HadleyFigueroa_7.3_RMarkdown_Images and Tables.RMD

Hadley Figueroa

2024-06-29

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.4.1
```

```
## Warning: package 'ggplot2' was built under R version 4.4.1
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
if(!require(colorspace)) {install.packages('colorspace')}; require(colorspace)
```

```
## Loading required package: colorspace
```

```
## Warning: package 'colorspace' was built under R version 4.4.1
```

```
library(readxl)
df <- read_excel("Sampledata2.xlsx")

view(df)

##Create a table
knitr::kable(head(df,10), caption = "\\label{tab:table1}The first 10 rows of the `Crime Rate` data")
```

Table 1: The first 10 rows of the `Crime Rate` data

State	Year	CrimeRate
Alabama	2007	448.0
Alabama	2008	452.8
Alabama	2009	450.1
Alabama	2010	377.8

State	Year	CrimeRate
Alabama	2011	419.8
Alabama	2012	449.9
Alabama	2013	431.0
Alabama	2014	427.4
Alabama	2015	473.0
Alabama	2016	532.3

```
##Create table with statistical values per year
grp_df <- df %>%
  group_by(Year) %>%
  summarize(
  mean_CrimeRate = mean(CrimeRate, na.rm = TRUE),
  max_CrimeRate = max(CrimeRate, na.rm = TRUE),
  min_CrimeRate = min(CrimeRate, na.rm = TRUE))

knitr::kable(head(grp_df), caption = "\\label{tab:table2}The mean/min/max of the `Crime Rate` data by Y
```

Table 2: The mean/min/max of the `Crime Rate` data by Year

Year	mean_CrimeRate	max_CrimeRate	min_CrimeRate
2007	407.460	788.3	118.0
2008	399.982	729.7	117.5
2009	387.688	705.2	119.9
2010	365.590	660.6	122.0
2011	354.274	610.1	123.3
2012	353.826	643.6	122.7

```
##Plot a grouped histogram for the variable "CrimeRate" based on Year, with the original data
n <- length(unique(df$Year))
col1=colorspace::diverge_hcl(n)

df %>%
  ggplot(aes(x=CrimeRate, color=as.factor(Year), fill=as.factor(Year))) +
  geom_histogram(position="dodge", alpha=0.5, bins=10) +
  scale_color_manual(values=col1) +
  scale_fill_manual(values=col1) +
  labs(fill="Year", color="Year", title="Crime Rate Histogram Plots", x="CrimeRate", y="Count") +
  theme_classic()
```

```
ggplot(grp_df, aes(x = Year, y = mean_CrimeRate)) +
  geom_point() +
  geom_line() +
  labs(title = "Mean Crime Rate by Year",
       x = "Year",
       y = "Mean Crime Rate") +
  theme_classic()
```

In Table 1 we see the first 10 rows of the Crime Rate data". In Table 2 we see the basic statistical values of crime rates over USA for each year. In Figure 1 we can see the grouped histogram for each year. In Figure 2 we can see the average crime rate changes along with years.

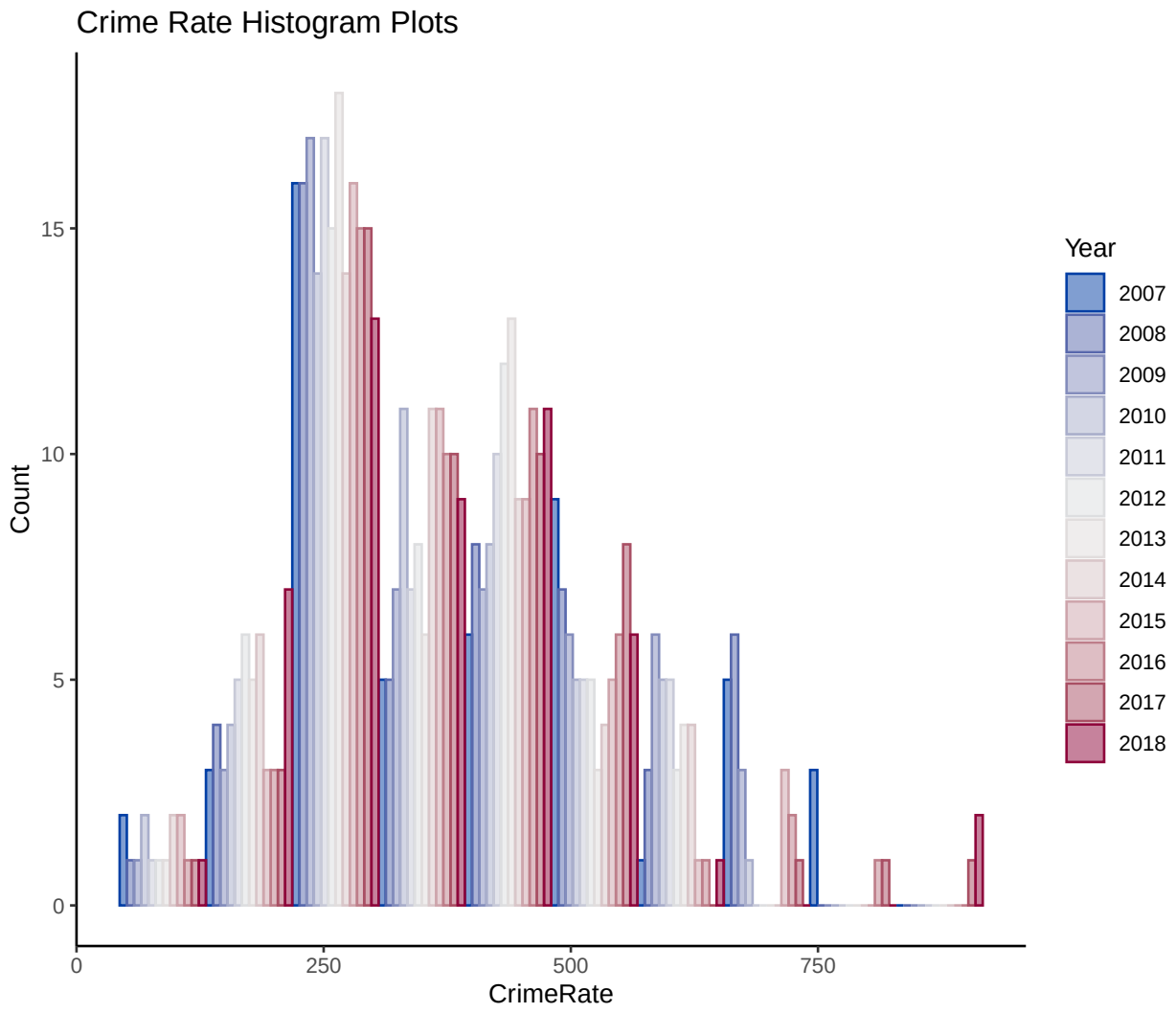

Figure 1: A Histogram of the Crime Rate in the US by Year

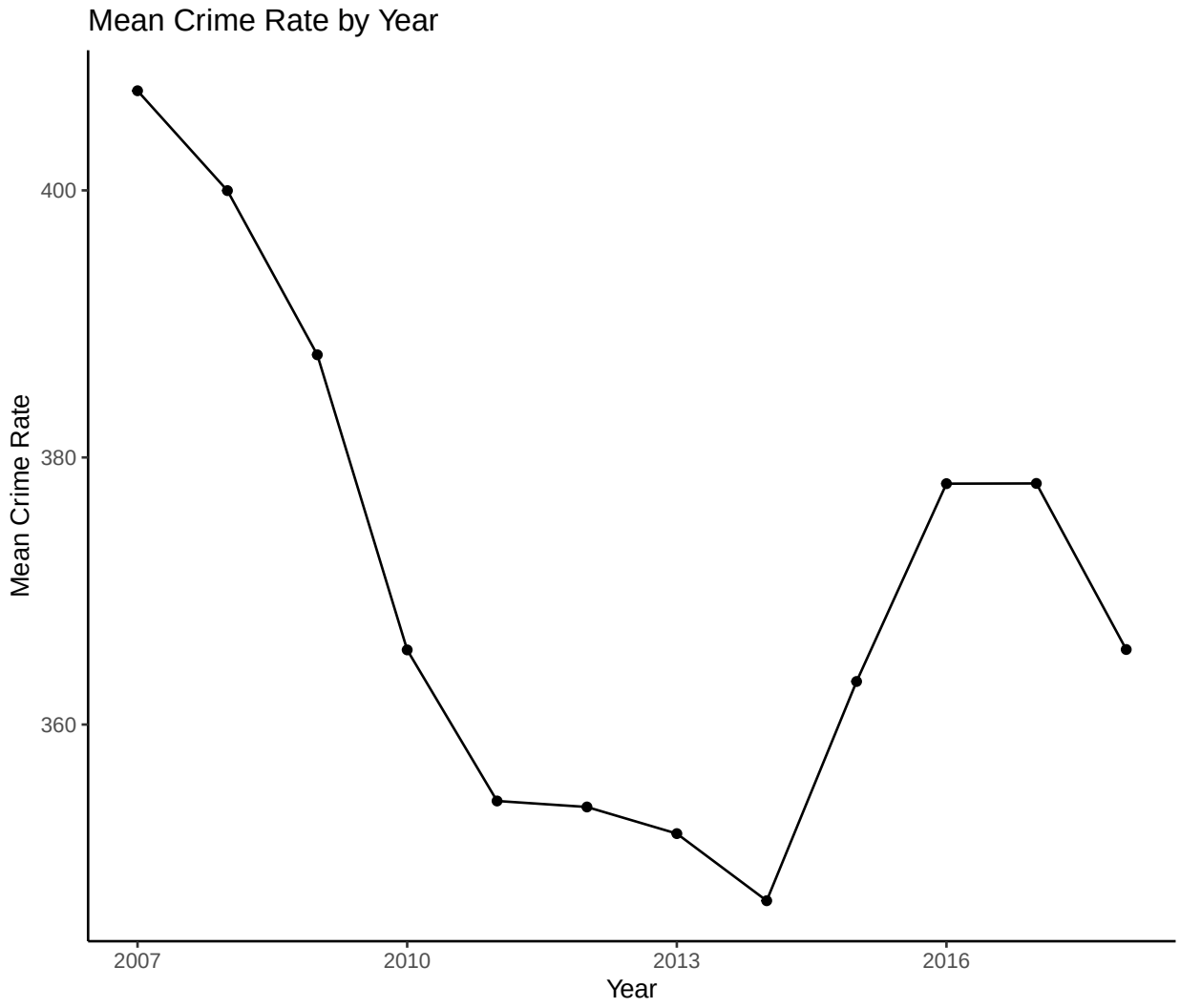

Figure 2: A Point-line Plot of the Mean Crime Rate in the US by Year