Graphics

Yehzee Ryoo

11/19/2020

Create the following graphs in `ggplot2`.

1. Check out the base R built-in dataset, `data("USArrests")`.

```
setwd("~/Library/Mobile Documents/com~apple~CloudDocs/1. PS materials/PS 811/Week7")

#library("here")
library("tidyverse")
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

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

```
library("tidylog") # gives feedback about data manipulation
```

```
##
## Attaching package: 'tidylog'
```

```
## The following objects are masked from 'package:dplyr':
##
##      add_count, add_tally, anti_join, count, distinct, distinct_all,
##      distinct_at, distinct_if, filter, filter_all, filter_at, filter_if,
##      full_join, group_by, group_by_all, group_by_at, group_by_if,
##      inner_join, left_join, mutate, mutate_all, mutate_at, mutate_if,
##      relocate, rename, rename_all, rename_at, rename_if, rename_with,
##      right_join, sample_frac, sample_n, select, select_all, select_at,
##      select_if, semi_join, slice, slice_head, slice_max, slice_min,
##      slice_sample, slice_tail, summarise, summarise_all, summarise_at,
##      summarise_if, summarize, summarize_all, summarize_at, summarize_if,
##      tally, top_frac, top_n, transmute, transmute_all, transmute_at,
##      transmute_if, ungroup
```

```
## The following objects are masked from 'package:tidyr':
##
##      drop_na, fill, gather, pivot_longer, pivot_wider, replace_na,
##      spread, uncount
```
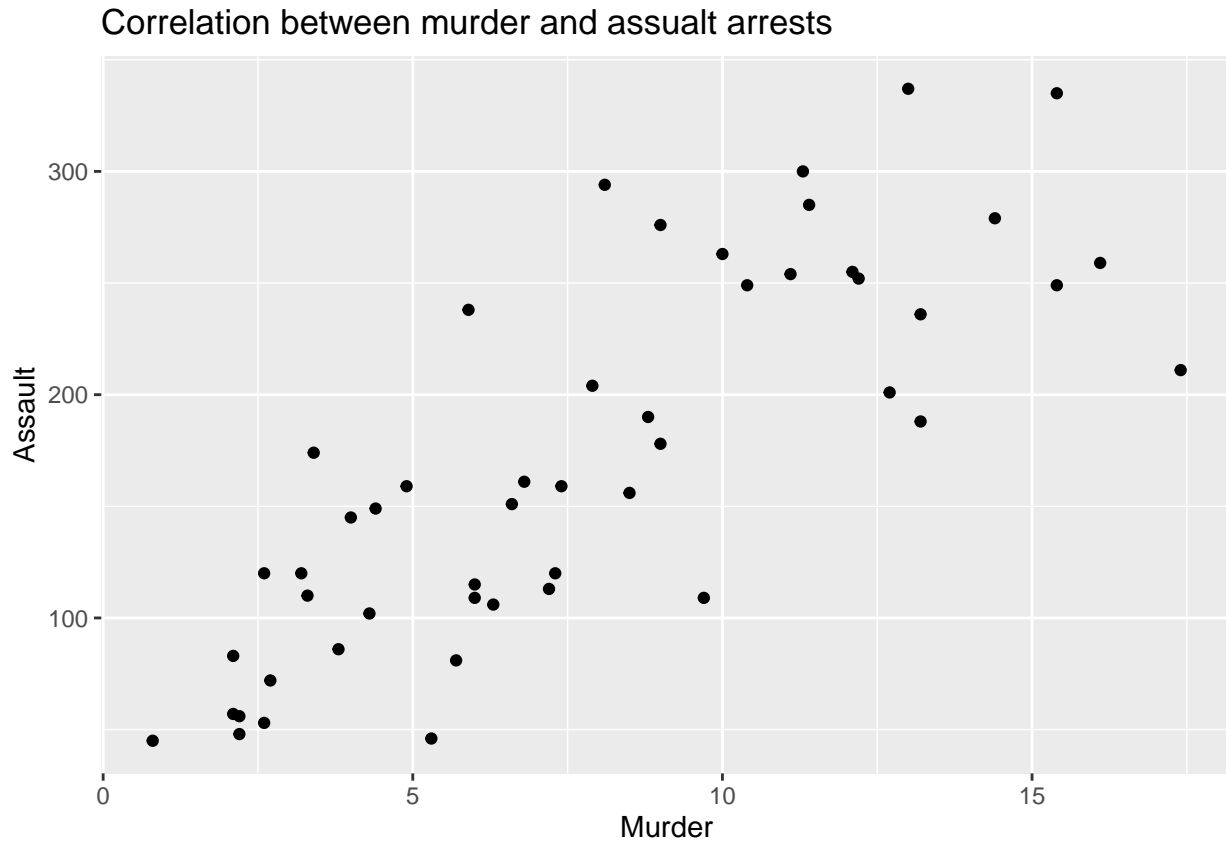
```
## The following object is masked from 'package:stats':
##
##      filter
```

```
US <- datasets::USArrests
```

2. Create a scatterplot that looks at the correlation between murder and assault arrests. Label the x and y axes and title the graph.

```
p_scatter <- ggplot(US, aes(x=Murder, y=Assault)) + geom_point() + labs(x= "Murder", y="Assault", title=

p_scatter
```

Correlation between murder and assualt arrests
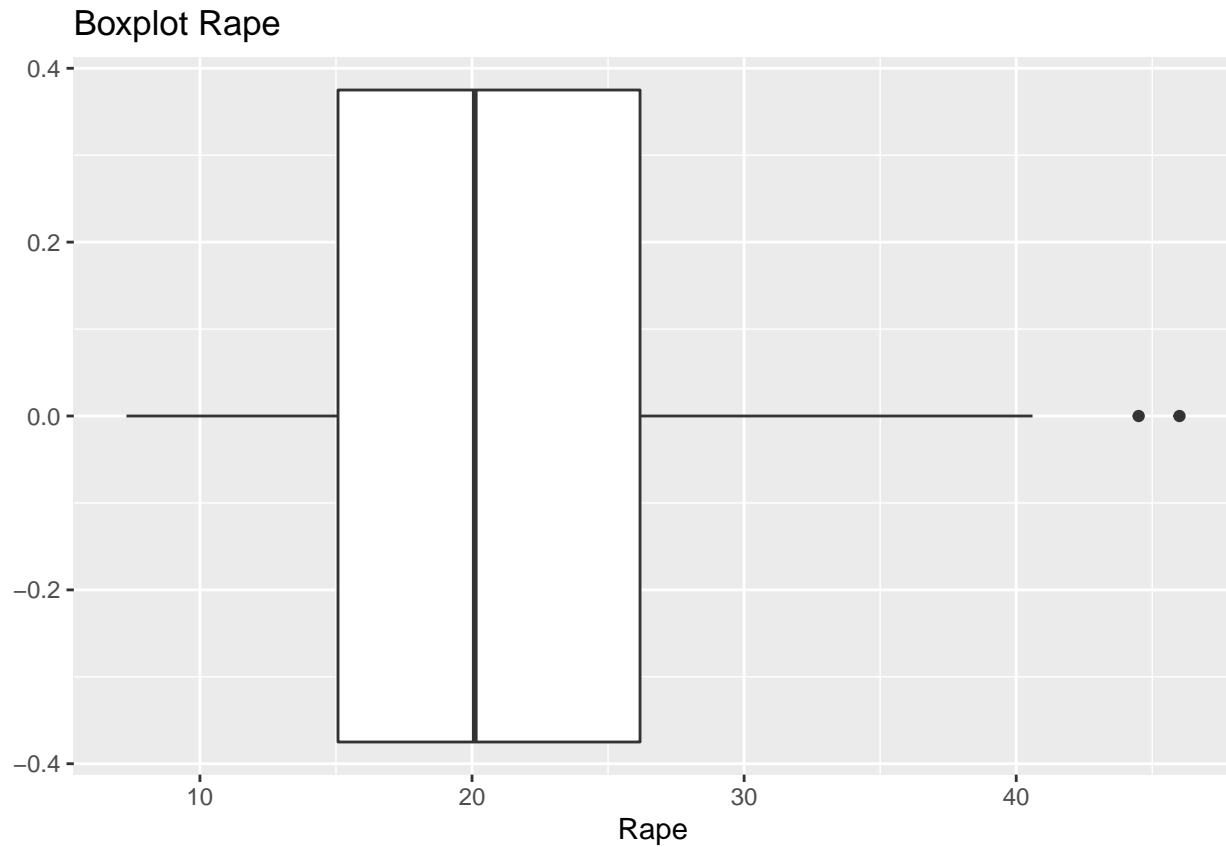

3. Create a boxplot of rape arrests. Label the plot.

```
p_box <-ggplot(US, aes(x=Rape)) + geom_boxplot() + labs(x= "Rape", title="Boxplot Rape")

p_box
```
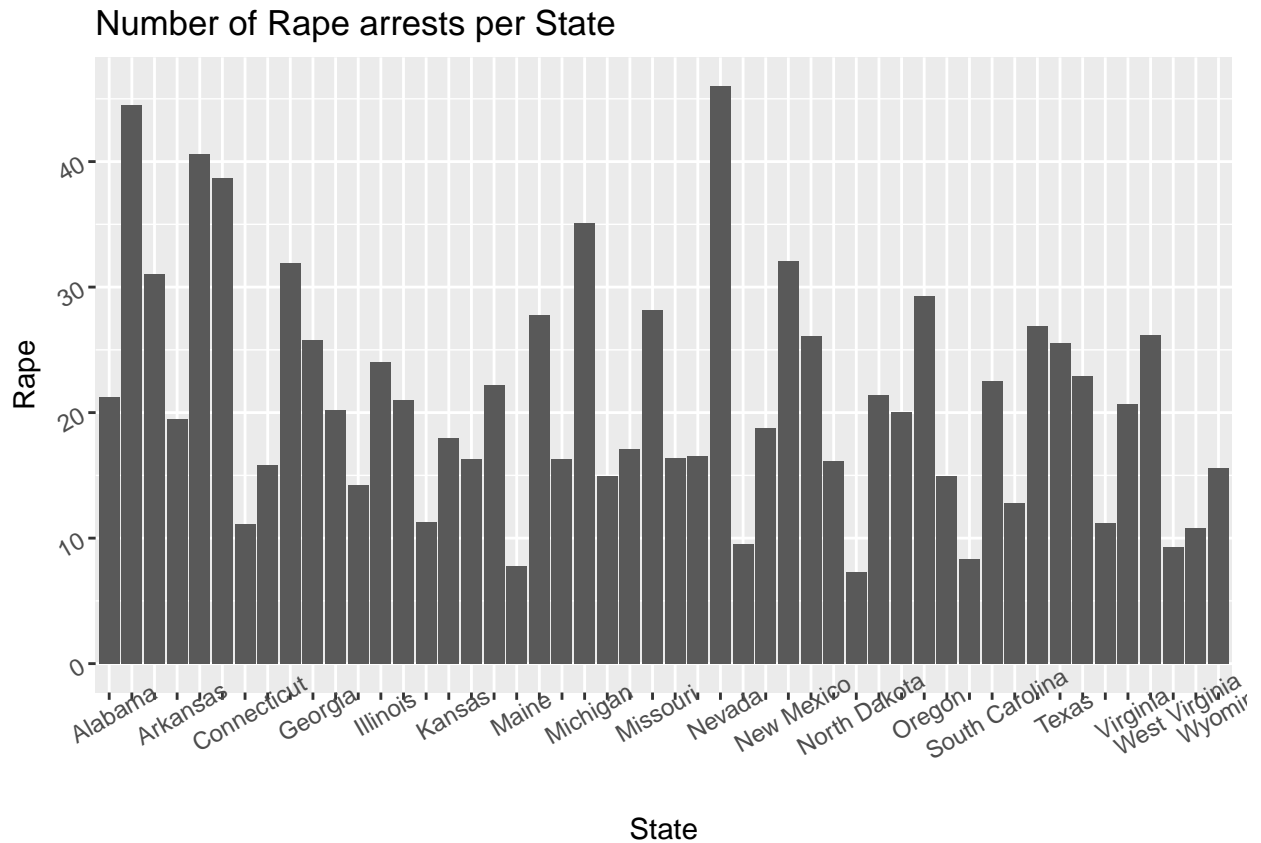
Boxplot Rape

4. Create a barplot of the number of rape arrests per state.

```r
US$State <- c(rownames(US))

p_bar <- ggplot(data= US, aes(x = State, y= Rape)) +
  geom_bar(stat = 'identity') +
  theme(axis.text = element_text(size = 9, angle = 30)) +
  scale_x_discrete(guide = guide_axis(check.overlap = TRUE))+
  labs(title = "Number of Rape arrests per State")
#scale_x_discrete(guide = guide_axis(n.dodge=2)) is also a good way to make the labels not overlap

p_bar
```
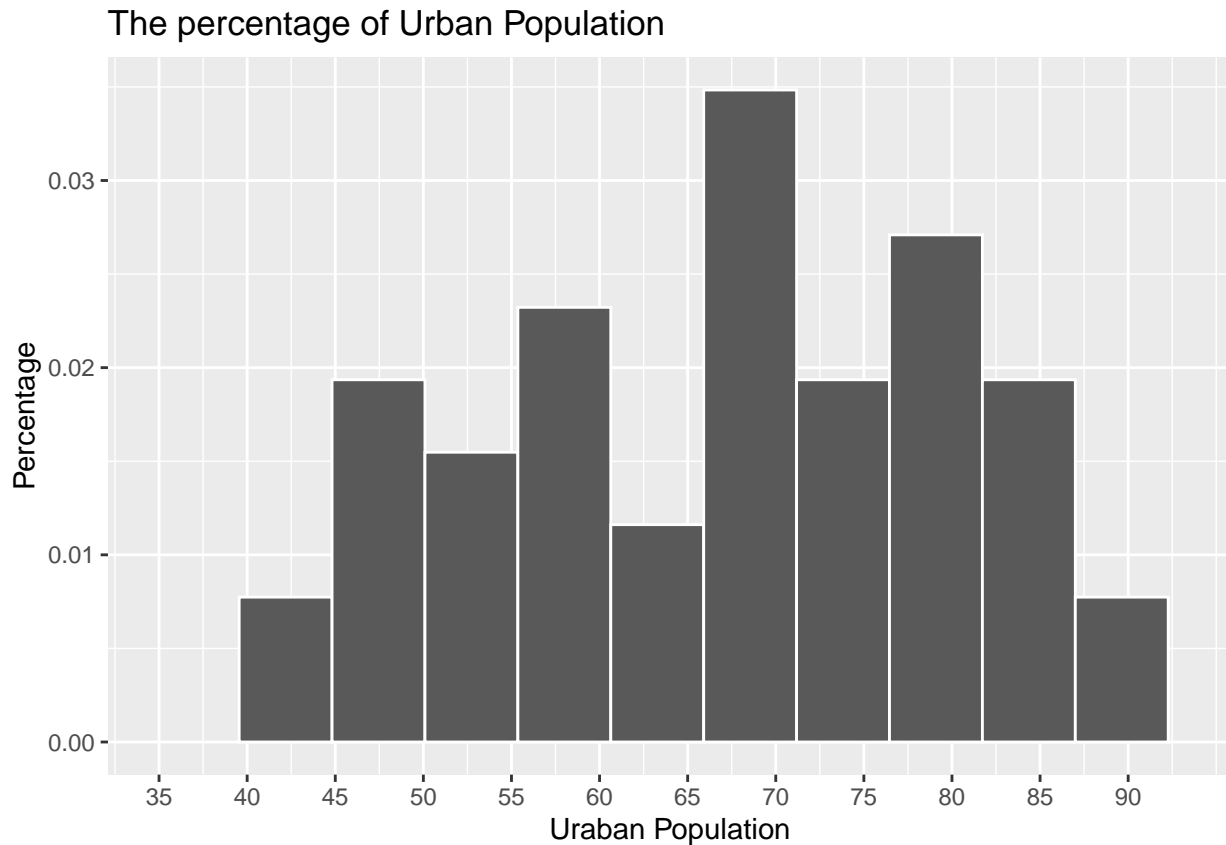
Number of Rape arrests per State

5. Create a histogram for the percent of urban population.

```
p_perc <- ggplot(US, aes(UrbanPop)) + geom_histogram(aes(y=stat(density)), bins = 12, color = "White")+

p_perc
```

```
## Warning: Removed 1 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 2 rows containing missing values (geom_bar).
```

The percentage of Urban Population

Your project

Now it's your turn. Use the `ggplot2` tools you used today to conduct data analysis for one of your final seminar papers.

1. Create a Git repository for your project.

2. Upload the dataset you are planning to use. If are you planning to collect original data, please provide all the variables that will be in that original dataseet.

3. How will these variables help answer your question?

4. Consider the variables. Which variables would you want to highlight? How would you visually represent them? Plan to create at least three descriptive graphs. Some suggestions:

   - If you have time-related variable, create a line graph showing changes over time!

   - If your observations can be separated by certain groups, create bar graphs or facets~

5. Given what you brainstormed in Question 4, create plots using the `ggplot2` package. Label all axes and title each graph. Provide descriptions for each graph.

Submit

Email me (laaker@wisc.edu) the link to your github when you are done.