HypothesisVisualizationHW4

Larri Miller

8/13/2020

HW # 4 This code outlines the creation of a relational hypothesis and visualization from the dataset used in previous HW assignments. The dataset used is scdb, which contains data on the Supreme Court’s decisions from its first in 1791 through 2018. More information about scdb can be accessed at <http://scdb.wustl.edu/>

This analysis uses the tidyverse package.

library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.0.2

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

## v ggplot2 3.3.2 v purrr 0.3.4  
## v tibble 3.0.1 v dplyr 1.0.1  
## v tidyr 1.1.0 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.5.0

## Warning: package 'ggplot2' was built under R version 4.0.2

## Warning: package 'dplyr' was built under R version 4.0.2

## -- Conflicts ------------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

First, read in the data.

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

Now, check to see the variables, or columns, in scdb to see what would be interesting to visualize.

colnames(scdb)

## [1] "caseId" "docketId"   
## [3] "caseIssuesId" "voteId"   
## [5] "dateDecision" "decisionType"   
## [7] "usCite" "sctCite"   
## [9] "ledCite" "lexisCite"   
## [11] "term" "naturalCourt"   
## [13] "chief" "docket"   
## [15] "caseName" "dateArgument"   
## [17] "dateRearg" "petitioner"   
## [19] "petitionerState" "respondent"   
## [21] "respondentState" "jurisdiction"   
## [23] "adminAction" "adminActionState"   
## [25] "threeJudgeFdc" "caseOrigin"   
## [27] "caseOriginState" "caseSource"   
## [29] "caseSourceState" "lcDisagreement"   
## [31] "certReason" "lcDisposition"   
## [33] "lcDispositionDirection" "declarationUncon"   
## [35] "caseDisposition" "caseDispositionUnusual"   
## [37] "partyWinning" "precedentAlteration"   
## [39] "voteUnclear" "issue"   
## [41] "issueArea" "decisionDirection"   
## [43] "decisionDirectionDissent" "authorityDecision1"   
## [45] "authorityDecision2" "lawType"   
## [47] "lawSupp" "lawMinor"   
## [49] "majOpinWriter" "majOpinAssigner"   
## [51] "splitVote" "majVotes"   
## [53] "minVotes"

## 1. Descriptions of variables

Two variables that stick out to me are decisionDirection, which codes the ideological direction of the decision (i.e. Liberal or Conservative), and issueArea, which codes the issue area of court cases (I used issueArea for the last HW)

First, I’m going to run some summary statistics of these two variables to see what they’re like.

summarize(scdb, mean.issueArea=mean(issueArea, na.rm=TRUE), mean.decisionDirection=mean(decisionDirection, na.rm=TRUE), median.issueArea=median(issueArea, na.rm=TRUE), median.decisionDirection=median(decisionDirection, na.rm=TRUE))

## mean.issueArea mean.decisionDirection median.issueArea  
## 1 5.206939 1.538255 4  
## median.decisionDirection  
## 1 2

Next I’m going to recode the numerical values for each variable to their corresponding category labels.

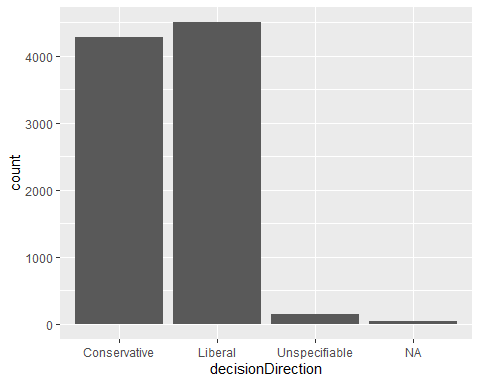
scdb <- scdb %>%  
 mutate(decisionDirection = recode(decisionDirection, `1` = "Conservative", `2` = "Liberal", `3` = "Unspecifiable")) %>%  
 mutate(issueArea = recode(issueArea, `1` = "Criminal Procedure", `2` = "Civil Rights", `3` = "First Amendment", `4` = "Due Process", `5` = "Privacy", `6` = "Attorneys", `7` = "Unions", `8` = "Economic Activity", `9` = "Judicial Power", `10` = "Federalism", `11` = "Interstate Relations", `12` = "Federal Taxation", `13` = "Miscellaneous", `14` = "Private Action"))   
  
head(select(scdb, decisionDirection, issueArea))

## decisionDirection issueArea  
## 1 Liberal Economic Activity  
## 2 Conservative Criminal Procedure  
## 3 Liberal Economic Activity  
## 4 Liberal Civil Rights  
## 5 Liberal Economic Activity  
## 6 Conservative Economic Activity

I’m going to do a quick visualization of each variable as well to get a visual sense for what they look like individually.

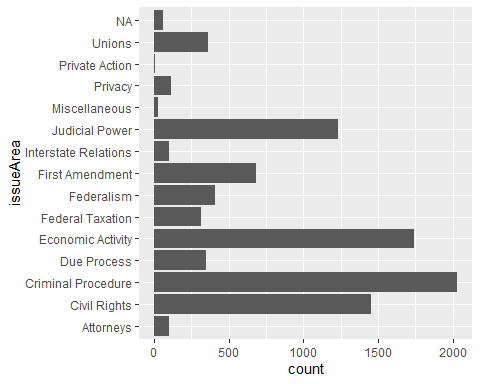
decisionDirection:

ggplot(scdb, aes(decisionDirection)) +  
 geom\_bar()



issueArea:

ggplot(scdb, aes(issueArea)) +  
 geom\_bar() +  
 coord\_flip()



## 2. Description of relationship between variables, hypotheses

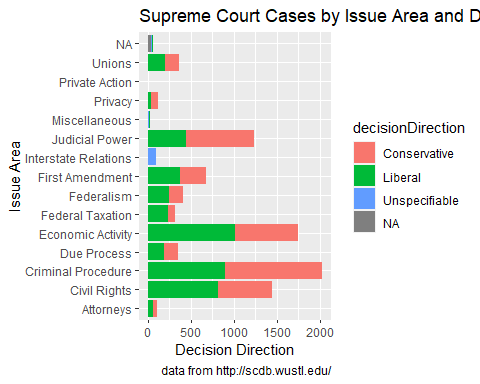
From these visualizations, I can see that the decisionDirection looks to be close between Conservative and Liberal (with Liberal having a slightly higher count) and that the top three issueAreas are Criminal Procedure, Economic Activity, and Civil Rights.

Based on my experiences with Conservative versus Liberal ideology, I’m going to make a general hypothesis a) that issue areas will differ in decision direction and more specific hypotheses that b) criminal procedure will lean conservative and c) civil rights will lean liberal.

## 3. Demonstration of relationship, visualization

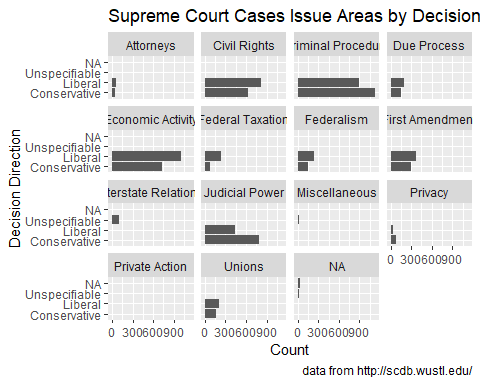
Now I’m going to plot these variables together to look for a relationship between them. I’m going to plot issueArea on the x axis and use color as an indicator of decisionDirection.

ggplot(scdb, aes(issueArea, fill = decisionDirection)) +  
 geom\_bar() +  
 coord\_flip() +  
 labs (  
 title = "Supreme Court Cases by Issue Area and Decision Direction",  
 caption = "data from http://scdb.wustl.edu/",  
 x = "Issue Area",  
 y = "Decision Direction"  
 )



I’m going to also try facet\_wrap() to look at individual issueAreas more closely.

ggplot(scdb, aes(decisionDirection)) +  
 geom\_bar() +  
 facet\_wrap(vars(issueArea)) +  
 coord\_flip() +  
 labs (  
 title = "Supreme Court Cases Issue Areas by Decision Direction Count",  
 caption = "data from http://scdb.wustl.edu/",  
 x = "Decision Direction",  
 y = "Count"  
 )



These graphs support hypothesis a in that individual issue areas seem to differ in decision direction. Similarly, the current graphs support hypotheses b and c in that criminal procedure leaned conservative and civil rights leaned liberal. Further analyses must be done to determine if these differences are statistically significant.