HomeworkFourRG

Ryan Gandolfi

8/14/2020

## Variables

The variables I chose to work with are “decisionDirection” and “chief”. The variable “decisionDirection” refers to which way the ruling went whether that be “conservative”(1), “liberal”(2), or inconclusive(3). The “chief” variable refers to the Chief Justice on each of the cases. This could present a unique representation of how the chief justice rules over time.

library(tidyverse)

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

## v ggplot2 3.3.2 v purrr 0.3.4  
## v tibble 3.0.3 v dplyr 1.0.1  
## v tidyr 1.1.1 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()

scdbData <- read.delim("scdb.csv", sep = ",", header= TRUE)

## First step

The first step I took was isolating the cases on a bar graph versus the variable “decisionDirection” to see what the overall graph looks like. The code to execute this is as follows ggplot(scdbData, aes(decisionDirection)) + geom\_bar( ) .

ggplot(scdbData, aes(`decisionDirection`)) + geom\_bar( )

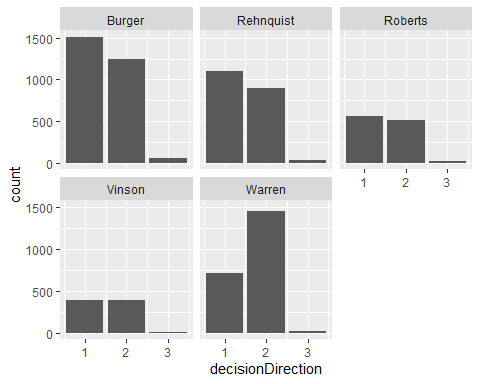
## Warning: Removed 39 rows containing non-finite values (stat\_count).



After I got comfortable with the data I decided to alter this visualization by organizing each of the decisionDirection values to the chief justice assigned to the case. I achieved this by using the facet function like so ggplot(scdbData, aes(decisionDirection)) + geom\_bar( ) + facet\_wrap(vars(chief)).

ggplot(scdbData, aes(`decisionDirection`)) + geom\_bar( ) + facet\_wrap(vars(`chief`))

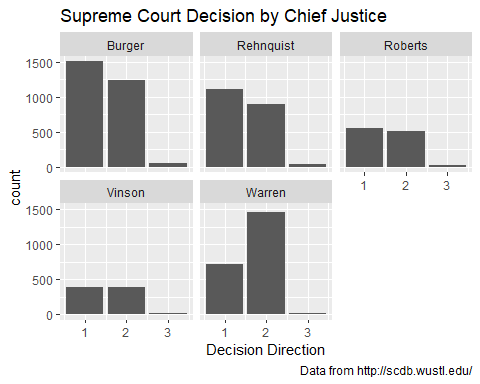
## Warning: Removed 39 rows containing non-finite values (stat\_count).



Now I can label this data properly to avoid confusion with the following line of code.

ggplot(scdbData, aes(`decisionDirection`)) + geom\_bar( ) + facet\_wrap(vars(`chief`)) +labs(title = "Supreme Court Decision by Chief Justice", caption = "Data from http://scdb.wustl.edu/", y = "count", x= "Decision Direction")

## Warning: Removed 39 rows containing non-finite values (stat\_count).



## Hypothesis

My thought process when choosing these variables is that the visualization of the Chief Justice versus the outcome of the decision could provide useful information on how a potential ruling could play out. Essentially, we can clearly see if there is any sort of bias in how a case is ruled based on the Chief Justice assigned to the case. A potential example of this can be seen above with Chief Justice Warren. Based on this data set it seems that Chief Justice Warren is very likely to have a more liberal outcome on her cases.

Looking further into some of the statistical functions available in R, I went ahead and retrieved the Mean, Median, Standard Deviation and Variance corresponding to the column value “decisionDirection”.

summarize(scdbData, mean.decisionDirection = mean(`decisionDirection`,na.rm = TRUE),summarize(scdbData, median.decisionDirection = median(`decisionDirection`,na.rm = TRUE),summarize(scdbData, var.decisionDirection = var(`decisionDirection`,na.rm = TRUE)),summarize(scdbData, sd.decisionDirection = sd(`decisionDirection`,na.rm = TRUE))))

## mean.decisionDirection median.decisionDirection var.decisionDirection  
## 1 1.538255 2 0.2823982  
## sd.decisionDirection  
## 1 0.5314115

At this point I felt the “max”, “min”, and “IQR” values didn’t quite help me in any way therefore I did not include them in this summary.