Lab 6

Morgen Kaufman

Math 241, Week 8

```
# Put all necessary libraries here
library(tidyverse)
library(leaflet)
library(tidycensus)
library(lubridate)
library(tmap)
```

Due: Friday, March 22nd at 8:30am

Goals of this lab

- Practice creating static and interactive choropleth maps.

Problem 1: Mapping Bike Rides in Portland

For this problem we will return to the biketown dataset.

a. Grab the code from activity 9, Problem 1 to read the data directly from Biketown's API- make sure to keep the longitude and latitude of the start of each ride (`StartLatitude`, `StartLongitude`).

```
biketown_data <- bind_rows(readr::read_csv("https://s3.amazonaws.com/biketown-tripdata-public/2017_01.cs
                           readr::read_csv("https://s3.amazonaws.com/biketown-tripdata-public/2017_07.cs
                           readr::read_csv("https://s3.amazonaws.com/biketown-tripdata-public/2017_11.cs

  select(StartDate, StartTime, EndDate, EndTime, Distance_Miles,
         BikeID, StartLatitude, StartLongitude)

biketown_data
```

```
## # A tibble: 75,322 x 8
##    StartDate StartTime EndDate  EndTime Distance_Miles BikeID StartLatitude
##    <chr>     <time>    <chr>    <time>           <dbl>  <dbl>         <dbl>
##  1 1/1/2017  01:36     1/1/2017 01:51             1.35   7398          45.5
##  2 1/1/2017  01:47     1/1/2017 02:00             1.91   7338          45.5
##  3 1/1/2017  01:50     1/1/2017 02:29             2.76   6442          45.5
##  4 1/1/2017  01:51     1/1/2017 02:28             2.77   6638          45.5
##  5 1/1/2017  01:52     1/1/2017 02:28             2.8    6399          45.5
##  6 1/1/2017  01:56     1/1/2017 02:08             1.31   6896          45.5
##  7 1/1/2017  02:08     1/1/2017 02:14             1.01   6831          45.5
```

```
##  8 1/1/2017  02:27     1/1/2017 02:41            1.34   6563          45.5
##  9 1/1/2017  02:42     1/1/2017 02:47            0.57   6152          45.5
## 10 1/1/2017  03:07     1/1/2017 03:26            1.73   6384          45.5
## # i 75,312 more rows
## # i 1 more variable: StartLongitude <dbl>
```

b. Create an interactive map of the start point of the rides using the `leaflet` package. Make sure to include a legend and a title. What do you notice about the distribution of rides?

NOTE: I am not adding a title, I have tried to figure out how to with Adrien but we decided it is easier to just not include one

```
biketown_data %>%
  leaflet() %>%
  addTiles() %>%
  addCircleMarkers(lng = ~StartLongitude, lat = ~StartLatitude)
```

```
leaflet(data = biketown_data) %>%
  addTiles() %>%
  addCircleMarkers(
    ~StartLongitude, ~StartLatitude,
    radius = 2
  ) %>%
  addLegend(
    position = "bottomleft",
    colors = "green",
    labels = "Start Point",
    title = "Ride Start Points"
  )
```

The the data distribution shows a heavy concentration in Downtown PDX and pearl district with decent distribution east of downtown. As we move farther away from the city center, distribution decreases.

c. Using the `lubridate` package, create a variable, `month`, indicating the month of each variable.

Add this variable to your interactive map using color. Make sure to include a legend and be mindful of your color palette choice. Do ride locations vary by months of the year?

```
biketown <- biketown_data %>%
  mutate(StartDate = mdy(StartDate),
         EndDate = mdy(EndDate)) %>%
  mutate(StartDateTime = ymd_hms(paste(StartDate, StartTime, sep = " ")),
         EndDateTime = ymd_hms(paste(EndDate, EndTime, sep = " ")))

biketown$month <- month(biketown$StartDate)


leaflet(data = biketown) %>%
  addTiles() %>%
  addCircleMarkers(
    ~StartLongitude, ~StartLatitude,
    color = ~ifelse(month == 1, "red",
```

```
            ifelse(month == 7, "blue",
            ifelse(month == 11, "green", NA))),
    radius = 2,
    fillOpacity = 0.2
  ) %>%
addLegend(
    position = "bottomleft",
    colors = c("red", "blue", "green"),
    labels = c("January", "July", "November"),
    title = "Months"
  )
```

We can see that the ride locations and ride frequency changed during differnt months. During January which is during winter very few rides can be seen and when they do occur they are closer to the city center. In contrast during July in the summer and Novermber in the fall when the weather is much more plesent, rides seem to be more spread out and occuring in higher numbers.

Problem 2: Choropleth Maps

For this problem, I want you to practice creating choropleth maps. Let's grab some data using `tidycensus`. Remember that you will have to set up an API key.

```
api_key <- "b3c5304d552e16e473b315db0ad7302455075d08"
```

a. Let's grab data on the median gross rent (`B25064_001`) from the American Community Survey for Multnomah county, Oregon. I want you to do data pulls at three geography resolutions: county subdivision, tract, and block group.

```
v21 <- load_variables(2021, "acs5", cache = TRUE)

count <- get_acs(geography = "county subdivision",
              variables = c(gross_rent = "B25064_001"),
              year = 2021,
              state = "OR",
              county = 'Multnomah County',
              survey = "acs5",
              output = "wide",
              geometry = T)
```

```
##   |                                                                              |
```

```
tract <- get_acs(geography = "tract",
              variables = c(gross_rent = "B25064_001"),
              year = 2021,
              state = "OR",
              county = 'Multnomah County',
              survey = "acs5",
              output = "wide",
              geometry = T)
```

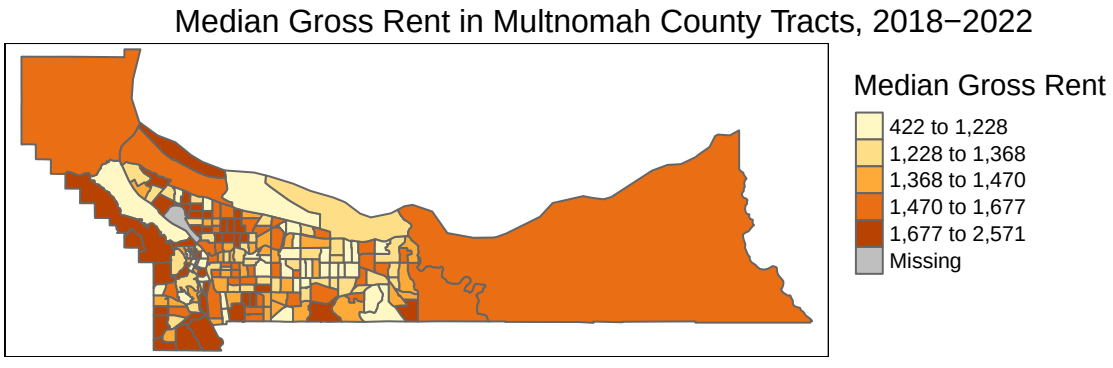
```
##   |                                                                              |
```

```
block <- get_acs(geography = "block group",
                 variables = c(gross_rent = "B25064_001"),
                 year = 2021,
                 state = "OR",
                 county = 'Multnomah County',
                 survey = "acs5",
                 output = "wide",
                 geometry = T)
```

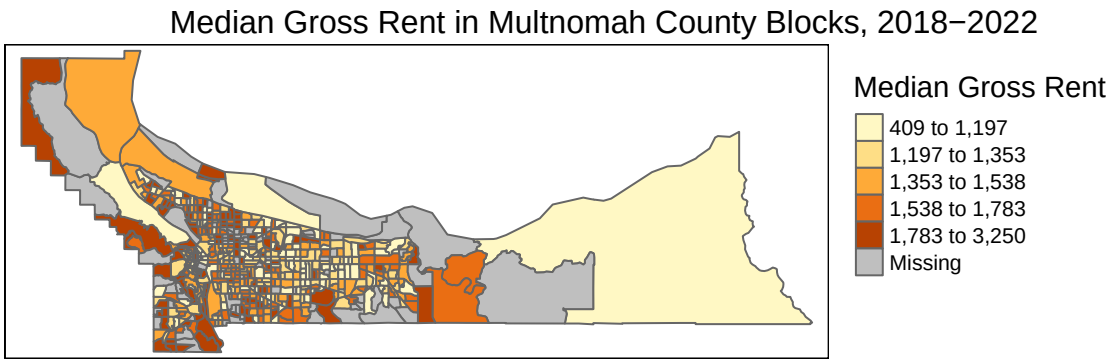
```
##   |                                                                              |
```

b. Create three choropleth maps of gross rent, one for each geography resolution. What information can we glean from these maps? Also, which resolution seems most useful for this variable? Justify your answer.

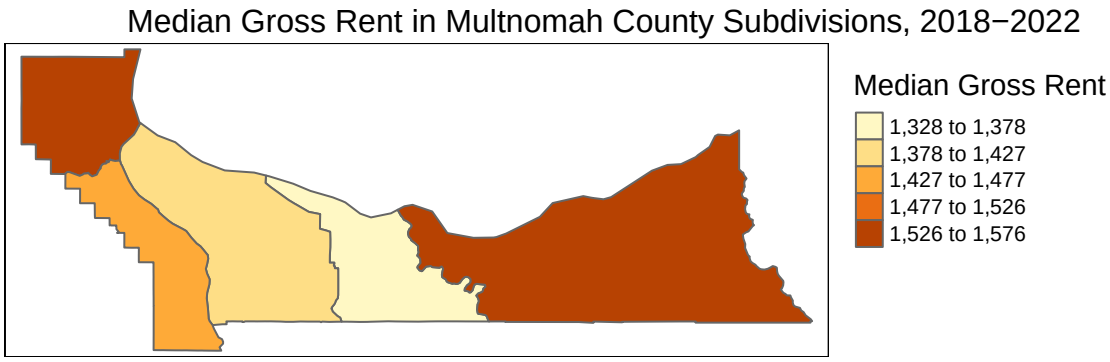
```
tm_shape(tract) +
  tm_polygons(col = "gross_rentE", fill = "gross_rentE", style = "quantile", title = "Median Gross Rent
  tm_layout(main.title = "Median Gross Rent in Multnomah County Tracts, 2018-2022",
            main.title.position = "center",
            legend.outside = TRUE,
            main.title.size = 1)
```



```
tm_shape(block) +
  tm_polygons(col = "gross_rentE", fill = "gross_rentE", style = "quantile", title = "Median Gross Rent
  tm_layout(main.title = "Median Gross Rent in Multnomah County Blocks, 2018-2022",
            main.title.position = "center",
            legend.outside = TRUE,
            main.title.size = 1)
```




```
tm_shape(count) +
  tm_polygons(col = "gross_rentE", fill = "gross_rentE", style = "quantile", title = "Median Gross Rent
  tm_layout(main.title = "Median Gross Rent in Multnomah County Subdivisions, 2018-2022",
            main.title.position = "center",
            legend.outside = TRUE,
            main.title.size = 1)
```



We can see that the county outskirts average rent is quite high. Additionally a number of tracts within central Portland have very high gross rents as well. The central north area of the county tends to be the least expensive in terms of rent however we are also missing a bit of data from this area and thus results may be skewed.

I think that the graphic divided by tract is the most effective at showing these trends. The county subdivision is also quite useful but it may lack some detail in the center and is mainly useful for a quick/brief analysis. The graphic divided by block is way too crowded and detailed and is very difficult to read.