

Week 48 assignment – Mathias Thomasen Madsen

2) Practice Web Scraping of Police Killing data from the Internet Archive, following the .Rmd file in this tutorial on web scraping data on US police killings <https://github.com/Digital-Methods-HASS/WebscrapingPoliceKillings>, Clone it and produce data visualisations.

I choose to do the 2.1 assignment.

2.1) produce data visualisations that shed light on another interesting aspect of the police killing data, expanding the present analysis to 2020 or exploring another aspect or type of visualisation. Declare your intentions up front, so it is clear how your work builds on this repository.

I used Adela's script to make this assignment. I wrote line 302-319.

In this assignment, I wanted to make a visualization, that shows, how the methods of police killings relate to race. I found that I could make a statistical graph, by using Adela's chunks, and altering them, in her rmarkdown script. First, I found a chunk, that would create a statistical graph:

```
261 - ```{r plot-method}
262 data %>%
263   filter(!is.na(Year)) %>%
264   filter(Method != "NA") %>%
265   filter(Gender %in% c("M", "F", NA)) %>%
266   group_by(Year,
267             Gender,
268             Method) %>%
269   tally() %>%
270   mutate(perc = n / sum(n) * 100) %>%
271   ggplot(aes(Method,
272             perc,
273             fill = Gender)) +
274   geom_col() +
275   facet_grid(Gender~Year) +
276   theme_minimal(base_size = 10) +
277   xlab("Method of killing") +
278   ylab("Percentage of all\npeople killed by police\nby gender")
279 - ```
```

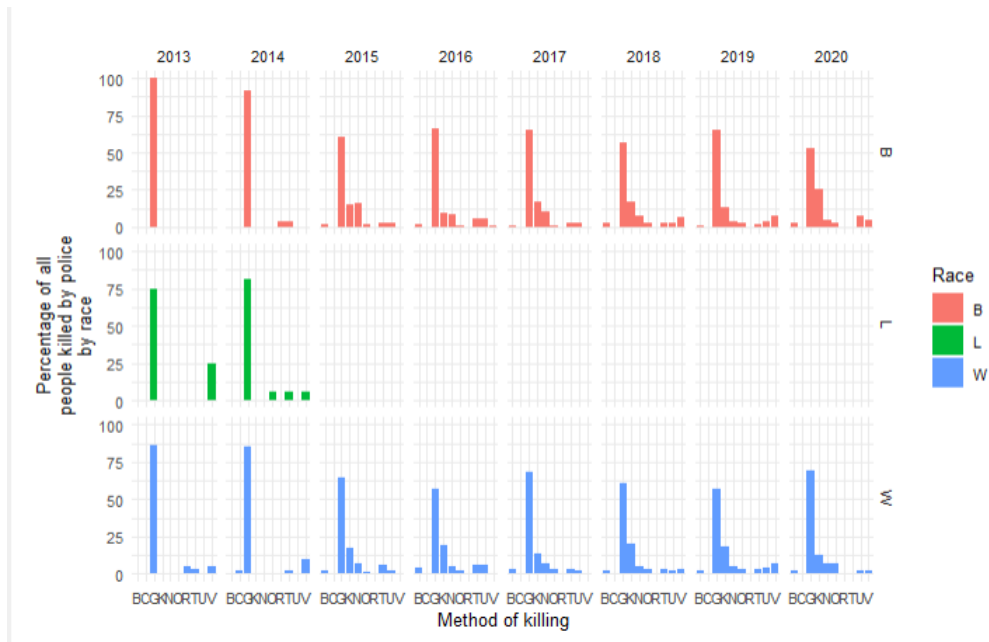
This chunk created a statistical graph showing, the percentage of all people killed by the police, based on what methods were used to kill them, and relating to gender. What I needed to do then, was to replace the gender fill, with race. In the below screenshot, you can see how I replaced gender with race (compare with the screenshot above). Notice, that in `filter(race%in%c("B", "W", "L"))` I wrote "B", "W", "L" which means I made the graph only show the percentage killed of blacks, whites and latinos, to make the graph tidier and better visually.

```

302 {r plot-method-race}
303 data %>%
304   filter(!is.na(Year)) %>%
305   filter(Method != "NA") %>%
306   filter(Race %in% c("B", "W", "L")) %>%
307   group_by(Year,
308     Race,
309     Method) %>%
310   tally() %>%
311   mutate(perc = n / sum(n) * 100) %>%
312   ggplot(aes(Method,
313     perc,
314     fill = Race)) +
315   geom_col() +
316   facet_grid(Race~Year) +
317   theme_minimal(base_size = 10) +
318   xlab("Method of killing") +
319   ylab("Percentage of all people killed by police by race")

```

This is the result of the chunk if you run it in the rmarkdown:



As can be seen, it is a statistical visualization, of what I wanted to show.

The screenshots can be found in my github repository, using my link:

https://github.com/Digital-Methods-HASS/au681088_Madsen_Thomassen_Mathias