

datablog

2024-05-14

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

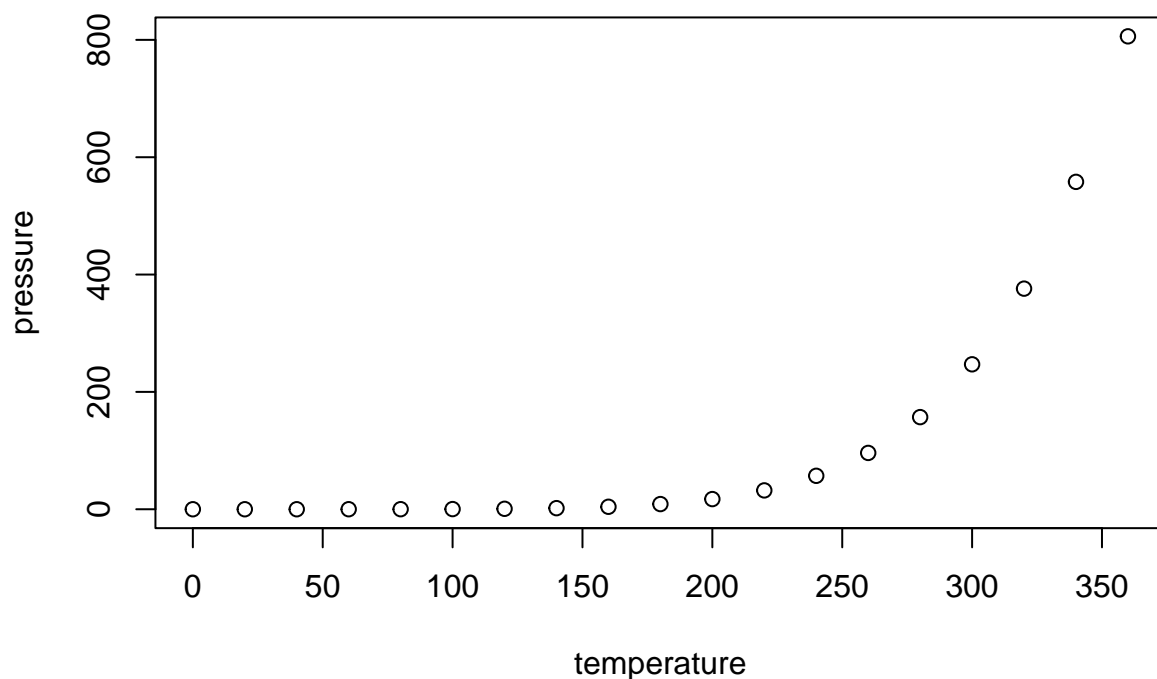
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean    : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.    :120.00
```

Including Plots

You can also embed plots, for example:



```
library(pacman)
p_load(tidyverse, ggplot2, ggrepel)
nom_dat <- read_csv("https://voteview.com/static/data/out/members/HSall_members.csv")
```

```
## Rows: 50488 Columns: 22
## -- Column specification -----
## Delimiter: ","
## chr (4): chamber, state_abbrev, bioname, bioguide_id
## dbl (17): congress, icpsr, state_icpsr, district_code, party_code, occupancy...
## lgl (1): conditional
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
south <- c(40:49, 51, 53)
polar_dat <- nom_dat %>%
  filter(congress > 45 &
         chamber != "President") %>%
  mutate(
    year = 2*(congress-1) + 1789,
  ) %>%
  group_by(chamber, congress, year) %>%
  summarize(
    party.mean.diff.d1 = mean(nominate_dim1[party_code==200], na.rm=T) -
      mean(nominate_dim1[party_code==100], na.rm=T),
```

```

prop.moderate.d1 = mean(abs(nominate_dim1)<0.25,na.rm=T),
prop.moderate.dem.d1 = mean(abs(nominate_dim1[party_code==100])<0.25,na.rm=T),
prop.moderate.rep.d1 = mean(abs(nominate_dim1[party_code==200])<0.25,na.rm=T),
overlap = (sum(nominate_dim1[party_code==200] <
               max(nominate_dim1[party_code==100],na.rm=T),na.rm=T) +
            sum(nominate_dim1[party_code==100] >
               min(nominate_dim1[party_code==200],na.rm=T),na.rm=T))/
            (sum(!is.na(nominate_dim1[party_code==100]))+
             sum(!is.na(nominate_dim1[party_code==200]))),
chamber.mean.d1 = mean(nominate_dim1,na.rm=T),
chamber.mean.d2 = mean(nominate_dim2,na.rm=T),
dem.mean.d1 = mean(nominate_dim1[party_code==100],na.rm=T),
dem.mean.d2 = mean(nominate_dim2[party_code==100],na.rm=T),
rep.mean.d1 = mean(nominate_dim1[party_code==200],na.rm=T),
rep.mean.d2 = mean(nominate_dim2[party_code==200],na.rm=T),
north.rep.mean.d1 = mean(nominate_dim1[party_code==200 &
                                   !(state_icpsr %in% south)],na.rm=T),
north.rep.mean.d2 = mean(nominate_dim2[party_code==200 &
                                   !(state_icpsr %in% south)],na.rm=T),
south.rep.mean.d1 = mean(nominate_dim1[party_code==200 &
                                   (state_icpsr %in% south)],na.rm=T),
south.rep.mean.d2 = mean(nominate_dim2[party_code==200 &
                                   (state_icpsr %in% south)],na.rm=T),
north.dem.mean.d1 = mean(nominate_dim1[party_code==100 &
                                   !(state_icpsr %in% south)],na.rm=T),
north.dem.mean.d2 = mean(nominate_dim2[party_code==100 &
                                   !(state_icpsr %in% south)],na.rm=T),
south.dem.mean.d1 = mean(nominate_dim1[party_code==100 &
                                   (state_icpsr %in% south)],na.rm=T),
south.dem.mean.d2 = mean(nominate_dim2[party_code==100 &
                                   (state_icpsr %in% south)],na.rm=T),
)

```

'summarise()' has grouped output by 'chamber', 'congress'. You can override
using the '.groups' argument.

```
head(polar_dat)
```

```

## # A tibble: 6 x 22
## # Groups:   chamber, congress [6]
##   chamber congress year party.mean.diff.d1 prop.moderate.d1
##   <chr>      <dbl> <dbl>          <dbl>          <dbl>
## 1 House      46  1879          0.786          0.130
## 2 House      47  1881          0.783          0.100
## 3 House      48  1883          0.724          0.216
## 4 House      49  1885          0.749          0.148
## 5 House      50  1887          0.765          0.161
## 6 House      51  1889          0.795          0.119
## # i 17 more variables: prop.moderate.dem.d1 <dbl>, prop.moderate.rep.d1 <dbl>,
## #   overlap <dbl>, chamber.mean.d1 <dbl>, chamber.mean.d2 <dbl>,
## #   dem.mean.d1 <dbl>, dem.mean.d2 <dbl>, rep.mean.d1 <dbl>, rep.mean.d2 <dbl>,
## #   north.rep.mean.d1 <dbl>, north.rep.mean.d2 <dbl>, south.rep.mean.d1 <dbl>,
## #   south.rep.mean.d2 <dbl>, north.dem.mean.d1 <dbl>, north.dem.mean.d2 <dbl>,

```

```
## # south.dem.mean.d1 <dbl>, south.dem.mean.d2 <dbl>
```

```
write_csv(polar_dat,path="Political Polarization in US Congress (Non-Assorted).csv")
```

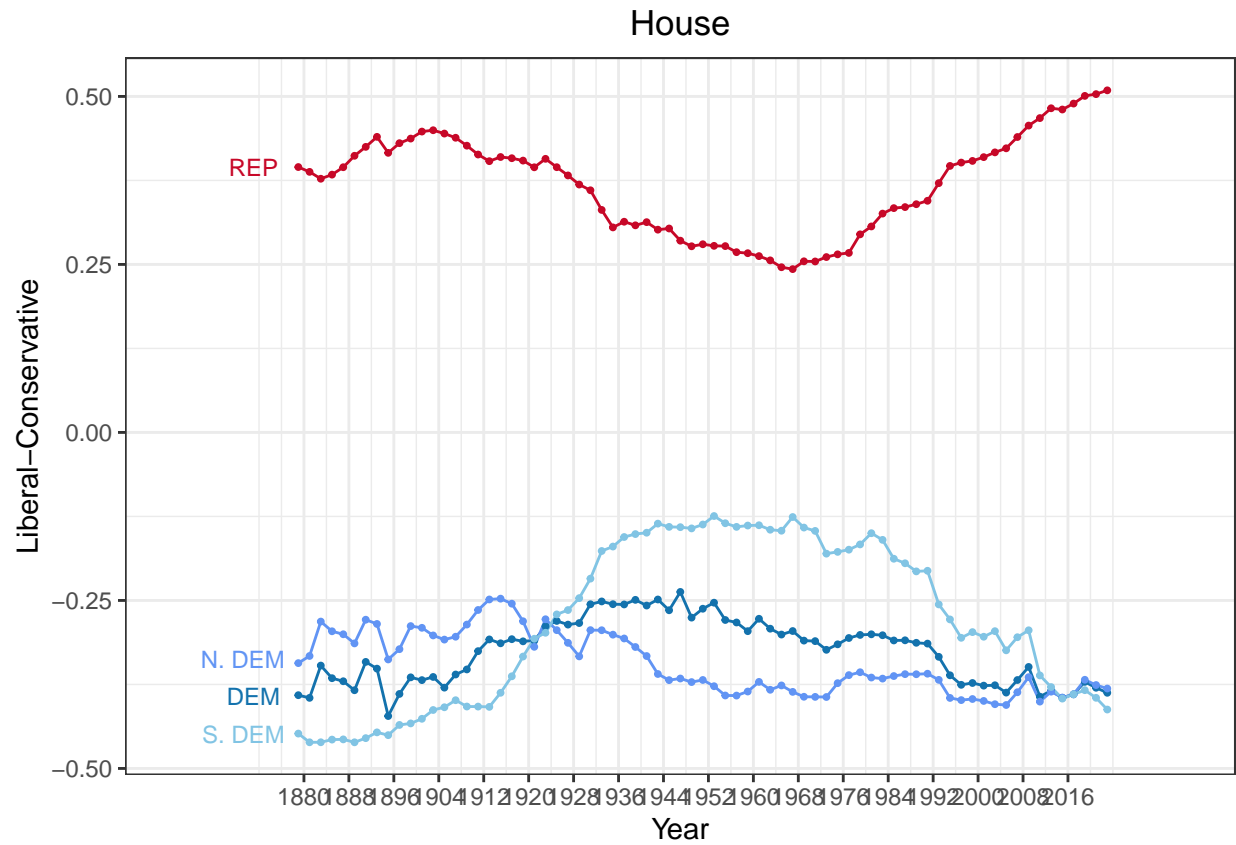
```
## Warning: The 'path' argument of 'write_csv()' is deprecated as of readr 1.4.0.  
## i Please use the 'file' argument instead.  
## This warning is displayed once every 8 hours.  
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was  
## generated.
```

```
# IMPORTANT: Make sure to download the non-assorted file first!
```

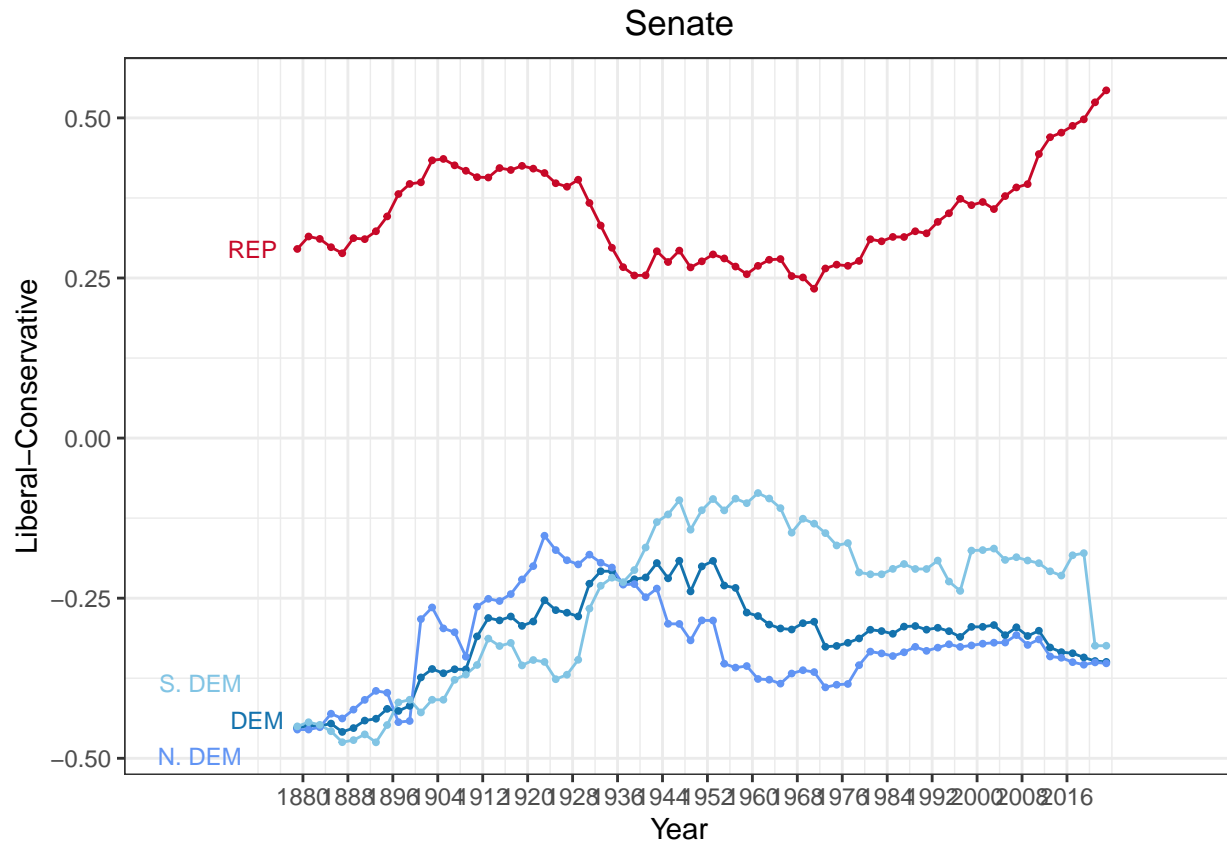
```
polar_dat_long <- polar_dat %>% gather(score,value,-chamber,-year,-congress)  
labels <- c("dem.mean.d1"="DEM",  
            "rep.mean.d1"="REP",  
            "north.dem.mean.d1"="N. DEM",  
            "south.dem.mean.d1"="S. DEM")
```

```
polarized_plot <- function(chamb) {  
  pdatl <- polar_dat_long %>%  
    filter(chamber==chamb,  
           score %in% c("dem.mean.d1","rep.mean.d1",  
                        "north.dem.mean.d1","south.dem.mean.d1")) %>%  
    mutate(party=labels[score]) %>%  
    ungroup()
```

```
gg <- ggplot(data=pdatl,  
             aes(x=year,y=value,group=party,col=party)) +  
  scale_x_continuous(expand = c(0.15, 0),  
                     breaks=seq(1880, max(pdatl$year), by=8)) +  
  geom_line() + geom_point(size=0.7) +  
  xlab("Year") + ylab("Liberal-Conservative") +  
  geom_text_repel(data=pdatl %>%  
                  filter(year == min(year)),  
                  aes(label = party, color = party),  
                  size = 3,  
                  nudge_x = -8,  
                  point.padding = 0.1,  
                  segment.color = NA,  
                  show.legend = FALSE) +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +  
  scale_color_manual(values=c("REP"="#c70828","DEM"="#1372ad",  
                              "N. DEM"="#6194F4",  
                              "S. DEM"="#81c4e4"),guide="none") +  
  theme_bw()  
gg  
}  
gg <- polarized_plot("House")  
gg + ggtitle("House") + theme(plot.title = element_text(hjust = 0.5))
```



```
gg <- polarized_plot("Senate")
gg + ggtitle("Senate") + theme(plot.title = element_text(hjust = 0.5))
```



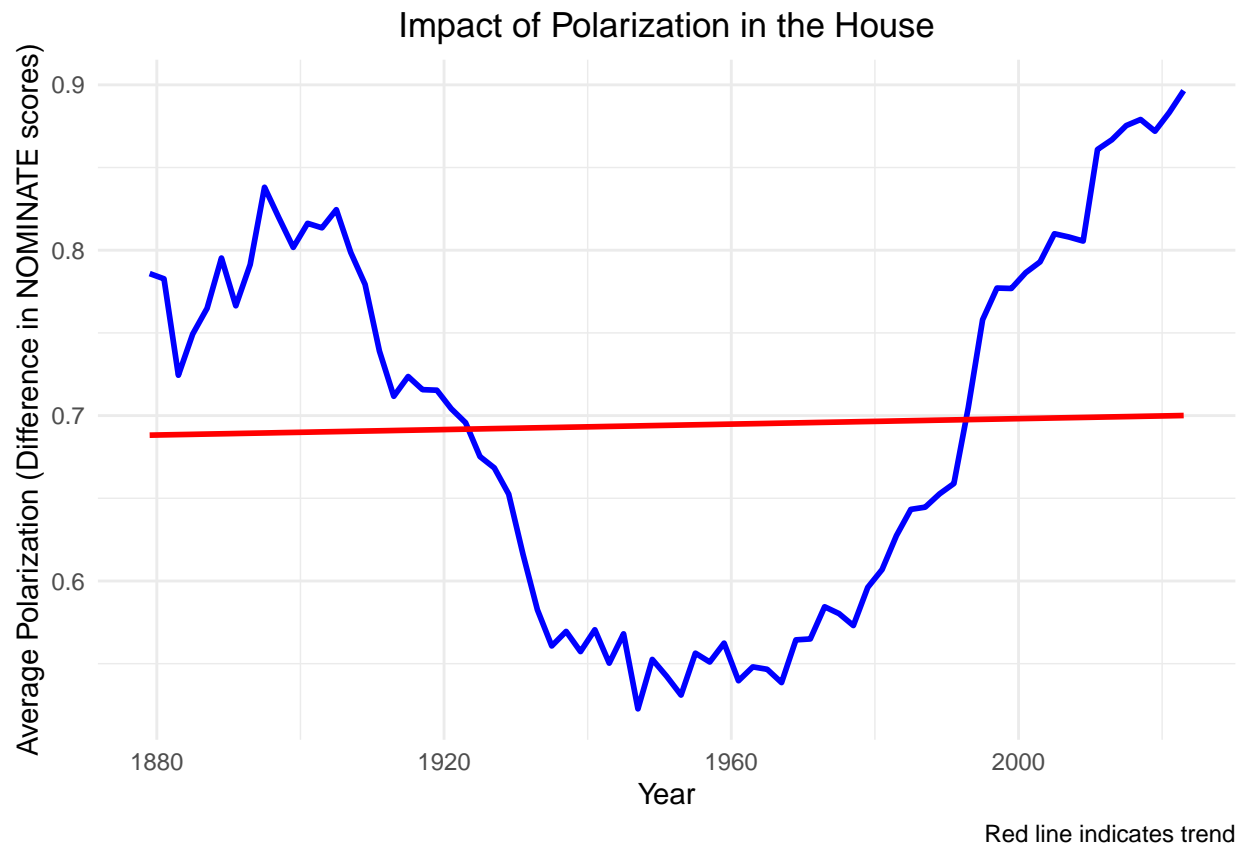
```
plot_polarization_impact <- function(chamber_name) {
  polar_dat %>%
    filter(chamber == chamber_name) %>%
    ggplot(aes(x = year, y = party.mean.diff.d1)) +
    geom_line(color = "blue", size = 1) +
    geom_smooth(method = "lm", se = FALSE, color = "red") +
    labs(title = paste("Impact of Polarization in the", chamber_name),
         x = "Year",
         y = "Average Polarization (Difference in NOMINATE scores)",
         caption = "Red line indicates trend") +
    theme_minimal() +
    theme(plot.title = element_text(hjust = 0.5))
}
```

```
# House
house_plot <- plot_polarization_impact("House")
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

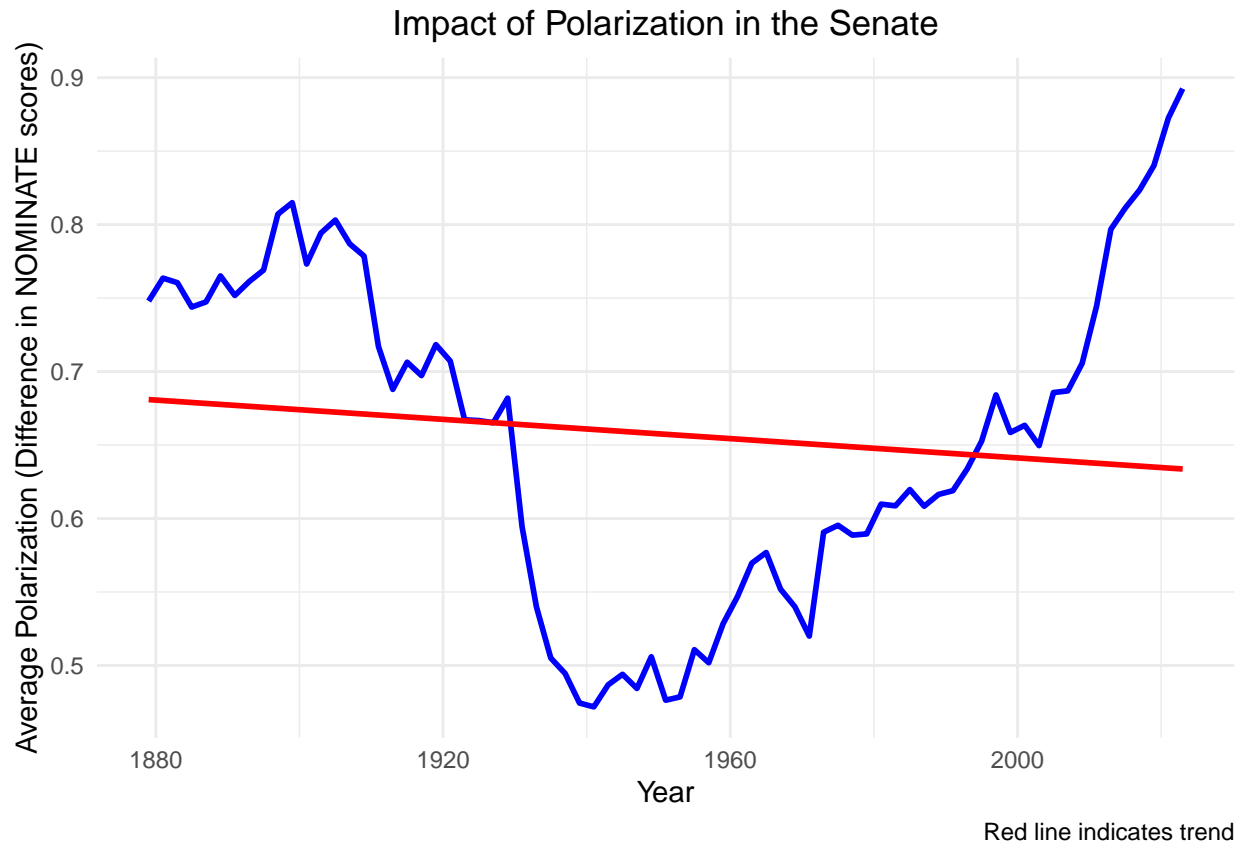
```
house_plot
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
# Senate
senate_plot <- plot_polarization_impact("Senate")
senate_plot
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



House Purpose The purpose of this graph is to visually depict the fluctuating levels of polarization between Democrats and Republicans in the House of Representatives over time. It aims to highlight periods of heightened division within American politics and the evolution of legislative polarization, offering insights into the development of increasingly partisan laws.

Name Sean Gee

Improvements Could be on more specific issues Show lines for both democrats and republicans instead of combining the two

Senate Purpose The purpose of the graph is to visually illustrate the historical trajectory of polarization within the U.S. Senate, employing NOMINATE scores as a metric for ideological differences between political parties. By charting these scores over time, the graph aims to highlight patterns of polarization, fluctuations, and notable shifts in political alignment. This visualizes how polarization has affected the Senate and therefore its decisions in its legislation in the country.

Same issue