STAT-S 670 Mini Project II: Political Spectrum

Yash Shah
MS Data Science
Indiana University, Bloomington
Indiana, USA
yashah@iu.edu

Ritwik Budhiraja
MS Data Science
Indiana University, Bloomington
Indiana, USA
rbudhira@iu.edu

Srimanth Agastyaraju
MS Data Science
Indiana University, Bloomington
Indiana, USA
sragas@iu.edu

Abstract— **This paper describes how to use R to perform exploratory data analysis to find the degree of polarization sought on a dataset containing legislation, amendments, roll call votes, and other key data about the United States Congress. The aim is to find trends and meaningful insights by using a few advanced data visualization techniques like Principal Component Analysis and Multidimensional Scaling, data transformations, and regression models amongst others.**

(Keywords — congress, polarization, votes, members, principal component analysis, multidimensional scaling, senators, bills, regressions, visualizations, effects, R)

I. INTRODUCTION

This project is an attempt to use Exploratory Data Analysis to investigate the political spectrum of the United States of America. The United States is growing increasingly politically divided: liberals are becoming more liberal, conservatives are becoming more conservative, and their opinions are only getting stronger by the day. Political scientists believe that the centre of the political spectrum is being hollowed out in either situation and examining the voting record of Congress is one way to find out if this is accurate. The main idea of this research is to examine individual senators' voting records for various bills throughout time to identify what sort of pattern the political spectrum has followed and whether certain factors, such as the senators' party or their state, have had an impact. To provide a more concise comprehension, the analysis has been broken down into three parts, each with subparts.

II. DESCRIPTION OF DATA

Beginning in 1989 and ending in 2014, the data source provides voting statistics for 26 years in a row. Each year is housed in its own folder, which contains two *.csv* files.

- members.csv: This file contains information about the senators who were involved in all of the bills that year (information such as the senator's name, id, state, and political party).

- voters.csv: This file contains the votes cast by senators for various bills throughout a given year.

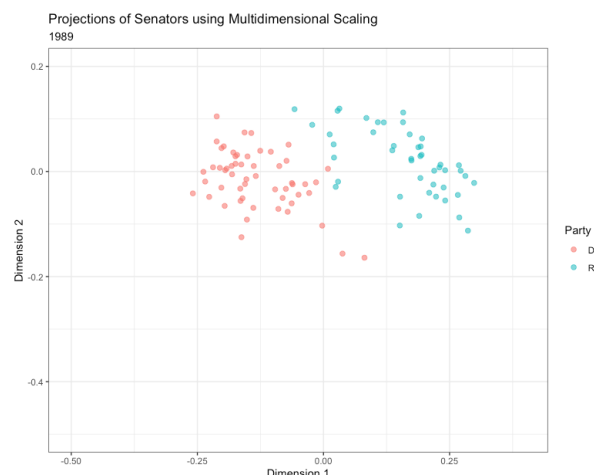## III. QUESTIONS & INFERENCES

A. Polarization in two years:

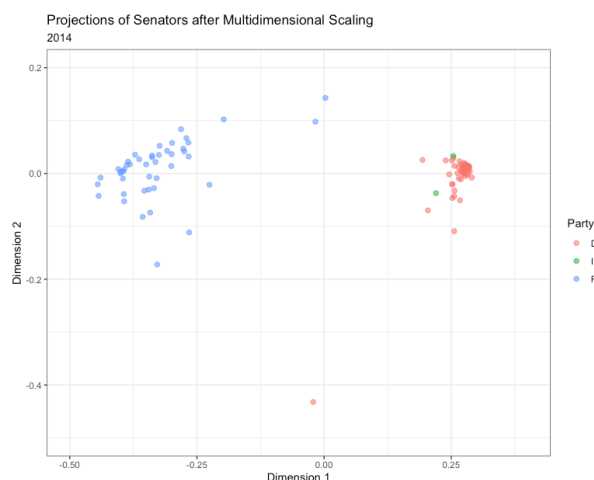Figure A1: Ideologies of Senators in 1989

Figure A2: Ideologies of Senators in 2014

Figure A1 and Figure A2 show scatterplots of MDS of 2 dimensions for the two years 1989 and 2014. To analyze the polarization over two years of voting data, we have reduced the dimensionality of the data using MDS (Multi-Dimensional Scaling) with pairwise distances using Gower's coefficient. Although a scatterplot depicts correlations or links between two variables, there is no certainty that the graphic depicts a cause-and-effect relationship. From the figures, the reduced data seems to fall on a one-dimensional axis, since the margin between the cluster of Democrats and Republicans is significantly higher in the year 2014, compared to the same in 1989. As a result, the reduced

data from 2014 is more linearly separable, contrary to the reduced data from 1989. A significantly farther outlier exists in the MDS data of the year 2014, who is a democrat. The distribution of the Republicans looks very similar in the reduced data of 1989 and 2014. However, we can see the following differences in 2014, compared to 1989 - The spread of the reduced data on the Republicans' cluster is relatively lower; the spread of the reduced data on the Democrats' cluster is significantly lower, except for an outlier in 2014; the distance between the clusters in the reduced space is significantly higher, which allows us to use a simple linear model for classification. The comparisons shown in the figures do not reveal anything about how the gap in the reduced dimension space between the democrats and republicans increased. To understand this, we must visualize how the gap in the reduced dimension space between the democrats and republicans changed over time.

Now that we've looked at the first and last years of data, we'll look at the differences in ideology across time in a more methodical way.
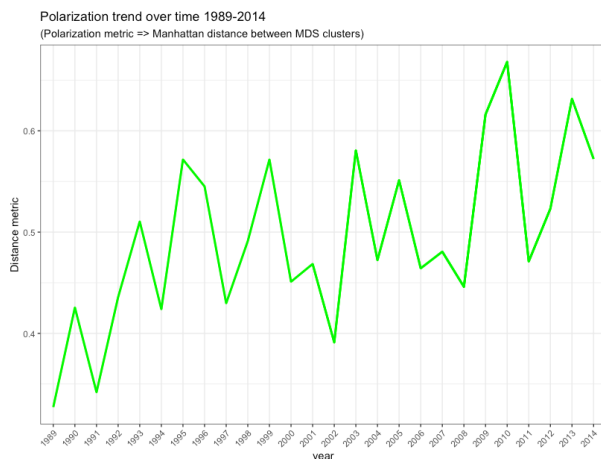
B. Polarization over time:

Figure B1: Polarization in ideologies of the Political Parties over time

We have created a line plot that shows the polarization between the ideologies of two political parties over time, as shown in Figure B1. Using the Manhattan Distance as a metric of polarization to measure the absolute levels of political differences between both the parties across the years. Now, simply based on the line plot shown above, we can see that polarization is increasing as we move ahead in time. There are a few noticeable valleys or dips in this plot, but they are followed by big peaks that push the overall trends of the line plot in an upward direction indicating an increase in polarization. As the distances tell us how close the ideologies of the two political parties were to each other, we can say that both the parties were quite like-minded in terms of ideologies and propaganda in the year 1989. This fact can be advocated by looking at the line graph and how the distance in the year 1989 was nearing approximately 0.325. On the contrary, both the political parties showed maximum opposing ideas in the

year 2010 when we can see the highest peak. With this contrasts we can say that both the parties are equally responsible for the polarization as with increments in time, both the parties become equally opposing when it came to pass bills, laws, and amendments. These ideological shifts in the congressional parties have occurred alongside – and, perhaps to some extent, because of – geographic and demographic shifts in their composition.

C. Ideological Position of one senator:

On the basis of a single senator's position in a party and over time, we may be able to talk about the absolute changes in the ideological positions of the parties and hold one responsible for polarization, but it can become a tricky question since the senators, their votes, and the bills on which they cast their votes are not constants over time.
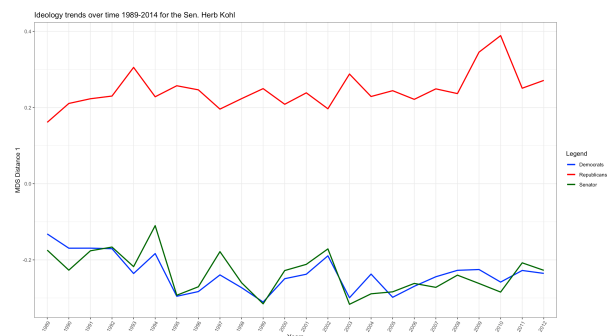

Figure C1: Polarization in ideologies of a Democratic Senator

From Figure C1, we can evidently see Sen. Herb Kohl's ideologies and how they went in tandem with the political party's propaganda over the years. It can be concluded that throughout his political career, he stayed a true Democrat.
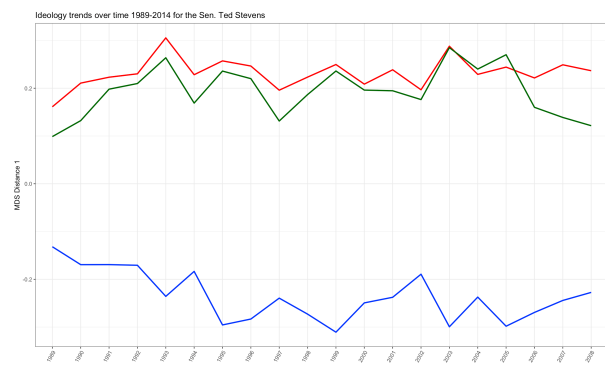

Figure C2: Polarization in ideologies of a Republican Senator

Similarly, we can see how the ideologies of Senator Ted Stevens were truly Republic until the year 2008, when we saw a small deviation in his beliefs as seen in Figure C2.

Lastly, Figure C3 shows the political stance of Sen. Arlen Specter starting from the year 1989 to 2010. Initially, his views closely followed the cohesive Republican ideologies which are portrayed by the closeness of the 'green' line plot to the 'red' line plot. Although throughout the time interval, we observe multiple crests and troughs, his ideologies remained more or less the same. A shocking and drastic plummet can be seen in his line plot with the beginning of the year 2008 which tells us that his ideologies took a turn and went with the overall

Democratic mindset. We cannot ignore the fact that he began his political career as a Republican and ended it as a Democrat. Since we neglect a senator's actual political party and only take into consideration his voting records, this deems as a possible limitation and can be a future improvement that we could work on in this project.
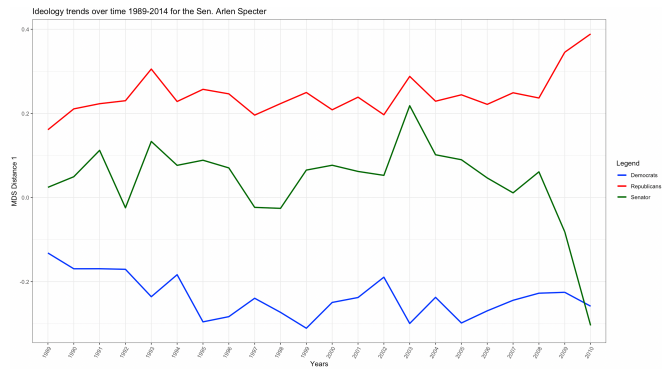

Figure C3: Polarization in ideologies of a Sen. Arlen Specter