

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

RODRIGO NUNES MONI DA SILVA

**Visualization of Roll Call Data for
Supporting Analyses of Political
Profiles**

Thesis presented in partial fulfillment
of the requirements for the degree of
Master of Computer Science

Advisor: Carla Maria Dal Sasso Freitas

Porto Alegre
September 2019

CIP — CATALOGING-IN-PUBLICATION

Nunes Moni da Silva, Rodrigo

Visualization of Roll Call Data for Supporting Analyses of Political Profiles / Rodrigo Nunes Moni da Silva. – Porto Alegre: PPGC da UFRGS, 2019.

104 f.: il.

Thesis (Master) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR-RS, 2019. Advisor: Carla Maria Dal Sasso Freitas.

1. Visualization, Political Data, Coordinated Views. I. Dal Sasso Freitas, Carla Maria. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Rui Vicente Oppermann

Vice-Reitor: Prof. Jane Fraga Tutikian

Pró-Reitor de Pós-Graduação: Prof. Celso Giannetti Loureiro Chaves

Diretor do Instituto de Informática: Prof. Carla Maria Dal Sasso Freitas

Coordenador do PPGC: Prof. Luciana Salete Buriol

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

AGRADECIMENTOS

Agradeço à minha família que esteve presente em todos os momentos difíceis nestes últimos anos, sempre me apoiaram e nunca me deixaram desistir. Um agradecimento especial à minha avó materna por sempre acreditar em mim e que, infelizmente, não pôde estar presente em mais um momento especial da minha vida. Por fim, gostaria de agradecer a ajuda da minha namorada Franciele Lobo que contribuiu com seus feedbacks sobre o trabalho, arrecadação de usuários e análise dos resultados.

Gostaria de agradecer por toda ajuda e paciência da minha orientadora Carla Maria Dal Sasso Freitas. Foi uma experiência excelente trabalhar com ela, além de ser uma ótima orientadora, também é uma ótima pessoa. Pude aprender muito com ela.

Gostaria de agradecer toda ajuda e atenção do Andre Suslik Spritzer, nosso colaborador que foi, na prática, meu co-orientador. Detentor de um vasto conhecimento, contribuiu tanto na concepção de visualizações, como no desenvolvimento da dissertação.

Agradeço ao autor do CivisAnalysis original, Francisco Gerdau de Borja, pelo tempo despendido em me ajudar e por toda base de trabalho que ele me disponibilizou.

Quero agradecer meus amigos e meus colegas, por toda a ajuda com dúvidas técnicas e feedbacks sobre a ferramenta.

Gostaria de agradecer a todos aqueles que participaram da avaliação da ferramenta e possibilitaram a coleta de dados importantes.

Também quero agradecer ao Instituto de Informática da UFRGS pela excelente estrutura fornecida, com ótimos equipamentos e ambientes de trabalho.

Agradeço também ao CNPq e à CAPES, agências brasileiras de fomento à pesquisa, pelo apoio financeiro.

Por fim, agradeço a Deus por me ajudar a alcançar mais um objetivo na minha vida.

ABSTRACT

This work presents a set of interactive visualization techniques that lets users explore the voting histories of elected representatives in Brazil’s lower legislative house (the Chamber of Deputies) by instantiating multiple, coordinated panels containing different views. Understanding the voting behavior of representatives is of value not only to professional political analysts, but also to common citizens, who can use this information to make better choices when casting their votes at the ballot box. The several interactive visualizations are integrated in a web-based application that allows users explore the Chamber of Deputies’ roll call dataset made available as part of the Brazilian government’s open data initiative. Users can select periods of time, legislatures, or presidential terms from an initial timeline that displays political parties’ voting patterns from 1991 to 2018 and launch panels containing visualizations from which other visualizations can be launched, creating a hierarchy of connected, coordinated visualizations that can be freely arranged on the workspace by moving, hiding, and resizing panels. For any chosen period, the techniques provided include visualizations of the representatives’ political spectrum (using three different dimensionality reduction methods), “cropped” timelines allowing for the comparison of representative and party behavior, a hybrid heatmap-histogram visualization of roll calls, a similarity force-directed graph visualization, an infographic of the distribution of parties and representatives in the Chamber of Deputies, and two visualizations of the composition of a selected cluster of representatives (a force-directed visualization and a bar chart). When appropriate, users are also provided with tools to filter, search, and customize the visualizations. The system was designed from a set of visual analytics questions based on the dataset and a list of requirements aimed at addressing these questions. We illustrate its use and potential through a case study on the cohesiveness of parties over time and evaluate it through user studies including both general public and expert users.

Keywords: Visualization, Political Data, Coordinated Views.

Visualização de dados de votações como suporte para análise de perfis políticos

RESUMO

Este trabalho apresenta um conjunto de técnicas interativas de visualização que permitem a usuários explorar o histórico de votações dos representantes eleitos na Câmara dos Deputados da legislatura brasileira. O entendimento dos padrões de votos dos deputados é importante tanto para analistas políticos como para cidadãos comuns, que podem usar essa informações para fazer melhores escolhas nas eleições. As várias técnicas de visualização estão integradas numa aplicação *web* que permite que usuários explorem os dados de votações da Câmara dos Deputados disponibilizados como parte da iniciativa de dados abertos do governo brasileiro. Usuários podem escolher períodos de tempo, legislaturas ou mandatos presidenciais a partir de uma linha de tempo que exhibe padrões de votos de partidos políticos de 1991 a 2018. Novas visualizações são criadas a partir da linha de tempo ou de outras, criando uma hierarquia de visualizações conectadas e coordenadas que podem ser livremente organizadas na área de trabalho através da movimentação, minimização e redimensionamento das janelas. Para qualquer período escolhido, pode-se obter visualizações do espectro político dos deputados (usando três diferentes métodos de redução de dimensionalidade), recortes da linha de tempo permitindo a comparação do comportamento de partidos e deputados, uma visualização de votações híbrida que combina mapa de calor e histograma, uma visualização baseada em forças do grafo de semelhança dos deputados, um infográfico da composição da Câmara e duas visualizações da composição de um agrupamento de deputados (*cluster*) selecionado (uma visualização baseada em forças e um gráfico de barras). Usuários podem filtrar, buscar deputados e personalizar as visualizações. O sistema foi planejado a partir de um conjunto de questões de análise visual (*visual analytics*) baseadas no conjunto de dados e de uma lista de requisitos que objetiva tratar dessas questões. O potencial do sistema é ilustrado através de um estudo de caso sobre a coesão partidária ao longo do tempo. Foram também realizadas avaliações com usuários envolvendo tanto cidadãos comuns quanto especialistas.

Palavras-chave: Visualização, Dados Políticos, Visualizações coordenadas.

LIST OF ABBREVIATIONS AND ACRONYMS

CMV	Coordinated Multiple Views
PCA	Principal Component Analysis
MDS	Multidimensional Scaling
t-SNE	t-Distributed Stochastic Neighbor Embedding
UFRGS	Universidade Federal do Rio Grande do Sul
VAQs	Visual Analytics Questions

LIST OF FIGURES

Figure 3.1 Connect 2 Congress.....	25
Figure 3.2 Ideological map of 116th Congress (2019-2021.....	27
Figure 3.3 Map of the deputies per party in the Chamber of Deputies and Map of Brazilian Senators.....	27
Figure 3.4 Map of legislators of the Municipal Chamber of Rio de Janeiro.....	28
Figure 3.5 Social Action.....	28
Figure 3.6 Co-authorship of proposals between deputies.....	29
Figure 3.7 Comparison between PT and PSL co-authorship of proposals networks.....	30
Figure 3.8 A fragment of the infochart Conspectus of the History of Political Parties and the Federal Government.....	31
Figure 3.9 Political Timeline of the U.S. Congress since 1788.....	31
Figure 3.10 Timeline of the behavior of U.S. senators for eight Congresses.....	32
Figure 3.11 Political Timeline in Voteview.....	33
Figure 3.12 Summary Module of the original CivisAnalysis.....	34
Figure 3.13 Inspection Module of the Original CivisAnalysis.....	34
Figure 3.14 Basômetro.....	36
Figure 4.1 The political timeline: the main visualization of CivisAnalysis 2.0.....	43
Figure 4.2 Tree structure of CivisAnalysis 2.0 user interface.....	44
Figure 4.3 Separate panels linked to the timeline show the political spectrum of different periods of time.....	45
Figure 4.4 Comparison of the cohesiveness of a selected party in two legislatures.....	47
Figure 4.5 Similarity of voting between deputies.....	48
Figure 4.6 Chamber of Deputies Infographic and Map of Roll Calls displayed in coordinated views.....	49
Figure 4.7 Example of Cropped Timeline visualization.....	51
Figure 4.8 Clustered Force Layout and Bar Chart visualizations of deputies per party.....	51
Figure 5.1 Fernando Collor de Mello (PRN) – 1990-1992.....	55
Figure 5.2 Itamar Franco (PMDB) – 1992-1995.....	56
Figure 5.3 Fernando Henrique Cardoso (PSDB) – 1995-1999.....	57
Figure 5.4 Fernando Henrique Cardoso (PSDB) – 1999-2003.....	58
Figure 5.5 Luís Inácio Lula da Silva (PT) – 2003-2007.....	58
Figure 5.6 Luís Inácio Lula da Silva (PT) – 2007-2011.....	59
Figure 5.7 Dilma Rousseff (PT) – 2011-2015.....	59
Figure 5.8 Dilma Rousseff (PT) – 2015-2016.....	60
Figure 5.9 Michel Temer (PMDB) – 2016-2019.....	62
Figure 6.1 Grid of bar charts showing the relative frequency of answers for each VAQ.....	68
Figure 6.2 Bar chart showing differences between the understanding of each VAQ, using absolute frequency.....	68
Figure 6.3 Summary of results from questions on the user understanding about Timeline visualizations.....	70
Figure 6.4 Summary of results from questions on the user understanding about Political Spectrum visualizations.....	70

Figure 6.5 Summary of results from questions on the user understanding about Map of Roll Calls, Chamber Infographic and Similarity Graph.	72
Figure 6.6 Summary of results from questions concerning the user understanding about the system features.	73
Figure 6.7 Summary of the results from questions on the level of user understanding of the different visualizations.	79
Figure 6.8 Summary of the results from questions on the utility of the different visualizations.	80
Figure 6.9 Summary of the results from questions on the ease of use of the different visualizations.	81
Figure 6.10 Summary of results from questions on the ease of use in different visualizations.	83

LIST OF TABLES

Table 4.1 Differences between the original CivisAnalysis and CivisAnalysis 2.0..	54
Table 6.1 Summary of participants.....	64
Table 6.2 Summary of VAQs results.....	67
Table 6.3 The table compares two different groups based on previous experience with visualization tools. The first column shows the users that do not have previous experience and the second column, the users that have previous experience.	74
Table 6.4 Summary of experts.....	76

CONTENTS

1 INTRODUCTION	13
1.1 Motivation	14
1.2 Objectives and Contribution	15
1.3 Organization of the Dissertation	16
2 BACKGROUND	17
2.1 Brazilian Political System	17
2.2 Roll Calls in the Chamber of Deputies	18
2.3 Roll Call Analysis.....	19
2.3.1 Principal Component Analysis	21
2.3.2 Multidimensional Scaling.....	22
2.3.3 t-Stochastic Neighbourhood Embedding	22
2.4 Coordinated Multiple Views	23
2.5 Final Comments	23
3 RELATED WORK	24
3.1 Political Spectra	24
3.2 Political Networks	27
3.3 Political Timelines	29
3.4 CivisAnalysis	32
3.5 Charts and Reports-Based Tools.....	35
3.6 Final Comments	35
4 VISUALIZATION OF POLITICAL PROFILES	37
4.1 Data Description	38
4.2 Political Spectrum of Deputies.....	38
4.2.1 Input Matrices	39
4.2.2 Generating the spectrum	40
4.3 Tasks and Overview	41
4.4 Visualization Techniques	46
4.4.1 Spectrum of Deputies	46
4.4.2 Deputy Similarity Graph.....	47
4.4.3 Chamber of Deputies Infographic.....	48
4.4.4 Map of Roll Calls	49
4.4.5 Cropped Timeline.....	50
4.4.6 Clustered Force Layout	50
4.4.7 Bar Chart	52
4.4.8 Additional Features	52
4.5 Comparing CivisAnalysis 2.0 with the original CivisAnalysis.....	52
4.6 Final Comments	53
5 CASE STUDY: PARTY COHESIVENESS OVER TIME	55
5.1 Collor (PRN) – 1990-1992.....	55
5.2 Itamar (PMDB) – 1992-1995	56
5.3 FHC (PSDB) – 1995-1999	57
5.4 FHC (PSDB) – 1999-2003.....	57
5.5 Lula (PT) – 2003-2007	58
5.6 Lula (PT) – 2007-2011	59
5.7 Dilma (PT) – 2011-2015.....	60
5.8 Dilma (PT) – 2015-2016.....	60
5.9 Temer (PMDB) – 2016-2019	61
5.10 Final Comments	62

6 EVALUATION WITH USERS	63
6.1 Remote Evaluation with General Public	63
6.1.1 Participants.....	64
6.1.2 Procedure.....	65
6.1.3 Results.....	66
6.1.3.1 Timeline and Cropped Timeline.....	69
6.1.3.2 Political Spectrum.....	69
6.1.3.3 Map of Roll Calls, Chamber Infographic and Similarity Graph.....	71
6.1.3.4 System Features.....	72
6.1.3.5 Final Analysis.....	74
6.2 Remote Domain Experts Study	75
6.2.1 Participants.....	75
6.2.2 Procedure.....	76
6.2.3 Results.....	77
6.2.3.1 Understanding the visualizations.....	78
6.2.3.2 Usefulness of the visualizations.....	79
6.2.3.3 Ease of use of the visualizations.....	80
6.2.3.4 General system features.....	82
6.2.3.5 Experts comments, suggestions and feedback.....	83
6.3 Final Comments	84
7 CONCLUSIONS AND FUTURE WORK	85
REFERENCES	86
APPENDIX A — QUESTIONNAIRES	91
A.1 General Public	91
A.2 Experts	98

1 INTRODUCTION

Several open data portals provide government information for the general public, researchers, and professionals as well as application developers. As examples, we can mention the sites that allow tracking information about the US Congress¹ and the Brazilian Senate² and Chamber of Deputies³. These portals let people obtain information on specific deputies, senators, and propositions through simple form-based queries. Developers and more advanced users can download datasets such as the detailed federal budget and the voting history of representatives in the National Congress.

Concerning the understanding of legislators' behavior, however, there is a considerable gap between the works developed by political scientists and those aimed at informing the general public. Political scientists over the past 28 years have used the analysis of roll call voting data for developing empirical models representing the spatial theory of voting (CARROLL; POOLE, 2014). Such models are also known as *ideal point estimators* because they infer the location of a legislator in an abstract space based on their roll call votes. Usually represented as a scatterplot, this space (or *political spectrum*) can be as broad as the whole set of roll calls from a legislature or as specific as a set of roll calls related to some particular issue. This type of data has been used in the study of the US Congress as well as in the analysis of legislative behavior in general (POOLE; ROSENTHAL, 1991; SCHONHARDT-BAILEY, 2003; ROSENTHAL; VOETEN, 2004; HIX; NOURY; ROLAND, 2007).

There has been some effort in communicating findings resulting from the methods mentioned above to the general public by providing visualizations of the political spectrum of legislators along with explanations or narratives (KINNAIRD; ROMERO; ABOWD, 2010; BASÔMETRO, 2019; MARINO, 2014; LEITE; TRENTO, 2016). However, these visualizations are either static, merely informative, or allow for little exploration because they rely on showing the political spectrum of legislators and not the roll call data set. Moreover, for the general public, such an abstract space may be too abstract for a clear understanding. This context motivated a project on the visualization of the Chamber of Deputies' open data.

¹<http://www.gov.track.us>, <http://www.gov.track.us/developers>

²<https://www12.senado.leg.br/transparencia/>

³<http://www2.camara.leg.br/transparencia/acesso-a-informacao/portal-da-camara-informacoes>

1.1 Motivation

It is said that information visualization is “the use of computer-supported, interactive, visual representations of abstract data to amplify cognition” (CARD, 1999, p. 8). The objective of tools and methods in information visualization is essentially to “help us speed up our understanding and action in a world of increasing information volumes” (CARD; JACKO, 2008, p. 542). In the field of political communication, science, and journalism, information visualization plays an essential role (WINDHAGER; SMUC, 2014). An ever-growing number of newsrooms and online newspapers use interactive charts and infographics to provide insight into political data, systems, and processes. It is vital for a healthy democracy to make this type of information widely available so that citizens can be properly informed about their representatives’ vote on the different roll calls that might impact their lives.

This scenario motivated the development of CivisAnalysis (BORJA; FREITAS, 2015; BORJA, 2017), a web-based tool for the visualization of roll call data of the Brazilian Chamber of Deputies. The main visualizations provided by this application were two kinds of political spectrum: one showing the deputies and another displaying the roll calls. When legislators are closely located in the political spectrum of deputies, it means that their voting patterns are similar; roll calls that are close in the political spectrum of roll calls mean that they have received similar votes from the same deputies.

CivisAnalysis used a data set of voting records of six legislatures (24 years, 513 deputies), comprising 2,458 roll calls (853,952 votes), as well as the information of six presidential elections, including election results and alliances made for the elections. Besides the two political spectra, CivisAnalysis also provided other visualizations: (i) a timeline, displaying the relative position of parties (computed from the deputies political spectrum) along the years, (ii) an infographic showing the composition of the Chamber in terms of parties, (iii) a map for filtering deputies per state, and (iv) a text panel for the text of propositions selected in the roll call political spectrum. It lets users observe the distribution of deputies on a political spectrum and compare the position of a specific legislator in relation to the others.

Still, it had limitations on how much users can learn about the political profiles of deputies and parties. It was not possible, for instance, to compare a deputy’s voting behavior to that of other deputies or the deputy’s party’s policies

across different legislatures, presidential terms, or periods of time. Besides, in an evaluation of the system, users reported having some difficulty understanding the application. The most problematic issue they cited was the spectrum of roll calls, because most of them were unable to extract any relevant information from that chart.

We summarize below the main limitations of CivisAnalysis:

1. Users can explore only a period at a time, making it impossible for them to compare different legislatures, terms, or years
2. It is only possible to analyze deputies individually or in contrast to other deputies while it would have been interesting to also be able to compare their patterns of behavior to those of political parties
3. The Spectrum of Roll Calls is not easy to interpret and users had some difficulties in exploring and getting useful information from it.

Based on CivisAnalysis, we devised the present work with the goal of both addressing the limitations cited above as well as supporting more types of visual analyses.

1.2 Objectives and Contribution

The main objective of this work is to provide an integrated set of visualization techniques to help users get insight into the changing profiles of deputies and parties through the exploration of the entire roll call voting dataset from 1991 to 2018, which includes seven legislatures and nine presidential terms. Visualizations are accessed through a user interface based on multiple coordinated panels that are organized hierarchically.

The main contributions of our work are:

- A robust system that allows users to create visualizations of political data based on any period of time. Users can create several coordinated panels, each displaying a (possibly different) visualization.
- A set of interactive techniques that let users compare deputy and party voting patterns over time.

We based our work on CivisAnalysis, but only used the original code for two of the visualization techniques (the Timeline and the Chamber of Deputies Infographic, described in Chapter 4) as well as the computation of the deputies' Political Spectrum, namely the Principal Component Analysis (PCA) (JOLLIFFE, 2011) and t-Stochastic Neighbourhood Embedding (t-SNE) (MAATEN; HINTON, 2008) implementations. We named our tool CivisAnalysis 2.0, with the agreement of the previous authors (BORJA; FREITAS, 2015; BORJA, 2017).

With this work, we expect to help professional political analysts and common citizens better understand how their elected representatives have voted over the course of their political history. Along with being useful for its historical and analytical value, this helps citizens not only hold politicians accountable for their actions, but also make better and more well-informed choices when they cast their votes at the ballot box.

1.3 Organization of the Dissertation

This dissertation is organized as follows:

- Chapter 2 gives a short introduction to Brazilian politics and roll call analysis.
- Chapter 3 reviews the visualization solutions adopted by political scientists to analyze roll call data. It also introduces CivisAnalysis (BORJA; FREITAS, 2015; BORJA, 2017).
- Chapter 4 presents the design and implementation details about CivisAnalysis 2.0, the visualization solution we propose for the exploration and analysis of the roll call dataset.
- Chapter 5 demonstrates the use of our system in a study of party cohesiveness over time.
- Chapter 6 presents the evaluation we conducted, which included users tests with common citizens as well as domain experts.
- Chapter 7 provides a summary of our work as well as comments on what is to be carried out in the future.

2 BACKGROUND

In this chapter, we briefly introduce topics that are important for giving the context and appropriate background for understanding our work. We shortly introduce the Brazil's political system and aspects of the roll calls in the Chamber of Deputies, and present concepts on roll call analysis, including dimensionality reduction techniques. We also comment on coordinated multiple views, an essential aspect of our interface.

2.1 Brazilian Political System

Brazil is a Presidential Federative Republic. The president is the Head of Government of the Union and Head of State, is elected for a four-year term, but may be re-elected for another term. Also, its important to mention that since Brazil is a Federative Republic, each state has political autonomy within its respective government, although this autonomy is limited by the principles established in the Federal Constitution. The State Power is divided into distinct political entities: Executive, Judiciary and Legislative, based on the "*Separation of Powers*" theory (MONTESQUIEU, 1989). The Executive power, represented by the figure of the President, is in charge of finishing the process of law creation, because h/she can veto or sanction any motions created by the Legislative. Besides that, the President can propose laws or post provisional measures (*Medidas Provisórias*), which should be approved by the Congress to become laws. The Legislative is represented by the National Congress, which is bicameral and is composed by two different Houses: the Chamber of Deputies and the Federal Senate. The Legislative has to create and describe the laws, besides to discuss and decide about the president's proposals. Every law has to pass through the Chamber of Deputies and then, if approved, it has to pass in the Senate. The Judiciary, almost exclusively exercise jurisdictional obligations, which means interprets the laws and enforces them.

All members of the executive and legislative powers are directly elected by Brazilian citizens. However, judiciary members are appointed by the President and, after that, they must be approved by the Senate. In most part of Brazil's democratic history, it has had a multi-party system with proportional representation. Participation in the elections is mandatory for the literate between 18 and 70 years

old, and optional for illiterates and those between 16 and 18 or older than 70.

General elections occur every four years, and in each election the Brazilian citizens elect a president and a legislature at the national level. The president is elected through a two-round system, where the candidate with absolute majority of votes is the new Head of State. Also, at the state level governors and local legislatures are elected. The National Congress, as mentioned before is divided in two Houses: 1) *Senate* has 81 members, elected for an eight-year term, with elections every four years for replace, alternatively, one-third and two-third of the seats. 2) *Chamber of Deputies* has 513 members, elected for a four year term by proportional representation, which means that the number of deputies elected by each electoral district is directly proportional to its inhabitants (minimum of 8, maximum of 70).

2.2 Roll Calls in the Chamber of Deputies

In the current legislature, the Chamber of Deputies has 513 representatives distributed over 30 parties, a record number of parties since 1986.

A roll call is the last stage of the legislative process that completes the cycle of the discussion of a proposition. A roll call will be valid and voted only if the quorum was verified. There must be at least 257 members in the plenary for a roll call.

To approve a motion in the Chamber of Deputies, the number of votes required varies according to the type of proposal. For example, to pass a Bill Project (*Projeto de Lei*) and a Provisional Measure (*Medida Provisória*), it is necessary a quorum of 257 representatives. Once obtained the minimum of members present, the proposal can be approved by simple majority, i.e. the major part of those deputies present in the session. In contrast, to approve a Proposal for Amendment to the Constitution (*Projeto de Emenda Constitucional*) and pass to Senate, it has to be approved by the plenary in a two-rounds system, with the votes of 3/5 of the deputies (308 votes).

The Brazilian Chamber of Deputies website maintains raw data related to the legislatures in archives (.csv, .json, etc), and provides an open web service to access this data through custom applications. The data includes legislature information, legislative motions (proposed, laws, procedures, etc.), roll calls, deputies information, office expenses, commissions information.

Basically, the data consists in a huge array of votes for each roll call, where each vote is represented by the name of the deputy and its vote *value*. The vote *values* are defined as follows:

1. *Yes* - Accepts the proposal
2. *No* - Declines the proposal
3. *Abstention* - No opinion, only counts for quorum purposes
4. *Obstruction* - Decided not to vote or directly expressed *in obstruction* and does not count for quorum purposes
5. *Null* - No data available
6. *Art. 17* - Value of neutral vote of current Chamber's president. The vote of president only has value if occurs a tie and its necessary to break it

All information about roll calls used in this work was accessed through the open web service of Chamber of Deputies.

2.3 Roll Call Analysis

Currently, most theories of legislatures and elections are somehow based on spatial models of the political system, where each political actor is positioned at its ideal point in a spectrum (ENELOW; HINICH, 1984). The results of such techniques may be interpreted both in probabilistic and in geometric terms.

Researchers use distinct sources to build political models. The most significant are: transcripts of representatives' discourses, motions co-authorships, public opinion surveys, committees seats, and the most valuable: roll call votes (FISHER; LEIFELD; IWAKI, 2013) (DESPOSATO, 2006) (GRIMMER; STEWART, 2013) (BARBERÁ et al., 2015).

Roll calls votes are commonly used to create spatial models due to their nature of faithfully expressing the political behavior of each representative. Clinton precisely defined the combination of the Spatial Theory of Voting with the roll call analysis: "In short, roll call analysis makes conjectures about legislative behavior amenable to quantitative analysis, helping make the study of legislative politics an

empirically grounded, cumulative body of scientific knowledge." (CLINTON; JACKMAN; RIVERS, 2004).

In the early 1980s, the political scientists Poole and Rosenthal developed a family of algorithms to analyze the legislative roll call voting behavior, the NOMINATE family (POOLE; ROSENTHAL, 1985). Every algorithm variation works based on the application of the multidimensional scaling technique and operates under the same fundamental assumptions: 1) project the data in a low-dimensional space (often two-dimensional); 2) within that space, representatives have utility functions normally distributed and maximized at their ideal point. The estimation of ideal points in spatial models is calculated by maximum-likelihood estimation (MLE) methods. The NOMINATE family of algorithms are widely used in roll call analysis and compare favorably to more modern algorithms (CARROLL et al., 2009).

To perform the roll call analysis in this work, considering prior techniques, we chose faster and less complex dimensionality reduction methods, such as Principal Component Analysis (PCA), classic version of Multidimensional Scaling (MDS) and t-Distributed Stochastic Neighbor Embedding (t-SNE).

PCA or Karhunen-Loeve transformation (JOLLIFFE, 2011) (RAO; YIP, 2000) creates a new space through dimensionality reduction while retaining the variance of the data. These dimension reduction technique tend to sparse different individuals and approximate similar ones by removing correlations. Despite the relative imprecision of PCA to find ideal points on the political space, the voting patterns observed are quite similar to those obtained with more sophisticated methods (JAKULIN et al., 2009).

MDS (TORGERSON, 1952) is a very flexible method that allows for the creation of a low-dimensional space that represents the similarities or dissimilarities between pairs of objects, in our case, political actors. Some significant works used MDS to gain insights into political science. This method helped to find coalition patterns on the Council of Ministers of the European Union (MATTILA; LANE, 2001). Also, a study compared six European countries to analyze the impact of globalization on the national political space. A relevant study analyzed in detail multidimensional scaling applied in United States House of Representatives roll call votes of 2005 (DIACONIS et al., 2008). Also, by using MDS to construct a graphical representation of the positions of parties and issues (Welfare, Economic liberalism,

Culture, Army, Security, etc), another work reported consistent similarities in the structure of the national political space between these countries (KRIESI et al., 2006).

Finally, the t-Stochastic Neighbourhood Embedding (MAATEN; HINTON, 2008) is now a popular dimensionality reduction technique, used for visualizing diverse kinds of data. Usually, data is embedded in two or three dimensions, creating interpretable visualizations of high dimensional spaces. In political science, t-SNE can be useful to cluster political actors, such as legislatures or parties, and provides meaningful views of a certain political space (BRIGADIR et al., 2016) (TAMAYO, 2014).

Previous studies compared PCA with the popular MDS method W-NOMINATE applied on roll calls data of the Brazilian Chamber of Deputies and Senate. The results showed that PCA is way faster than W-NOMINATE and gives similar results (LEITE; TRENTO, 2016). Another work compared MDS with Factor Analysis, which is a related method to principal component analysis (PCA), but they are not identical (BRAZILL; GROFMAN, 2002).

2.3.1 Principal Component Analysis

The principal component analysis (PCA) is one of the most popular methods for the dimensional reduction of large data sets (KANTARDZIC, 2011), made available in several statistical software and libraries. It is an unsupervised technique that simplifies multidimensional data to a low-dimension space where it is easier to interpret using visualizations, improving the understanding of the studied subject. PCA creates new variables, the principal components, each one being a linear combination of the original variables in order to keep the amount of data variance that they explain. The results of Principal Component Analysis are displayed as projections of multidimensional scores (the amount of the principal component present in a particular sample) and loadings (the contribution that each variable makes to the principal component) in two or three-dimensional spaces.

2.3.2 Multidimensional Scaling

Classic Multidimensional Scaling (MDS) is a broadly used method for analyzing the similarities between elements in a data set, approximating the interpoint distances, or dissimilarities, of points in a high-dimensional space by actual distances between points in low dimensional Euclidean space. It makes the visual data analysis and exploration easier than using a collection of tabular data (BORG; GROENEN, 2003). MDS is a widely known and referenced technique, for early and precise references, see Young and Householder (YOUNG; HOUSEHOLDER, 1938) and Torgerson (TORGERSON, 1952). It receives as input a distance matrix with the distances between each pair of elements in a data set, and the selected number of dimensions. The resultant coordinate matrix can be derived by eigenvalue decomposition from $\mathbf{B}=\mathbf{X}\mathbf{X}'$, and the matrix \mathbf{B} can be calculated from the input matrix by using double centering (WICKELMAIER, 2003).

2.3.3 t-Stochastic Neighbourhood Embedding

The t-SNE is a modification of Stochastic Neighbor Embedding (ROWEIS; SAUL; HINTON, 2002), which is much easier to optimize, and produces significantly useful visualizations.

This method represents each object as a point in a two-dimensional space and manages the points in such a way that similar objects are positioned in nearby points and dissimilar objects are positioned in points further apart. A map built using t-SNE typically gets better results than when it is constructed by using other techniques, like principal components analysis or classical multidimensional scaling. This effect occurs due to two factors: 1) t-SNE principally focuses on appropriately modeling small pairwise distances (local structure) in the map and 2) t-SNE has a procedure to correct for the enormous difference in the volume of high-dimensional and two-dimensional space. As an outcome of these two features, t-SNE commonly produces maps that provide much clearer insight into the underlying (cluster) structure of the data than the alternative techniques cited before.

Two main stages describe the algorithm: 1) It constructs a probability distribution over pairs of high-dimensional space in such a way that similar objects have a high chance of being picked while different points have a tiny chance of being

selected. 2) It defines a similar probability distribution over the points in the low-dimensional map, and it minimizes the Kullback–Leibler divergence, a measure of how one probability distribution is distinct from another (KULLBACK; LEIBLER, 1951) (KULLBACK, 1997), between the two distributions concerning the locations of the points in the map.

2.4 Coordinated Multiple Views

Coordinated Multiple Views (CMV) is a visualization approach for exploring the data (usually a large volume of data) utilizing different views that are coordinated, i.e., interaction with data in a view triggers some change in the related views. (BALDONADO; WOODRUFF; KUCHINSKY, 2000)(ROBERTS, 2007). The use of CMV can be useful in data analysis, because, it can make straightforward to uncover relationships that may be hard or impossible to notice with only a single visualization. The general assumption for the technique is that users comprehend their data better if they interact with the exhibited information, and view it through different representations, perceiving distinct aspects of the data.

2.5 Final Comments

In this chapter we introduced the different methods we used in our work to generate the spatial model of the Chamber of Deputies: PCA, MDS, and t-SNE. Each technique provides a different view of the political spectrum due to its intrinsic characteristics. We provide these three methods to allow the user to find a better representation of the set of deputies, and then get the best understanding of their behavior in a given legislature, term, and period.

The user interface we developed for CivisAnalysis is strongly based on coordinated multiple views. From a visualization, users can derive other views, and selection of elements in one view might modify other views accordingly. For example, the user is able to create three different political spectra based on different dimensionality reduction techniques, and with the CMV support can compare them.

3 RELATED WORK

In this chapter we review the related works on visualization of political data. There are few academic projects devoted to that, but we found some applications used for communication in news media. Although most part of these works explore the U.S data (Senate and House of Representatives), there are some relevant works on Brazilian Political data as well.

3.1 Political Spectra

In their analyses of roll call data, many political scientists have used spatial models to depict the position of legislators in a political space (CARROLL; POOLE, 2014) and to measure and compare the ideology and heterogeneity of political parties (CARROLL; KUBO, 2017). Systems based on such models try to provide insights about the political variation in a certain abstract space. The first systems were usually built from regression analysis and questionnaires. The main problem with these kind of systems was the high bias, frequent modifications were needed to adapt them, and so represent different political contexts (LESTER, 1994).

There are some very popular tools on the internet that creates political spectra based on questionnaire responses, such as Political Compass¹ and Political Coordinates². Basically, both systems work in the same way: the user has to answer a set of questions related to economics, religion, human rights, etc; after that, a political spectrum is displayed with the user positioned inside it and some historical personalities to compare.

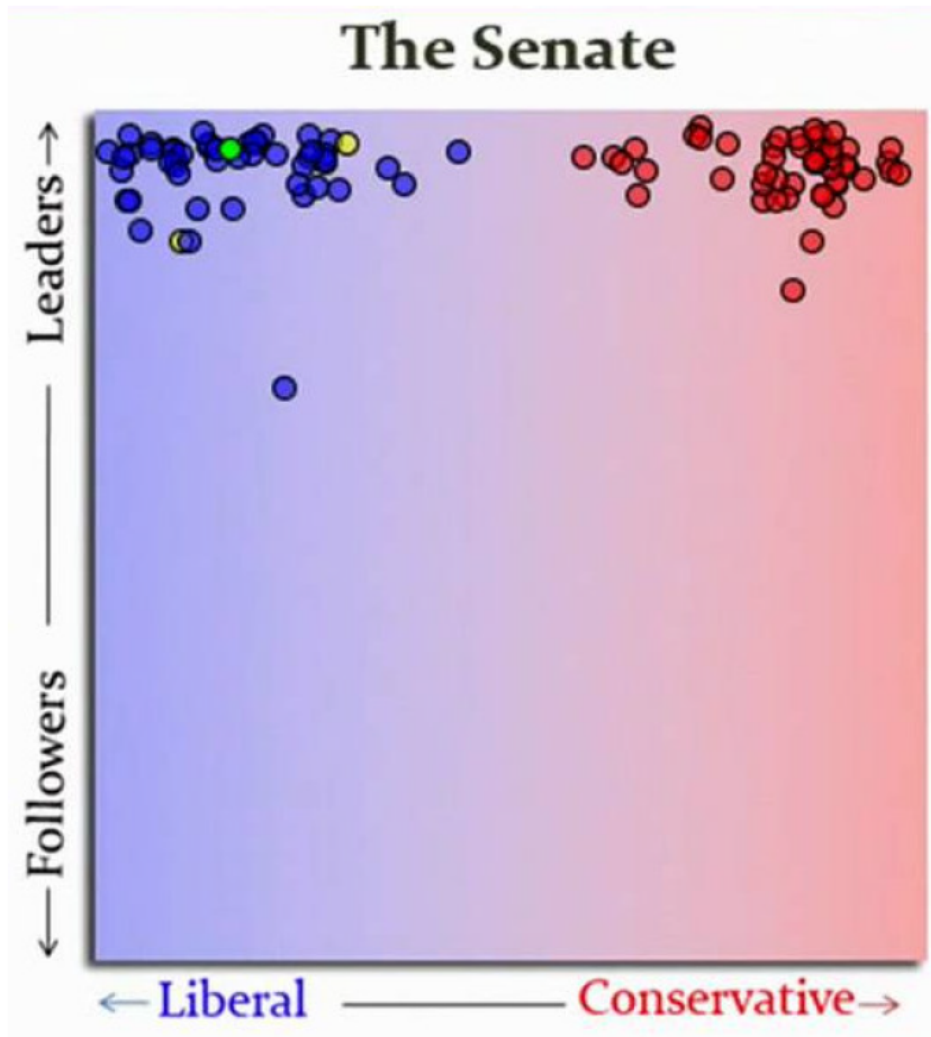
Aiming at reducing bias, current tools chose to use only quantitative information about representatives, primarily from recorded votes. In this scenario, a political spectrum is built based on the results from roll call analysis methods. Generally, scatterplots are used to represent the spatial models created by roll calls analysis techniques. Each representative is displayed as a point whose position represents its value associated to the variable axes. Thus, the closeness of two legislators on the map shows how similar their voting are.

Using NOMINATE scores, Connect 2 Congress (C2C) (KINNAIRD; ROMERO;

¹<<http://www.politicalcompass.org/>>

²<<http://www.celebritytypes.com/political-coordinates/test.php>>

Figure 3.1: Connect 2 Congress. Source:(KINNAIRD; ROMERO; ABOWD, 2010)



ABOWD, 2010) (Fig. 3.1) creates a two-dimensional political spectrum of the U.S. Congress for arbitrary time frames within a two-year period (2007-2008). Users can inspect roll call data as tabular data and filter and highlight representatives by name, state, party, religion, and gender. The time frame can be dynamically modified, resulting in an animation where representatives are continuously repositioned on the spectrum according to their behavior. As it only covers two years, C2C does not have inter-legislature analysis.

Using DW-NOMINATE to calculate ideological positions, Voteview³ (Fig. 3.2) allows users to inspect every congressional roll call vote in American history supported by a U.S. map, which shows how votes are distributed across the territory. Also, users can explore the data through an ideological map that includes information about the ideological positions of voting Senators and Representatives. The

³<<https://voteview.com/>>

system uses scatterplots to represent such maps for an entire period or a selected roll call, but also uses spatial visualization techniques to display the proportion of votes across the country and temporal visualization techniques to show the parties ideology variation across the time.

As for the Brazilian Congress, Marino (MARINO, 2014), Radar Parlamentar (LEITE; TRENTO, 2016) and Atlas Político⁴ create two-dimensional spectra of deputies. Marino displays the political spectrum of Deputies as a scatterplot, where the diagonal divides the government coalition and the opposition.

Radar Parlamentar shows the political spectrum in a radial format and only allows for the analysis of distances, as party positions change significantly over short periods of time. Parties appear in clusters, with the radius of the party's corresponding circle being proportional to its number of deputies, and its position being the average position of its deputies, as one can see in Fig 3.3(a). Parties can be expanded to show the original positions of their deputies.

Atlas Político is a platform focused on data intelligence, which provides several features, such as: (i) rankings for senators and deputies, according to five indexes created by the authors and based on a methodology that aims the competence of representatives on defending a political agenda independent of their ideological content; (ii) monitoring tweets about relevant political topics like government approval, military coup, gun control, with help of visualization techniques; (iii) content aggregating of relevant news in politics; (iv) map of Senate and Chamber of Deputies, using a political spectrum, (Fig 3.3(b)). Moreover, it provides a set of exclusive features for premium users, such as daily and weekly reports about leading trends in politics on the web.

Both Marino's work and Radar Parlamentar are based on PCA, and Atlas Político is based on DW-NOMINATE. Recently, MDS was used on the roll calls of deputies in the Chamber of Rio de Janeiro to build an ideological map of the representatives, which allowed the analysis of the cohesiveness of the most relevant parties (NERY, 2018) (Fig. 3.4).

⁴<http://atlaspolitico.com.br/>

Figure 3.2: Ideological map of 116th Congress (2019-2021). Source: (VOTEVIEW, 2019).

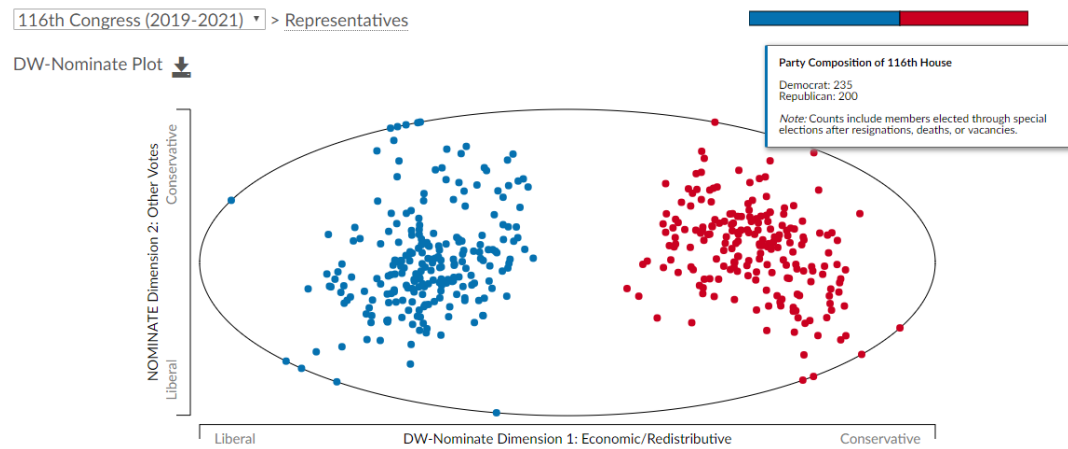
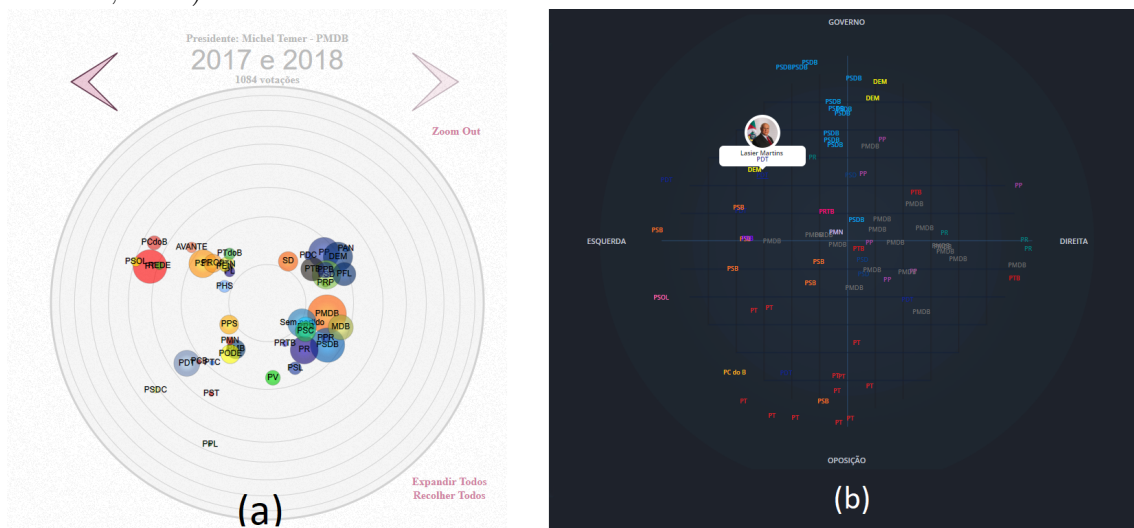


Figure 3.3: Map of the Chamber of Deputies per party (a) in Radar Parlamentar (LEITE; TRENTO, 2016). (b) Map of Brazilian Senators in Atlas Político (ATLAS... , 2019).



3.2 Political Networks

Social networks are commonly used in several fields of science: anthropology, biology, communication, economics, geography, information science, social psychology and sociology. Basically, a social network is a system composed by a set of social actors (can be individuals or entire organizations) and a set of connections between these actors. Applying this approach to political science, the political actors are the representatives, congress committees, political issues, etc., and the relationships are roll calls votes, speeches, co-authorship in motions, committees seats, etc.

Social Action (PERER; SHNEIDERMAN, 2008) builds a visualization of the behavior of the U.S. senators through a social network (Fig. 3.5) that represents

Figure 3.4: Map of legislators of the Municipal Chamber of Rio de Janeiro. Source: (NERY, 2018)

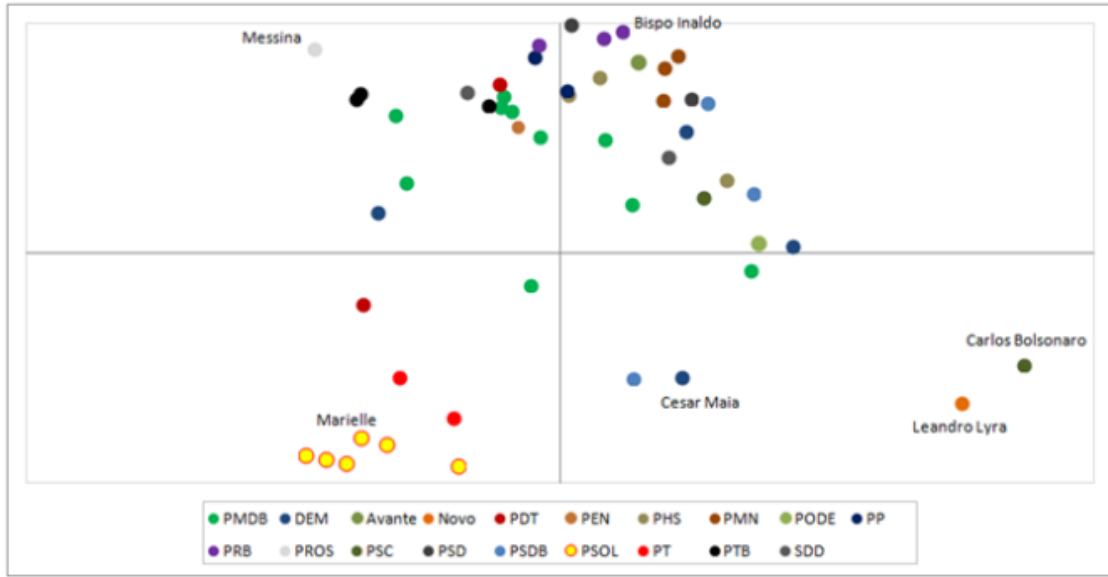
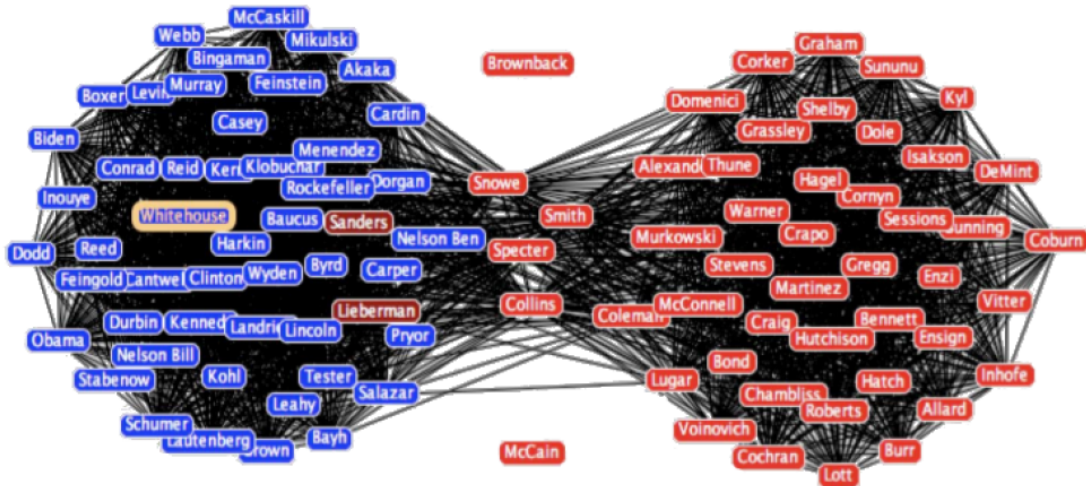


Figure 3.5: Social Action. Source: (PERER; SHNEIDERMAN, 2008).



the correlation of votes between senators. The network is displayed using a force-directed layout algorithm, and users can explore the data by interactively applying filters and statistical tools to uncover patterns of voting groups at specific points in time.

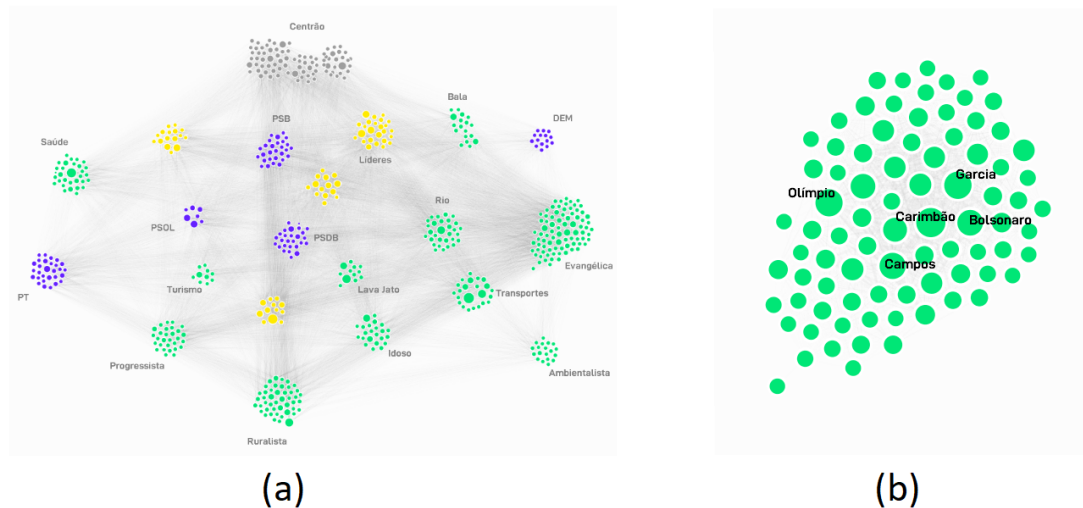
As for the Brazilian Congress scenario, some relevant articles were published, using representatives of the Chamber of Deputies as social actors and co-authorship of motions as a link between them.

The first study explored the data of past legislatures and revealed that some Parliament Fronts⁵ are more cohesive and bigger than some parties. Furthermore,

⁵groups of representatives tied by some common interest and not necessarily by party: defense of civil armament, defense of agriculture, evangelical, etc.

the author discusses the most influential representatives: Jair Bolsonaro, who would win the general elections in 2018, was one of them (Fig. 3.6)(NERY, 2019b). The second work explores the current Legislature, analyzing mainly the current President’s party (PSL) and the largest opposition party (PT) in the Chamber of Deputies. The study showed that PSL is very sparse and not structured as the opposition. Fig. 3.7 shows the network of both parties (NERY, 2019a).

Figure 3.6: Co-authorship of proposals between deputies: Representatives are political actors and co-authorship of proposals are links. The deputies are grouped by Parliament Fronts and other well-known groups. (b) The Evangelical group is more cohesive and larger than other parties. The most influential deputies are highlighted (NERY, 2019b)

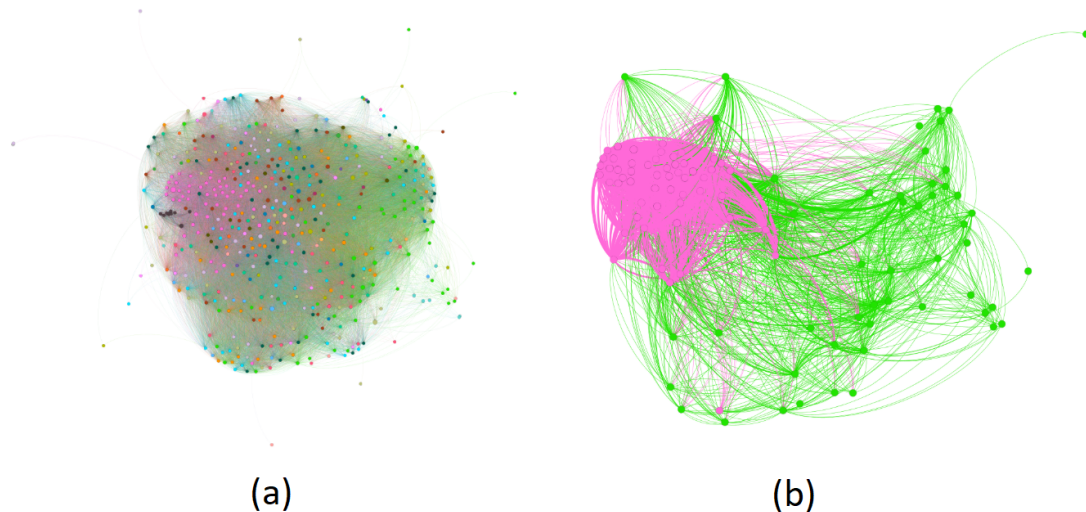


3.3 Political Timelines

Using temporal visualization techniques to represent variation of political positions can be very useful to inspect the political trajectories of individuals or parties over time. These political timelines provide a long-term overview, usually depicted as infocharts.

The “Conspectus of the History of Political Parties and the Federal Government” (HOUGHTON, 1880) shows the U.S. parties political position as a flowchart in the period of 1700 to 1880. The x-axis represents the time and y-axis represents which party has the president seat. The infochart has several annotations of historical facts and highlights main seats in executive, legislative and judiciary sectors (Fig. 3.8).

Figure 3.7: Comparison between PT and PSL co-authorship of proposals networks: (a) each deputy is represented as a circle and is linked to other if they proposed a motion together. The colors represent their parties. (b) PT (pink) and PSL (green) are the two largest parties in the current legislature, they lead opposition and governmental base, respectively. PT is way more cohesive than PSL, which is very sparse. Source: (NERY, 2019a)



Using DW-NOMINATE scores, the webcomic Xkcd (MONROE, 2012) created a political timeline of the U.S. Congress since 1788, low and upper house presented as large infographic. Similar to work cited before, this visualization also shows annotations of historical events, presidents, control of the chamber, new and leaving members, and relevant members represented apart from their party. The quantitative roll call analysis provided by DW-NOMINATE defines the position of parties in the x-axis. The authors used colors to paint only the two major parties, to reinforce the idea of U.S. two-party system. The other parties are represented as undefined clouds above the main parties, as can be seen in Fig. 3.9.

Friggeri and Fleury’s visualization (FRIGGERI; FLEURY, 2012) shows the behavior of U.S. senators through agreement groups for eight Congresses. To create this timeline, it was used a custom clustering algorithm applied to roll call data, with all Senators displayed. It allows users to interact with visualization and select a senator to track his group alignment (Fig. 3.10).

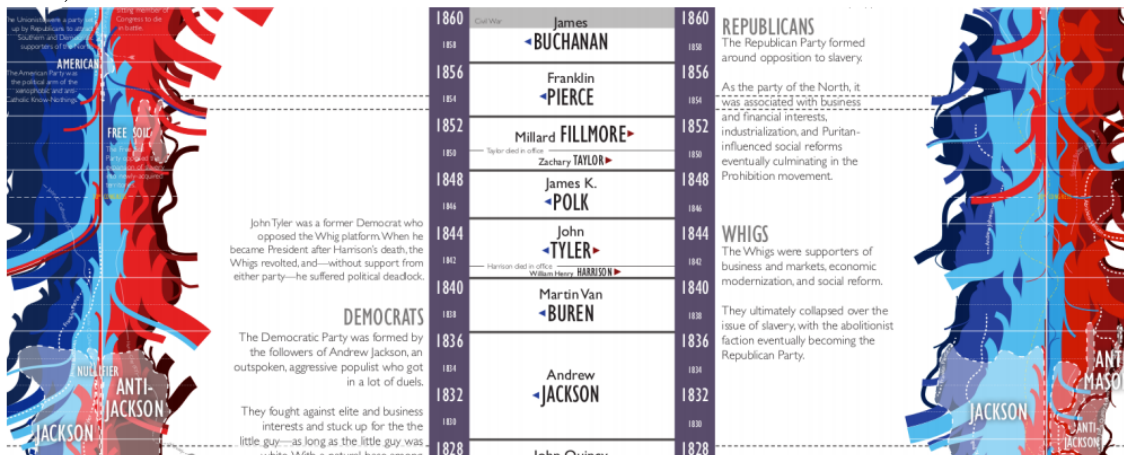
The Voteview⁶ system provides a political timeline that covers the period from the 2nd to the 116th Congress. It uses DW-NOMINATE scores to get an ideology value, that means how much a party is liberal or conservative, and then depicts the ideology trajectory of most relevant parties in history. Also, the chart

⁶<<https://voteview.com/parties/all>>

Figure 3.8: Covering the period from 1829 to 1860, the Conspectus of History of Political Parties and the Federal Government depicts the president in the first line, followed by the ministers, judiciary presidents and legislative presidents. The parties are represented as flows over time, and the y-axis represents how much they are aligned to the presidential government. Source: (HOUGHTON, 1880)

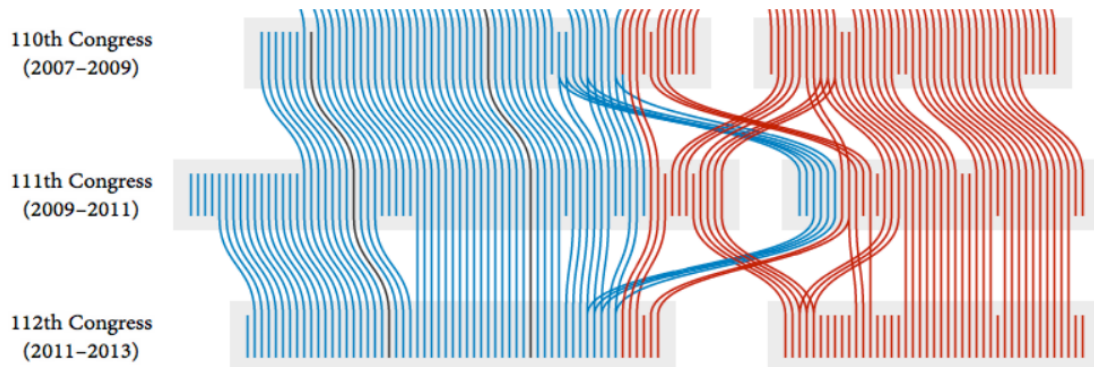


Figure 3.9: Political Timeline of the U.S. Congress since 1788. Source: (MONROE, 2012)



shows the average ideology of Congress in a certain period combined with the range of Party ideology, allowing the user to see which are the ideology limits of each party, as you can see in Fig. 3.11. The timeline uses as support a stacked bar chart

Figure 3.10: Timeline of the behavior of U.S. senators through agreement groups for eight Congresses. Source: (FRIGGERI; FLEURY, 2012)



to shows the size of parties across time. The user can select any party and get a timeline focused on it with a map of the U.S showing the distribution of members across the territory.

3.4 CivisAnalysis

The first version of CivisAnalysis was the main inspiration and starting point of our work. It is an open source project developed for the web, using JavaScript, and meant to be used through a browser.

Using roll calls voting data of six legislatures and six presidential elections, the original CivisAnalysis provides a unique view of the political history of Brazil. It incorporates classical roll calls visualization techniques, as scatterplots for political spectrum, with techniques for the visualization of temporal data. The system is divided into two principal modules: *Summary Module* and *Inspection Module*.

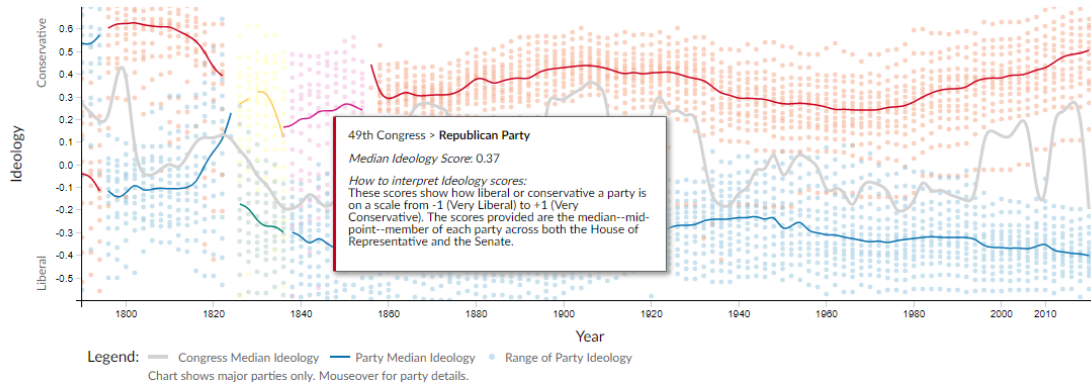
The *Summary Module* corresponds to a political timeline that presents a summary of 24 years of the Brazilian Chamber of Deputies (1991-2014). It displays in chronological order presidents, legislatures, general elections, and positions of parties in the spectrum, where Y-axis represents government (top) vs. opposition (bottom), and the X-axis time. To obtain details on roll calls and deputies data, the users have to select a period, which can be year, legislature, term, or arbitrary period by brushing. After the period selection, the system shows the *Inspection Module*.

The *Inspection Module* provides a set of coordinated views of the roll call data set. The Chamber Infographic depicts deputies sorted by 39 parties. The semi-circle

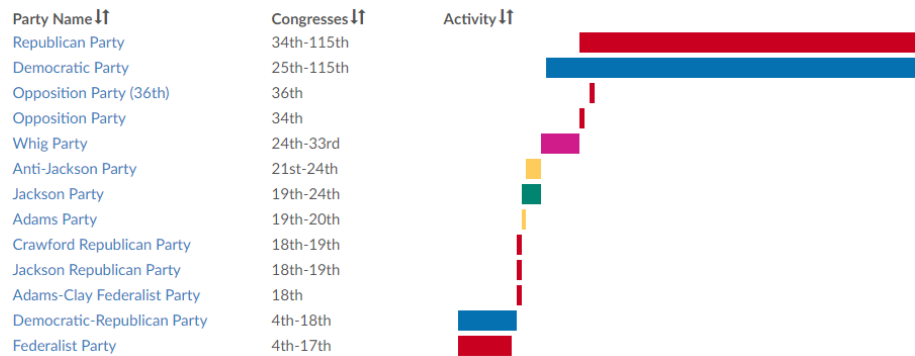
Figure 3.11: Political timeline covering the periods from the 2nd to 116th US Congress. Source: (VOTEVIEW, 2019)

Parties > Parties Overview

Congress at a Glance: Major Party Ideology



Parties Throughout History



design was adopted because it is a traditional representation of assemblies, chambers, or legislative houses.

To analyze the individual behavior of deputies, the user can create a political spectrum. The more similar deputies vote, the closer their respective circles are. CivisAnalysis includes a spectrum of roll calls, i.e., a scatterplot that shows how roll calls relate to one another depending on how deputies voted on them. This view allows users to directly inspect roll calls and make it easier for them to find related ones. A map of Brazil divided by electoral districts is depicted to represent the geographic properties of the data set. The users can click on a region to select it.

The extensive use of techniques of multiple coordinated views in the *Inspection Module* clearly shows to the user how roll calls, deputies, parties, and districts are related, every action on one visualization are reflected in all the others.

Figure 3.12: Original CivisAnalysis’s *Summary Module*: political spectrum of parties are traces across 24 years and six legislatures. Each party is colored with a different color, and parties in gray are the ones that do not exist anymore (BORJA; FREITAS, 2015).

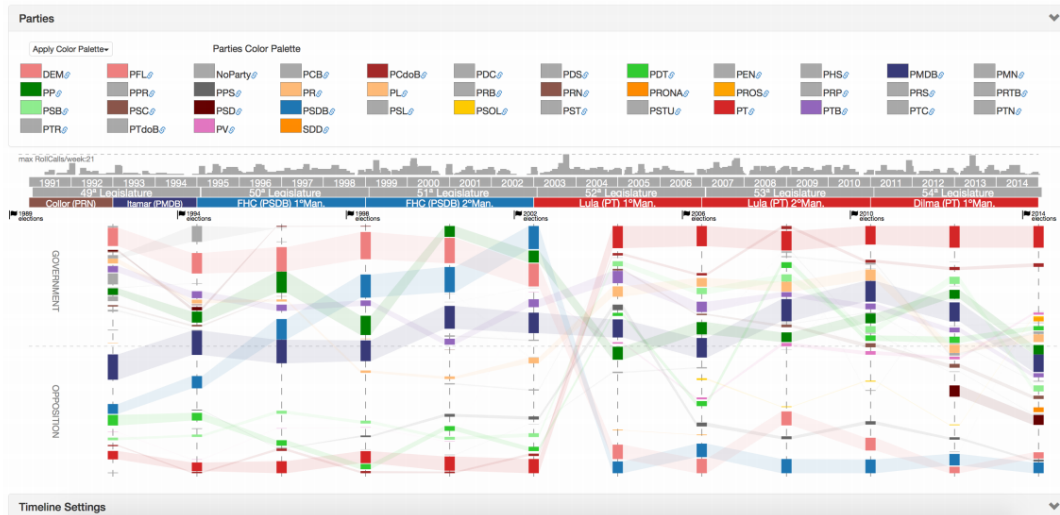
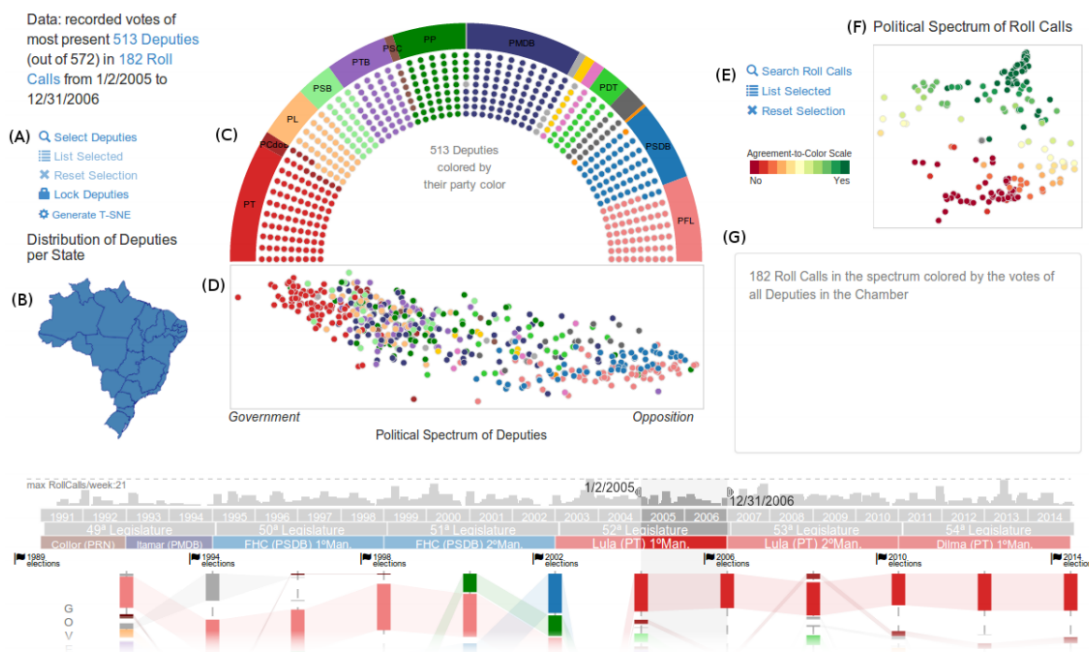


Figure 3.13: Original CivisAnalysis’s *Inspection module*: the Brazilian Chamber of Deputies in 2005-2006 (selected in the timeline of *Summary Module*, displayed at the bottom). (A) is an advanced menu for searching and filtering deputies; (B) the map of Brazil allows showing the deputies of each federal district and their voting pattern; (C) infographic of the distribution of deputies grouped by the party and sorted by deputy positions in (D), which represents the political spectrum of deputies. On the right: (E) is the advanced menu to search and filter the roll calls, which can be visualized in (F), the political spectrum of roll calls. (G) is a text window which provides details about the selected roll call (BORJA; FREITAS, 2015).



3.5 Charts and Reports-Based Tools

Basômetro (BASÔMETRO, 2019) is an interactive tool for monitoring roll call votes and measure how much the representatives are pro-government or opposition from 2003 until now. It provides two different visualizations intending to represent the size of the pro-government and opposition bases using a stacked chart (Fig. 3.14(a)). A second visualization uses a bar chart to show how much the pro-government and opposition bases change along the time in a certain government (Fig. 3.14(b)). In both visualizations, the user can filter by a specific party or an selected deputy.

There are some commercial applications that follow the daily procedures in both the Legislative Houses and Executive. Radar Governamental⁷ provides a business intelligence tool to support the exploration and analysis of the motions and other projects in real time. Sigalei⁸ is a system for monitoring the Legislative, by providing data about representatives and motions, also with a real-time feedback on the law-making processes and political information about the National Congress, States Assemblies and Municipal Chambers. Besides allowing users to organize and filter relevant data, it allows users to create their own reports. Finally, there are some smartphone applications, for example, InfoJá⁹, which provides information about the representatives and processes in the Brazilian Chamber of Deputies.

3.6 Final Comments

In this chapter, we presented a set of related studies in the area of visualization of political data. Most of them adopt the classical approach of representing the legislators in a political spectrum, a scatterplot, using some method to convert the original multidimensional data to a 2D space. These works combine different visualization techniques, such as spatial and temporal, with user interaction, to provide analytical tools.

It is worthy to note that C2C is an inspiration for original CivisAnalysis, and consequently for our version. C2C is an analytic web application, which provides

⁷<https://radargovernamental.com.br/>

⁸<https://sigalei.com.br/>

⁹<https://play.google.com/store/apps/details?id=zimmermann.larissa.legislativoapp>

Figure 3.14: (a) Stacked chart: Each point represents a deputy, those who vote the according to government orientation are positioned to the right (\rightarrow) of the chart. The higher the opacity of the circle, the more present the deputy was in roll calls. (b) For a specific deputy, this visualization uses bar charts to show the changes of pro-government and opposition ratios over the months in a certain government (BASÔMETRO, 2019).



a political spectrum based on roll call analysis for an arbitrary period selected by the user. C2C uses the W-NOMINATE method to calculate representative scores. However, for more extensive periods it is too costly, due to the heavy computation of W-NOMINATE.

CivisAnalysis (BORJA; FREITAS, 2015) was the primary and most important inspiration for our work. It allows users to inspect roll calls and deputies behavior with the support of a political timeline, fully interactive. The main objective is to make the user able to verify which events defined the political variations. In CivisAnalysis 2.0, we decided to keep and implement only light-weight roll call analysis, choosing fast dimensionality reduction methods and enabling real-time processing of any period in 28 years.

The political spectra are very useful to display the behavior of representatives. However, for small periods, the changes are hard to notice, and with this in mind, we wanted to provide alternatives using different temporal visualizations techniques, as the party ideology timeline in Voteview and timeline for senators in Friggeri and Fleury's visualization (FRIGGERI; FLEURY, 2012).

4 VISUALIZATION OF POLITICAL PROFILES

In this chapter we present the design and implementation of CivisAnalysis 2.0, a system whose goal is to help users identify parties and deputies' profiles through analyses of the Brazilian roll call data set, supported by a set of different visualization techniques. Our approach combines interactive visualization techniques for the exploration of political data with a system that provides a workspace where it is possible to instantiate different and coordinated panels each one containing possibly different views of the dataset. The content of this chapter has already been published (SILVA; SPRITZER; FREITAS, 2018).

Like the first version (<<https://github.com/fgborja/CivisAnalysis>>), CivisAnalysis 2.0 is an open source project hosted at GitHub¹ (<<https://github.com/rodrimoni/CivisAnalysis2>>). To implement the visualizations we used D3.js², a JavaScript library for manipulating data-based documents. The windowing system was implemented with support of jQuery³.

Being a web-based application written mostly in Javascript, CivisAnalysis 2.0 can run on any modern browser. It provides three techniques for dimensionality reduction, along with spectrum creation, search by deputies or roll calls, and filtering, all running on the client side, freeing the server from intensive computing. The roll call data are loaded on demand using XMLHttpRequest objects. CivisAnalysis 2.0 keeps the scalable characteristic of the original version, not requiring powerful web servers or too much bandwidth. The application is available in <<http://www.inf.ufrgs.br/~rnmsilva/CivisAnalysisDoc/>>

Since it is devoted to the general public but has some inherent complexity of use, we provide a complete tutorial for each visualization, explaining the features with texts and videos with a maximum duration of 1 minute. The first tutorial is triggered when the application starts and the others are loaded on the go according to the visualization created. In order to reach a greater public, we provide the interface in Portuguese as well as in English.

To a certain extent, in the following sections we follow Munzner's "What-Why-How" framework (MUNZNER, 2014). First, we describe the data coming from the Chamber of Deputies's web service (the "What" part of the framework)

¹<<https://github.com/>>

²<<https://d3js.org/>>

³<<https://jquery.com/>>

and then we present the user tasks that we aimed to address with this work (the “Why”). We subsequently describe the visualization techniques (the “How”) along with implementation details.

4.1 Data Description

We used the data (about 60 MB) from the Brazilian Chamber of Deputies’s open data portal (<https://dadosabertos.camara.leg.br/>), which provides information about all deputies (name, party, terms, etc.), propositions (abstract, keywords, the whole text, etc.) and roll calls (votes by each deputy for each roll call, timestamp of each roll call). These data can be obtained through RESTful API, accessing a specific topic, such as proposals or deputies information using GET or POST methods. As an alternative, the data can be also obtained through files, including all data related to a selected legislature.

Deputy Set: Brazil’s Chamber of Deputies has 513 seats for the elected representatives. To pursue other political activities (e.g., join the government as a minister), some deputies sometimes quit their seats, giving them away to new deputies. It is thus not unusual for a deputy not to participate in an entire legislature. As such, when we consider the whole legislature, we sometimes have to deal with a number of deputies that is larger than 513. Therefore, to obtain a better representation of the Chamber, for each legislature we decided to include the 513 most active deputies—i.e., the ones who are most often present in roll calls.

Deputies and Votes Set: The political spectrum of deputies is based on how each deputy voted in a roll call. The set of votes is thus represented by a matrix $R(M \times N)$ where $M = 513$ deputies and $N =$ the number of roll calls, so that entry $R_{m,n}$ represents the vote of the m^{th} deputy in the n^{th} roll call. Vote values are defined as follows: 1 for *Yes*, -1 for *No*, and 0 when the deputy was absent, obstructed, or there is no data.

4.2 Political Spectrum of Deputies

A political spectrum of deputies is generated from a time period selected by users. The application loads the deputies and the roll call data relative to this date

range and applies a dimensionality reduction (DR) method to obtain the scatterplot that represents the similarity of voting behavior between deputies. Users can choose from three DR methods: PCA (JOLLIFFE, 2011) by Singular Value Decomposition (SVD), Multidimensional Scaling (MDS)(TORGERSON, 1952), and t-SNE (MAATEN; HINTON, 2008).

PCA is a widely known and referenced method for dimensionality reduction. Moreover, it provides similar patterns to those identified by NOMINATE, a method used by social scientists. However, its result is not always easy to interpret and visualize, so we provided MDS and t-SNE as alternatives.

MDS is a classical method for visually representing the distances or dissimilarities between objects. The algorithm takes as input a symmetric matrix containing the pairwise distances between data points (e.g., deputy dissimilarities). In contrast to PCA, which retains the variance of the data, MDS preserves the distance between data points. Its main disadvantage is that the scatterplot generated by the algorithm may be distorted and induce a wrong interpretation of the data.

t-SNE (t-Distributed Stochastic Neighbor Embedding) is a modern method used for the visualization of high-dimensional data. The main advantage in comparison to PCA is the reduction of the tendency to crowd points together in the center of the graph, providing well-separated clusters and thus helping users identify them. However the algorithm is non-deterministic due to its use of random samples and local optimization, so users may have to run it multiple times until they get a good representation.

4.2.1 Input Matrices

The DR methods are based on matrices built from recorded votes. Furthermore, we have to create a dissimilarity matrix of deputies to serve as input for the MDS algorithm.

For PCA and t-SNE, the input is the matrix $R(M \times N)$ described before: $M = 513$ deputies and $N =$ number of roll calls, each cell containing the deputy's vote on the respective roll call.

MDS requires a dissimilarity matrix as input, so we have to apply a function to calculate the distance between representatives. Our distance function was directly inspired by the work of Diaconis (DIACONIS et al., 2008). The result is a

dissimilarity matrix $D(N \times N)$ where $N = 513$ deputies and the $D_{m,n}$ value represents the dissimilarity between the m^{th} and the n^{th} deputies. We first count the number of matching votes between each pair of deputies, i.e., if the deputies' votes for a determined roll call are equal, we increase the number of matching votes for the pair. Matching votes are saved in a matrix $N \times N$, then we calculate the correspondent percentage of each match according to the total number of votes in the given period, which yields a matrix $A(N \times N)$, where $A_{m,n}$ element is the percentage of matching votes between the m^{th} and n^{th} deputies. Finally, the "distance" between any two deputies is calculated according to the following equation:

$$D_{m,n} = \begin{cases} 0, & \text{if } m = n. \\ 100 - A_{m,n}, & \text{otherwise.} \end{cases} \quad (4.1)$$

4.2.2 Generating the spectrum

To obtain PCA-based spectrum, the SVD method is applied to matrix R (recorded votes), producing the following matrices:

$$\underbrace{R}_{513 \times N} = \underbrace{U}_{513 \times 513} \times \underbrace{\Sigma}_{513 \times N} \times \underbrace{V^T}_{N \times N} \quad (4.2)$$

The matrix \mathbf{U} is a $513 \times N$ real unitary matrix, \mathbf{V} is an $N \times N$ real unitary matrix, and Σ is a rectangular diagonal matrix of singular values. To calculate the bidimensional *deputies spectrum* we multiply the two largest singular values found in Σ by the left-singular vectors of \mathbf{U} .

t-SNE uses the matrix R , along with the following parameter values, as proposed by Borja and Freitas (BORJA; FREITAS, 2015): *learning rate* of 10, *perplexity* of 30, and 10 seconds for the *iteration time*.

Finally, MDS takes as input the dissimilarity matrix of deputies D that was obtained from the recorded votes. We used the classical version of the algorithm. We use matrix \mathbf{D} in some transformation until obtaining a matrix we can call \mathbf{B} to use the eigendecomposition, and then get the matrix of coordinates in the low-dimensional space \mathbf{X} . First, we raise every cell of \mathbf{D} to the power of 2, transforming

\mathbf{D} into $\mathbf{D}^{(2)}$:

$$\mathbf{D}^{(2)} = [d_{i,j}^2] \quad (4.3)$$

A matrix \mathbf{J} is then computed by subtracting the identity matrix \mathbf{I} by a matrix of ones $\mathbf{1}\mathbf{1}'$ divided by the number of deputies N (WICKELMAIER, 2003).

$$\mathbf{J} = \mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}' \quad (4.4)$$

The \mathbf{B} matrix is then defined as below:

$$\mathbf{B} = -\frac{1}{2}\mathbf{J}\mathbf{D}^{(2)}\mathbf{J} \quad (4.5)$$

Now, we can get the eigendecomposition of the \mathbf{B} matrix. We extract the m largest positive eigenvalues $\lambda_1 \dots \lambda_m$ of \mathbf{B} and the corresponding m eigenvectors $e_1 \dots e_m$, where m is the number of desired dimensions for the low-dimensional space.

Finally, an m -dimensional spatial configuration of the N deputies can be derived from the coordinate matrix \mathbf{X} :

$$\mathbf{X} = \mathbf{E}_m \mathbf{\Lambda}_m^{\frac{1}{2}} \quad (4.6)$$

\mathbf{E}_m is the matrix of m eigenvectors and $\mathbf{\Lambda}_m$ is the diagonal matrix of m eigenvalues of \mathbf{B} .

4.3 Tasks and Overview

The design of CivisAnalysis 2.0 was based on analyses that the users might want to perform on roll call data from the Brazilian Chamber of Deputies. Although Borja and Freitas (BORJA; FREITAS, 2015) based their work on what they called “Citizens Tasks,” we adopted “Visual Analytics Questions” (VAQs) to describe the issues that we aimed to address with our techniques. As we did not provide the political spectrum of roll calls in our approach (tests performed by Borja and Freitas showed that the general public was not always able to understand that visualization), we kept only three of the eight tasks they addressed in their work.

We have targeted the following visual analytics questions:

VAQ1 Verify how each deputy voted in individual cases.

- VAQ2 Verify how parties vote: are they cohesive?
- VAQ3 Verify which deputies voted alike and if they continued to do so over time.
- VAQ4 Verify which parties voted alike and if they continued to do so over time.
- VAQ5 Check which deputies are politically divergent from their parties.
- VAQ6 Verify the behavior of parties and deputies across different periods of time.
- VAQ7 Check how deputies switched parties over time.
- VAQ8 Verify the activity of the Chamber of Deputies (i.e., the number of propositions voted in a given period of time).
- VAQ9 Verify how a deputy behaves after switching parties (i.e., is the deputy more aligned with the new party or the old one?).

In order to support these questions, the main requirement is that visualizations must be created on demand, and the exploratory process must allow users to compare data shown in different, simultaneous views. It is important to provide a highly flexible application, allowing users to adjust the use of the tool to their needs. Working with only one visualization instance may not be enough for a complete data analysis. For example, in the original CivisAnalysis, it is only possible to display data from one period at a time, and all visualizations were shown to the user simultaneously, in a single dashboard. As such, we established the following requirements for the design of CivisAnalysis 2.0:

- R1 Provide **new and different visualization techniques**.
- R2 Allow **multiple instances** of the same visualization technique (e.g., for different time periods or data selections).
- R3 Allow for the creation of an **unlimited number of simultaneous views**.
- R4 **Resizable** views that can be **freely repositioned** in the workspace for a better exploration flow.
- R5 Explicit display of the relationships between views (shown as lines), so that users can **keep track of the steps taken in their exploration**.
- R6 Let users **show and hide** views to allow for a more customized workspace.

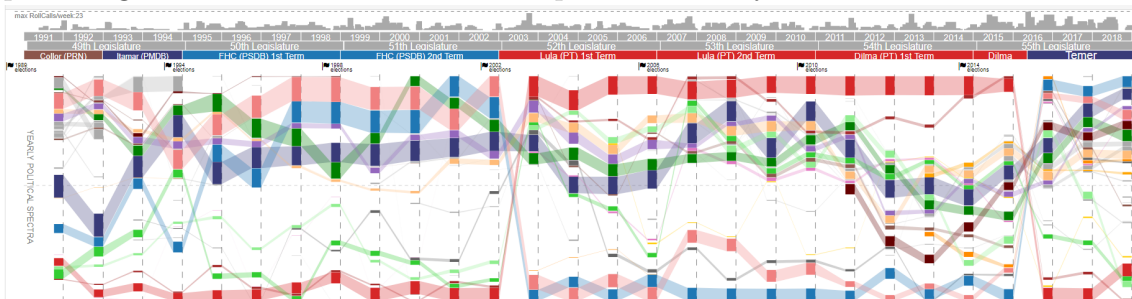
As with the original CivisAnalysis, our application initially displays the political timeline (Fig.4.1), which provides an overview of the parties' voting behavior in the spectrum over 28 years of Chamber of Deputies history (helping address VAQ6). The periods are divided into years, presidential terms, and legislatures. It represents the parties as area elements in the spectrum (thicker lines with varying width in Fig. 4.1).

CivisAnalysis 2.0 shows parties using a *Major Parties Differences* palette, which is a selection of colors that considers some of the parties' actual adopted colors (e.g., red for PT, blue for PSDB) and makes the most relevant parties visually distinguishable from one another.

The timeline serves as the starting point of the exploration process since users can launch other visualizations from it (Figs. 4.2(b) and 4.3), which is done by selecting a period and clicking on it to open a menu of techniques. Visualizations are displayed in panels (supporting requirement R1) and, depending on the technique, new views can be created from these visualizations. Users can create as many instances of a technique as they desire (R2). Thus, we based the user interface on multiple coordinated views that let users create visualizations of different periods of time.

We take a hierarchical approach to manage all connected views: we use a tree structure that stores the necessary data regarding the visualizations and their corresponding windows, with each node representing one panel with a unique identifier (*id*). The interface reflects the tree structure: nodes are panels and a line connecting two panels depicts a parent-child relationship between them (R5). Fig-

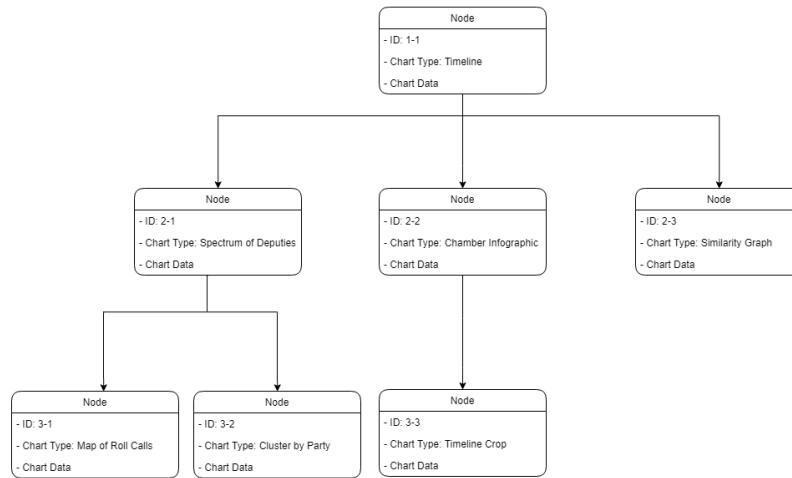
Figure 4.1: The political timeline: the main visualization of CivisAnalysis 2.0 showing an overview of the voting behavior of parties across the spectrum. Each party is depicted in a different color. Proximity means similarity of voting behavior. Parties closer to the top tend to be pro-government whereas those towards the bottom are in opposition. The timeline covers seven legislatures and nine presidential terms, providing an overview of Brazil's recent political history.



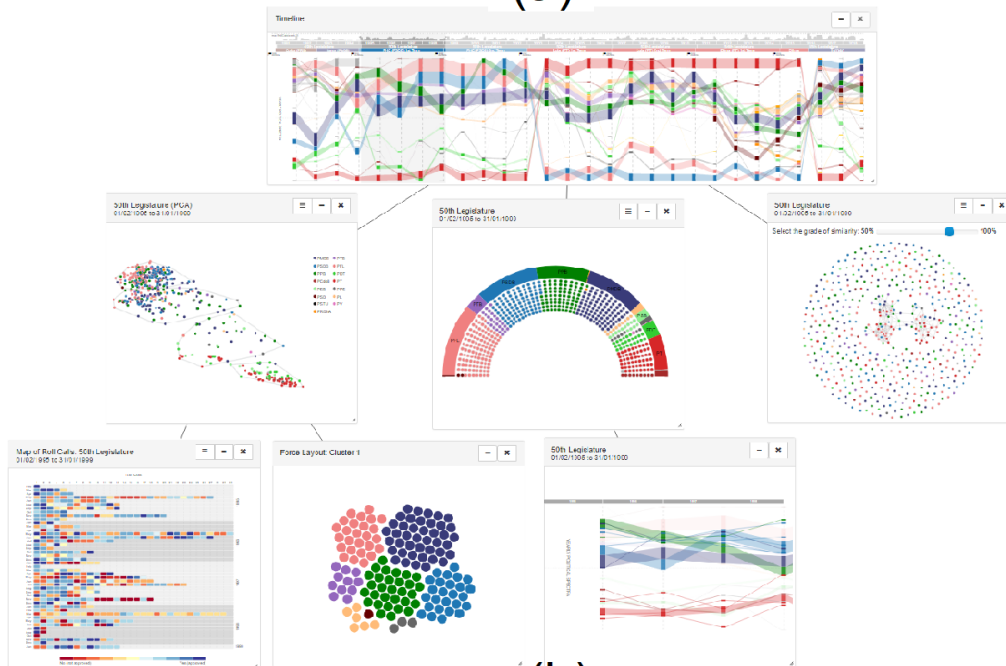
ure 4.2 clarifies the relationship between the tree structure and the user interface, showing the exact resulting UI in Fig. 4.2(b) of the tree structure defined in our system (Fig. 4.2(a)), with the correspondent *ids*, *Chart Types* and *Chart Data*.

Figure 4.2: Tree structure of the user interface (a), where each node stores the essential information regarding the managing of the UI panels and the data used in each visualization. (b) The corresponding user interface, where the user can create or remove panels, each interaction being reflected in the tree structure.

UI Tree Structure



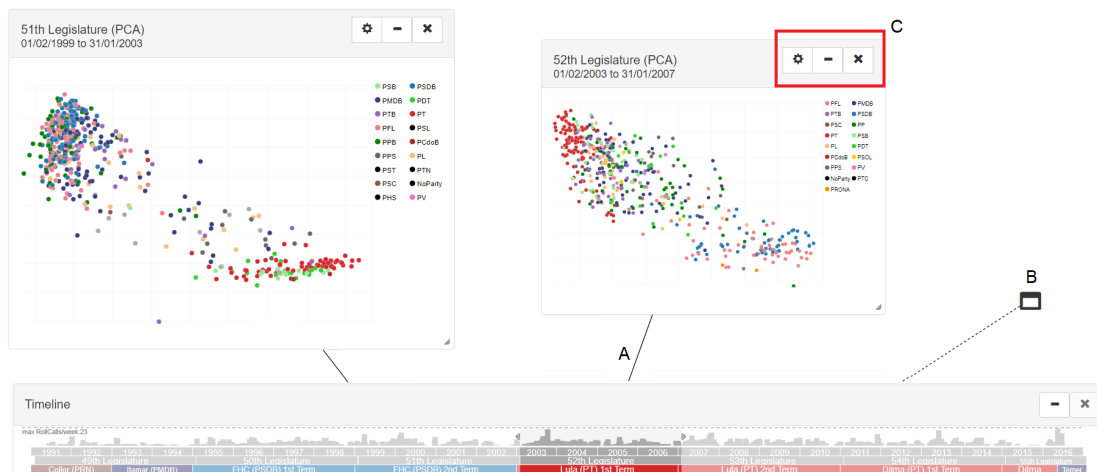
(a)



(b)

Users can customize the layout of the workspace by creating an unlimited number of views (R3), which are displayed in panels that can be *resized* and *moved*

Figure 4.3: Three panels are linked (A) to the timeline (cropped at the bottom): two scatterplots depicting different legislatures and one minimized panel, shown as an icon (B). Users can interact with the panel using buttons (C): settings, minimize, and close.



by drag-and-drop (R4). Except for the timeline, all panels can be minimized, maximized, and removed (R6)—actions which are performed through buttons on each panel’s title bar (see Fig. 4.3). When a panel is minimized, it is replaced with a small icon. Minimized panels are treated just like full-sized ones: they are also *draggable* and their hierarchical relations to other panels are also shown as lines (the only difference being the use of dotted lines instead of solid ones). Double-clicking on a minimized panel restores it to its full size. Depending on the visualization, panels may include a settings menu that lets users interact with the data shown in the view.

CivisAnalysis 2.0 uses the CMV idea, introduced in Sect. 2.4: when the user interacts with a view, using brushing, hovering, or selection, it is reflected in the other visualizations. The design of this windowing system was based on two similar works (DUNNE et al., 2012)(CAVA, 2017). The design based on multiple coordinated panels is most helpful for comparing different periods of the Chamber of Deputies’s history. Its flexibility also helps users to resize, minimize and reposition their visualizations however they like, which can facilitate the data exploration process.

4.4 Visualization Techniques

Three visualizations can be launched from the timeline view: the *Spectrum of Deputies*, the *Similarity Graph* and the *Chamber of Deputies Infographic*. The period to be visualized is specified by selecting a year, legislature, or presidential term, or by brushing the timeline. Right-clicking the selected period triggers a menu from which a visualization type can be chosen. Clicking on the desired type will launch the visualization, which will be displayed in a new panel on the workspace.

In the next subsections we describe CivisAnalysis 2.0 using the Visual Analytics Questions (VAQs) as arguments in the discussion of each visualization component.

4.4.1 Spectrum of Deputies

Regardless of the dimensionality reduction method used, the spatial proximity represents the vote similarity between deputies (thus addressing VAQ4). By default, deputies are represented as circles colored according to their respective parties. However, if the Map of Roll Calls (described in subsection 4.4.4) is instantiated and a single roll call is selected, deputies' colors correspond to their votes, according to a Vote-to-Color map (["Yes", "No", "Obstruction", "Abstention", Chamber President, absence] are respectively mapped to [blue, red, green, purple, yellow, gray]). Users select deputies by clicking on them or by using the brushing tool. Hovering over deputies with the cursor shows more information (e.g., the deputy's name and state). The Spectrum of Deputies addresses VAQ3, VAQ4, and VAQ5.

Fig. 4.4 shows how CivisAnalysis 2.0 deals with VAQ3 and VAQ4. Note that users can also cluster the deputies using K-Means by selecting the parameter k (i.e., the number of clusters) on the view's settings. A convex hull is displayed to identify each cluster. Two additional views (described later in this section) can be launched for each subset created by this clustering: the *Clustered Force Layout* (Sect.4.4.6) and the *Bar Chart* (Sect.4.4.7).

4.4.2 Deputy Similarity Graph

This visualization provides a distinct perspective of deputy voting behavior, based on a degree of vote similarity. We used a graph to represent deputies as nodes that are connected by edges whenever a pair achieves a minimum level of similarity, i.e., when they have voted similarly in a certain number of roll calls. We used the D3 library’s force-directed layout algorithm (FRUCHTERMAN; REINGOLD, 1991) to draw this chart. It is based only on the graph’s topology, ignoring contextual information. In this algorithm, each node exerts a repulsive force on all the others at the same time that it attracts nodes to which it is linked by an edge. An aesthetically pleasing layout that minimizes edge crossings is reached when the physical simulation reaches stability.

Users can interactively select the degree of similarity by moving a slider. The graph will be automatically updated as this value is changed. The similarity between representatives can be obtained from the matrix A, which calculates the similarity between representatives for the given period. It is the same we used in processing MDS input (see 4.2.1). User actions include the selection of deputies and hovering over items for more information (as in the Political Spectrum).

Figure 4.4: Comparison of the cohesiveness of a selected party in two legislatures: all deputies of the PP party in the 53rd legislature are selected. One deputy, Jair Bolsonaro, has a divergent behavior. Addressing VAQ3 and VAQ4, we instantiated the 54th legislature: his behavior became more divergent and the party as a whole moved to the center of the spectrum.

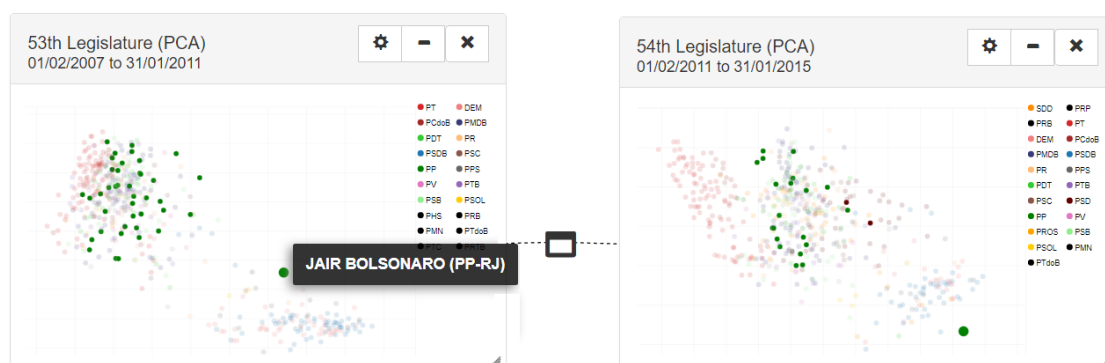
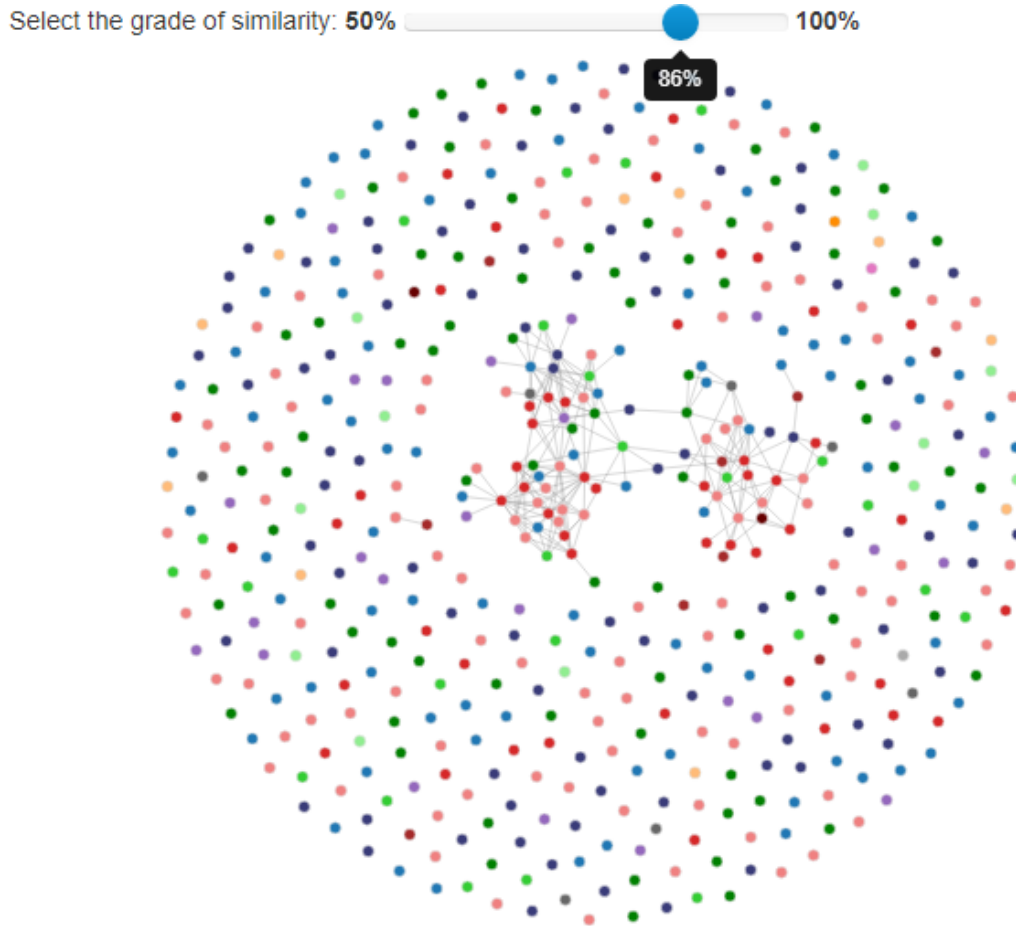


Figure 4.5: The graph shows groups of deputies that voted with a certain degree of agreement. In this graph, a pair of deputies is connected when the two deputies achieved 86% or more of agreement in voting. The slider that can be used to control the degree of similarity can be seen at the top.



4.4.3 Chamber of Deputies Infographic

Often used to depict the distribution of seats in legislatures, this visualization shows deputies and parties in a semi-circle (Fig. 4.6-left). Deputies are positioned according to their positions in the political spectrum and parties, according to the average of their deputies, which results in a party-based clustering of the representatives. Additionally, a half-donut chart wraps the seat representations to display the proportion of each party in the legislature. User actions include the selection of deputies and parties and hovering over items for more information (as in the Political Spectrum).

In the standard visualization of the Chamber Infographic, deputies are colored according to their parties. In Fig.4.6-left, however, they are colored depending on their individual votes, which is a color palette called Vote-to-Color map.

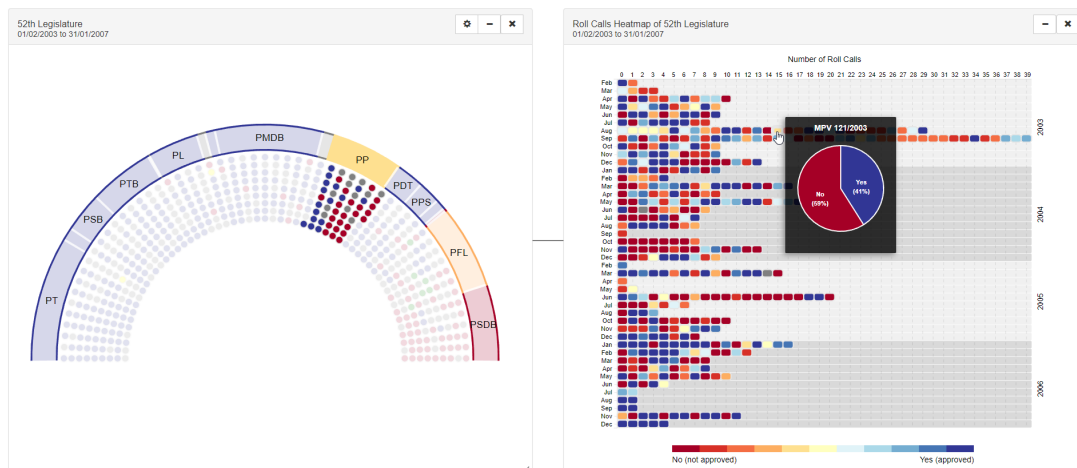
From the Political Spectrum, the Deputy Similarity Graph and the Chamber of Deputies Infographic, two other visualizations can be launched: the *Map of Roll Calls* and the *Cropped Timeline*, which are described in the following sections.

4.4.4 Map of Roll Calls

This visualization, shown in Fig. 4.6-right, helps users inspect the roll calls and provides them with an idea of how many motions were voted during a given time period. The data is shown as a horizontal histogram: a horizontal stack of roll calls (represented as rectangular cells) is displayed for each month (the y-axis). Roll calls are selected by clicking on them and hovering over them displays a tooltip containing a pie chart of the proportion of *Yes* and *No* votes in that roll call. This visualization addresses VAQ8, as users can identify the most active periods and see what was happening in the country at the time in order to relate them to their historical context.

Roll call cells are colored according to a colorblind-safe adaptation of the Agreement-to-Color scale of the first version of CivisAnalysis: the vote agreement scale ([100% “No” votes to 100% “Yes” votes]) is mapped to a color scale ([red, yellow, blue]) in which brightness corresponds to the number of votes.

Figure 4.6: Chamber of Deputies Infographic and Map of Roll Calls displayed in coordinated views. The PP party was selected on the Infographic and the Map of Roll Calls reflects its votes using the Agreement-to-Color scale. The cursor is hovering over a roll call, triggering a tooltip that shows the results as a pie chart. This is reflected back on the infographic, where the deputies’ colors represent their individual votes based on the Vote-To-Color map.



The Map of Roll Calls is directly connected to its parent visualization. The

default cell colors represent the voting ratio of selected deputies in the parent visualization, which can be the Political Spectrum, the Chamber of Deputies infographic or the Similarity Graph. Colors change dynamically according to the selected deputies and when only one deputy is selected, the roll call cells are colored according to the Vote-To-Color map. Along with the Chamber of Deputies Infographic, this technique addresses VAQ1 and VAQ2.

4.4.5 Cropped Timeline

The Cropped Timeline (Fig. 4.7) presents a slice of the main timeline. It is very similar to its parent view: it displays the parties in the timeline as area elements (i.e., thicker lines with varying width) anchored on the y-axis and divided into one-year intervals. Additionally, this visualization lets users add specific deputies to the timeline. The behavior of the deputies (their paths on the timeline) is represented by simple (i.e., thin, constant-width) lines instead of area elements. As in other views, users can hover over the lines to obtain more information about the deputies and parties.

This visualization allows for the observation of party switching by the displayed deputies, which addressed VAQ9. This is noticeable through the changes in the color of their respective lines, supporting VAQ7. This visualization is useful for analyzing specific deputies' voting trajectories, which can be compared to that of other deputies or parties, helping address VAQ6. Fig. 4.7 illustrates this by comparing the behavior of deputy Jair Bolsonaro, a frequent party switcher, to that of some of the parties he belonged to.

4.4.6 Clustered Force Layout

The clustered force-based layout (Fig. 4.8-A) provides a view of the composition of a cluster selected in a political spectrum that has been clustered with k -means (refer to Sect.4.4.1). Deputies are shown as circles colored according to their respective parties and a force-directed layout uses only the party information to determine the attraction between them. As usual, users can hover over circles to obtain more details about deputies.

Figure 4.7: Example of Cropped Timeline visualization: voting patterns of deputy Jair Bolsonaro and the parties he has belonged to from 2005 to 2010. (A) and (B) show, respectively, the patterns of his first and second parties (PTB and PP). (C) shows his voting behavior when he was a member of PTB and (D) depicts his votes after switching to PP.

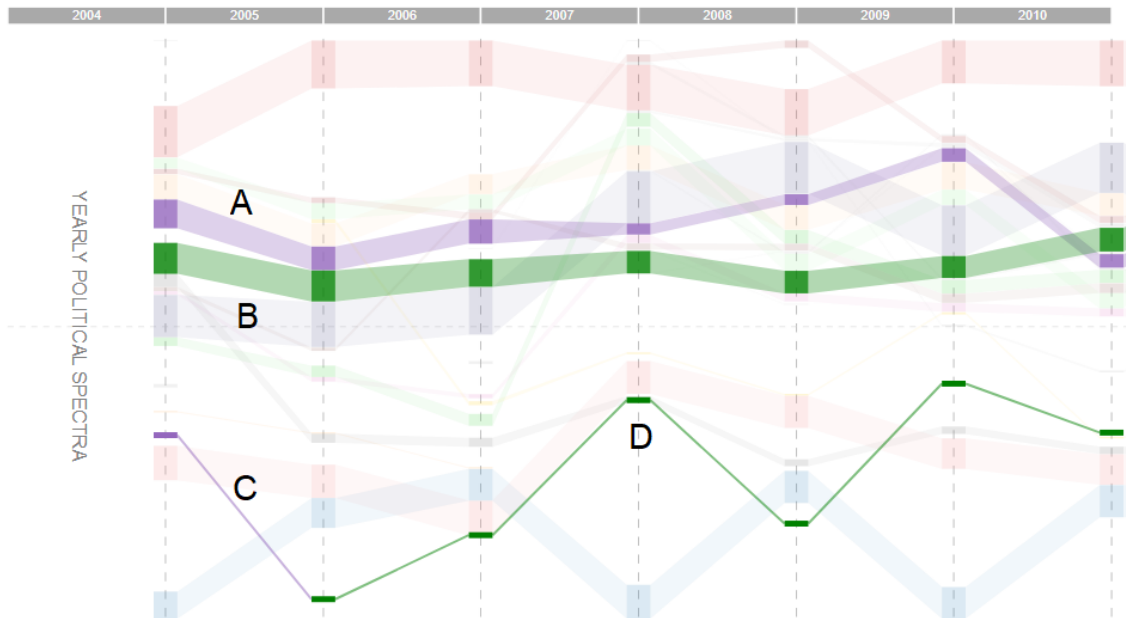
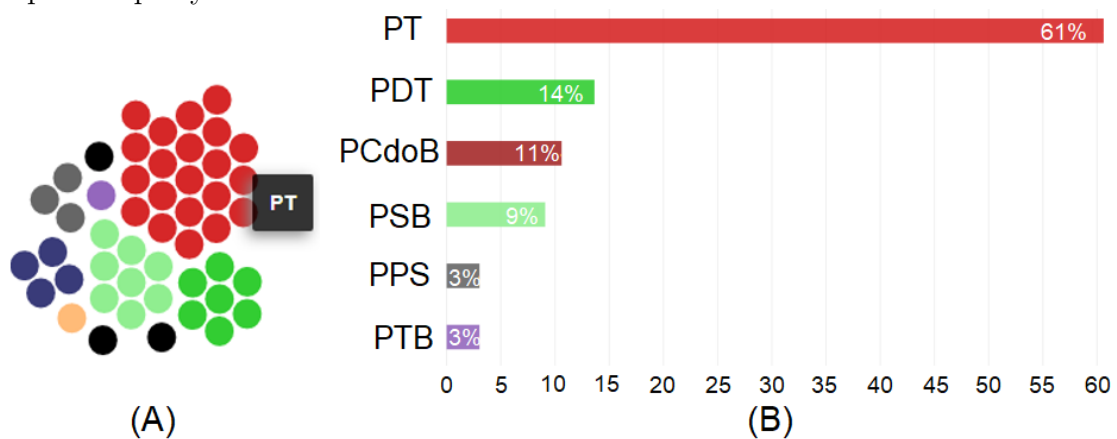


Figure 4.8: Clustered Force Layout (A) and Bar Chart (B) visualizations of the distribution of deputies by party: (A) uses a force-directed layout in which deputies that belong to the same party attract one another, forming clusters that help users see the number of deputies of each party in a given subset (we can see here that most deputies belong to PT). (B) is a traditional bar chart, with the percentage of deputies that belong to a bar's corresponding party being displayed inside each bar. Hovering over a bar displays a tooltip containing more information about its respective party.



4.4.7 Bar Chart

This visualization (Fig. 4.8-B) is similar to the Clustered Force Layout, and displays the composition of a cluster selected in a clustered political spectrum. Each bar corresponds to a party and its size represents the percentage of deputies of that party in the cluster.

Both the bar chart and the clustered force layout support the analysis of subsets of interest.

4.4.8 Additional Features

Some other features are provided to enhance users' exploratory experience.

Visual Filter: Filtering is supported by clicking, hovering, or brushing deputies or roll calls and is coordinated between views (i.e., actions taken in one view will also be applied on the others).

District Filter: One or more districts (i.e., states) can be selected from a combo box and deputies corresponding to the chosen districts will be highlighted.

Deputies Filter: A text input box accessible from the panels' settings menu can be used to search for deputies based on their names. Found deputies are highlighted.

Roll Calls Filter: In the Map of Roll Calls, we provide three different filtering options that can be accessed on the panel's settings menu: 1) a text box input used to search for a roll call based on its identifier (if found, the roll call will be highlighted); 2) a text box input used to filter roll calls based on their types (e.g., PL, PEC, etc.); and 3) two date inputs to filter roll calls that took place within a user-given time period.

4.5 Comparing CivisAnalysis 2.0 with the original CivisAnalysis

Since CivisAnalysis 2.0 was heavily inspired by CivisAnalysis (BORJA; FREITAS, 2015), it is crucial to highlight the differences between them and illustrate the main contributions of our work to the application. Besides providing new analytics capabilities, based on the evaluation of the original version, we aimed to solve the

most critical problems found by users.

In the original, users found it difficult to interpret the meaning of the Spectrum of Roll Calls, so we decided to replace it with a new visualization: the Map of Roll Calls (see Sect. 4.4.4). Furthermore, we provided a whole new interaction system based on a set of coordinated panels that let users compare different periods as well as freely organize visualizations on the workspace by moving, hiding, showing, and resizing panels. To better acquaint users with the system and all its features, we also included a tutorial that explains the different visualization techniques and guides them in their exploration of the data. Additionally, we also made the application available in a second language (Portuguese), which users can choose when the system starts. Table 4.1 summarizes the differences between our work (CivisAnalysis 2.0) and the previous version.

4.6 Final Comments

In this chapter, we have described CivisAnalysis 2.0 in detail, showing examples of all its features. The system was based on tasks that the general public as well as more specialized users may want to perform, so we described how our tool can address these tasks through the interactive visualization techniques it provides.

Using a workspace with multiple coordinated panels, we add a new dimension to the original approach, giving users more flexibility and increasing the tool's exploratory power. We provided an entirely new user interface along with several new visualizations, a tutorial, and support for Portuguese and English languages.

In CivisAnalysis 2.0, users will be able to compare how deputies behaved in different periods of time. Users can also simultaneously display parties' and deputies' timelines to verify how aligned deputies are to their respective parties and how often they witch parties. We also provided a new way to investigate roll calls and individual votes with the Map of Roll Calls.

Table 4.1: Differences between the original CivisAnalysis and CivisAnalysis 2.0

	Original	Our tool
System Interaction	Only one instance of each visualization Visualizations are always visible	Multiple instances of visualizations Customizable workspace
Visualization Techniques	Spectrum of Roll Calls	Map of Roll Calls Cropped Timeline Similarity Graph
Dimensionality Reduction Methods	PCA and t-SNE	PCA, MDS, and t-SNE
Date Coverage	From 1991 to 2014	From 1991 to 2018
Language	English	English and Portuguese
Tutorial	No	Yes

5 CASE STUDY: PARTY COHESIVENESS OVER TIME

To illustrate how CivisAnalysis 2.0 can be used in practice, we conducted an informal case study analyzing party cohesiveness over time¹. The analysis covered each presidential term included in our data and was based on each legislature’s t-SNE scatterplot visualization and Chamber of Deputies composition infographic. t-SNE was chosen because it provides a much clear view of local structures (i.e., clusters) in the the dataset (refer to Sect. 2.3.3), which allows perceiving cohesiveness. The analysis was also complemented by historical information that helps contextualize and understand what is being seen.

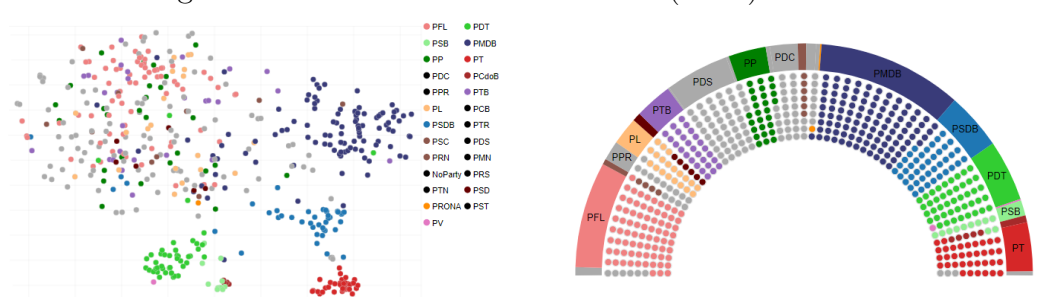
5.1 Collor (PRN) – 1990-1992

Main parties (at least 20 deputies): PFL (76), PL (20), PTB (27), PDS (46), PP (27), PDC (20), PMDB (104), PSDB (41), PDT (44), PT (35)

The first election after the military period resulted in a very fragmented congress and a president, Fernando Collor de Mello, from a very minor party, PRN, which had only four seats in the Chamber of Deputies. The two largest parties were the democratic reinventions of the two parties that existed during the dictatorship: the rightwing PFL (Liberal Front Party), which emerged from the conservative, pro-regime ARENA, and the PMDB (Brazilian Democratic Movement Party), a rebranding of opposition party MDB (Democratic Movement Party), which served as an umbrella for ideologies covering the entire political spectrum. Other sizable parties include right-leaning ARENA offshoots PDS (Social Democratic Party), PP (Progressive Party), and PDC (Christian Democratic Party), the social-liberal

¹The study was mostly performed by André Spritzer, PhD in Computer Science, but currently a student of Political Science and International Relations.

Figure 5.1: Fernando Collor de Mello (PRN) – 1990-1992



PL (Liberal Party), the centrist MDB spin-off PSDB (Party of the Brazilian Social Democracy), the centrist PTB (Brazilian Labor Party), and the leftwing PDT (Democratic Labor Party) and PT (Workers' Party).

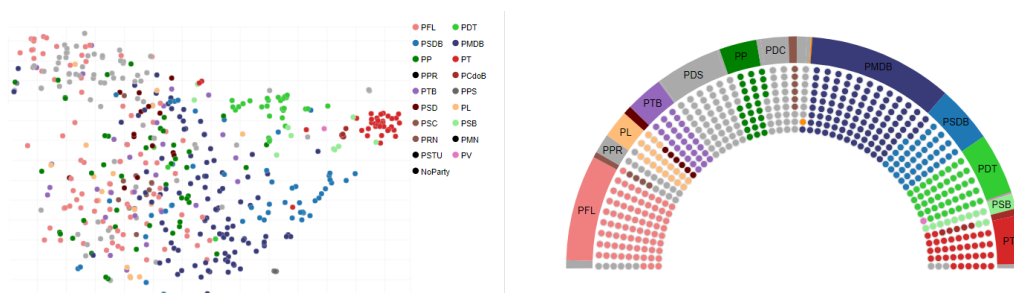
In CivisAnalysis 2.0 (Figure 5.1), we can see that some deputies form well-defined clusters, revealing the cohesiveness of their parties in terms of how similarly deputies voted. From this, we can see that PDT, PSDB, PT, and PMDB were very cohesive in this first legislature of Brazil's new democracy. PFL was slightly less cohesive and most of the other parties were a lot less so while not straying too far from PFL's positions (their deputies are spread over a large area overlapping PFL's, but they remain far from the other major parties).

5.2 Itamar (PMDB) – 1992-1995

Main parties (at least 20 deputies): PPR (60), PFL (85), PTB (29), PP (46), PMDB (101), PSDB (51), PDT (32), PT (36)

1992 was a turbulent year in Brazilian politics, culminating in the impeachment of president Collor following a corruption scandal. Vice-president Itamar Franco (PMDB) took over the presidency amidst an economic crisis (with annual inflation at over 1000%) and built a government based on a broad coalition. The political and economic turmoil deeply affected the behavior and voting patterns of the deputies, as can be clearly seen in CivisAnalysis 2.0 (Figure 5.2). Although leftwing PDT and PT managed to maintain party unity, centrist PSDB became considerably less cohesive and the previously close-knit PMDB and the somewhat cohesive PFL became completely loose. Also notable were the growth of PSDB, which gained 10 seats and the 1993 incorporation of PDS and PDC into the previously tiny PPR (Progressive Reformist Party), which grew from 14 to 66 deputies and behaved in a

Figure 5.2: Itamar Franco (PMDB) – 1992-1995



mostly united way.

5.3 FHC (PSDB) – 1995-1999

Main parties (at least 20 deputies): PFL (110), PTB (25), PSDB (91), PPB (79), PMDB (86), PDT (24), PT (53)

The success of Fernando Henrique Cardoso's (PSDB) term as finance minister under president Itamar Franco brought him to the presidency and restored political and economic order, with most of the major parties' deputies voting in cohesive patterns (Figure 5.3). The exception to the rule was PMDB, whose deputies divided into smaller clusters. Notable changes include the shrinking of PMDB and the growth of PFL (which became the largest party in the Chamber of Deputies), PSDB (which almost doubled in size), and PT. Also notable was the fusion of PP and PPR into PPB (Brazilian Progressive Party), which formed a mostly cohesive block of deputies.

5.4 FHC (PSDB) – 1999-2003

Main parties (at least 20 deputies): PPB (50), PFL (96), PSDB (95), PMDB (86), PTB (35), PL (23), PT (63)

Changes in voting patterns during president Fernando Henrique Cardoso's second term were not too drastic (Figure 5.4). PFL and PSDB behaved almost like a single party, with the latter being more cohesive. PMDB became more united than it was in the previous term, but many of its deputies still did not vote together. Also notable were the continued growth of PT and the considerable decreases in size of both PPB and PFL, which lost its place as largest party in the Chamber of Deputies to PSDB. PDT also continued on the downward trend that took it from 44 deputies

Figure 5.3: Fernando Henrique Cardoso (PSDB) – 1995-1999

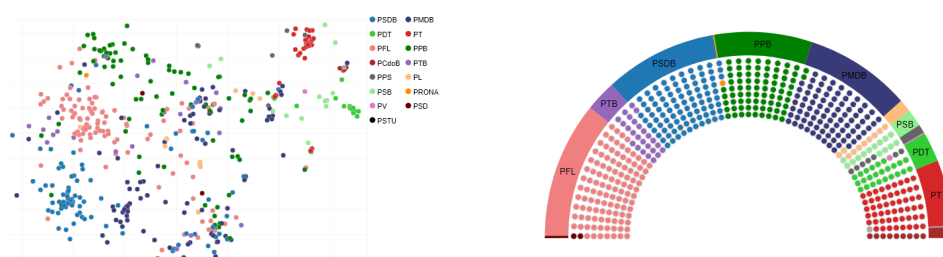
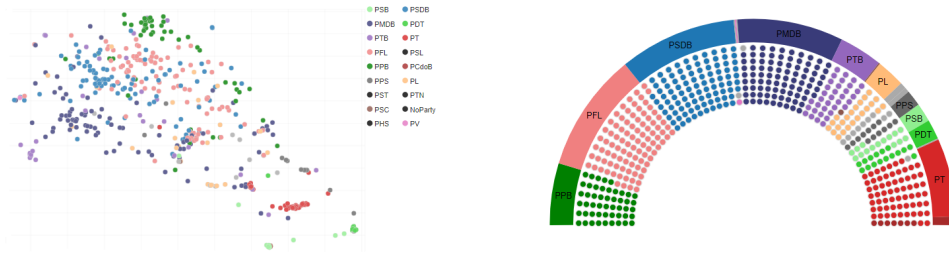


Figure 5.4: Fernando Henrique Cardoso (PSDB) – 1999-2003



in the Collor government to only 16 in Fernando Henrique's second term. PTB grew a little in size but remained very loose. PPB, in turn, remained mostly the same in both size and cohesiveness.

5.5 Lula (PT) – 2003-2007

Main parties (at least 20 deputies): PT (89), PSB (28), PTB (46), PL (36), PMDB (76), PP (51), PDT (20), PFL (59), PSDB (53)

The election of Lula (PT) followed another economically turbulent period that resulted in the growth of left-wing and populist-leaning parties, such as PT (now the largest party in the chamber of deputies), PSB, PTB, and PDT, and the considerable shrinking of PFL and PSDB, which became the only substantive opposition parties, forming two similar and close-knit opposition blocs (Figure 5.5). Of the remaining parties of significant size, PT was the only one whose voting behavior was very cohesive, although both PL and PTB had sizable cohesive clusters along with many deputies that voted their own way. Also notable is the change in PPB, which was renamed PP and became a lot less cohesive despite retaining the same size.

Figure 5.5: Luís Inácio Lula da Silva (PT) – 2003-2007

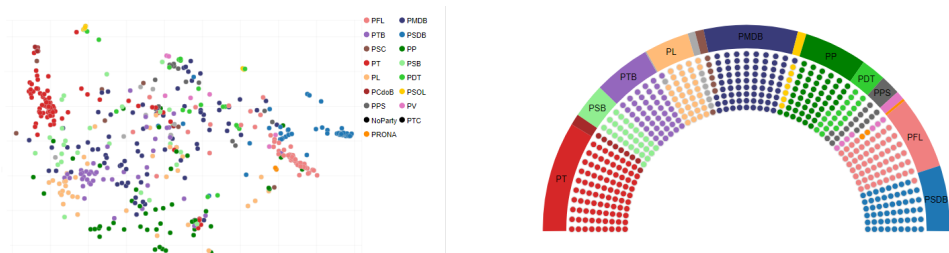
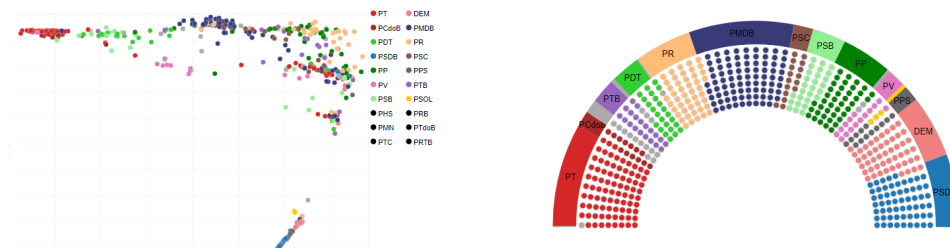


Figure 5.6: Luís Inácio Lula da Silva (PT) – 2007-2011

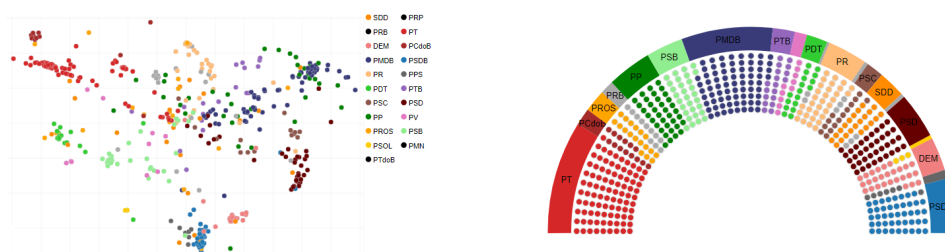


5.6 Lula (PT) – 2007-2011

Main parties (at least 20 deputies): PT (83), PTB (20), PDT (24), PR (46), PMDB (85), PSB (27), PP (41), DEM (50), PSDB (59)

Despite a vote-buying scandal that threatened to bring down Lula’s (PT) government during his first term in office, an economic boom led to his reelection and to very high approval ratings. This was reflected in the voting behavior of deputies and can be very clearly seen in CivisAnalysis 2.0 (Figure 5.6): the opposition, made up mostly of PSDB and DEM (Democrats—a rebranded PFL), forms a small and cohesive but completely isolated bloc. All other parties are bundled together, although roughly overlapping and cohesive clusters can still be seen. It is interesting to note that PT is still especially cohesive, but there is now a breakaway cluster that is very far from the main “body” of the party. Closer to PT’s main clusters are PDT and about half of PSB. After years of being very loose, PMDB is again very cohesive. Slightly less so are PP and PR (Party of the Republic), a new party resulting from the fusion of PL and the tiny PRONA (Party of the Reconstruction of the National Order). Also notable during this period was the beginning of the proliferation of very small parties after the Supreme Court struck down an electoral threshold as unconstitutional.

Figure 5.7: Dilma Rousseff (PT) – 2011-2015



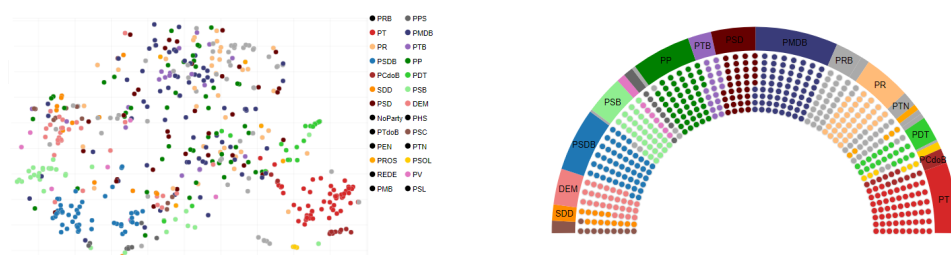
5.7 Dilma (PT) – 2011-2015

Main parties (at least 20 deputies): PT (91), PP (36), PSB (30), PMDB (74), PR (32), SDD (22), PSD (37), DEM (27), PSDB (44)

Lula's high approval ratings and the continued economic growth led to the election of his successor, Dilma Rousseff (PT), and the further erosion of the opposition. However, being less skilled as a negotiator than her predecessor and facing harsher economic conditions and a growing number of parties, Dilma had more difficulty in maintaining cohesiveness between the parties that made up her ample and mostly non-ideological coalition (Figure 5.7). The two main opposition parties, PSDB and DEM, remained as two isolated and very cohesive clusters. DEM, however, lost about half its deputies to offshoot PSD (Social Democratic Party), a roughly cohesive, sizable group (37 deputies) that joined the government coalition. Among Dilma's supporters, the largest and most cohesive group is still PT. PMDB is still second in the alliance and is also cohesive, although not very tightly so. Other considerable government allies include the smaller but reasonably cohesive PR and the much looser PSB and PP. Traditional ally PDT shrunk to 17 deputies but remained very close-knit. Also notable is the emergence of the Solidarity, a reasonably-sized (22 deputies) but very divided party. A noteworthy political event that might have shaped voting patterns during this period were mass protests held in 2013 that reflected a growing but still very vague dissatisfaction of the population with the political class as a whole.

5.8 Dilma (PT) – 2015-2016

Main parties (at least 20 deputies): DEM (29), PSDB (51), PSB (31), PP (48), PSD (36), PMDB (67), PRB (22), PR (38), PDT (20), PT (57)



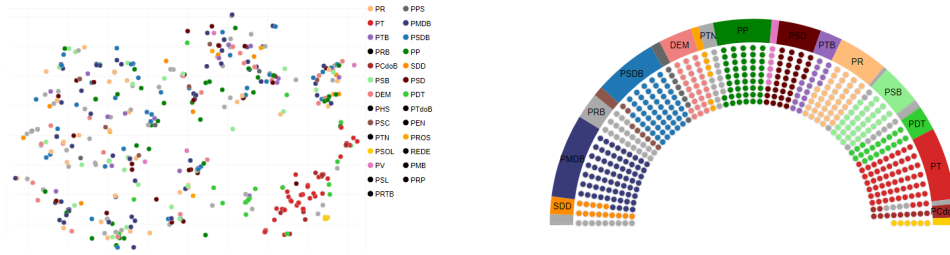
Dilma (PT) was reelected following a very hotly contested election and a controversial campaign, in terms of both funding and method. Facing an unprecedented economic crisis, Dilma did the opposite of what she promised during the campaign. What's more, PT found itself at the center of another major corruption scandal (uncovered by the now famous Operation Car Wash, which would soon also engulf the two other larger parties—PMDB and PSDB). Dilma and PT's newfound unpopularity and her difficulty in dealing with congress resulted in a breakdown of her coalition, which, along with fiscal mismanagement by her government, culminated in her impeachment and removal from office after a series of large-scale protests all over the country. Throughout this very turbulent period, PT not only lost several of its major allies, such as PSB and PMDB, but also found itself a little less cohesive than it used to be. Of the two remaining larger parties, PSDB managed to stay slightly cohesive while PMDB imploded, with its deputies rebelling against the government (theoretically their ally, as they held the vice-presidency) and voting all over the political space (Figure 5.8). DEM and PSB (which moved to the opposition during the election) managed to stay reasonably cohesive, but the same is not true for almost all other medium-sized parties (PDT being a notable exception). Also worth noting is the continued growth in the number of parties.

5.9 Temer (PMDB) – 2016-2019

Main parties (at least 20 deputies): PMDB (67), PRB (21), PSDB (51), DEM (27), PP (47), PSD (34), PR (40), PSB (33), PDT (20), PT (58)

With Michel Temer (PMDB), who was vice-president under Dilma, PMDB found itself again in the presidency after an impeachment. Once in power, Temer managed to form a broad and loose coalition including PSDB, DEM, and most of PT's old allies (with the notable exception of the most leftwing parties, which stuck with PT). Temer was able to break the parliamentary gridlock of Dilma's second term, and also put together a new team, trying to get the economy back on track. However, along with prominent cabinet members and influential leaders of both PMDB and PSDB, he soon found himself engulfed by corruption accusations and investigations partly related to the Car Wash scandal, eroding his already feeble popularity and political capital. The effect of these events on party cohesiveness can be clearly seen in the voting patterns: the only significant party that managed

Figure 5.9: Michel Temer (PMDB) – 2016-2019



to remain reasonably cohesive was PT—in the visualization, deputies of all other parties are scattered all over the place (Figure 5.9).

5.10 Final Comments

As one observes, several online newspapers use interactive charts and infographics to provide insight into political data, systems, and processes. In this chapter we reported a case study developed for illustrating the utility of CivisAnalysis as a presentation tool. The visualizations provided in our work can be used to build narratives regarding different aspects, situations and facts.

6 EVALUATION WITH USERS

In the previous chapter, we presented a case study to demonstrate how our system can be used in a specific analysis. However, it is essential to validate our work with actual users. Therefore, we conducted a remote survey with two distinct users groups, which are the ones that we targeted in developing our visualization and the application: the general public and experts.

While the objective of the test with the general public¹ is to measure the user's understanding of the general flow of the system, with tutorial assistance, the evaluation with experts² aims to measure the utility and difficulty of using CivisAnalysis 2.0. The complete survey is available in Appendix A (A.1 and A.2). Since measuring system usability is not our current objective, we decided for now not to evaluate its usability, which is left for future work. We used R language³ with support of RStudio⁴ to analyze our results, performing statistical analysis and generating plots.

6.1 Remote Evaluation with General Public

The survey was centered on a community of users that were potentially interested in investigating the behavior of representatives from the Brazilian Chamber of Deputies. Our focus was to assess the understanding of our system, based on our design choices concerning both interactive features and visualization, in general tasks like:

- Understanding the parties' flow across time;
- Comprehending the political spectrum of deputies;
- Understanding the map of roll calls;
- Assessing the ease of selecting deputies and roll calls;
- Evaluating the efficiency to generate new views;

¹General public survey available in <<https://forms.gle/XGXetSCXxhh62j939>>

²Experts survey available in <<https://forms.gle/CFRz4eDNh8kSYXLY8>>

³<<https://www.r-project.org/>>

⁴<<https://www.rstudio.com/>>

- Understanding the relationship between views;
- Evaluating the searching and filtering selection controls;

6.1.1 Participants

We sent the survey with instructions to professors, students, and researchers of different areas, such as computer science and statistics, via the university’s mailing lists. After a week, the test involved 39 participants, 28 (71.79%) males, and 11 (28.21%) females, with their ages ranging from 20 to 42 years old ($\bar{x} = 26.33$ and $\sigma = 5.46$). 66.7% of these participants are familiar with visualization tools such as: D3.js⁵ (50%), Plot.ly⁶ (32.1%) and Power BI⁷ (32.1%). About 90% are from the field of computer science, computer engineering, and statistics. The remaining users are from other fields, such as health and social sciences. Table 6.1 summarize data about the participants.

Table 6.1: Summary of participants

N=39	
Age:	
Mean (SD)	26.33 (5.46)
Gender :	
Female	11 (28.21%)
Male	28 (71.79%)
Experience with Visualization Tools:	
Yes	26 (66.67%)
No	13 (33.33%)

⁵<<https://d3js.org>>

⁶<<https://plot.ly/>>

⁷<<https://powerbi.microsoft.com>>

6.1.2 Procedure

The survey was organized in three parts, each represented as a section of the questionnaire. First, users were introduced with a short description of the system's purposes and the information about the used data set. Afterwards, users were asked to answer some personal information to characterize and understand their profiles. We asked about their age and gender, their professional profile, and experience with visualization tools.

After collecting the user's information, we provide the link to access our tool. We asked them to use CivisAnalysis 2.0 freely with the help of tutorials that would eventually appear, and then answer our questionnaire. The users performed the evaluation remotely, so the only instructions on how to use the system are provided on the website by tutorials.

The questionnaire had 20 statements with a five-point Likert scale of agreement (from 1 meaning strongly disagree to 5, strongly agree)(LIKERT, 1932). We grouped the statements according to specific parts of the system, and some of them were defined to verify explicitly if our Visual Analytics Questions, described in section 4.3, can be answered by the general public. The statements are described as follows:

- **Timeline (TL):**

TL1- I can verify how the parties have voted over the years (VAQ4).

TL2- I can select a period in the Timeline.

TL3- I can identify the size of each party.

- **Political Spectrum (PS):**

PS1- I can understand the spectrum of deputies.

PS2- I can check if the parties are cohesive (VAQ2).

PS3- I can check which deputies are politically divergent from their parties (VAQ5).

- **Chamber Infographic (CI):**

CI1- I can identify the distribution of seats in the Chamber.

- **Map of Roll Calls (MR):**

MR1- I can verify how each deputy voted in individual cases (VAQ1).

MR2- I can verify the most active periods in the Chamber of Deputies (VAQ8).

MR3- I can select motions and apply filters on them.

- **Cropped Timeline (CTL):**

CTL1- Verify the behavior of parties and deputies across different periods of time (VAQ3/VAQ6).

CTL2- I can check how deputies switched parties over time (VAQ7).

CTL3- I can verify how a deputy behaves after switching parties (VAQ9).

- **Similarity Graph (SG):**

SG1- I can see which deputies have voted in a similar way.

- **General System (GS):**

GS1- I can create new views from others.

GS2- I can understand how visualizations are related.

GS3- I can organize and position the views according to what I want to see.

GS4- I can apply the filter by districts.

GS5- I can reset all the previous selections.

GS6- I can search deputies by their names.

In the last question, we optionally asked users to give some feedback, critics, and suggestions about the system. Notice that we did not get information about the computational platform that volunteers used neither how much time they spent using CivisAnalysis 2.0.

6.1.3 Results

The main goal of this questionnaire was to get feedback from users regarding if they think they would be able to explore the Chamber of Deputies data with the system. Additionally, we aimed at evaluating if the visualization techniques helped

them to understand the deputies and parties behavior. Since our system’s design was highly based on VAQs, we will analyze in detail the results of questionnaire’s statements which can be interpreted as VAQs. Note that VAQ3 and VAQ6 can be mapped to the same statement.

With the help of Figs. 6.1 and 6.2, and Table 6.2, we can verify that in general the users could answer the VAQs.

Table 6.2: Summary of VAQs results

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
VAQ1	4 (10.26%)	1 (2.56%)	7 (17.95%)	11 (28.21%)	16 (41.03%)
VAQ2	3 (7.69%)	2 (5.13%)	5 (12.82%)	13 (33.33%)	16 (41.03%)
VAQ3/VAQ6	1 (2.56%)	5 (12.82%)	5 (12.82%)	11 (28.21%)	17 (43.59%)
VAQ4	5 (12.82%)	3 (7.69%)	8 (20.51%)	7 (17.95%)	16 (41.03%)
VAQ5	3 (7.69%)	0 (0%)	4 (10.26%)	8 (20.51%)	24 (61.54%)
VAQ7	8 (20.51%)	3 (7.69%)	2 (5.13%)	16 (41.03%)	10 (25.64%)
VAQ8	2 (5.13%)	2 (5.13%)	6 (15.38%)	12 (30.77%)	17 (43.59%)
VAQ9	8 (20.51%)	3 (7.69%)	5 (12.82%)	10 (25.64%)	13 (33.33%)

However, we can notice a certain difficulty to perform VAQ7 and VAQ9, for example. Both VAQs are related to Cropped Timeline and both showed the highest value of disagreement (21%), probably due to the task complexity.

Analyzing users comments, seven users (17.94%) related issues with Timeline and its variant Cropped Timeline, mainly with difficulty to interpret the meaning of Y-axis. Also, users showed problems in generating Cropped Timeline view and its interaction with parties and deputies.

The second higher value of the disagreement was observed in VAQ4 (20.5%), which is directly related to the main Timeline. Users had some difficulties in interpreting the parties behavior across time, including a user asked about the meaning of the drastic changes on parties behavior between the end and beginning of a term.

The VAQ5 was the most straightforward question to answer, based on its agreement score: 61.5% marked “Strongly agree” and 20.5% marked “Agree”, with a total of 82% of agreement. The users were asked to find the deputies politically divergent of their parties in Spectrum of Deputies. Apparently, the majority was able to select a party in legend and then verify which are the differing representatives. Still regarding Spectrum of Deputies, another VAQ has high values of agreement:

Figure 6.1: Grid of bar charts showing the relative frequency of answers for each VAQ.

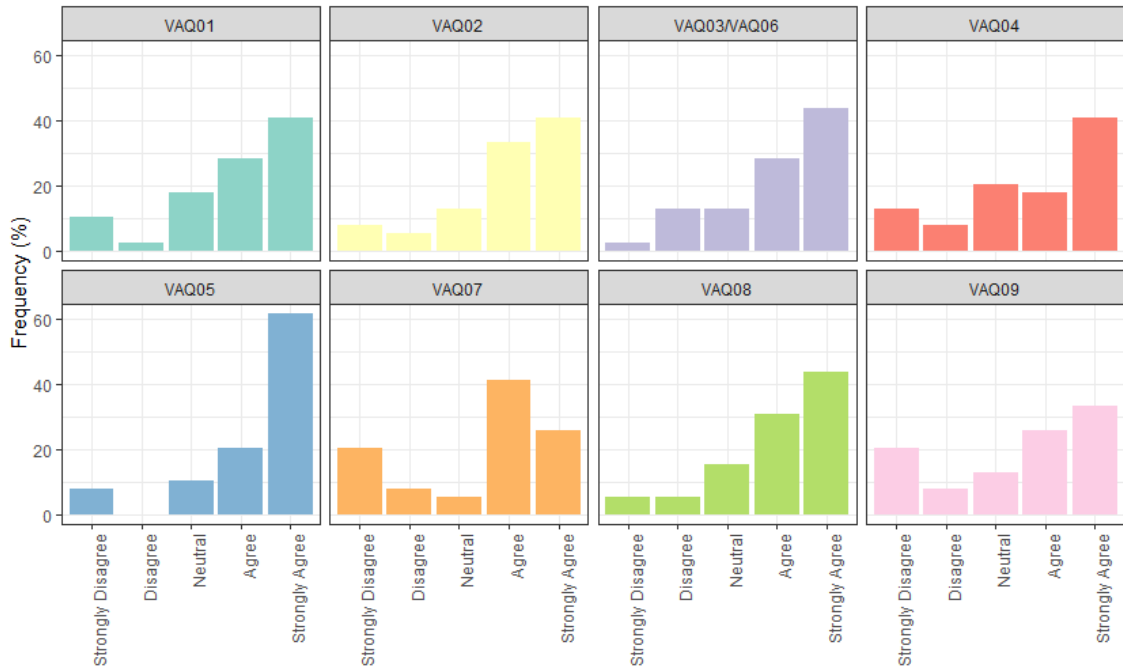
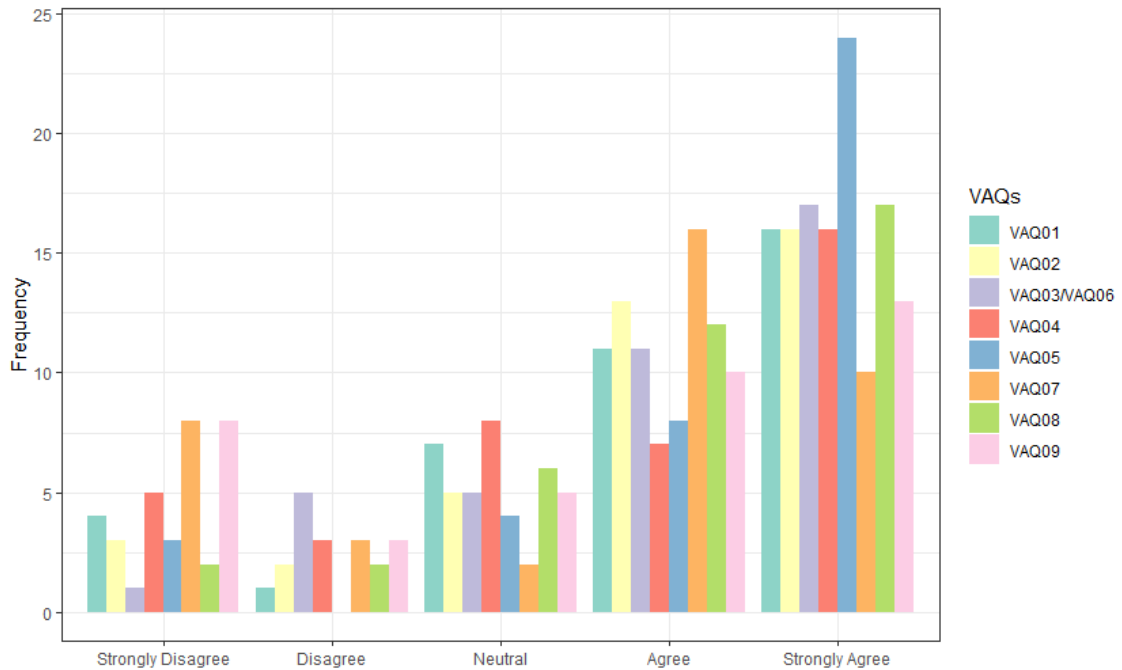


Figure 6.2: Bar chart showing differences between the understanding of each VAQ, using absolute frequency.



VAQ2 with 74.3% of agreement, revealed that users were able to investigate the cohesiveness of parties in the spectrum. Finally, with 74.36% of agreement, users

were capable of verifying the activity of several periods in the Chamber of Deputies, addressed by VAQ8.

Considering the specific groups created for the questionnaire, in the next sections we will perform a detailed analysis and discussion about the results obtained in each group.

6.1.3.1 Timeline and Cropped Timeline

As Fig. 6.3 shows, the majority of users (92.31%) could select a period in the main timeline, either by clicking or brushing, with no difficulties (TL2). Also, we got some reasonable results to TL3 and CTL1 (both with 71.19% of agreement), which evaluate the understanding of users to verify the parties' size and verify the behavior of deputies and parties across time, respectively. Regarding the perception of deputies size, a volunteer give us an interesting idea. For instance, today the PSOL is a small party and, consequently, is harder to find in our Timeline. However, this party has considerable visibility on media and in our society. So, it is essential to find a way that gives more visibility, in our temporal visualizations, to this kind of parties.

The main issue with these visualizations was the difficulty of users in understanding the Y-axis, which positions the pro-government parties on top and the ones that belong to the opposition, on the bottom. This problem affected directly the understanding of Timeline, which can be noticed by the high level of disagreement in TL1 (20.51%). Analyzing the Cropped Timeline, these high disagreement results can be explained due to the complexity of CTL2 (28.21%) and CTL3 (28.21%), which demands the user to pick a particular deputy that changes his/her party on a selected period. A user suggests using a zoom feature to explore the Cropped Timeline, since the user can select an extensive period, and the space between years can be minimal and hard to interpret. Also, it was cited a certain complexity in distinguishing the parties and deputies traces in Cropped Timeline.

6.1.3.2 Political Spectrum

In general, Political Spectrum showed acceptable results with users agreement: PS1 (74.36%), PS2 (74.36%), PS3 (82.05%), as you can see in Fig. 6.4. The high level of agreement in PS3 showed that most users could check which deputies

Figure 6.3: Summary of results from questions on the user understanding about Timeline visualizations.

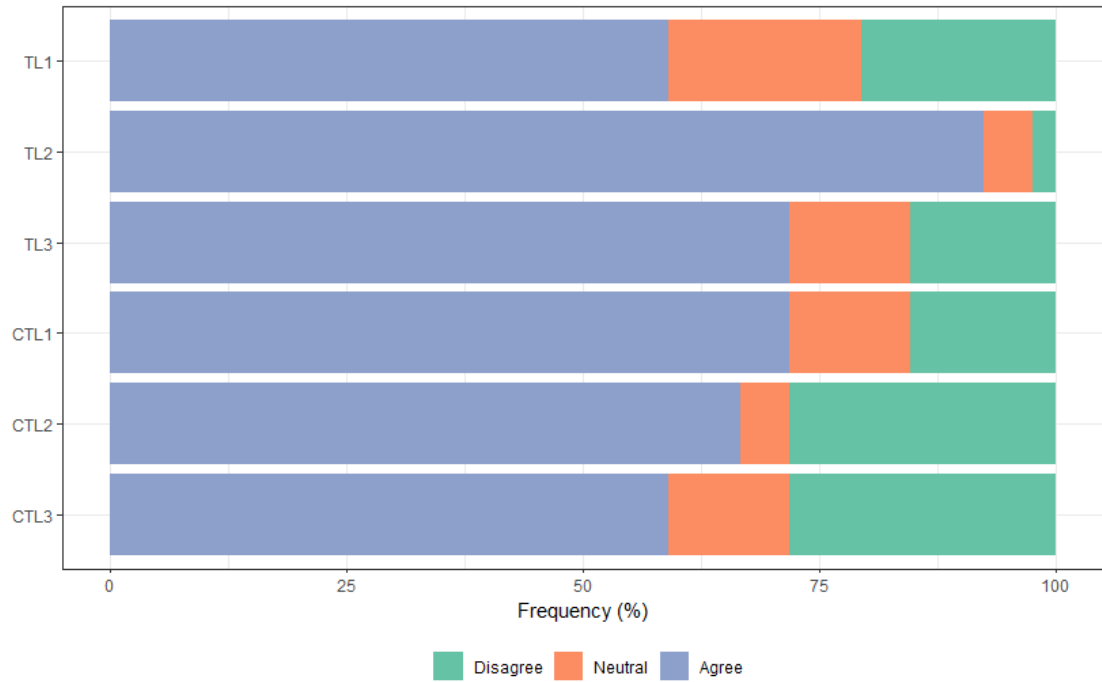
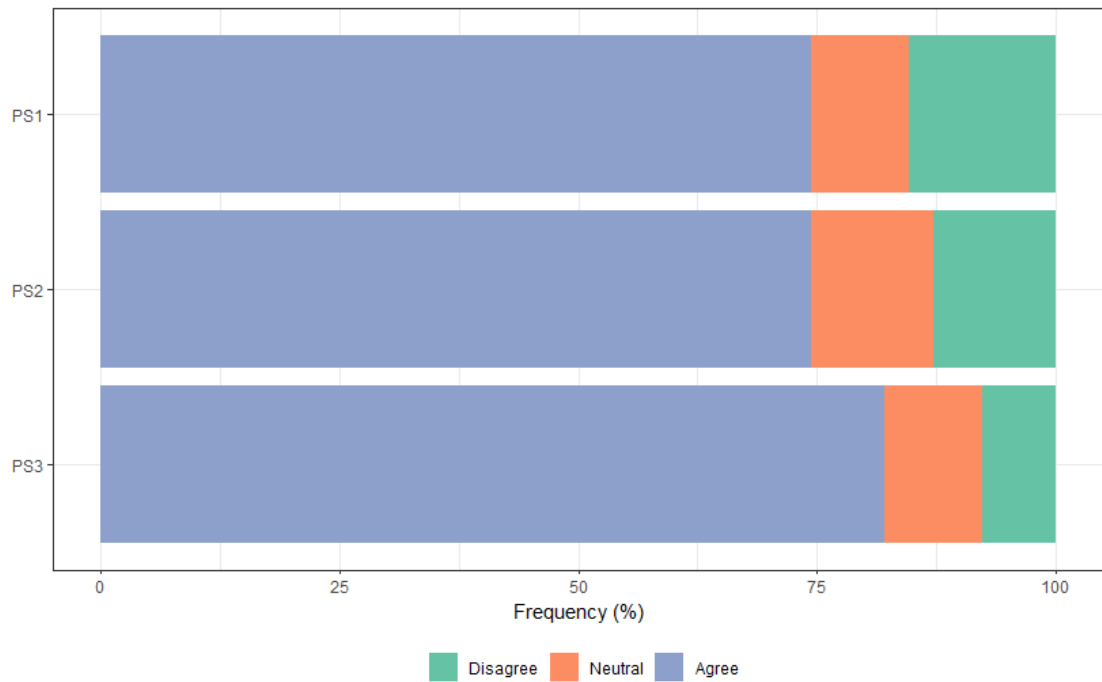


Figure 6.4: Summary of results from questions on the user understanding about Political Spectrum visualizations.



are politically divergent from their parties without difficulties. Some users asked about the meaning of three different dimensionality reduction techniques. It is rel-

evant to discuss it since our system can be used by the general and specific public. For the general public, we may need to redesign the interface to use familiar terms to describe these techniques and provide details on demand to experts, that may be interested in which method is being used.

Also, some users noticed the missing identification on the axis. Since our spectrum is a low-dimensional space created by these techniques, the axis has no semantical meaning of our data. Considering this, we choose to hide the axis values. In the future, we want to give an ideological sense for axis, such as conservative and liberal.

6.1.3.3 Map of Roll Calls, Chamber Infographic and Similarity Graph

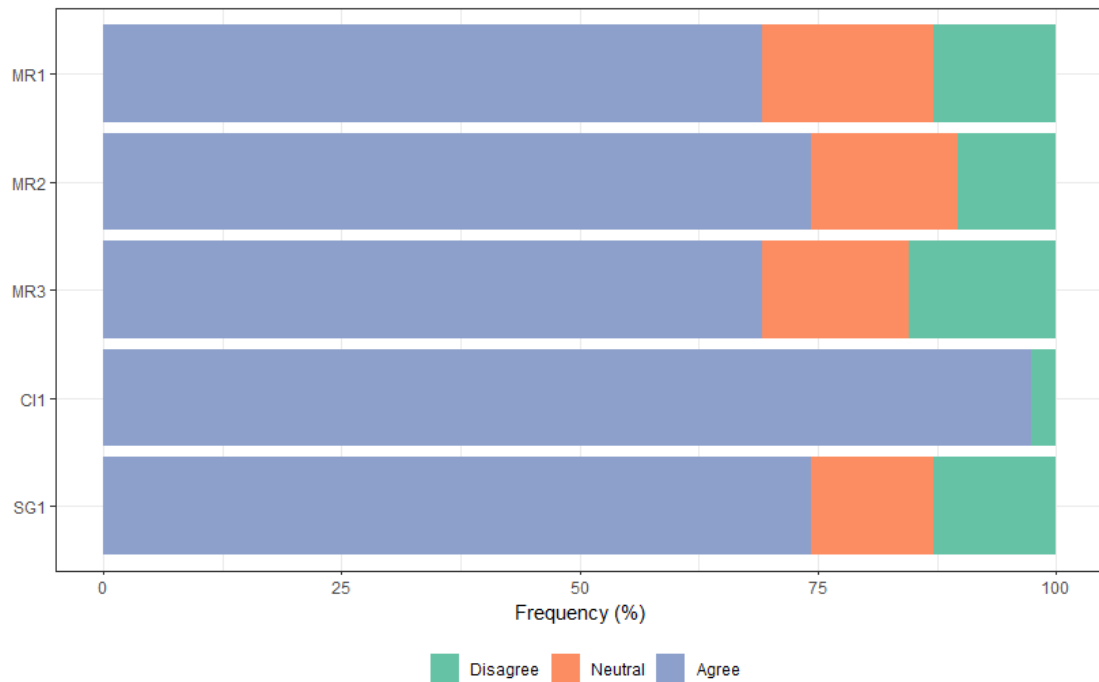
Overall, Map of Roll Calls showed reasonable results (see Fig. 6.5). The statement MR2 showed the best result, with 74.36% of agreement. Users could verify the most active periods of Chamber of Deputies with no problems. The use of histogram elements in our Map of Roll Calls may have contributed to a better understanding of this visualization, since histogram is a widely used visualization technique. With 69.23 % of agreement, MR1 may be complicated to some users. To identify the individual votes, the user has to be aware of how coordinated views works. The user has to select a deputy in some visualization and then verify in Map of Roll Calls his/her votes, or oppositely: select a roll call and see on parent visualization the individual votes. Also, it was verified by a user the lack of a legend for individual votes.

A volunteer suggests us to cluster roll calls by subjects, which can be a powerful tool to investigate the behavior of deputies based in particular areas, such as agriculture, sports, health, etc.

The CI1 showed one of the highest values of agreement (97.44%). The Chamber Infographic used a semi-circle representation with deputies grouped by parties to presents the distribution of seats in a simple way. The excellent result can be attributed to the adopted design since it is frequently used to depict Legislative Houses.

The Similarity Graph (SG1) showed a reasonable result, with 74.36% of agreement. Some users compared the similarity graph with the spectrum and affirmed that in some cases, it could be easier to interpret than the political spectrum. Another issue reported by several users was the low performance of the system when

Figure 6.5: Summary of results from questions on the user understanding about Map of Roll Calls, Chamber Infographic and Similarity Graph.



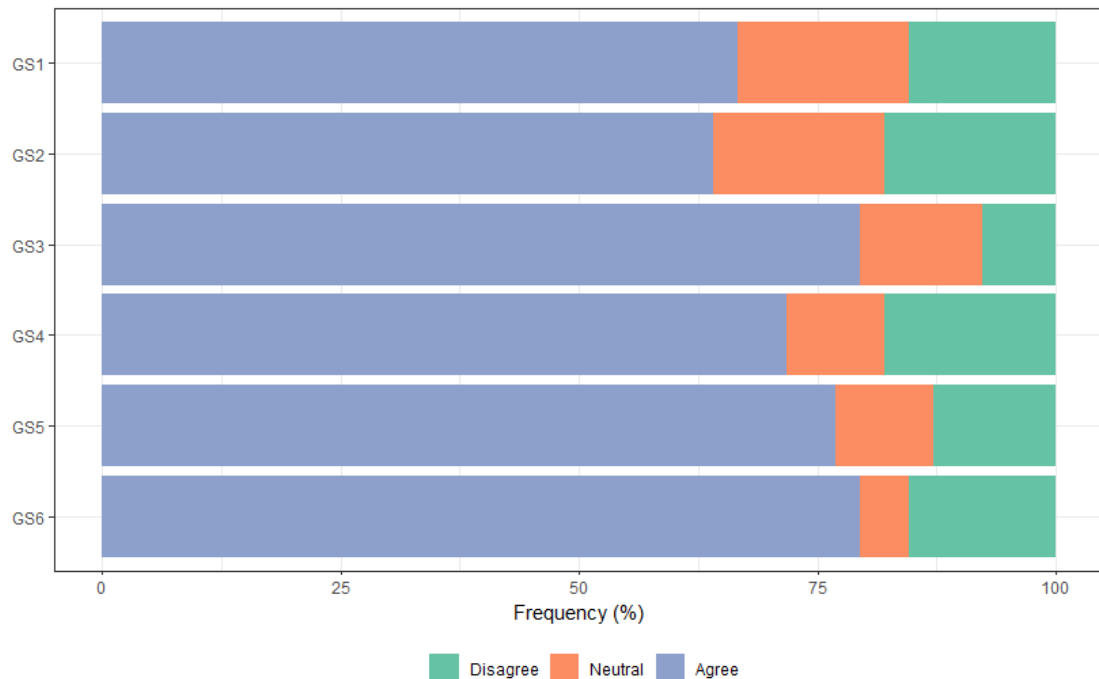
the graph is too dense, with a large number of edges and points. It is crucial to find good thresholds for the similarity measure in different periods to solve this performance problem.

6.1.3.4 System Features

Considering the general features of the system, the best results were obtained in GS3 (79.49%), GS5 (76.92%), GS6 (79.49%), as you can see in Fig. 6.6. Users could use the windows system in the workspace, organizing and positioning the views according to what they want to analyze. A relevant interface problem cited by a user was the lack of identification in minimized windows, in a scenario with several minimized panels can be hard to find a specific view. Regarding searching deputies by name and resetting all selections, the users could perform these actions without any difficulties. To improve the workflow, a user recommends using a reset selection button for each panel instead of using a single button to reset all selections.

The process of generating views and the relationship between them were not well understood by users, which can be verified by lower values of agreement, GS1 (66.67%) and GS2 (64.10%), respectively. Several users reported problems with the

Figure 6.6: Summary of results from questions concerning the user understanding about the system features.



use of the context menu to generate new views. Some users were not able to create specific visualizations; others wasted a considerable time to realize how it works. They complained mainly about the inconsistent interaction with elements of the interface, some components opens a context menu, and others do not.

Another critical issue reported by users is regarding the District filter (GS4), with 71.79% of agreement. The filter is available as soon as application loads, and if the user interacts with it before generating any visualization from the Timeline, nothing will happen, and the user may be confused with this behavior. The district filter is applied only on views that show deputies, such as Spectrum of Deputies, Chamber Infographic and Similarity Graph. A possible solution would be repositioning this filter input to each panel, in its settings menu.

The tutorial was not evaluated on questionnaire statements, but users give some feedback on it: 1) Must have an alternative way to trigger the tutorials, such as buttons in the interface. In some cases, the user may forget how to do something and have to reload the entire application to replay the tutorial. 2) Since we used GIFs in the tutorial to show how to use our system, the user can not play/pause the video, also they do not have feedbacks of how long it will last.

Finally, a user considered the parties' color palette inadequate. Our idea was

to represent parties with actual colors, for example, red for PT, blue for PSDB, etc. But, to improve the understanding of the system, we can investigate another palettes.

6.1.3.5 Final Analysis

Intending to investigate if the previous use of a visualization tool influences the understanding of our system, we divided our sample into two different groups: those who used some visualization tool before and those who did not. We calculated the agreement mean for each questionnaire statement and used a T-test to compare them. The results are summarized in Table 6.3.

The result of the comparison of the two groups showed no significant difference ($p > 0.05$) between them. Each statement was analyzed individually and all p-values are greater than 0.05. The statement that showed the greater difference, but still not significant, was the CTL2: verify how deputies switched parties over time, which can be a complicated task to those are not accustomed to temporal visualizations.

Table 6.3: The table compares two different groups based on previous experience with visualization tools. The first column shows the users that do not have previous experience and the second column, the users that have previous experience.

	No experience N=13	Experience N=26	p.overall
TL1	3.15 (1.68)	3.92 (1.23)	0.158
TL2	4.62 (1.12)	4.54 (0.65)	0.822
TL3	3.77 (1.64)	4.08 (0.98)	0.542
PS1	4.00 (1.35)	4.04 (1.18)	0.931
PS2	3.62 (1.39)	4.12 (1.11)	0.271
PS3	4.23 (1.24)	4.31 (1.16)	0.853
CI1	4.92 (0.28)	4.65 (0.85)	0.150
MR1	3.85 (1.21)	3.88 (1.34)	0.929
MR2	4.15 (1.34)	3.96 (1.04)	0.656
MR3	3.23 (1.64)	3.88 (0.86)	0.198
CTL1	3.46 (1.51)	4.23 (0.86)	0.107
CTL2	2.77 (1.88)	3.77 (1.14)	0.096

continued on next page

Table 6.3 – *continued from previous page*

	No experience N=13	Experience N=26	p.overall
CTL3	2.85 (1.86)	3.73 (1.28)	0.141
SG1	4.00 (1.22)	4.04 (1.08)	0.924
GS1	3.46 (1.71)	4.04 (1.08)	0.283
GS2	3.23 (1.36)	3.77 (1.03)	0.224
GS3	4.00 (1.35)	4.35 (0.98)	0.422
GS4	3.62 (1.76)	4.27 (1.19)	0.242
GS5	3.92 (1.50)	4.38 (1.06)	0.334
GS6	4.46 (1.33)	4.08 (1.38)	0.409

6.2 Remote Domain Experts Study

The second evaluation was based on a survey centered on users who are familiar with visualization tools and have professional experience with political analysis, being thus interested in exploring the Chamber of Deputies roll calls dataset. The focus of this study was on how useful these domain experts found our system to be for the purposes of professional political analysis.

6.2.1 Participants

We sent the survey with instructions to experts via email. The test involved 6 participants, 5 male and 1 female, their ages ranging from 30 to 45 years old ($\bar{x} = 34$ and $\sigma = 5.51$). All participants are familiar with visualization tools such as Tableau (3), Qlik (3), and Power BI (4). Three of them have an MSc degree and the other three have PhDs. Regarding their professions, 2 are economists, 2 are professors (one of them of political science), 1 is an economic analyst, and 1 is a journalist. Table 6.4 summarizes the information about our participants.

Table 6.4: Summary of experts

N=6	
Age:	
Mean (SD)	34.00 (5.51)
Gender:	
Female	1 (16.67%)
Male	5 (83.33%)
Education:	
PhD	3 (50.00%)
MSc	3 (50.00%)
Profession:	
Economic Analyst	1 (16.67%)
Economist	2 (33.33%)
Journalist	1 (16.67%)
Professor	2 (33.33%)

6.2.2 Procedure

The survey was organized in 11 steps, each represented as a section of a questionnaire. The first three steps are the same used in the general public study: 1) a short presentation of the system; 2) questions about basic personal information, such as age, gender, and profession; and 3) instructions and link to access Civis-Analysis 2.0. In this study, users performed the evaluation remotely. The only usage instructions they received were the content of the embedded tutorials.

The next six sections concerned different visualization techniques, namely: Timeline (TL), Political Spectrum (PS), Chamber Infographic (CI), Map of Roll Calls (MR), Cropped Timeline (CTL), and Similarity Graph (SG). Our focus was on their utility for the expert users, so we first had to establish whether they understood the techniques before asking them if they found them useful and were happy with their usability. Each visualization section therefore included the following three statements to which participants had to indicate their level of agreement using a five-point Likert scale in which 1 indicated strong disagreement and 5, strong

agreement (LIKERT, 1932). In general, the statements are described as follows:

- I understand the visualization.
- I think the visualization is useful.
- I think the visualization is easy to use.

Each section also included a paragraph input where users could provide commentary on the corresponding visualization. Due to the simplicity of Chamber Infographic, we did not ask about its ease of use.

The 10th section evaluated the system in general, with statements designed to assess whether users understood and found useful the windowing system in both its interactive (e.g., moving, hiding, showing, and resizing panels) (GS-W) and hierarchical aspects (GS-H). These statements are defined as follows:

- I understand the feature.
- I think the feature is useful.

Finally, the 11th section asked users for general comments, suggestions, and feedback about the system.

6.2.3 Results

We decided to discuss and analyze the results by comparing the same statements in different visualizations. Using this approach, we can identify the visualizations that are most understandable, most useful, and easiest to use.

In the following subsections, we will describe and discuss the users' answers and feedback on several aspects of CivisAnalysis 2.0. The users were identified as follows:

- (*ECO1*) Economist 1
- (*ECO2*) Economist 2
- (*PROF1*) Professor 1
- (*PROF2*) Professor 2

- (*EA*) Economic Analyst

- (*JOUR*) Journalist

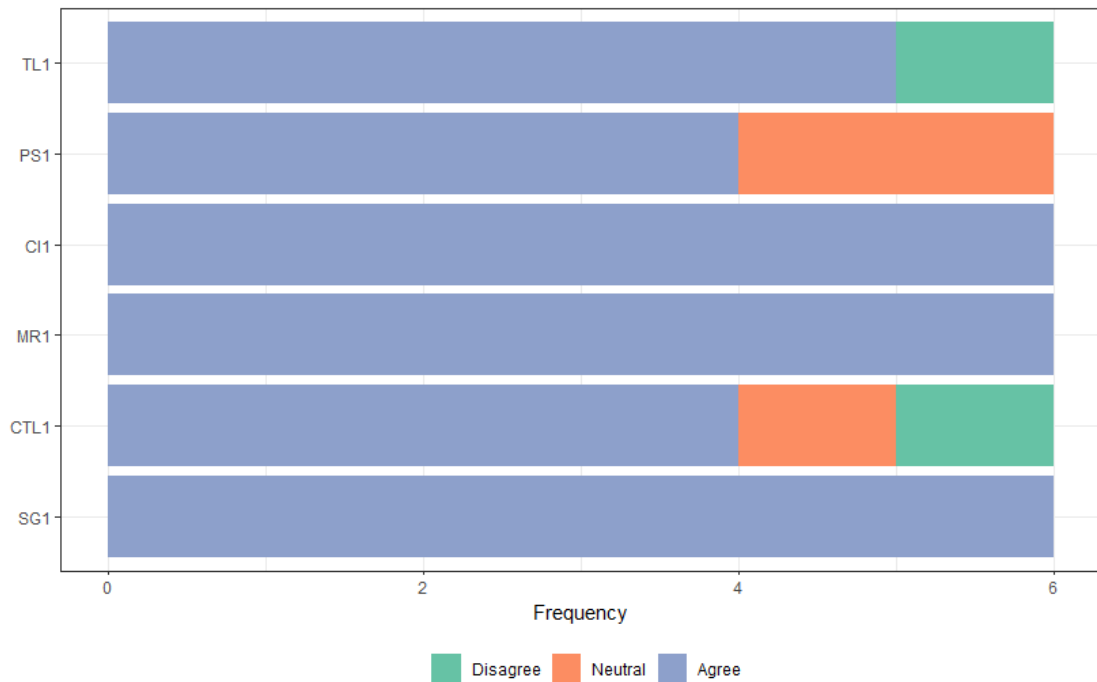
6.2.3.1 Understanding the visualizations

As evidenced in Fig. 6.7, all expert users understood the Chamber Infographic, the Map of Roll Calls and the Similarity Graph. Considering the understanding of Map of Roll Calls, the *EA* suggested making the distinction between a deputy's vote and the roll call's result clearer. Despite their agreement with the statements with regards to the Similarity Graph, three users (*ECO 2*, *EA*, and *JOUR*) asked for more explanations to be included about the similarity degree.

Users manifested less agreement with the statements regarding the temporal visualizations (TL1 and CTL1). *PROF1* (the only disagreement shown in Fig. 6.7) was not able to understand the Y-axis on TL1. Even though the other experts understood the meaning of the vertical axis, *EA* and *JOUR* complained of a lack of explanation about it. The absence of a label explaining the government-opposition behavior also affected results. Regarding the main timeline, *ECO2* suggested the use of zooming to obtain another perspective of a particular period, similar to interactive plots available in popular systems. Concerning CTL1, the only neutral assessment came from *EA*, who reported difficulties in distinguishing between parties and deputies. Since the Cropped Timeline is quite similar to the Main Timeline, *PROF1* also disagreed with the statement with respect to it, mentioning again difficulties in understanding the Y-axis.

As for the the understanding of Political Spectrum of Deputies, no user disagreed with the statement, but two provided a neutral answer (*ECO2* and *EA*) and the other four had a lower level of agreement. *EA* suggested we provide more explanations on how the spectrum is generated. *JOUR*, in turn, suggested we remove the names of reduction dimensionality methods, as technical terminology may create a considerable distance between our system and the general public. One alternative would be to keep method names hidden by default, but make them available for experts. The users' responses concerning this visualization were also interesting in that some experts were also evaluating the statement with the general public in mind (e.g., *JOUR*), which can even be an explanation for their lower levels of agreement.

Figure 6.7: Summary of the results from questions on the level of user understanding of the different visualizations.



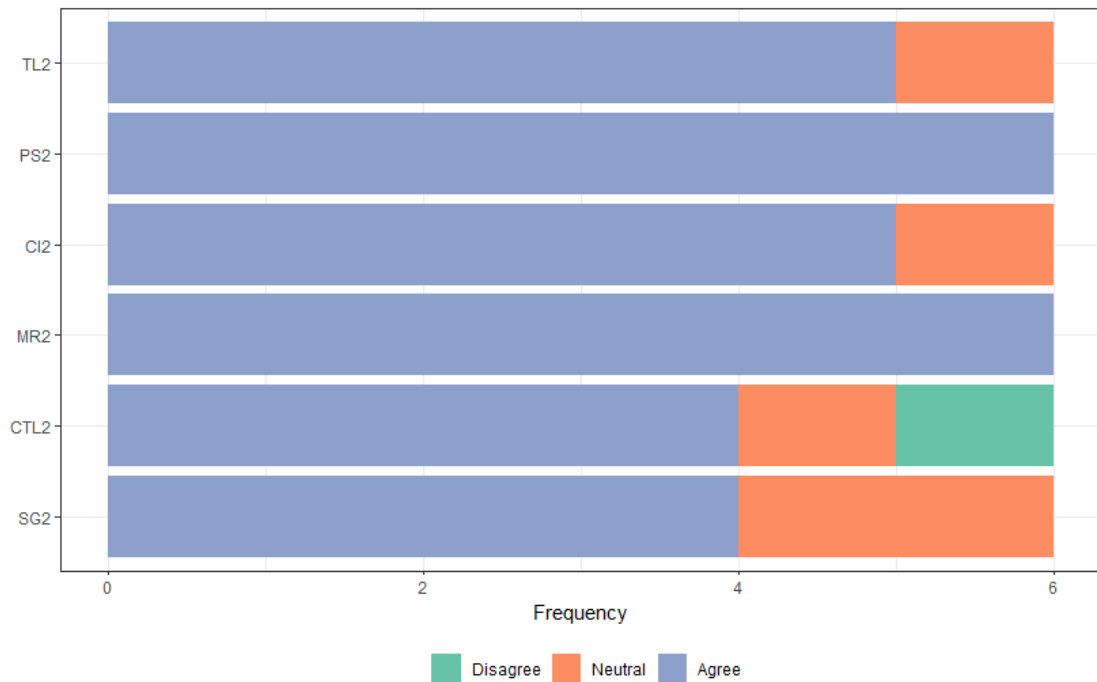
6.2.3.2 Usefulness of the visualizations

Fig.6.8 shows relevant results about the utility of our visualizations. Users did not disagree with the statement regarding any of the visualizations, with the exception of Cropped Timeline (CTL2), which had one disagreement.

Again, the Map of Roll Calls (MR2) obtained six agreement responses, but this time regarding its utility. The *ECO2* suggested including the popular names some of the roll calls have so that they can be more easily found. He also suggested placing roll calls in thematic categories, such as economy, civil rights, property rights, defense, etc..

All six users also agreed with the statement regarding the utility of the Political Spectrum of Deputies (PS2). As mentioned by *ECO2*, the spectrum can be useful to uncover groups and verify how similarly deputies vote. However, he also pointed out that the analytical power of this visualization could be improved if we also provided spectra that considered semantic information (e.g., the liberal-conservative dimension). It is important to note that even though *EA* and *ECO2* gave neutral responses regarding the understanding of the spectrum (PS1), they recognized its importance and utility.

Figure 6.8: Summary of the results from questions on the utility of the different visualizations.

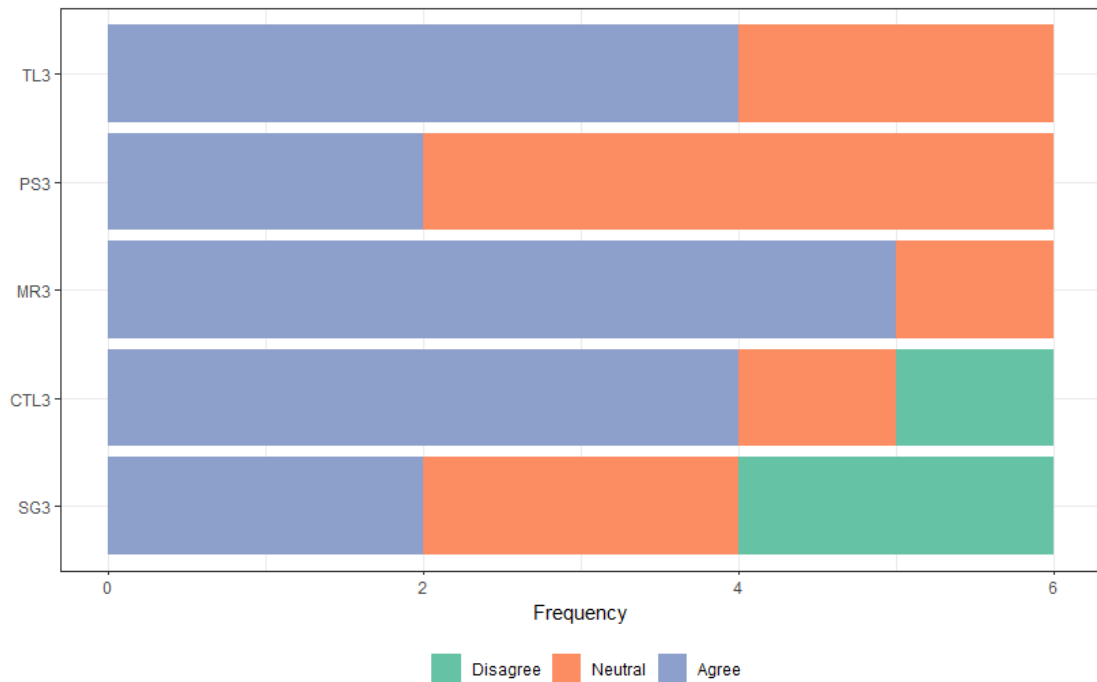


According to the experts, the least useful visualizations are Cropped Timeline (CTL2) and Similarity Graph (SG2), both of which obtained four agreement responses. The only disagreement concerning CTL2 was given by *PROF1*. This can probably be explained by his disagreement response in the previous subsection, as it is unlikely that someone would find useful something that he was unable to understand. Two users gave two neutral responses to the statement with respect to SG2: *PROF1* and *PROF2*. On the other hand, *EA* and *PROF2* commented that the Similarity Graph is an excellent visualization, although *PROF2* also mentioned that regular users may have trouble understanding it.

6.2.3.3 Ease of use of the visualizations

Concerning the Map of Roll Calls (MR3), no users disagreed with the statement and five agreed with it, as seen in Fig. 6.9. We found these results are positively surprising, as the Map of Roll Calls has a complex interaction scheme, using extensively coordinated features. The choice of a histogram to represent roll call votes seems to have been a good design decision, probably contributing to these results. The only neutral response was given by the *EA*, who manifested some difficulty

Figure 6.9: Summary of the results from questions on the ease of use of the different visualizations.



in distinguishing individual votes from roll call results, as mentioned in subsection 6.2.3.1.

Concerning TL3 and CTL3, users gave the same answers as they had for ease of use (four agreement responses). As previously discussed (in subsections 6.2.3.1 and 6.2.3.2), they complained that the lack of an explanation for the Y-axis made these temporal visualizations hard to use and understand.

Regarding the Political Spectrum of Deputies (PS3), we found a peculiar result: four neutral and two agreement responses. The political spectra can be created considering different criteria and we did not make clear what criteria we used, as reported by *EA*. Therefore, as mentioned by *ECO2*, the axes can have multiple interpretations and we did not provide indication of their meaning. These issues might have influenced the results.

The Similarity Graph (SG3) divided our users in three same-sized groups, with two responses for each category (disagreement, neutral, agreement). *PROF1* and *PROF2* disagreed with the statement. Although *PROF2* understood the meaning of the graph, he evaluated it with the general public in mind and thought it could be hard to use for those unfamiliar with it. This answer could explain his neutral response in the utility of the Similarity Graph (SG2). If the general public wouldn't

be able to use the view, it loses its utility. The neutral responses were given by *EA* and *JOUR*, both of whom asked for more explanations about the similarity degree. Finally, even though *ECO2* gave us an agreement response, he wished more details had been provided about the definition of “similarity.”

6.2.3.4 General system features

To evaluate the general features of the application, we wanted to assess how useful users found our windowing system to be. This involved assessing not only its utility, but also users’ understanding of it. Seen in Fig. 6.10, GS-W1 concerns the understanding of this system’s features (e.g., moving, hiding, resizing, and showing panels, etc.) by users, whereas GS-W2 refers to whether users found these features useful for the exploratory analysis of the data. As seen in the figure, users showed a high level of understanding of the windowing system, with five agreement responses, but only four experts considered it useful. *JOUR* suggested that we provide users with a list of all available visualizations when the application loads, as this could make the exploration process easier and reduce the mental effort of remembering all the options that are hidden in context menus.

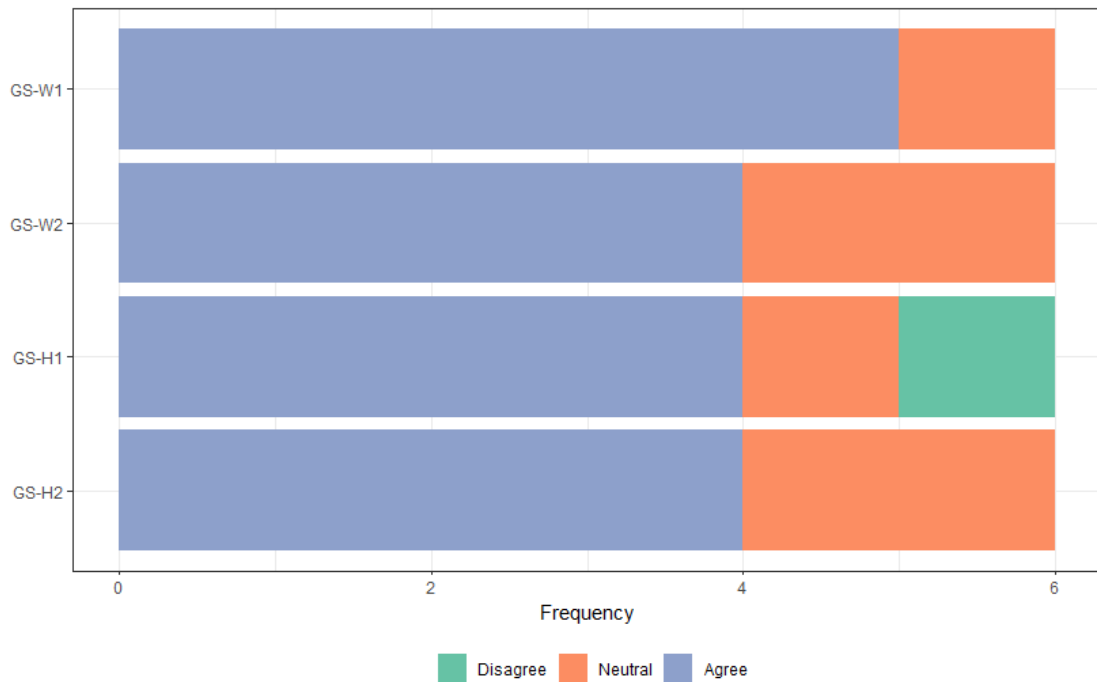
Another aspect of the user interface we wanted to evaluate was the utility (and therefore also the users’ understanding) of the hierarchical aspect of the windowing system. We therefore asked experts if they considered linked and coordinated views easy (GS-H1) to understand and useful (GS-H2) to explore the dataset.

As seen in Fig. 6.10, both statements (GS-H1 and GS-H2) had the same level of agreement: 4 responses. It is also noteworthy that although *ECO1* did not understand the hierarchical approach, he considered it useful.

Furthermore, *EA* suggested that we make the relationship between the views more evident. We believe this might have been because when a user picks a deputy in one view, this deputy will be selected in other views regardless of the hierarchical structure. In this case, the relationships are not explicit as child-parent relationship. Visualizations in which this happens include the Spectrum of Deputies and the Map of Roll Calls.

Some users also commented that it was not very intuitive to generate views from context menus. *EA* mentioned that the process of clicking on a period and then right-clicking to open a context menu is not intuitive.

Figure 6.10: Summary of results from questions on the ease of use in different visualizations.



6.2.3.5 Experts comments, suggestions and feedback

Overall, the experts were very excited about the system and its potential. They gave some compliments about the work, providing comments such as: *ECO1*: “Amazing tool,” *JOUR*: “The system has a huge potential to become a competitive platform for study, analysis, and discussion in the public sphere (press, etc.),” *EA*: “Excellent tool with precious information.”

We also received feature suggestions: to use our tool to explore the Brazilian Federal Senate and include a liberal x conservative dimension in our Political Spectrum (*ECO2*). *EA* was concerned about the system intuitiveness and suggested that we included more definitions (e.g., definition of similarity).

The users’ overall main concern was the high complexity of the system. Considering that the general public, experts suggested some improvements: display all visualizations options in a way (*JOUR*) that makes the exploration process more straightforward; identify the Y-axis in temporal visualizations (*PROF1*, *EA*, *JOUR*); remove the nomenclature of reduction dimensionality methods (*JOUR*).

6.3 Final Comments

Our two experiments involved 45 individuals, 39 from the general public and six experts. In general, both groups showed reasonable results. None of our statements had a high-level of disagreement, and most users were able to use the system and answer almost all questions.

In the general public study, we evaluated all statements that were directly connected to our VAQs. We have observed that VAQ5 was the most straightforward question to answer, with 82% of agreement. This VAQ asks users to find deputies politically divergent of their parties, which may be an interesting exercise to predict parties switching since the deputy is not aligned with the current party. Also, we evaluated each statement separately, observing high agreement values for Chamber Infographic, Political Spectrum and Map of Roll Calls.

With experts, the highlight was the Map of Roll Calls, with 6, 6 and 5 agreement responses, for understanding, utility and ease of use, respectively. However, in contrast to the general study, the Political Spectrum of Deputies showed lower agreement levels.

We can observe issues reported either by the general public or by experts, such as problems to create new visualizations with context menus, not understanding the relationship between some views, a difficulty in understanding the Y-axis of temporal visualizations, doubts about reduction dimensionality methods. These issues must be prioritized in the next implementation cycles.

We also received a lot of compliments from both groups. The comments are a mixture of suggestions to adjust our system, with motivating messages to continue our research in the area, highlighting the importance to make this data accessible and understandable for all kinds of users. We will summarize all the issues and, based on the users' feedback, we will update our system to provide a better experience in future.

7 CONCLUSIONS AND FUTURE WORK

We presented a set of interactive visualization techniques integrated in a web-based application for the exploration of roll calls data aiming at supporting users in gaining insights about the voting behavior of legislators and political parties in Brazil. Inspired by a previous solution, we designed CivisAnalysis 2.0 to address some limitations of that application by providing a set of techniques integrated through a user interface based on hierarchically organized multiple coordinated views. From a political timeline that covers 28 years of Brazil's Chamber of Deputies activity, users can generate different views that depict the voting behavior of parties and deputies over time. The application can be of use not only by citizens who want to be better informed but also by journalists interested in telling data-driven stories. A case study on party cohesiveness illustrates such kind of application.

The approach was evaluated through a remote users study divided in two groups: general public and experts. While the evaluation with the general public was planned to measure the user's understanding of the general flow of the system, with tutorial assistance, the evaluation with experts aimed at assessing the utility and difficulty of using our system.

As future work, we aim to improve the system usability based on the feedback from both groups of users, and perform a final usability evaluation. Also, we aim to generate political spectra of deputies and parties based on subsets of roll calls to allow for the analysis of deputy votes on specific topics such as education and health. Other possible features include creating co-authorships networks, using a similar approach to Similarity Graph; analyzing deputies behavior considering Parliament Fronts and comparing with Parties behavior. Furthermore, we aim to include the roll calls voted in 2019. Since it is a new legislature and president, new patterns could emerge and the tool can be useful for political analysts.

REFERENCES

- ATLAS Político. 2019. <<http://atlaspolitico.com.br>>.
- BALDONADO, M. Q. W.; WOODRUFF, A.; KUCHINSKY, A. Guidelines for using multiple views in information visualization. In: ACM. **Proceedings of the working conference on Advanced visual interfaces**. [S.l.], 2000. p. 110–119.
- BARBERÁ, P. et al. Tweeting from left to right: Is online political communication more than an echo chamber? **Psychological science**, Sage Publications Sage CA: Los Angeles, CA, v. 26, n. 10, p. 1531–1542, 2015.
- BASÔMETRO. 2019. <<http://estadaodados.com/basometro/>>.
- BORG, I.; GROENEN, P. Modern multidimensional scaling: Theory and applications. **Journal of Educational Measurement**, Wiley Online Library, v. 40, n. 3, p. 277–280, 2003.
- BORJA, F. G. **CivisAnalysis: exploring representatives' voting behaviour**. Programa de Pós-Graduação em Computação: UFRGS, 2017. Disponível em: <<http://hdl.handle.net/10183/157571>>.
- BORJA, F. G. de; FREITAS, C. M. Civisanalysis: Interactive visualization for exploring roll call data and representatives' voting behaviour. In: IEEE. **Graphics, Patterns and Images (SIBGRAPI), 2015 28th SIBGRAPI Conference on**. [S.l.], 2015. p. 257–264.
- BRAZILL, T. J.; GROFMAN, B. Factor analysis versus multi-dimensional scaling: binary choice roll-call voting and the us supreme court. **Social Networks**, Elsevier, v. 24, n. 3, p. 201–229, 2002.
- BRIGADIR, I. et al. Dimensionality reduction and visualisation tools for voting record. In: CEUR WORKSHOP PROCEEDINGS. **24th Irish Conference on Artificial Intelligence and Cognitive Science (AICS'16), University College Dublin, Ireland, 20-21 September 2016**. [S.l.], 2016.
- CARD, A. S. S.; JACKO, J. **"Information visualization" in The human-computer interaction handbook: Fundamentals evolving technologies and emerging applications**. [S.l.]: New York:Lawrence Erlbaum, 2008.
- CARD, M. **Readings in information visualization: using vision to think**. [S.l.]: Morgan Kaufmann, 1999.
- CARROLL, R.; KUBO, H. Measuring and comparing party ideology and heterogeneity. **Party Politics**, 2017.
- CARROLL, R. et al. Comparing nominate and ideal: Points of difference and monte carlo tests. **Legislative Studies Quarterly**, Wiley Online Library, v. 34, n. 4, p. 555–591, 2009.
- CARROLL, R.; POOLE, K. Roll call analysis and the study of legislatures. In: _____. **The Oxford Handbook of Legislative Studies**. [S.l.]: Oxford University Press, 2014. cap. 5, p. 103–125.

CAVA, R. A. **Abordagens heterogêneas para a exploração interativa de grafos multivariados**. Programa de Pós-Graduação em Computação: UFRGS, 2017. Disponível em: <<http://hdl.handle.net/10183/157522>>.

CLINTON, J.; JACKMAN, S.; RIVERS, D. The statistical analysis of roll call data. **American Political Science Review**, Cambridge University Press, v. 98, n. 2, p. 355–370, 2004.

DESPOSATO, S. W. Parties for rent? ambition, ideology, and party switching in brazil's chamber of deputies. **American Journal of Political Science**, Wiley Online Library, v. 50, n. 1, p. 62–80, 2006.

DIACONIS, P. et al. Horseshoes in multidimensional scaling and local kernel methods. **The Annals of Applied Statistics**, Institute of Mathematical Statistics, v. 2, n. 3, p. 777–807, 2008.

DUNNE, C. et al. Graphtrail: Analyzing large multivariate, heterogeneous networks while supporting exploration history. In: ACM. **Proceedings of the SIGCHI conference on human factors in computing systems**. [S.l.], 2012. p. 1663–1672.

ENELOW, J. M.; HINICH, M. J. **The spatial theory of voting: An introduction**. [S.l.]: CUP Archive, 1984.

FISHER, D. R.; LEIFELD, P.; IWAKI, Y. Mapping the ideological networks of american climate politics. **Climatic change**, Springer, v. 116, n. 3-4, p. 523–545, 2013.

FRIGGERI, A.; FLEURY, E. Agreement groups in the united states senate. In: **IOGDC-International Open Government Data Conference-2012**. [S.l.: s.n.], 2012.

FRUCHTERMAN, T. M.; REINGOLD, E. M. Graph drawing by force-directed placement. **Software: Practice and experience**, Wiley Online Library, v. 21, n. 11, p. 1129–1164, 1991.

GRIMMER, J.; STEWART, B. M. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. **Political Analysis**, Cambridge University Press, v. 21, n. 3, p. 267–297, 2013.

HIX, S.; NOURY, A. G.; ROLAND, G. **Democratic politics in the European Parliament**. [S.l.]: Cambridge University Press, 2007.

HOUGHTON, W. R. **Conspectus of the History of Political Parties and the Federal Government**. [S.l.]: Granger, Davis & Company, 1880.

JAKULIN, A. et al. Analyzing the u.s. senate in 2003: Similarities, clusters, and blocs. **Political Analysis**, [Cambridge University Press, Oxford University Press, Society for Political Methodology], v. 17, n. 3, p. 291–310, 2009. ISSN 10471987, 14764989.

JOLLIFFE, I. Principal component analysis. In: **International encyclopedia of statistical science**. [S.l.]: Springer, 2011. p. 1094–1096.

- KANTARDZIC, M. **Data mining: concepts, models, methods, and algorithms**. [S.l.]: John Wiley & Sons, 2011.
- KINNAIRD, P.; ROMERO, M.; ABOWD, G. Connect 2 congress: visual analytics for civic oversight. In: ACM. **CHI'10 Extended Abstracts on Human Factors in Computing Systems**. [S.l.], 2010. p. 2853–2862.
- KRIESI, H. et al. Globalization and the transformation of the national political space: Six european countries compared. **European Journal of Political Research**, Wiley Online Library, v. 45, n. 6, p. 921–956, 2006.
- KULLBACK, S. **Information theory and statistics**. [S.l.]: Courier Corporation, 1997.
- KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. **The annals of mathematical statistics**, JSTOR, v. 22, n. 1, p. 79–86, 1951.
- LEITE, L.; TRENTO, S. Análise de votações nominais do legislativo brasileiro utilizando componentes principais. **Leviathan (São Paulo)**, n. 12, p. 120–163, 2016.
- LESTER, J. The evolution of the political compass (and why libertarianism is not right-wing). **Journal of Social and Evolutionary Systems**, Elsevier, v. 17, n. 3, p. 231–241, 1994.
- LIKERT, R. A technique for the measurement of attitudes. **Archives of psychology**, 1932.
- MAATEN, L. van der; HINTON, G. Visualizing data using t-sne. **The Journal of Machine Learning Research**, Wiley Online Library, v. 9, n. 11, p. 2579–2605, 2008.
- MARINO, R. **A valsa dos partidos, de Collor a Dilma**. 2014. <<http://www.todasasconfiguracoes.com/2014/04/24/a-valsas-dos-partidos/>>.
- MATTILA, M.; LANE, J.-E. Why unanimity in the council? a roll call analysis of council voting. **European Union Politics**, SAGE Publications, v. 2, n. 1, p. 31–52, 2001.
- MONROE, R. **A History of The United States Congress, Partisan and Ideological Makeup**. 2012. Disponível em: <<http://xkcd.com/1127/large>>.
- MONTESQUIEU, C. D. **Montesquieu: The spirit of the laws**. [S.l.]: Cambridge University Press, 1989.
- MUNZNER, T. **Visualization analysis and design**. [S.l.]: AK Peters/CRC Press, 2014.
- NERY, P. F. **Mapa da atuação da bancada do PSOL na Câmara do RJ, até o atentado contra Marielle**. 2018. Disponível em: <<https://www.jota.info/opiniao-e-analise/artigos/mapa-psol-camara-atentado-marielle-07052018>>.

- NERY, P. F. **Atuação do PSL é difusa e desorganizada, mostram dados**. 2019. Disponível em: <<https://www.jota.info/opiniao-e-analise/artigos/psl-atuacao-difusa-dados-22052019>>.
- NERY, P. F. **Em alta, bancada evangélica é mais numerosa e coesa que grandes partidos**. Gazeta do Povo, 2019. Disponível em: <<https://www.gazetadopovo.com.br/vozes/pedro-fernando-nery/em-alta-bancada-evangelica-e-mais-numerosa-e-coesa-que-grandes-partidos/>>.
- PERER, A.; SHNEIDERMAN, B. Integrating statistics and visualization: Case studies of gaining clarity during exploratory data analysis. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. New York, NY, USA: ACM, 2008. (CHI '08), p. 265–274. ISBN 978-1-60558-011-1.
- POOLE, K. T.; ROSENTHAL, H. A spatial model for legislative roll call analysis. **American Journal of Political Science**, JSTOR, p. 357–384, 1985.
- POOLE, K. T.; ROSENTHAL, H. Patterns of congressional voting. **American journal of political science**, JSTOR, p. 228–278, 1991.
- RAO, K. R.; YIP, P. C. **The transform and data compression handbook**. [S.l.]: CRC press, 2000.
- ROBERTS, J. C. State of the art: Coordinated & multiple views in exploratory visualization. In: IEEE. **Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)**. [S.l.], 2007. p. 61–71.
- ROSENTHAL, H.; VOETEN, E. Analyzing roll calls with perfect spatial voting: France 1946–1958. **American Journal of Political Science**, Wiley Online Library, v. 48, n. 3, p. 620–632, 2004.
- ROWEIS, S. T.; SAUL, L. K.; HINTON, G. E. Global coordination of local linear models. In: **Advances in neural information processing systems**. [S.l.: s.n.], 2002. p. 889–896.
- SCHONHARDT-BAILEY, C. Ideology, party and interests in the british parliament of 1841–47. **British journal of political science**, Cambridge University Press, v. 33, n. 4, p. 581–605, 2003.
- SILVA, R. N. M.; SPRITZER, A.; FREITAS, C. D. S. Visualization of roll call data for supporting analyses of political profiles. In: **2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)**. [S.l.: s.n.], 2018. p. 150–157. ISSN 2377-5416.
- TAMAYO, L. F. T. **SmartParticipation: a fuzzy-based recommender system for political community-building**. [S.l.]: Springer, 2014.
- TORGERSON, W. S. Multidimensional scaling: I. theory and method. **Psychometrika**, Springer, v. 17, n. 4, p. 401–419, 1952.
- VOTEVIEW. 2019. <<https://voteview.com>>.

WICKELMAIER, F. An introduction to mds. **Sound Quality Research Unit, Aalborg University, Denmark**, Citeseer, v. 46, n. 5, p. 1–26, 2003.

WINDHAGER, F.; SMUC, M. The arts of the possible information visualization in the field of politics. In: **Conference for E-Democracy and Open Government**. [S.l.: s.n.], 2014. p. 375.

YOUNG, G.; HOUSEHOLDER, A. S. Discussion of a set of points in terms of their mutual distances. **Psychometrika**, Springer, v. 3, n. 1, p. 19–22, 1938.

APPENDIX A — QUESTIONNAIRES

A.1 General Public

Avaliação do CivisAnalysis 2.0

CivisAnalysis 2.0 é uma aplicação Web que usa um conjunto de diferentes visualizações coordenadas para explorar o comportamento dos deputados, através dos seus votos. Para atingir este objetivo foi criado um sistema onde o usuário possui uma "Área de trabalho" e pode decidir quais visualizações devem ser ou não mostradas.

Os dados utilizados neste trabalho abrangem: informações sobre deputados, proposições, votações, partidos, etc. Os dados foram obtidos do Portal de Dados Abertos da Câmara dos Deputados (<https://dadosabertos.camara.leg.br/>).

Este trabalho foi desenvolvido como Dissertação de Mestrado de Rodrigo Moni, aluno do Programa de Pós-Graduação em Computação do Instituto de Informática da UFRGS. O projeto foi inspirado e dá continuidade ao trabalho iniciado na Dissertação de Mestrado de Francisco Gerdau de Borja, também no PPGC-INF/UFRGS.

Caso você concorde em avaliar a aplicação, agradecemos!

Para iniciar, basta clicar no botão abaixo e avançar para a próxima seção.

* Required

Caracterização do usuário

Primeiro gostaríamos de saber um pouco mais sobre você!
Responda essas perguntas para prosseguir com a avaliação.

1. Idade *

2. Gênero *

Mark only one oval.

- Feminino
- Masculino
- Não desejo informar
- Other: _____

3. Formação *

Mark only one oval.

- Sem escolaridade
- Ensino Fundamental
- Ensino Médio
- Ensino Superior
- Especialização
- Mestrado
- Doutorado

4. Profissão *

5. **Você já utilizou alguma ferramenta para visualização de dados ? Se sim, quais?**

Check all that apply.

- Tableau
- Power BI - Microsoft
- Infogram
- Qlik
- D3.js
- Plot.ly
- Other: _____

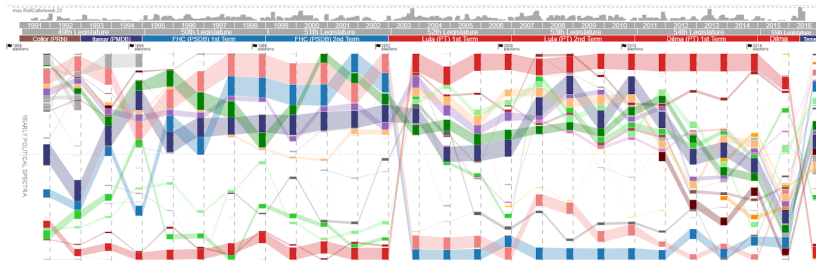
Utilização do CivisAnalysis

Você deverá utilizar a ferramenta CivisAnalysis 2.0 através do seguinte link:
<http://www.inf.ufrgs.br/~rmsilva/CivisAnalysis2/>

Um tutorial será disparado assim que a página for aberta. Leia e veja os exemplos com atenção. Como dito anteriormente, o sistema possui diferentes visualizações e para cada uma será mostrado um diferente tutorial.

Você deve usar a ferramenta livremente e quando sentir-se confortável, responda as questões abaixo. Cada questão se refere a uma parte do sistema (indicada no início da questão) e consiste de uma afirmação, que deverá ser respondida através de notas de 1 a 5 (1- Discordo totalmente e 5 - Concordo totalmente).

Linha do tempo



6. **Linha do tempo - "Eu consigo verificar como os partidos votaram ao longo dos anos." ***

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

7. **Linha do tempo - "Eu consigo seleccionar um período histórico na visualização." ***

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

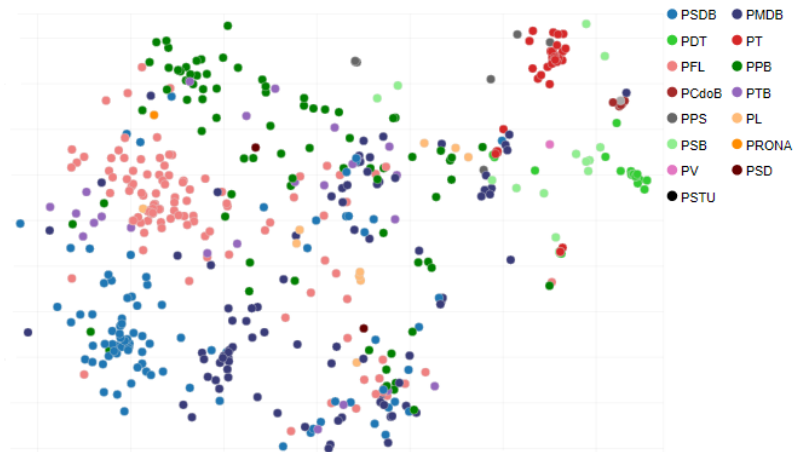
8. Linha do tempo - "Eu consigo identificar o tamanho de cada partido." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

Espectro Político



9. Espectro Político - "Eu consigo compreender o espectro de deputados." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

10. Espectro Político - "Eu consigo investigar se os partidos são coesos." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

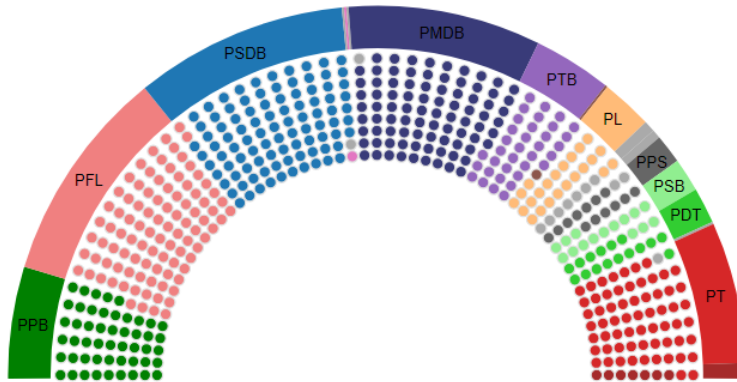
11. Espectro Político - "Eu consigo encontrar quais são os deputados divergentes de seus respectivos partidos." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

Infográfico da Câmara

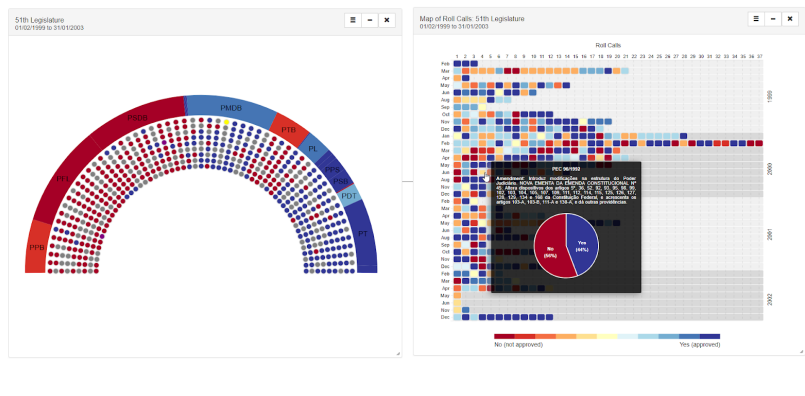


12. Infográfico da Câmara - "Eu consigo identificar a distribuição de assentos na Câmara." *
 Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

Mapa de votações (direita). Lembrando que: o mapa de votações está diretamente ligado a sua visualização pai.



13. Mapa de votações - "Eu consigo identificar os votos individuais de cada deputado." *
 Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

14. Mapa de votações - "Eu consigo identificar os períodos mais ativos da Câmara dos Deputados." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

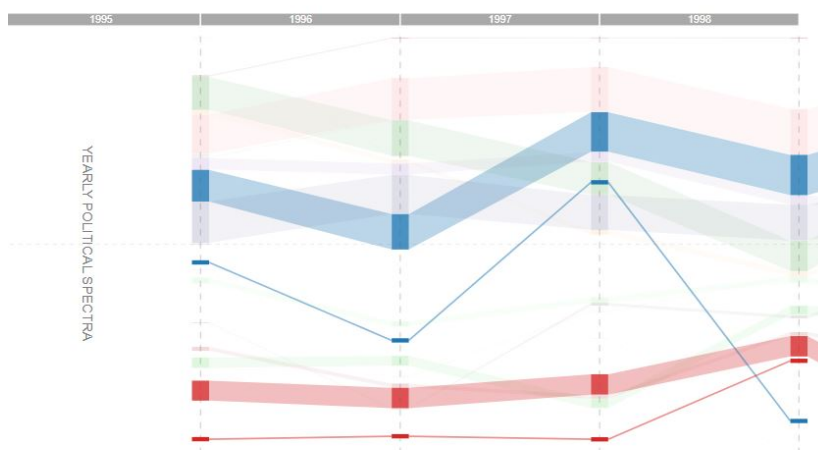
15. Mapa de votações - "Eu consigo selecionar proposições e aplicar filtros sobre elas." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

Linha do tempo com deputados



16. Linha do tempo com deputados - "Eu consigo verificar como partidos e deputados se comportam ao longo do tempo." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

17. **Linha do tempo com deputados - "Eu consigo verificar como os deputados trocam de partido ao longo do tempo."** *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

18. **Linha do tempo com deputados - "Eu consigo verificar como os deputados se comportam após uma mudança de partido."** *

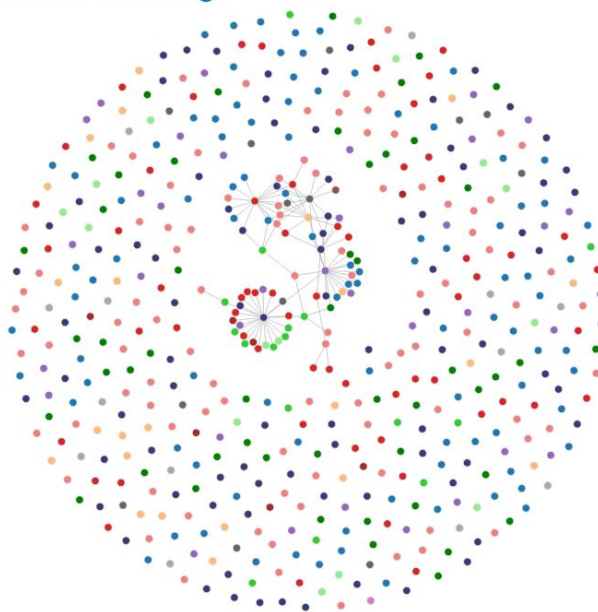
Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

Grafo de similaridade

Select the grade of similarity: 50% 100%



19. **Grafo de similaridade - "Eu consigo verificar quais deputados votaram de forma mais similar (tiveram um maior índice de similaridade)."** *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

20. Sistema Geral - "Eu consigo criar novas visualizações a partir de outras." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

21. Sistema Geral - "Eu consigo entender como as visualizações se relacionam." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

22. Sistema Geral - "Eu consigo organizar e mover as visualizações de acordo com o que desejo visualizar melhor." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

23. Sistema Geral - "Eu consigo aplicar o filtro por Estados." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

24. Sistema Geral - "Eu consigo reiniciar todas as seleções previamente feitas." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

25. Sistema Geral - "Eu consigo realizar buscas de deputados por nome." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

26. Deixe aqui sua opinião sobre o CivisAnalysis 2.0, suas críticas e sugestões.

Fim da Avaliação!

A.2 Experts

Avaliação do CivisAnalysis 2.0

CivisAnalysis 2.0 é uma aplicação web que usa um conjunto de diferentes visualizações coordenadas para explorar o comportamento dos deputados, através dos seus votos.

Os dados utilizados neste trabalho abrangem: informações sobre deputados, proposições, votações, partidos, etc. Eles foram obtidos do Portal de Dados Abertos da Câmara dos Deputados (<https://dadosabertos.camara.leg.br/>).

Este trabalho foi desenvolvido como Dissertação de Mestrado de Rodrigo Moni, aluno do Programa de Pós-Graduação em Computação do Instituto de Informática da UFRGS. O projeto foi inspirado e dá continuidade ao trabalho iniciado na Dissertação de Mestrado de Francisco Gerda de Borja, também no PPGC-INF/UFRGS.

Caso você concorde em avaliar a aplicação, agradecemos!

Para iniciar, basta clicar no botão abaixo e avançar para a próxima seção.

* Required

Caracterização do usuário

Primeiramente, gostaríamos de saber um pouco mais sobre você.
Por favor, responda as perguntas abaixo para seguir com a avaliação.

1. Idade *

2. Gênero *

Mark only one oval.

- Feminino
 Masculino
 Não desejo informar
 Other: _____

3. Formação *

Mark only one oval.

- Sem escolaridade
 Ensino Fundamental
 Ensino Médio
 Ensino Superior
 Especialização
 Mestrado
 Doutorado

4. Profissão *

5. **Você já utilizou alguma ferramenta para visualização de dados ? Se sim, quais?**

Check all that apply.

- Tableau
- Power BI - Microsoft
- Infogram
- Qlik
- D3.js
- Plot.ly
- Other: _____

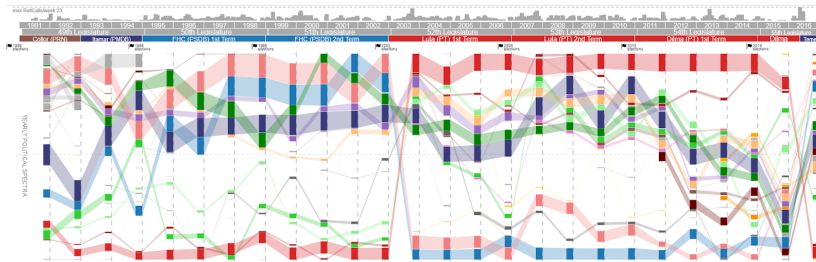
Utilização do CivisAnalysis

Você deverá utilizar a ferramenta CivisAnalysis 2.0 através do seguinte link:
<http://www.inf.ufrgs.br/~rmmsilva/CivisAnalysis2/>

Um tutorial será disparado assim que a página for aberta. Leia e veja os exemplos com atenção. Como dito anteriormente, o sistema possui diferentes visualizações e para cada uma será mostrado um diferente tutorial.

Você deve usar a ferramenta livremente e quando sentir-se confortável, responda as questões abaixo. Cada questão se refere a uma parte do sistema (indicada no início da questão) e consiste de uma afirmação, que deverá ser respondida através de notas de 1 a 5 (1- Discordo totalmente e 5 - Concordo totalmente).

Linha do tempo



6. **Linha do tempo - "Eu consigo verificar como os partidos votaram ao longo dos anos." ***

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

7. **Linha do tempo - "Eu consigo selecionar um período histórico na visualização." ***

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

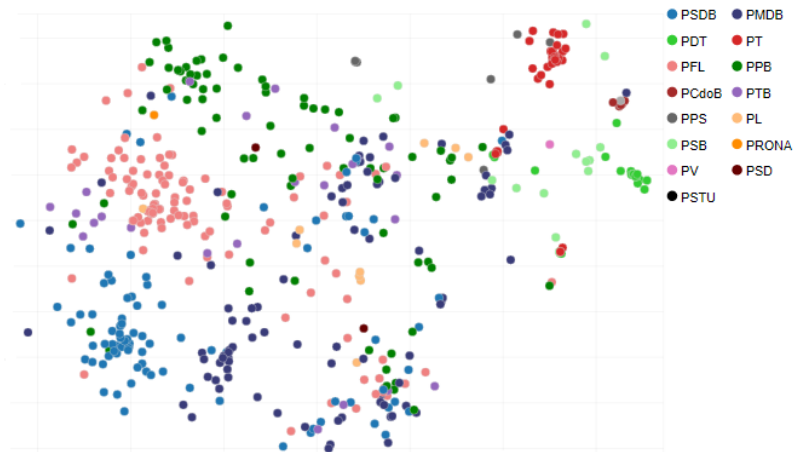
8. Linha do tempo - "Eu consigo identificar o tamanho de cada partido." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

Espectro Político



9. Espectro Político - "Eu consigo compreender o espectro de deputados." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

10. Espectro Político - "Eu consigo investigar se os partidos são coesos." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

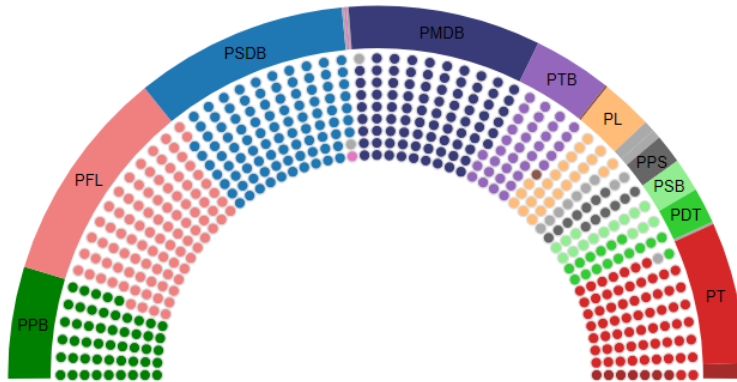
11. Espectro Político - "Eu consigo encontrar quais são os deputados divergentes de seus respectivos partidos." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

Infográfico da Câmara

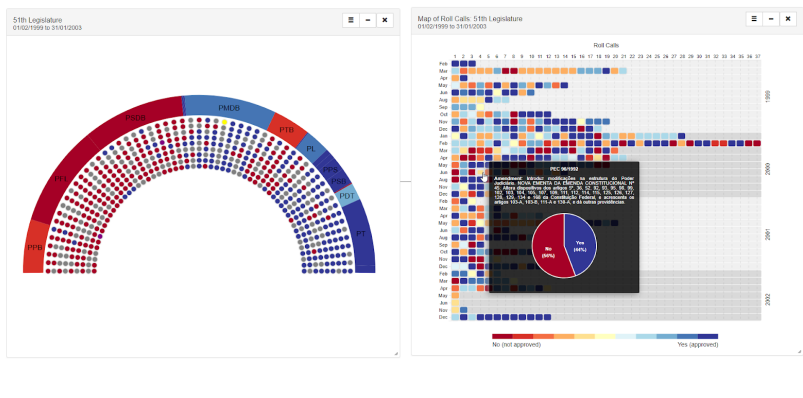


12. Infográfico da Câmara - "Eu consigo identificar a distribuição de assentos na Câmara." *

1 2 3 4 5

Discordo totalmente Concordo totalmente

Mapa de votações (direita). Lembrando que: o mapa de votações está diretamente ligado a sua visualização pai.



13. Mapa de votações - "Eu consigo identificar os votos individuais de cada deputado." *

1 2 3 4 5

Discordo totalmente Concordo totalmente

14. Mapa de votações - "Eu consigo identificar os períodos mais ativos da Câmara dos Deputados." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

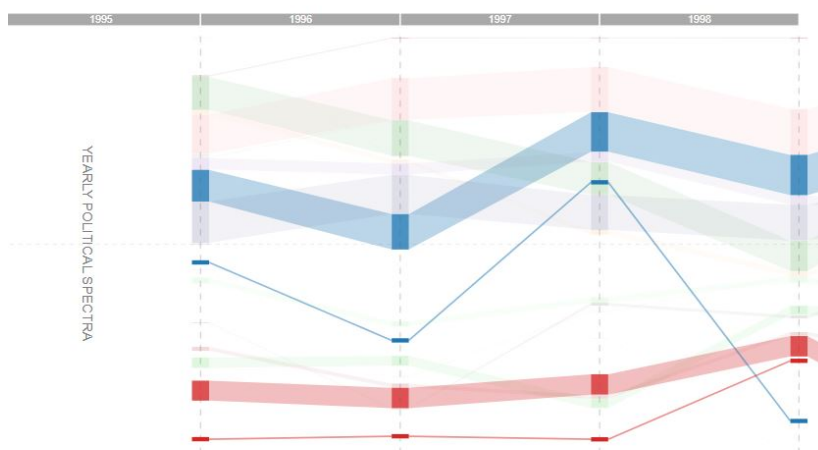
15. Mapa de votações - "Eu consigo selecionar proposições e aplicar filtros sobre elas." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

Linha do tempo com deputados



16. Linha do tempo com deputados - "Eu consigo verificar como partidos e deputados se comportam ao longo do tempo." *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

17. **Linha do tempo com deputados - "Eu consigo verificar como os deputados trocam de partido ao longo do tempo."** *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

18. **Linha do tempo com deputados - "Eu consigo verificar como os deputados se comportam após uma mudança de partido."** *

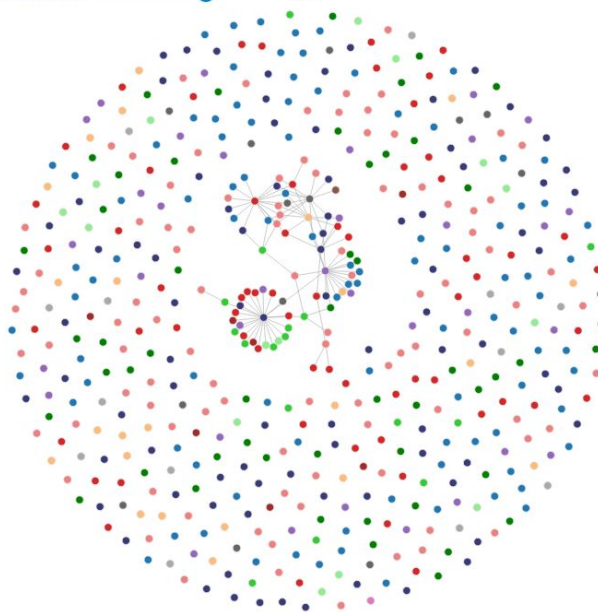
Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

Grafo de similaridade

Select the grade of similarity: 50% 100%



19. **Grafo de similaridade - "Eu consigo verificar quais deputados votaram de forma mais similar (tiveram um maior índice de similaridade)."** *

Mark only one oval.

1 2 3 4 5

Discordo totalmente Concordo totalmente

20. Sistema Geral - "Eu consigo criar novas visualizações a partir de outras." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

21. Sistema Geral - "Eu consigo entender como as visualizações se relacionam." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

22. Sistema Geral - "Eu consigo organizar e mover as visualizações de acordo com o que desejo visualizar melhor." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

23. Sistema Geral - "Eu consigo aplicar o filtro por Estados." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

24. Sistema Geral - "Eu consigo reiniciar todas as seleções previamente feitas." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

25. Sistema Geral - "Eu consigo realizar buscas de deputados por nome." *

Mark only one oval.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

26. Deixe aqui sua opinião sobre o CivisAnalysis 2.0, suas críticas e sugestões.

Fim da Avaliação!