual number of cases in most counties in the state without having to do a lot of mental math. There's more explanation in the comments on that post.

For the entire U.S., the median county population is 25,000 (when rounded to the nearest 5,000). However, the mean county population in the U.S. is about 102,000, which is very close to a power of 10 that would customarily be used for public health reporting. As such, I would have a harder time justifying a design choice different than what's customary. So, this animation uses the customary 100,000 figure for per capita calculations.

What do you think about this animation?

The code for the post follows:

```r
library(tidyverse)
library(lubridate)
library(plotly)
library(gganimate)
library(tidycensus)
library(transformr)
library(ggthemes)
library(viridis)

options( scipen = 10 ) # print full numbers, not scientific notation

covid_cases <- read_csv("https://raw.githubusercontent.com/
CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/
time_series_covid19_confirmed_US.csv")
covid_cases <- pivot_longer(covid_cases, 12:length(covid_cases),
names_to = "date", values_to = "cases") %>%
  mutate(date = lubridate::as_date(date, format = "%m/%d/%y", tz =
"UTC"))

population <- tidycensus::get_estimates(geography = "county",
"population") %>%
  mutate(GEOID = as.integer(GEOID)) %>%
  pivot_wider(
    names_from = variable,
    values_from = value
  )
```

```r
# Per capita calculation is to nearest 5k of median county population
per_capita <- population %>%
  summarize(avg = mean(POP)) %>%
  unlist() %>%
  plyr::round_any(., 1e4)

roll_us_cases <- covid_cases %>%
  filter(`Country_Region` == "US" | `Country_Region` == "United
States") %>%
  filter(Province_State != "Puerto Rico") %>%
  filter(FIPS < 80000) %>%
  # filter(Province_State != "Alaska" & Province_State != "Hawaii") %>%
  filter(Admin2 != "Unassigned") %>%
  arrange(date) %>%
  group_by(UID) %>%
  mutate(prev_count = lag(cases)) %>%
  mutate(prev_count = ifelse(is.na(prev_count), 0, prev_count)) %>%
  mutate(new_cases = cases - prev_count) %>%
  mutate(roll_cases = round(zoo::rollapply(new_cases, 7, mean, fill =
0, align = "right", na.rm = T)))%>%
  ungroup() %>%
  select(-prev_count) %>%
  left_join(
    population %>% select(-NAME),
    by = c("FIPS" = "GEOID")
  ) %>%
  mutate(
    cases_capita = round(cases / POP * per_capita), # cases per 100,000
residents
    new_capita = round(new_cases / POP * per_capita), # cases per
100,000 residents
    roll_capita = round(roll_cases / POP * per_capita) # rolling new
cases per 100,000 residents
  )

# tidycensus version
# Includes Alaska and Hawaii as rescaled and shifted
data("county_laea")
data("state_laea")

first_date <- min({ roll_us_cases %>%
    group_by(date) %>%
    summarize(roll_cases = sum(roll_cases)) %>%
    ungroup() %>%
    filter(roll_cases > 0) %>%
    select(date)
}$date)

temp <- roll_us_cases %>%
  filter(date >= first_date) %>%
  mutate(roll_capita = ifelse(roll_capita <= 0, 1, roll_capita)) %>% #
log10 scale plot
```

```r
  mutate(roll_cases = ifelse(roll_cases <= 0, 1, roll_cases)) # log10
scale plot

data("county_laea")
data("state_laea")

temp_sf <- county_laea %>%
  mutate(GEOID = as.numeric(GEOID)) %>%
  mutate(GEOID = ifelse(GEOID == 46113, 46102, GEOID)) %>% # SD Oglala
Lakota County name change
  mutate(GEOID = ifelse(GEOID == 2270, 2158, GEOID)) %>% # AK Kusilvak
census area
  inner_join(temp, by = c("GEOID" = "FIPS"))

days <- NROW(unique(temp$date))

p <- ggplot() +
  geom_sf(data = temp_sf, aes(fill = roll_capita), size = 0) +
  # geom_sf(data = temp_sf, aes(fill = roll_cases), size = 0) +
  geom_sf(data = state_laea, fill = "transparent", color =
alpha("gray70", 0.25), size = 0.75) +
  # scale_fill_gradient(
  #   name = "7-day rolling cases: ",
  #   trans = "log10",
  #   high = "red",
  #   low = "blue"
  # ) +
  scale_fill_viridis(
    name = "7-day rolling cases: ",
    trans = "log10",
    option = "plasma",
  ) +
  ggthemes::theme_map() +
  theme(legend.position = c(0.5, 0.01), legend.direction =
"horizontal") +
  labs(
    title = paste0("US 7-day rolling average of new COVID cases per ",
scales::comma(per_capita), " residents"),
    subtitle = "Date: {frame_time}"
  ) +
  transition_time(date)

anim <- animate(
  p,
  nframes = days + 10 + 30,
  fps = 5,
  start_pause = 10,
  end_pause = 30,
  res = 96,
  width = 800,
  height = 600,
  units = "px"
```

```
)

anim_save("images/us_covid_rolling_cases_plasma.gif", animation = anim)
anim…
```