homework 5

Maelle Coupannec

17/11/2021

LOADING LIBRARY

We are loading some library to clean the data and create the map.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.3     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(forcats)
library(tigris)
```

```
## To enable
## caching of data, set `options(tigris_use_cache = TRUE)` in your R script or .Rprofile.
```

```
library(viridis)
```

```
## Loading required package: viridisLite
```

```
library(sf)
```

```
## Linking to GEOS 3.9.1, GDAL 3.2.1, PROJ 7.2.1
```

```
library(ggthemes)
library(tinytex)
```

Reading And Cleaning The Data

First we read the data file which is in the data sub directory of the project file. Using fct_lump() we only keep the three races with the highest number of homicides for that city, the rest are assigned to others and are filtered out. Finally, the disposition factor are reassigned to solved or unsolved for uniformity.

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

```
## Rows: 52179 Columns: 12
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (9): uid, victim_last, victim_first, victim_race, victim_age, victim_sex...
## dbl (3): reported_date, lat, lon
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
homicides <- homicides %>%
  unite(col = city_name, c(city, state), sep=", ") %>%
  filter(city_name=="Las Vegas, NV") %>%
  mutate(victim_race= fct_lump(victim_race, n=3)) %>%
  filter(!is.na(lon)|!is.na(lon)) %>%
  filter(victim_race!="Other") %>%
  mutate(disposition=fct_recode(disposition,
                                solved = "Closed by arrest",
                                unsolved = "Closed without arrest",
                                unsolved ="Open/No arrest"))


head(homicides)
```

```
## # A tibble: 6 x 11
##   uid        reported_date victim_last victim_first victim_race victim_age victim_sex
##   <chr>              <dbl> <chr>       <chr>        <fct>       <chr>      <chr>
## 1 Las-000001      20070103 PICKFORD    NOEL         Black       19         Male
## 2 Las-000002      20070104 AFSHARE     ELHAM        White       27         Female
## 3 Las-000003      20070105 HICKMAN     JAMAR        Black       23         Male
## 4 Las-000005      20070107 TURNER      ASHLEY       Black       20         Female
## 5 Las-000006      20070109 HENDERSON   MICHAEL      Black       25         Male
## 6 Las-000007      20070114 SCHMALFELDT BRENDA       White       46         Female
## # ... with 4 more variables: city_name <chr>, lat <dbl>, lon <dbl>,
## #   disposition <fct>
```

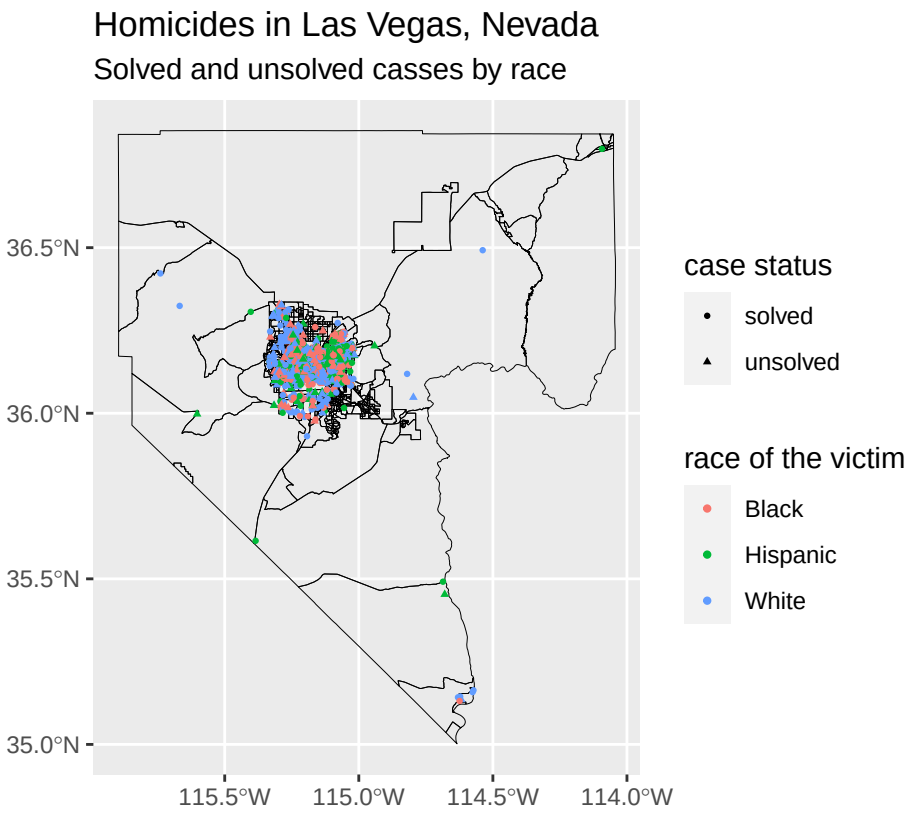
Creation Of The Map

The map SF shape is downloaded from the Tigris package, it was chosen to upload the block contour of Clark county as boundaries for Las vegas geography. Points showing homicides were added on the map. Different facets for solved versus unsolved homicides and different colors to show the three race groups were used.

```
map_lv <- block_groups( state="NV", county = "Clark county",
                  year = NULL, class="sf")
```

```
##   |                                                                      |
```

```
ggplot()+
  geom_sf(data = map_lv, color = "black",
          alpha = 0.1, size = 0.1)+
  geom_point(data = homicides,
             aes(x = lon, y = lat,
                 shape = disposition,
                 color = victim_race), size = 0.8) +
  ggtitle("Homicides in Las Vegas, Nevada",
          subtitle = "Solved and unsolved casses by race")+
  labs(x = "", y = "", shape= "case status",
       color= "race of the victim")
```



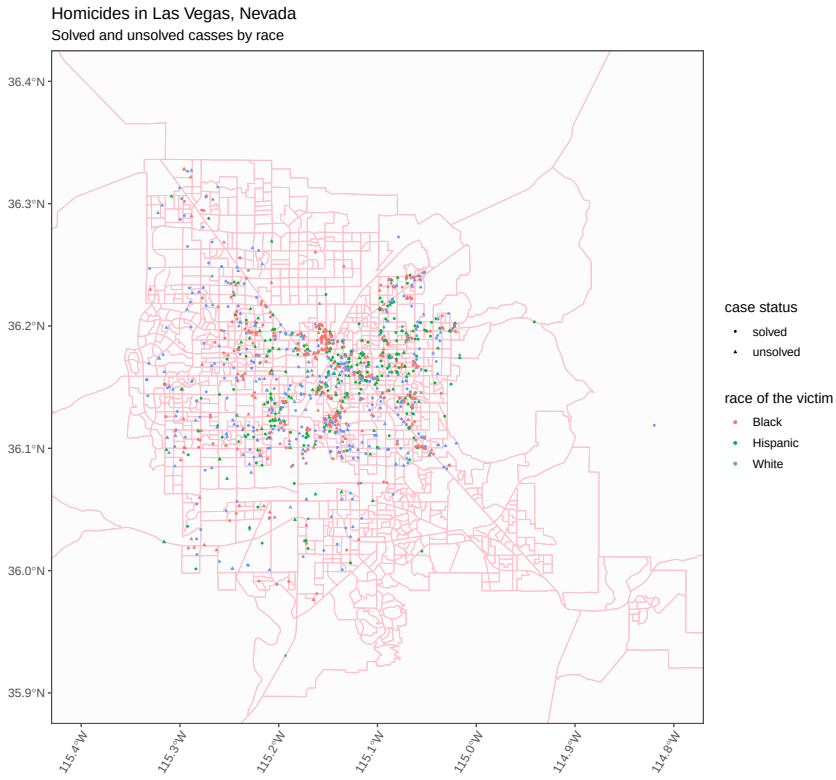
Second Map Sowing only the city area

Since the first map shown the whole county and points where on top of each other, the geographic coordinate were restricted to the city area.

```
ggplot()+
  geom_sf(data = map_lv, color = "pink",
          alpha = 0.1, size = 0.1)+
  geom_point(data = homicides,
```

```
            aes(x = lon, y = lat,
               shape = disposition,
               color = victim_race), size = 0.8)+
  xlim(c(-115.4, -114.8)) + ylim(c(35.9, 36.4))+
  ggtitle("Homicides in Las Vegas, Nevada",
          subtitle = "Solved and unsolved casses by race")+
  theme_few() +
  labs(x = "", y = "", shape= "case status",
       color= "race of the victim")+
  theme(axis.text.x = element_text(angle = 60, hjust = 1))
```

```
## Warning: Removed 16 rows containing missing values (geom_point).
```


Homicides in Las Vegas, Nevada
Solved and unsolved casses by race
36.4°N
36.3°N
36.2°N
36.1°N
36.0°N
35.9°N
115.4°W
115.3°W
115.2°W
115.1°W
115.0°W
114.9°W
114.8°W
case status
solved
unsolved
race of the victim
Black
Hispanic
White