

DATA 606 - Project Proposal

Joshua Sturm

October 31, 2017

<https://www.kaggle.com/shoklan/which-year-produced-the-most/notebook> <https://github.com/fivethirtyeight/data/blob/master/congress-age/congress-terms.csv> http://online.wsj.com/public/resources/documents/info-CONGRESS_AGES_1009.html

Data Preparation

```
library(tidyverse)
library(ggplot2)

# Load the files from the working directory
amendments_raw <- read.csv("amendment_list.csv")
members_raw <- read.csv("congress_terms.csv")

# Tidy the datasets

# Select the columns relevant to my research
amendments <- amendments_raw %>%
  select(5, 7:ncol(amendments_raw)-1)

# Use regex to shift errant data to their appropriate columns
for (i in 1:(length(amendments$year))) {
  pat <- "\\D{3,}"
  if (grepl(pat, amendments[i, "month"])) == 1)
  {
    amendments[i, "year"] <- amendments[i, "month"]
    amendments[i, "month"] <- amendments[i, "day"]
    amendments[i, "day"] <- amendments[i, "congress"]
    amendments[i, "congress"] <- amendments[i, "congressional_session"]
    amendments[i, "congressional_session"] <- amendments[i, "joint_resolution_chamber"]
  }
}
amendments$year <- gsub("\\D{4}$", "", amendments$year)
```

Research question

Is there a relationship between the average age of congress (members) and the number of bills proposed?

The average age of congressional representatives has been steadily climbing since the second world war. The current (115th) one is among the oldest in its history. How has this affected the effectiveness of congress? Are older more representatives more or less active?

I plan to explore this via proxy, by taking a look at all the *constitutional amendments* proposed since the first congress through the 113th, and recording the age of each of the bill's sponsors. Additionally, I will seek any interesting tidbits in the data, such as the most active years, as well as which state representatives propose the most legislation.

Cases

Each case represents a constitutional amendment proposed by congress. There are a total of 11797 cases in this dataset.

Data collection

The list of 11,000+ amendments was compiled by staff and volunteers of the National Archives and Records Administration. The list of representatives was compiled by The UnitedStates Project (House members), and The New York Times Congress API (senate).

Type of study

This is an observational study.

Data Source

The amendment list was retrieved from Kaggle, while the members list was taken from FiveThirtyEight. Another source is from the Wall Street Journal.

Response

The response variable is legislative activity and is numerical.

Explanatory

The explanatory variable is median age of congressional representatives and is numerical.

Relevant summary statistics

```
lvls <- levels(amendments$year)

ggplot(amendments, aes(year)) +
  geom_bar() +
  scale_x_discrete(breaks=seq(1788, 2014, 20))
```

