

ANALYTICS AND VISUALIZATION OF POLITICAL DATA

MITCH HASLEHURST, ANIKET JOSHI, GUOJUN MA, SAM SIMON, SHEN-NING TUNG

1. INTRODUCTION

In this report, we analyze the voting records and visualize the behaviour of politicians from both Canada and the United States of America. The data analyzed here includes details of the legislators, such as education level, place, active year in politics, the topics of bills passed by the legislators, historical voting record, and so forth. The goal is to build metrics to quantitatively measure the performance of politicians using these data. One obvious and commonly employed approach to measuring performance is via approval rating, but we seek an approach which is more rigorous. Alternatively (as a starting point), the quality of a politician can be measured by the number of bills they sponsored, or the number of bills passed as laws align with a their political belief. It is reasonable to assume when a politician vote “yes” to a bill, it means the bill aligns with their political belief, or vice versa.

The idea of measuring a politician’s performance is similar to sport analytics, which has become a crucial part of almost every major sport in recent decades. For example, in the current NBA, the team management often analyzes the player’s strengths and weaknesses by looking at the historical data such as where the players most often take shots in the field, points per game, assists per game, re-bounces per game, and so forth. There is little exaggeration in claiming that a given player’s market worth is entirely determined by their statistics. In recent years, many different kind of statistical methods have been applied to politics, such as using the poll before a major election to predict the result (perhaps the most well-known one), conducting a survey to obtain approval ratings, analyze politicians’ speech to find topic distribution, and many more.

We approach the search for such methods of measurement through several lenses. In the first section we examine how often, in Canadian Parliament, the outcome of a vote fares in favour of a given legislator. In the third and fourth sections the number of bills sponsored by both Canadian and American legislators is tracked, along with the topics of bills. In the fifth section we delve further into bills sponsored, examining the topics closely and comparing their global versus national impact. Lastly, we discuss the issue of bipartisanship in the United States. To achieve this we construct a graph to describe the interaction between politicians based on their political party.

All data used in this report is courtesy of IOTO International, and was cleaned and analyzed using Python.

2. CANADIAN PARTY LEADERS

One thing that we can measure for Canadian politicians in Parliament is how often a vote goes in favour of how they vote. That is, if they vote yes on the bill and it passes or they vote no on the bill and it doesn't pass. This could be comparable to scoring in various sports or for example hits in baseball. In this section we look at this voting record data for Canada's current three majority party leaders, i.e. Justin Trudeau, Jagmeet Singh, and Erin O'Toole. We compare them against each other for different parliamentary sessions and also look at individual performance over different parliamentary sessions.

We may consider each party as a "team," each Parliamentary session as a "season" and each party leader as "team captain." In this regard, there is a natural correlation between the "team's performance" and the "captain's performance". That is, when your team is doing well in the season, i.e. the Prime Minister is a member of your party, then your team captain is also doing well. Similarly if your team is not doing well that season, i.e. the Prime Minister is not a member of your party, then your team captain also does not have the best performance. In the following figures, "yy" means "voted yes and passed", "yn" means "voted yes and didn't pass", "ny" means "voted no and passed", and "nn" means "voted no and didn't pass". Thus, the sections of blue and red are those that correspond to a vote going in a politicians favour.

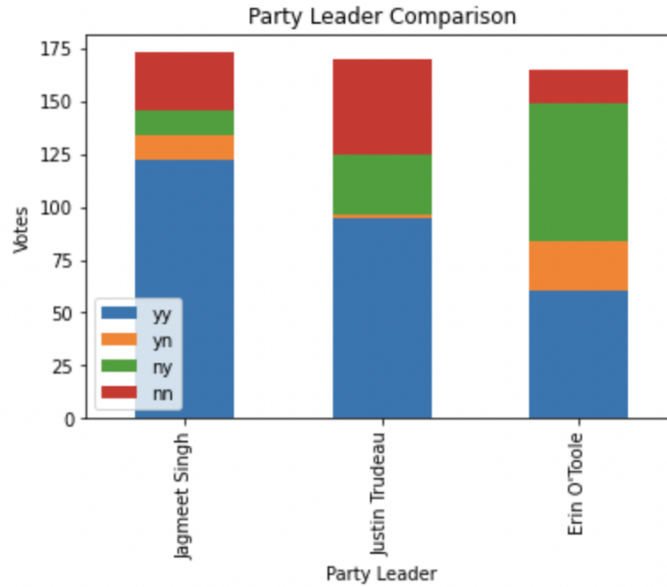


FIGURE 1

In Figure 1 we compare the voting record for the three party leaders over the parliamentary sessions 43-1 and 43-2. Of note, there is only one time that Justin Trudeau votes "yea" for a bill that didn't pass. Conversely, Erin O'Toole had the largest amount of bills that didn't go in his favour. Both of these trends are expected to some extent. Justin Trudeau is the Prime Minister and leader of the

Liberal party, so it is natural that many of the bills go in his voting favour. However, Erin O'Toole is the leader of the Conservative party which has historically differed with the Liberal party on many issues.

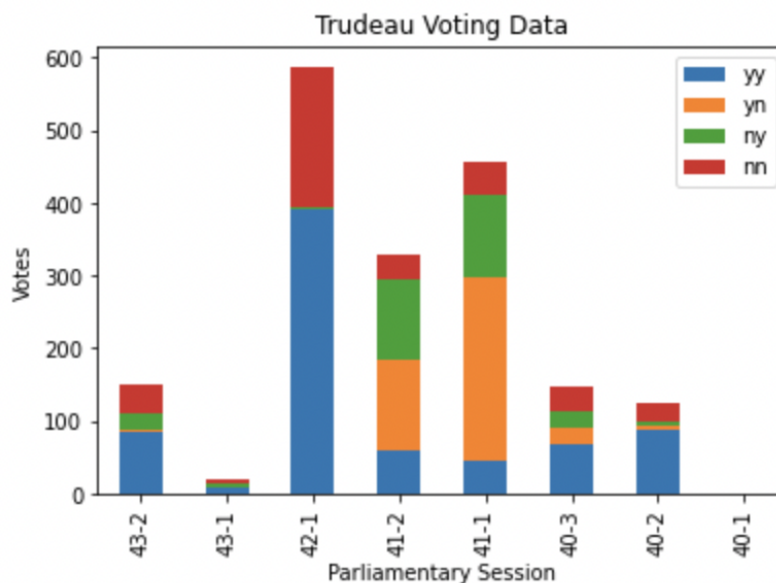


FIGURE 2

Figure 2 shows Trudeau's voting performance from Parliamentary session 40-1 to 43-2. We see that in the 41st parliament, most of the votes didn't go in a favourable way for him. Contrast this to the Parliament immediately after, where almost all of them did. Naturally, this is also expected to some extent as the Prime Minister transitioned from Harper (Conservative) to Trudeau himself (Liberal).

Figure 3 shows O'Toole's voting record. Comparing to Trudeau's we see a mirrored picture in the sense that the sessions where O'Toole performs well are where Trudeau performs poorly, and the reverse also holds. We also note that during the 41-1 and 41-2 Parliamentary sessions that almost every bill went in O'Toole's favor. In fact, only 3 votes out of a total of 602 votes went against his own vote. This was during Harper's term as Prime Minister, so one might naturally question the similarities between the two. However, when comparing them directly, we see that they voted differently on 131 bills out of the 591 that they both voted on. Therefore, O'Toole more closely resembles the political policies but in place through Parliament during the 41st session rather than Harper's personal voting record.

We remark that this method of comparison and "scoring" is rather naïve. Although the Prime Minister is from one of the main parties, the composition of the Parliament is also important to consider.

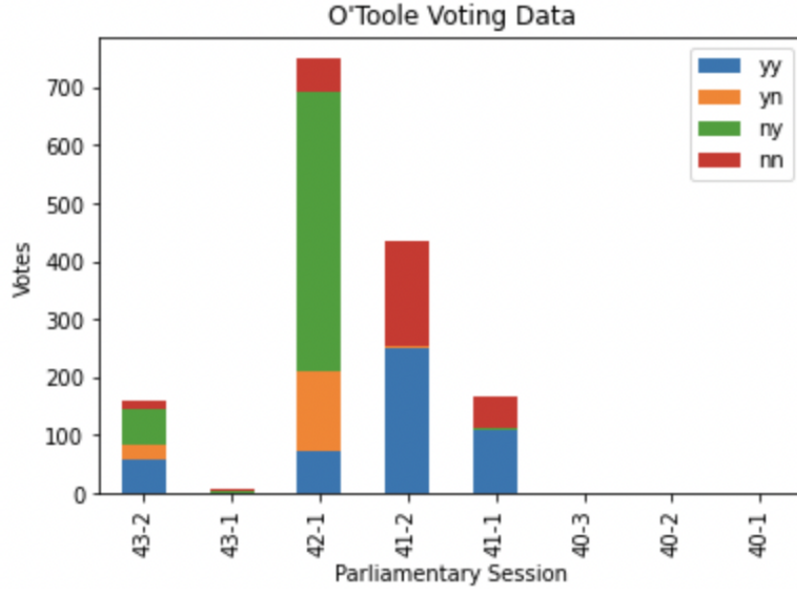


FIGURE 3

3. CANADIAN LEGISLATORS

Another plausible measurement for the performance of Canadian politicians in Parliament is to count how many bills they sponsored. To be more precise, we count the total number of sponsored bills on each topic for individual legislators. This could be an analogue of the contribution of a soccer player in a game by passing, scoring, and so on.

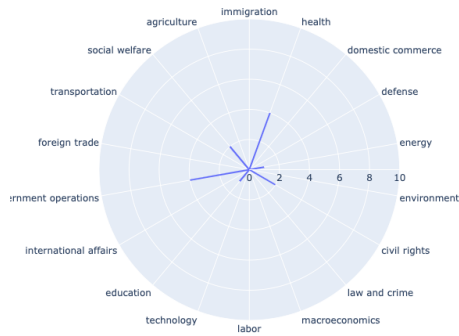
After collecting these data, we find that most legislators have completely no record. Thus this might not be a satisfying metric due to a lack of available Canadian legislation data. To visualize the performance, we draw radar plots. In Figure 4, we display the top six legislators who sponsored the most bills.

4. USA STATE BILL TOPICS

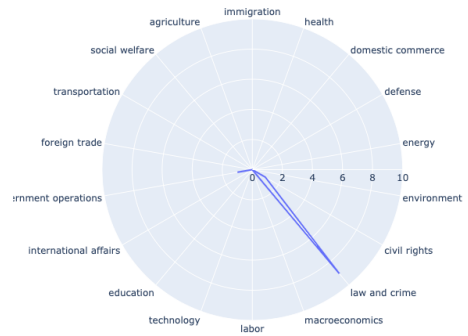
Given that the Canadian legislation data is limited, we move our focus to legislation data from the United States. But instead of considering individual legislators, this time we count the the total number of bills for each topic state by state. Due to the number of legislators might differ in different states, we use the total number of bills divided by the total number of legislators as our metric. This could be an analogue of the average performance of players in a particular position in a sport.

Doing so, from Figure 5, we see that most of states have very small graph, hence this metric is not very satisfying as well.

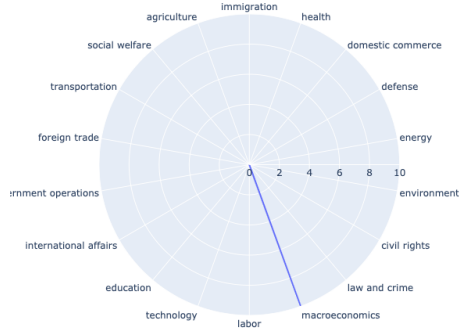
Nevertheless, we may still consider the weight (or percentage) for each bill topic. This measures the focus of legislation for each state, which could be an analogue of analyzing how a player in a particular position fares in a game. We use pie charts to visualize the difference between topic weights. In Figure 6, we see that



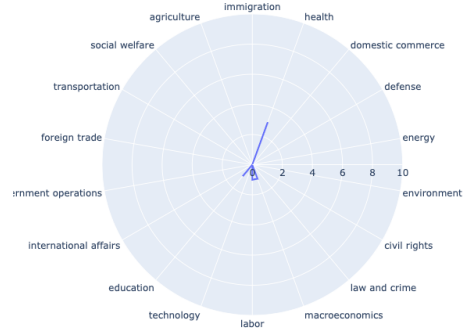
(A) Don Davies



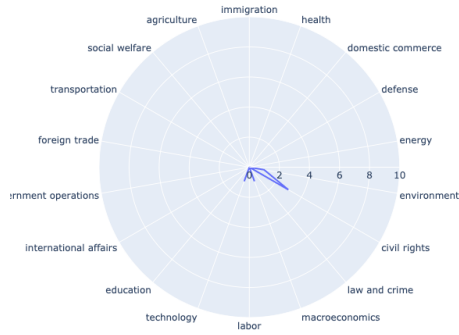
(B) David Lametti



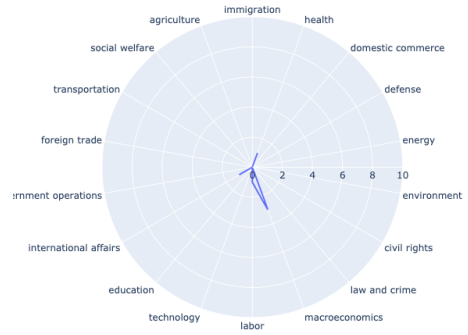
(C) Jean-Yves Duclos



(D) Carla Qualtrough



(E) Brian Masse



(F) Chrystia Freeland

FIGURE 4. Canadian Legislators' Radar Plot

the top two topics are always “government operations” and “law and crime”, while the remaining topics differ from state to state.

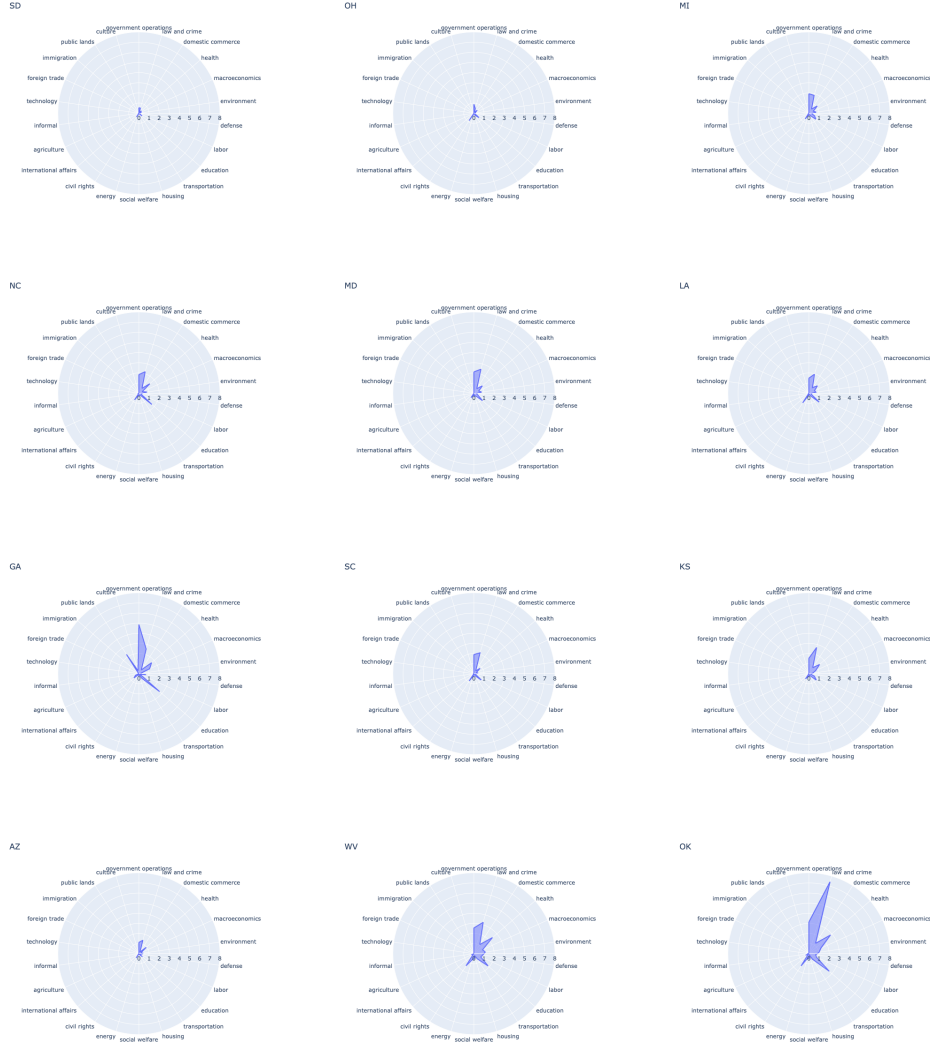


FIGURE 5. US State Bill Mean Radar Plot

5. GLOBAL INTERESTS VERSUS NATIONAL INTERESTS

In further pursuit of performance metrics, we turn our attention directly to issues addressed by legislators. Specifically, we focus on whether or not the issues a given legislator is concerned with is global in extent, or is confined to one's country alone.

Due to data shortage we consider two select states in the United States: Georgia and Ohio. We begin this analysis by extracting the data containing the names of legislators (state senators), the bills they sponsored, and the topics of the bills. Next, to achieve an experimental partition of the topics into those of global concern and those of national concern, we compare the topics with global issues currently

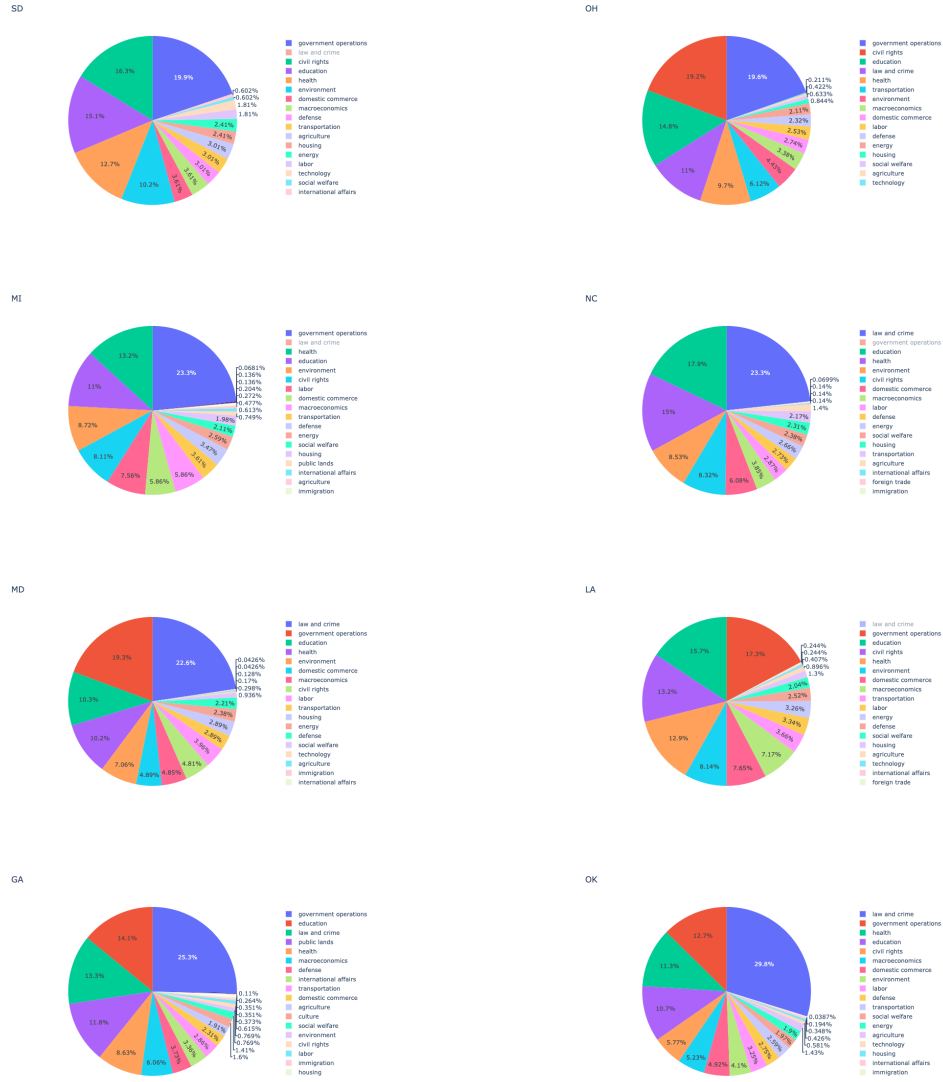


FIGURE 6. US State Bill weight Pie Plot

of interest to the United Nations. For example, topics such as “government operations”, “defense”, and “law and crime” are labelled as national issues, while “environment”, “education”, and “health” are labelled as global issues.

In Figure 7 we have four bar charts, each associated to a state senator (Butch Miller and Sandra Scott from Georgia, and Matt Dolan and Tina Maharath from Ohio), that track the number of bills principally sponsored by that senator under a given bill topic. The bars themselves have been labelled red and green; red indicates a topic that has been associated to a global interest, and green to a global interest.

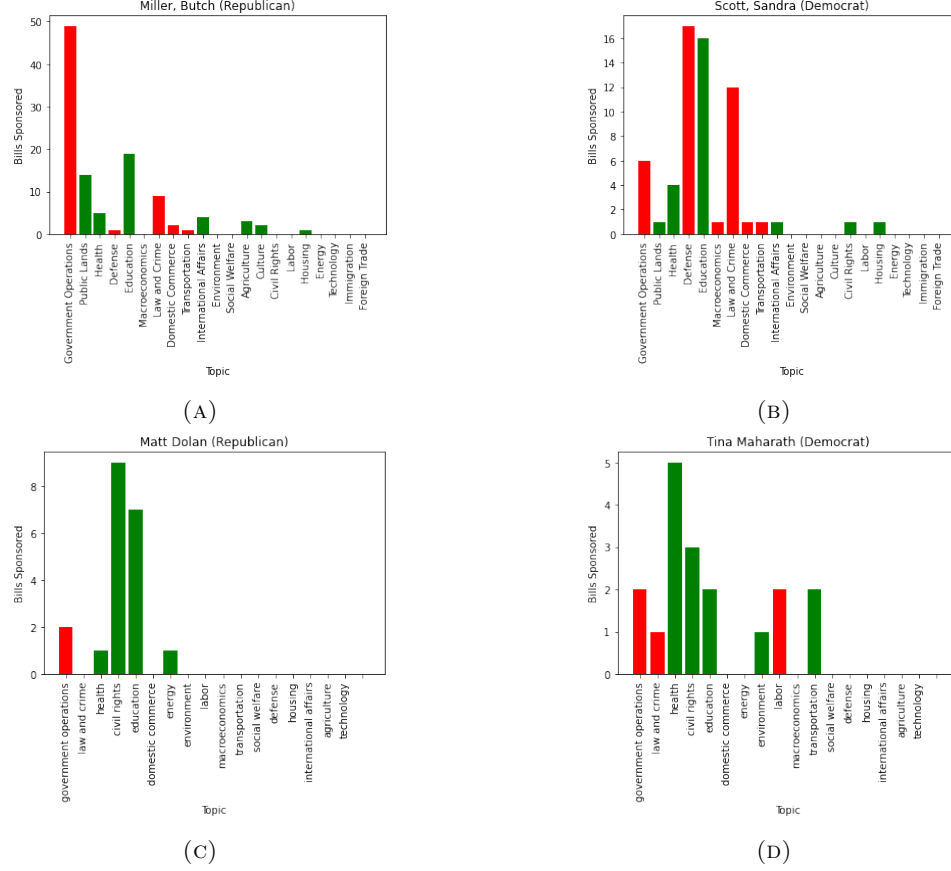


FIGURE 7. Global interests versus national interests in Georgia (top row) and Ohio (bottom row)

A naïve association of political party to interests' extent might associate the Republican party to national interests and the Democratic party to global interests. This notion is challenged by this data. Indeed, the amount of red and green in the bar charts does not correlate with political party in the way that one might expect, therefore a further investigation into whether or not a metric such as this will provide a more quantitative and rigorous measure of performance may be merited.

6. BIPARTISANSHIP IN THE UNITED STATES

We can visualize the interaction of legislators from the Democratic party and the Republican party by using network graphs. Such graphs are commonly used by social media such as Facebook, to analyze the relationships between its users. We can model each legislator as a node in a graph. Two nodes are connected by an edge if they cosponsored a bill together. In Figure 8 we look at the network graph for North Carolina legislators.

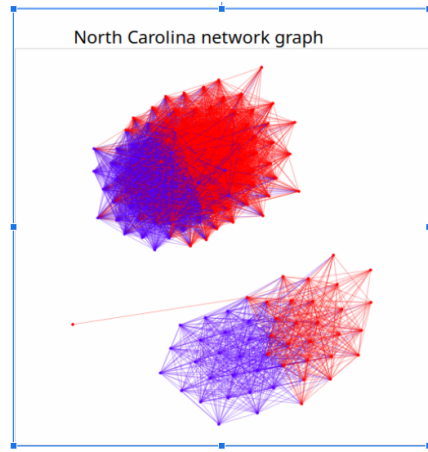


FIGURE 8. North Carolina Network Graph

The colour blue represents the Democratic party and the colour red represents the Republican party. We see two clusters of nodes, one where the connectivity is high within the clusters, and one where it is low across the clusters. On average, each node connects to 34.04 other nodes for members of the Democratic party, and each node connects to 31.07 other nodes for the Republican party. Democrats also have cross ratio of 30.7 and the Republicans have 23.1. All in all, this means that 30.7% of bills sponsored by Democrats are cosponsored with Republicans, and 23.1% of bills sponsored by Republicans are co-sponsored with Democrats. Upon analyzing the data sets from other states, a similar conclusion may be found: that Democrats have higher cross over ratio. This data suggests that Democrats tend to be more bipartisan.

7. CONCLUSION

It is fair to say that the content of this report is mostly descriptive, that is, an examination of and speculation upon historical data. Indeed, potential further work would be to gather more data and extend the preliminary metrics outlined above to rigourously evaluate performance, and perhaps even build predictive models that could illuminate directions in which politicians would carry their followers into the future. Needless to say, this is will be an extremely difficult task and it certainly requires more information and expertise, not least due to the immeasurable complexities of real world politics, where there are a lot of variables to affect politicians' votes and public opinions. Indeed, there are many things in the real world that are simply hard to quantify.

8. ACKNOWLEDGEMENTS

The authors would like to thank the organizers of the Math to Power Industry workshop for the incredible opportunity to train and acquire true industry experience. In addition, we thank our mentors Adam Kashlak of the University of Alberta, and William Spat and Samuel Leung of IOTO International for their helpful advice during the compilation of this project.