hw_5.Rmd

Group 1

*November 5, 2018*

HW5 Part 1

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


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")
```

```r
ggplot(chicago_homicides) +
  geom_sf(data = chicago_subdiv) +
  geom_sf(data = chicago_homicides, aes(color = top_3_races, fill = top_3_races), size = 1.5) +
  xlim(c(-87.8, -87.5)) +
  ylim(c(41.6, 42)) +
  facet_wrap(~ unsolved)
```
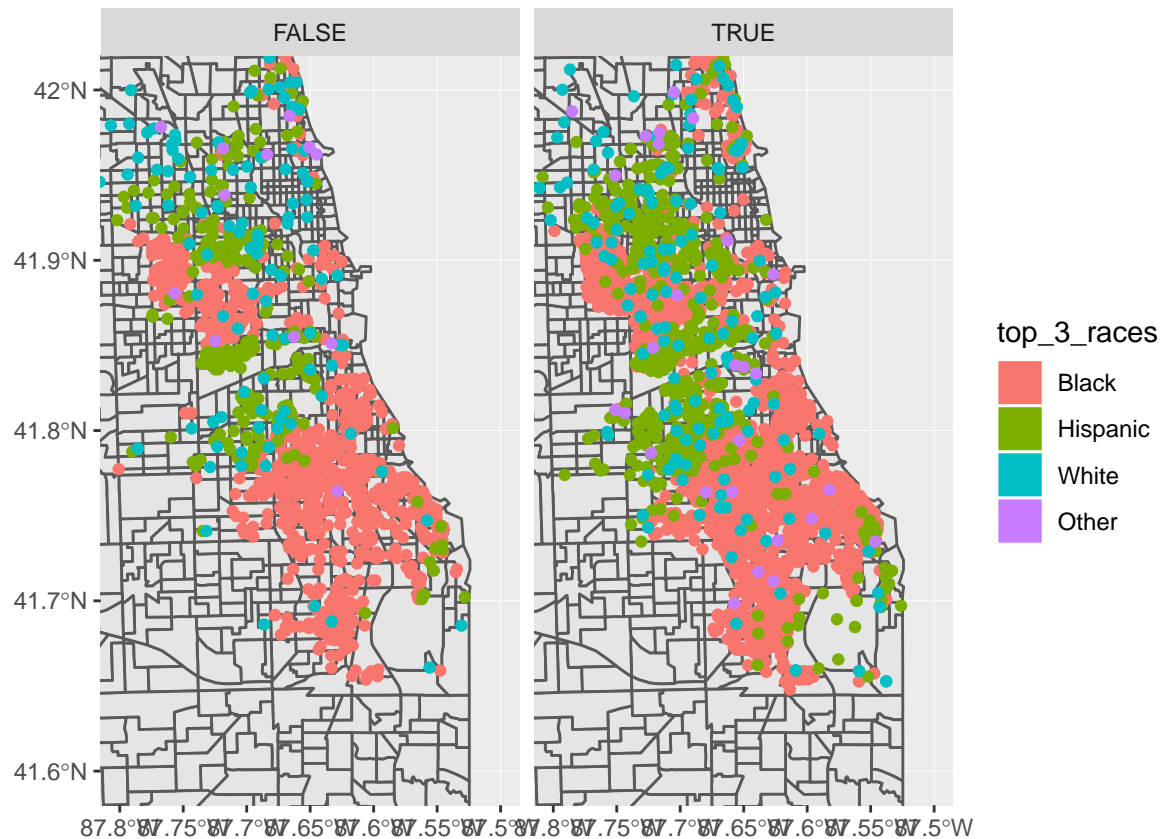
HW5 Part 2

The team was intersted in understanding how the average number of homicides per month may have changed after the arrest of Freddie Gray. A bar graph with the number of homicides per month was used to depict how the number of homicides changes across time. The arrest of Freddie Gray was marked by a red dashed line. It is apparent that the number of homicides increased on average each month after the arrest of Freddie Gray.

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

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") +
  geom_smooth(aes(x = reported_date, y = homicide_count), method = "loess", se = FALSE, span = 0.1) +
  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", size = 1) +
  geom_text(x = 16400, y = 35, label = "Arrest of \nFreddie Gray", color = "lightgrey", size = 4) +
  theme(legend.position='bottom')
```
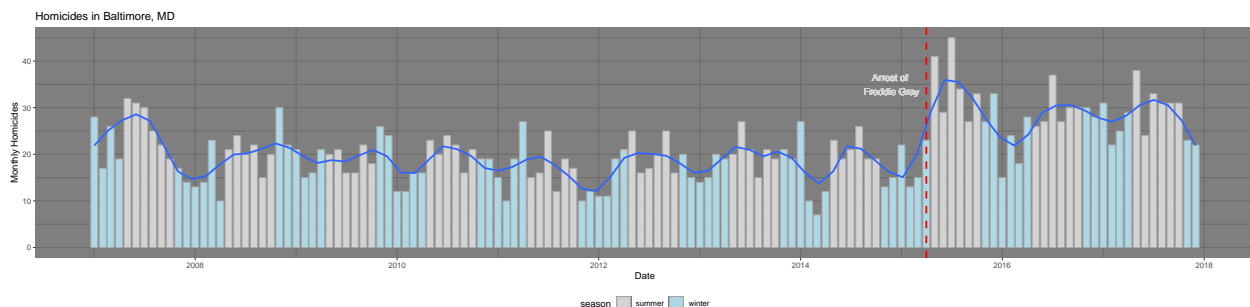
```
homicideplot
```


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