hw_5.Rmd

Group 1

November 5, 2018

HW5 Part 1

```
chicago_homicides <- read_csv("../Data/homicide-data.csv")
```

```
FALSE Parsed with column specification:
FALSE cols(
FALSE   uid = col_character(),
FALSE   reported_date = col_integer(),
FALSE   victim_last = col_character(),
FALSE   victim_first = col_character(),
FALSE   victim_race = col_character(),
FALSE   victim_age = col_character(),
FALSE   victim_sex = col_character(),
FALSE   city = col_character(),
FALSE   state = col_character(),
FALSE   lat = col_double(),
FALSE   lon = col_double(),
FALSE   disposition = col_character()
FALSE )
```

```
chicago_homicides <- chicago_homicides %>%
  filter(city == "Chicago") %>%
  mutate(unsolved = disposition %in% c("Closed without arrest",
                                       "Open/No arrest")) %>%
  mutate(top_3_races = fct_lump(victim_race, n = 3)) %>%
  filter(!is.na(lat)) %>%
  filter(!is.na(lon))

chicago_homicides <- st_as_sf(chicago_homicides, coords = c("lon", "lat")) %>%
  st_set_crs(4267)

chicago_subdiv <- tracts(state = "IL", cb = TRUE, class = "sf")
```

```
FALSE
  |
  |                                                                      |   0%
  |
  |=                                                                     |   1%
  |
  |=                                                                     |   2%
  |
  |==                                                                    |   3%
  |
  |===                                                                   |   4%
  |
  |===                                                                   |   5%
  |
  |====                                                                  |   6%
```

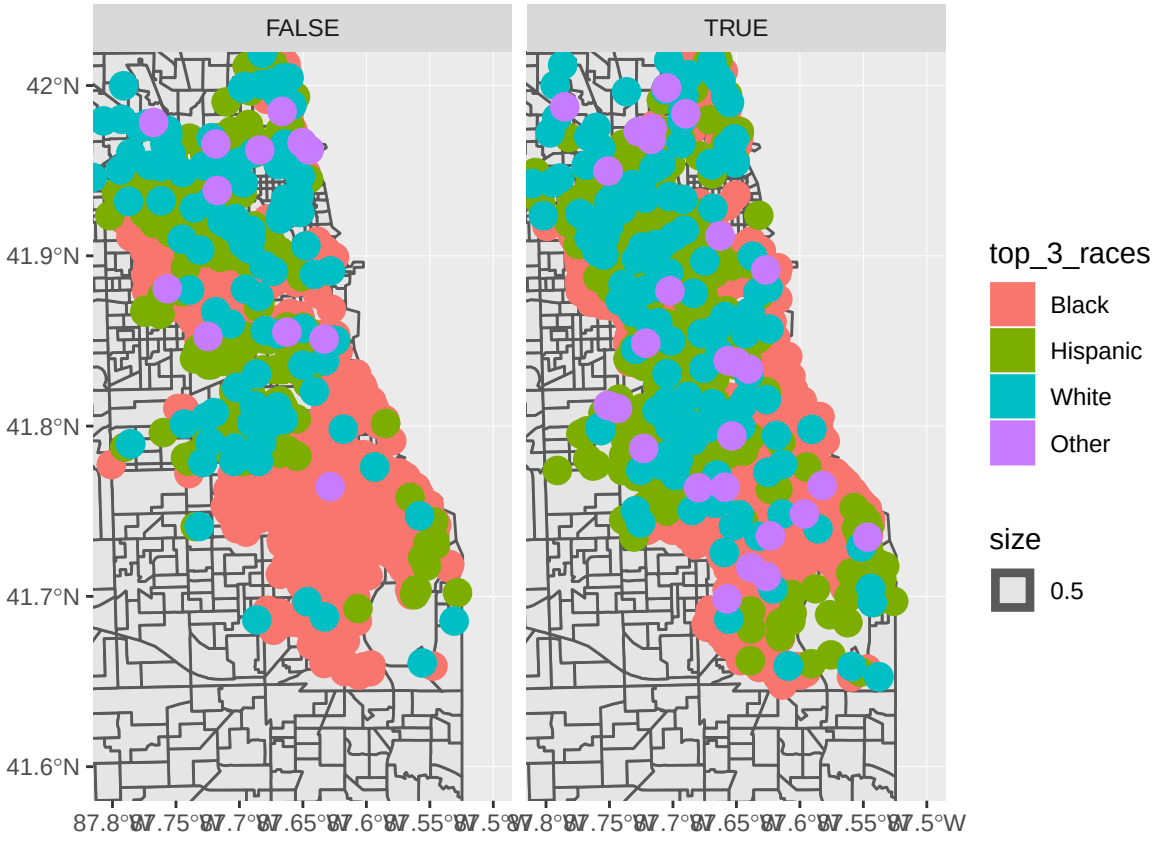
```
|
|=====                                                                 |   7%
|
|=====                                                                 |   8%
|
|======                                                                |   9%
|
|=======                                                               |  10%
|
|=======                                                               |  11%
|
|========                                                              |  12%
|
|=========                                                             |  13%
|
|=========                                                             |  14%
|
|==========                                                            |  15%
|
|===========                                                           |  16%
|
|===========                                                           |  17%
|
|============                                                          |  18%
|
|============                                                          |  19%
|
|=============                                                         |  20%
|
|==============                                                        |  21%
|
|==============                                                        |  22%
|
|===============                                                       |  23%
|
|================                                                      |  24%
|
|=================                                                     |  25%
|
|=================                                                     |  26%
|
|==================                                                    |  28%
|
|===================                                                   |  30%
|
|====================                                                  |  30%
|
|====================                                                  |  32%
|
|======================                                                |  34%
|
|=======================                                               |  36%
|
|=======================                                               |  37%
```

```
|
|========================                                 |  38%
|
|========================                                 |  39%
|
|=========================                                |  40%
|
|=========================                                |  41%
|
|==========================                               |  42%
|
|==========================                               |  43%
|
|===========================                              |  44%
|
|===========================                              |  45%
|
|============================                             |  46%
|
|============================                             |  47%
|
|=============================                            |  49%
|
|=============================                            |  50%
|
|==============================                           |  51%
|
|==============================                           |  52%
|
|===============================                          |  53%
|
|===============================                          |  54%
|
|================================                         |  55%
|
|================================                         |  56%
|
|=================================                        |  57%
|
|=================================                        |  58%
|
|==================================                       |  59%
|
|====================================                     |  61%
|
|=====================================                    |  63%
|
|=====================================                    |  64%
|
|======================================                   |  65%
|
|======================================                   |  66%
|
|=======================================                  |  67%
```

```
|
|=========================================                      |  68%
|
|==========================================                     |  69%
|
|==========================================                     |  70%
|
|===========================================                    |  71%
|
|===========================================                    |  72%
|
|============================================                   |  73%
|
|============================================                   |  74%
|
|=============================================                  |  75%
|
|=============================================                  |  76%
|
|==============================================                 |  77%
|
|==============================================                 |  78%
|
|===============================================                |  79%
|
|===============================================                |  80%
|
|================================================               |  81%
|
|================================================               |  82%
|
|=================================================              |  83%
|
|=================================================              |  84%
|
|==================================================             |  85%
|
|==================================================             |  86%
|
|===================================================            |  87%
|
|=====================================================          |  89%
|
|======================================================         |  91%
|
|========================================================       |  93%
|
|========================================================       |  93%
|
|==========================================================     |  95%
|
|============================================================   |  97%
|
|============================================================   |  98%
```

```
  |
  |======================================================================| 99%
  |
  |======================================================================| 100%
```

```
ggplot(chicago_homicides) +
  geom_sf(data = chicago_subdiv) +
  geom_sf(data = chicago_homicides, aes(color = top_3_races, fill = top_3_races,
                                        size = 0.5)) +
  xlim(c(-87.8, -87.5)) +
  ylim(c(41.6, 42)) +
  facet_wrap(~ unsolved)
```



HW5 Part 2

```
homicide <- read_csv('../Data/homicide-data.csv')
```

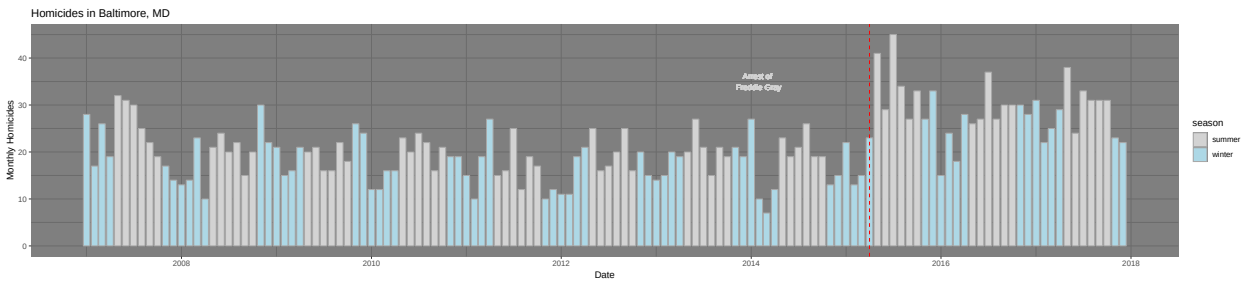
```
## Parsed with column specification:
## cols(
##   uid = col_character(),
##   reported_date = col_integer(),
##   victim_last = col_character(),
##   victim_first = col_character(),
##   victim_race = col_character(),
##   victim_age = col_character(),
##   victim_sex = col_character(),
##   city = col_character(),
##   state = col_character(),
```

```
##   lat = col_double(),
##   lon = col_double(),
##   disposition = col_character()
## )
```

```
homicide <- homicide %>%
  unite(cityname, city, state, sep = ", ") %>%
  select(reported_date, cityname, victim_last, victim_first) %>%
  filter(cityname == "Baltimore, MD") %>%
  mutate(reported_date = ymd(reported_date)) %>%
  mutate(reported_date = floor_date(reported_date, unit = "month")) %>%
  mutate(year = year(reported_date),
         month = month(reported_date)) %>%
  mutate(season = recode(month,
                         '5' = "summer",
                         '6' = 'summer',
                         '7' = 'summer',
                         '8' = 'summer',
                         '9' = 'summer',
                         '10' = 'summer',
                         '11' = 'winter',
                         '12' = 'winter',
                         '1' = 'winter',
                         '2' = 'winter',
                         '3' = 'winter',
                         '4' = 'winter')) %>%
  group_by(reported_date, year, month, season) %>%
  count() %>%
  ungroup () %>%
  rename("homicide_count" = 'n') %>%
  arrange(year, month) %>%
  mutate(date_numeric = as.numeric(reported_date))

homicideplot <- homicide %>%
  ggplot() +
  geom_col(mapping = aes(x = reported_date, y = homicide_count, fill = season), color = "darkgrey") +
  theme_dark() +
  scale_fill_manual(values=c("lightgrey", "lightblue")) +
  labs(x = "Date", y = "Monthly Homicides") +
  ggtitle("Homicides in Baltimore, MD") +
  scale_x_date(date_breaks = "2 year", date_labels = "%Y") +
  geom_vline(xintercept = 16526, colour = "red", linetype = "dashed") +
  geom_text(x = 16100, y = 35, label = "Arrest of \nFreddie Gray", color = "lightgrey", size = 3)

homicideplot
```



HW5 Part 3

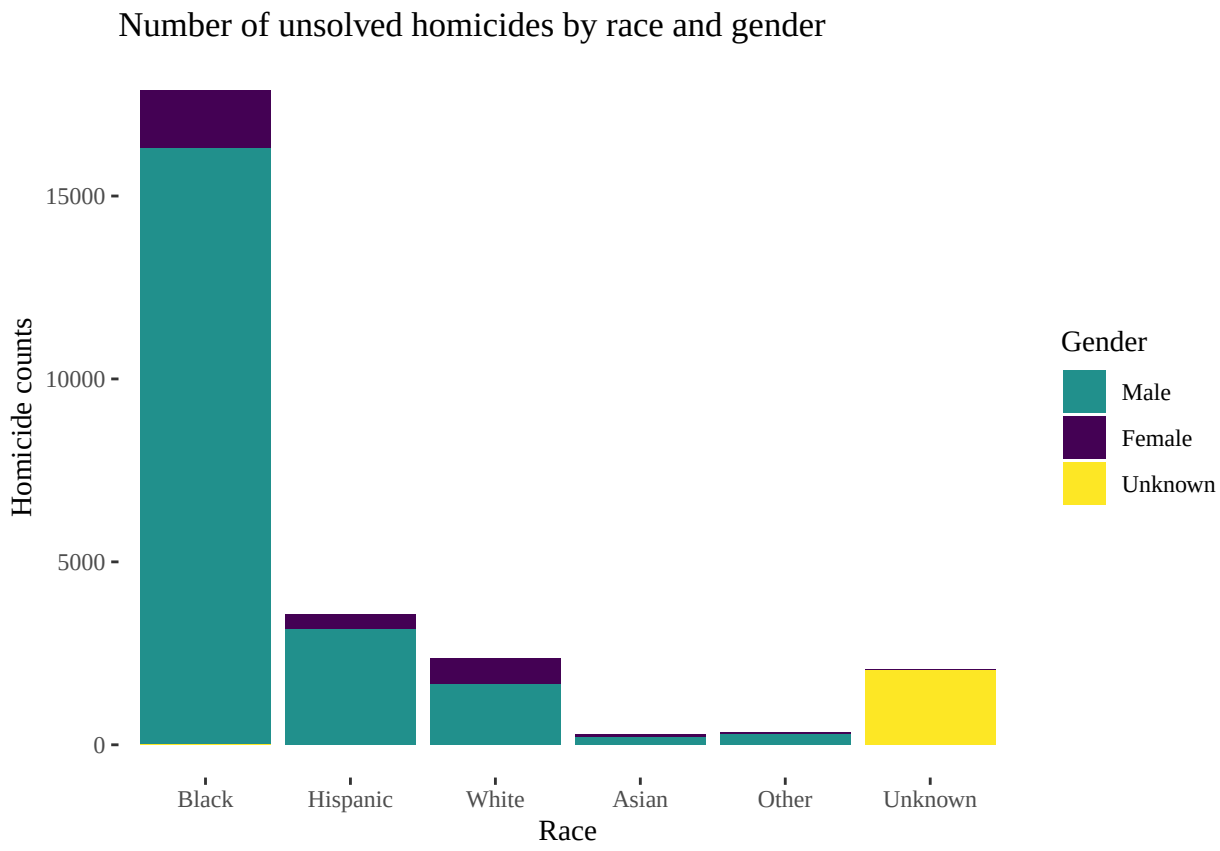
Group 5 was curious about the amount of unsolved homicides in regards to race and gender. The below graph was generated to explore this question. The data was ordered from the highest count of unsolved homicides to the lowest count of unsolved homicides by race. A stacked bargraph was used to depict the relative amount of female to male unsolved homicides for each race. The results of the graph are very striking in that unsolved homicides of black men appear to far outnumber any of the other races or female homicides. Furthermore, males made up the majority of unsolved homicides for each race. It is interesting to note that outside of the three highest races (black, hispanic, and white), asian and other races returned very minimal counts of unsolved homicides. Further investigation may be necessary to control for confounding including the relative population of each race by city.

```
homicides2 <- read_csv("../Data/homicide-data.csv")

homicides2 <- homicides2 %>%
  mutate(unsolved = disposition == "Closed without arrest" | disposition == "Open/No arrest",
         reported_date = date(ymd(reported_date))) %>%
  unite(city_name,
        city,
        state,
        sep = ", ") %>%
  select(-uid,
         -victim_last,
         -victim_first,
         -disposition)
```

```
homicides2plot <- homicides2 %>%
  filter(unsolved == TRUE) %>%
  ggplot() +
  geom_bar(aes(x = victim_race,
               fill = victim_sex)) +
  scale_x_discrete(limits = c("Black", "Hispanic", "White", "Asian", "Other", "Unknown")) +
  labs(x = "Race",
       y = "Homicide counts",
       title = "Number of unsolved homicides by race and gender") +
  theme_tufte() +
  scale_fill_viridis_d(name = "Gender",
                       breaks = c("Male", "Female", "Unknown"))

homicides2plot
```


Number of unsolved homicides by race and gender
Homicide counts
15000
10000
5000
0
Black
Hispanic
White
Asian
Other
Unknown
Race
Gender
Male
Female
Unknown