SAPIENZA
UNIVERSITÀ DI ROMA

European Parties Explorer: a Visual Analytics system for European politics

Faculty of Information Engineering, Informatics and Statistics
Master of Science in Engineering in Computer Science

Stefano Bonetti
ID number 1764624

Advisor
Prof. Simone Lenti

Co-Advisor
Prof. Giuseppe Santucci

Academic Year 2024/2025

European Parties Explorer: a Visual Analytics system for European politics
Master Degree thesis. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: bonetti.1764624@studenti.uniroma1.it

To all those who have
always pushed me
to move forward

Abstract

The complexity of the European political landscape has grown significantly over the last twenty-five years, with the Old Continent facing several challenges such as the 2008 economic crisis, the migrant crisis, Brexit, climate change, the Covid pandemic, the rise of authoritarian governments and the return of war on European soil. The European Union and, more in general, European politics have certainly changed dramatically during this period of time, and in order to study this evolution there exist many datasets offered by the world of political science research, but this data remains mostly inaccessible to non-expert users.

This thesis presents *European Parties Explorer*, a Visual Analytics system created to allow the study of European politics. Adopting a user-centred design approach, we identified the analytical needs of distinct categories of users and designed this system with the intention of displaying complex data about European parties in the most accessible way possible. The resulting application facilitates the discovery of significant insights such as the evolution of political families and the analysis of major political topics.

Contents

Chapter 1

Introduction

In the field of political science, there exist many datasets designed for studying the evolution of party ideologies in Europe. Among these, the **Chapel Hill Expert Survey** (**CHES**) is a dataset whose objective is allowing the analysis of European national parties, mainly for political science researchers, and it offers estimations of party stances on many political topics like the European Union, economy, environment, immigration and LGBT rights, starting from 1999. However, the data is high-dimensional, temporally fragmented, and distributed in raw formats that are difficult for non-experts to utilise. Therefore, finding insights about the evolving European political landscape is something that remains mostly inaccessible to the general public, and researchers too have no way to directly explore this data in an interactive way.
This thesis presents the whole design process of **European Parties Explorer**, a Visual Analytics system built upon the CHES data. The objective of this application is to present the CHES dataset in an interactive environment where voters, journalists and political scientists can explore the history of European parties; the project follows a user-centred design approach, which means that every step during design takes into consideration the defined users and their needs. By analysing the different user tasks we were able to identify the application's functional requirements, which guided us in choosing what visualisations to use and how to integrate them in the political analysis process.

The thesis is structured as follows:

- **Chapter 2** provides the theoretical background on Visual Analytics in general and existing approaches to data-driven political analysis, contextualising the work within the current state of the art and presenting what its objectives are;

- **Chapter 3** analyses the CHES dataset in detail, discussing its structure, how the data is gathered, the main political topics covered, the dataset's reliability and examples of how it is used in related work;

- **Chapter 4** presents the design process, defining the user personas, the tasks they can perform, and the functional requirements of the system.

- **Chapter 5** describes the technical implementation, explaining the data preprocessing step, the system architecture, the chosen visualisation techniques and in what ways they support the political analysis;
- **Chapter 6** presents how the user tasks can be accomplished on the actual system implementation together with real-world usage scenarios, demonstrating how the system can be used to extract political findings;
- **Chapter 7** summarises the results achieved, discusses the limitations of the current system, and outlines directions for future work.

Chapter 2

Background

Before going into the specifics of this thesis' project, we should first describe the theoretical context in which it is situated and present existing visual tools dedicated to the political analysis of European national parties. This chapter introduces what the field of Visual Analytics is, its fundamental concepts that were essential during the design process of European Parties Explorer, how it can be adapted to political science and the current approaches to data-driven political analysis.

2.1 Introduction to Visual Analytics

The field of Visual Analytics was originally defined by Thomas and Cook as "the science of analytical reasoning facilitated by interactive visual interfaces" [25]; building on this, Keim et al. provided a more specific definition, stating that Visual Analytics "combines automated analysis techniques with interactive visualisations for an effective understanding, reasoning and decision making on the basis of very large and complex datasets" [18]. As such, according to the authors, the goal is the creation of tools to enable the following:

- synthesise information and derive insight from massive, dynamic, ambiguous, and often conflicting data;
- detect the expected and discover the unexpected;
- provide timely, defensible, and understandable assessments;
- communicate these assessments effectively for action.

In practice, Visual Analytics is a discipline that deals with representing great amounts of data effectively through many visualisations, so that one can not only analyse and explore it, but also accomplish other tasks such as explaining the data, solving problems and taking decisions.

2.1.1 The Visual Analytics process

It is essential to notice that Visual Analytics applications should not merely just "display" information: they should most importantly provide the instruments to

discover **insights**, that is extracting facts that otherwise would have been very difficult, if not impossible, to find by just observing the data without the help of visual tools; however, gaining knowledge from data is not a linear and trivial procedure. The so-called **Visual Analytics process**, illustrated in Figure 2.1, describes how automatic and visual analysis methods coupled with human interaction allow the discovery of insights.
The Visual Analytics cycle typically consists of four parts:

- the **data** is preprocessed (e.g. cleaned, normalised) to make it suitable for analysis;
- automated analysis techniques (like data mining) can be applied to the data to generate a **model** of it; this can be refined, in case interacting with visualisations displaying the model itself;
- the data can be explored with **visualisations**, allowing direct user interaction, with the automated algorithms confirming the generated hypotheses;
- **knowledge** is obtained thanks to the continuous interaction between user, views and automatic analysis.

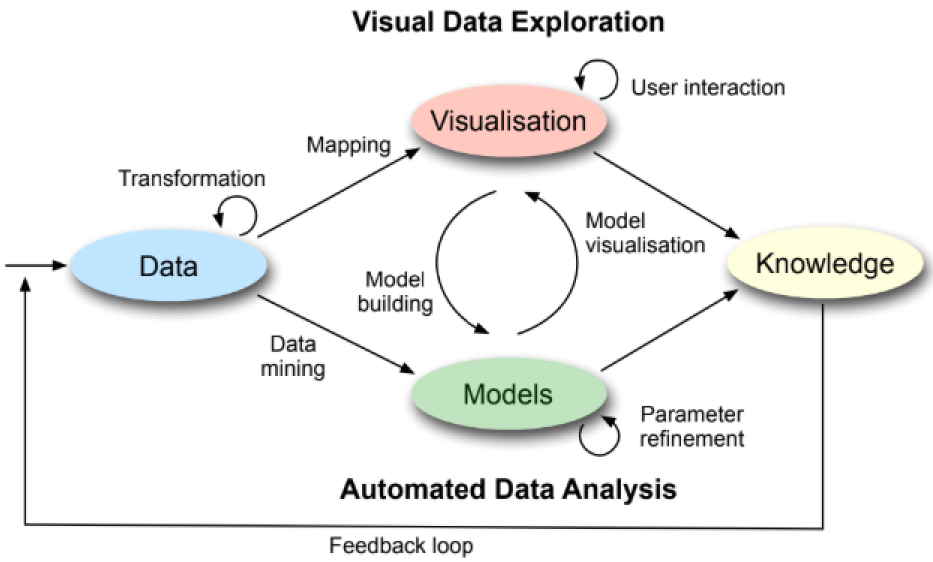


Figure 2.1. The Visual Analytics process. Source: Keim et al. [18]

The user is involved in every step of this loop: they can interact directly with the visualisations to explore the data, but they can also interact with the model (e.g. by changing parameters) or the data processing stage. This workflow was summarised by Keim as:

> “Analyse first, show the important, zoom/filter, analyse further, details on demand” [18].

Here lies the difference with traditional Information Visualisation: in Visual Analytics, retrieving and showing the data is not enough. Instead, the data has to be analysed, displayed showing its most relevant features, and the system must provide interaction to extract details from the data.

2.1.2 Multiple coordinated views

Visual Analytics systems rely on interaction techniques that allow the user to manipulate the data visualisations directly. In most cases, utilising only one chart for displaying the data and/or interacting with the system is not sufficient: complex datasets are generally multidimensional, which means that the data possesses a great number of dimensions (or "attributes"), thus making it hard to represent it in a single 2D chart. **Multiple coordinated views** is a paradigm that aims at solving this issue by utilising two or more distinct visualisations, synchronised among each other, to display simultaneously in different ways the same data or distinct sections of it. As described by Baldonado et al. [30], this approach is particularly useful when there is a great diversity among the attributes that a single chart could not support, and when one view alone is not enough to bring out correlations or "hidden" relationships in the data.

Of course, the effectiveness of multiple coordinated views relies heavily on the coordination between visualisations; this is typically achieved through **brushing and linking**, originally introduced by Becker and Cleveland [7]:

- brushing refers to a form of interaction with the system where the user selects a subset of the data, thus highlighting it, typically by dragging a "box" over the data;
- linking ensures that the selection made via brushing is immediately propagated to all the other views. When a user highlights some data in one chart, the corresponding data points are automatically highlighted in all other active visualisations.

This strategy greatly helps the user when exploring complex datasets, allowing the search for correlations among views.

2.2 Existing approaches in data-driven political analysis

To understand the contribution of this thesis, whose objective is creating a Visual Analytics system for data-driven political analysis, let us have a look at some examples of how political data is currently visualised. It is possible to categorise the existing approaches into two main groups: academic analysis and election results charts.

2.2.1 Academic statistical studies

In the field of political science research, datasets like the Chapel Hill Expert Survey one are used as a source for many research works. We have already mentioned this dataset in the introduction, but we have not yet explained what it is; this will be done in detail in Chapter 3 from a "conceptual" point of view together with a presentation of the CHES academic project, and in Chapter 5 from a more practical coding perspective. For now, it is enough to know that the CHES dataset contains data across the years of European national parties and their opinions on a diverse array of political topics, and as mentioned it is often used as a source for political science research articles: the following list contains examples of such works.

- Whenever new data from the CHES project is published, its authors (Bakker et al.) also publish papers analysing the political situation in Europe and the evolving positioning of European national parties, with the most recent one being *The 2024 Chapel Hill Expert Survey on political party positioning in Europe: Twenty-five years of party positional data* [6]. These papers present interesting political findings supported by the data, for instance differences between Eastern and Western Europe parties, the fact that opposition to the European Union is concentrated in radical groups (whereas moderate parties are mostly pro-EU), and how Green parties have started off as mildly Eurosceptic but have become increasingly Europeanist during the years [3][4].
- Hobolt and De Vries [15] analysed how the 2008 financial crisis increased the support for Eurosceptic parties in the following years, and they used the CHES data to identify such parties and position them on the left-right axis.
- Carrieri, Conti and Morini [9] combined CHES data and voter surveys to show that party opinions on the EU affect citizens' voting intention.

Most studies like these utilise charts to display the data and possible correlations, but their authors have to manually construct each time such charts in order to show their findings, instead of using a "unified" and already available system. Furthermore, the researchers are the only real users of the data, since the readers can rely only on the visualisations available "on paper" and have no way to further explore the data on their own.
Aside from the examples presented in Chapter 3 (which will be described after a more in-depth explanation of the Chapel Hill Expert Survey project), a unified and highly interactive tool displaying information like the CHES data was not found; that is the gap that European Parties Explorer aims to fill.

2.2.2 Election results charts

When talking about political parties in general, one of the first things that often come to mind are electoral results. For this reason, visualisations about elections are numerous and widely used, and were among the first results when looking for reference material during the early stages of this thesis' project. A great example is *Politico*'s charts for the 2024 EU Parliament election results[1] (Figure 2.2); Politico's visualisation shows the distribution of seats in the European parliament, but it is highly interactive too. The user can compare the 2024 results with those from 2019, check which and how many seats belong to each country, and see the exact election results in single nations.

[1]https://www.politico.eu/europe-poll-of-polls/european-parliament-election/

These tools excel at showing electoral results, they are very popular and simple to use; however, they generally ignore the parties' positions and ideologies, at best just placing them on the left-right axis. Still, given the high quality of Politico's visualisation, it was a good reference and source of inspiration for European Parties Explorer, even though they have different objectives.

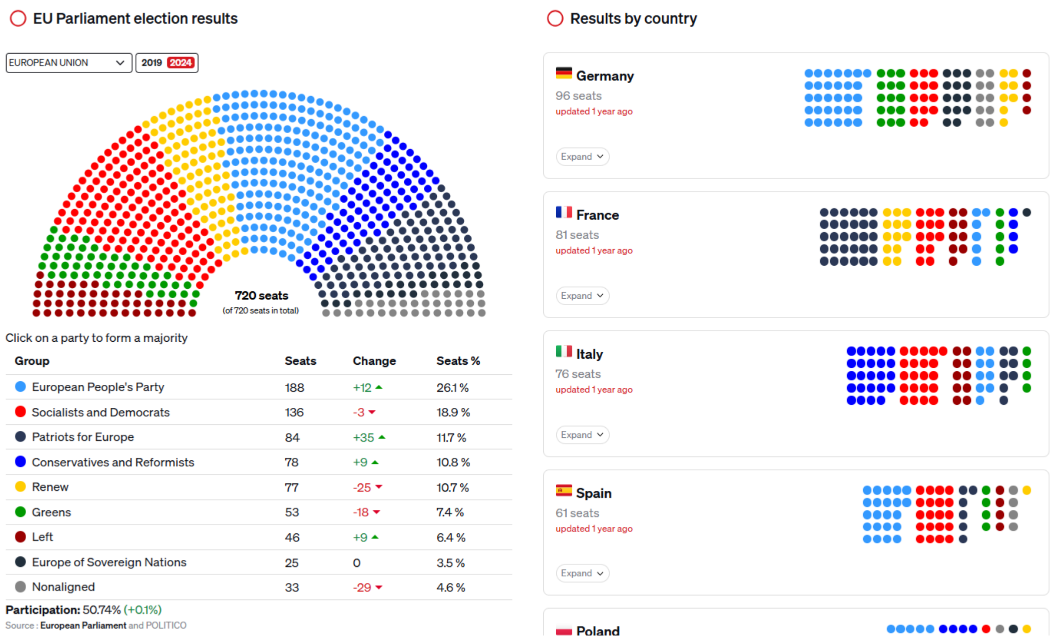


Figure 2.2. 2024 European parliament election results on Politico.eu

2.3 Applying Visual Analytics to political science

While Visual Analytics is generally used in computer science domains like cybersecurity or network management, European Parties Explorer is an attempt at creating a visual system to support data-driven political analysis starting specifically from the CHES dataset. This is not a too far-fetched task, as shown by the rise in recent years of **data journalism**. This discipline was defined by Veglis and Bratsas [29] as:

> the process of extracting useful information from data, writing articles based on the information, and embedding visualisations (interacting in some cases) in the articles that help readers understand the significance of the story or allow them to pinpoint data that relate to them.

Data journalism combines the traditional storytelling of journalism with data analysis and visualisations, recognising that what is happening beyond what the eye can see (i.e. the data) has a growing value [8]; it has a similar pipeline to the one of Visual Analytics, giving great importance to retrieving the data, understanding and cleaning it, and then visualising it with techniques borrowed from computer science. Just like data journalism, European Parties Explorer aims at translating a

complex dataset into an understandable and interactive tool for data exploration.
The probably best example of data-driven approach for analysing politics was the work of Nate Silver and his website **FiveThirtyEight** [14]. Originally an independent blog, later acquired by ABC News and then discontinued in early 2025 [13], FiveThirtyEight was notorious for its election forecasting models making impressively accurate predictions of United States elections, and for the use of original visualisations in its articles about politics, economics, science etc. According to Silver [24], aggregating lots of data is essential, thus putting the data at the centre of this approach; still, it is necessary to distinguish between the "signal" (the real information, allowing us to make analysis, find patterns and predictions) and the "noise" (random variability).

FiveThirtyEight popularised the use of Visual Analytics in the political discourse, and it demonstrated how politics can be visualised not only in the academic world, but also in the real world for the greater public. Even though nowadays most visualisations are finalised to predicting and analysing electoral results, European Parties Explorer shares many features with data journalism and Silver's work: moving away from static data to provide an interactive and visual exploration of the political landscape and communicate data-driven insights, possibly even to non-experts. Also, having explained earlier the concepts of Visual Analytics, the Chapel Hill Expert Survey data itself can be considered fit for an application of this field: not only does it possess great dimensionality, its attributes are also very diversified, justifying the need for multiple views; even when used in research papers, political scientists generally create many different charts to prove their findings. Furthermore, the dataset contains data spanning over many years, so a form of temporal analysis should be supported too.

Chapter 3

Dataset

To fully comprehend the objectives of this project, how it supports them and how the final application was implemented, we must know and understand the dataset it was built on; after all, data is the foundation of any Visual Analytics system, as we could observe in the previous chapter. This section will deal only with explaining the data itself at a high level and what parts of it are of our interest, whereas the practical part of how the dataset was processed and in general all the operations that were made on it will be described in Section 5.1.

3.1 The Chapel Hill Expert Survey

As anticipated, European Parties Explorer is based on the dataset provided by the Chapel Hill Expert Survey (in short, CHES).
CHES is an academic project created by political science researchers, professors and experts, which aims at evaluating ideologies of national parties from many areas around the world. As stated on the CHES website [10], the mission of this project is to "estimate party positioning on ideology and policy issues, and international relations for national parties in countries across the world", and it has received funding from the University of North Carolina at Chapel Hill and both American and European research programmes. The biggest project and main focus of the Chapel Hill Expert Survey is CHES-Europe[1], monitoring and evaluating national parties from European countries during many years spanning from 1999 to 2024, but in recent years CHES members have started collecting data on political parties of Latin America[2], Canada[3] and Israel[4] too.
Since this thesis' work focuses on European parties, from now on we will only deal with CHES-Europe, whose authors' information (Ryan Bakker, Liesbet Hooghe, Seth Jolly, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen, Milada Anna Vachudova) can of course be found on the appropriate section of the CHES website[5].

[1] https://www.chesdata.eu/ches-europe
[2] https://www.chesdata.eu/chesla
[3] https://www.chesdata.eu/chescanada
[4] https://www.chesdata.eu/chesisrael
[5] https://www.chesdata.eu/our-team

3.1.1 Methodology and evolution

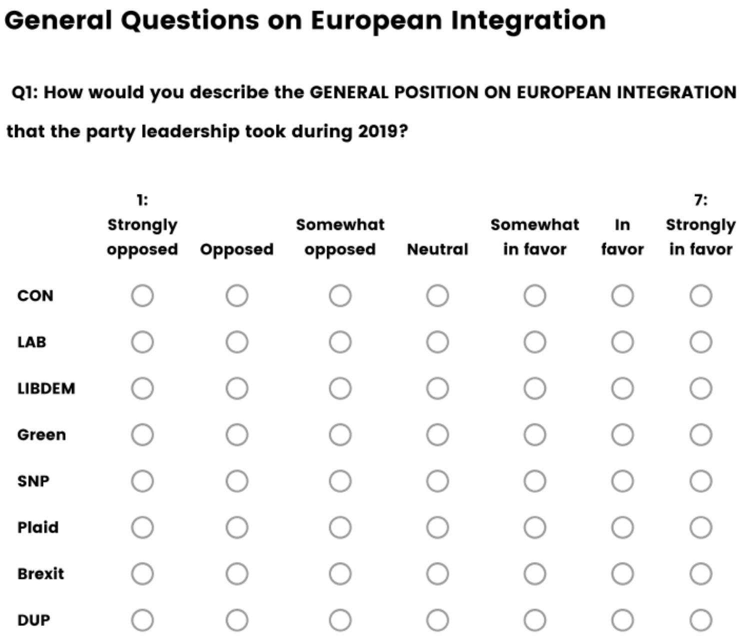
Let us now better explain how the Chapel Hill Expert Survey works and gathers data. As the name may suggest, the core methodology involves distributing surveys to a carefully selected pool of political scientists and experts specialising in the politics of European countries. Many surveys were created during the years, with questions asking the experts to evaluate national parties from European nations on various topics, using numbered scales. The CHES-Europe website explains briefly in what years the questionnaires were carried on, from which nations the political parties were evaluated, and what kind of questions were asked:

> the first survey was conducted in 1999, with subsequent waves in 2002, 2006, 2010, 2014, 2019, 2024. The number of countries increased from 14 Western European countries in 1999 to 24 current or prospective EU members in 2006 to 31 countries in 2024. In this time, the number of national parties grew from 143 to 279. [...] Questions on parties' general position on European integration, several EU policies, general left/right, economic left/right, and social left/right are common to all surveys. More recent surveys also contain questions on non-EU policy issues, such as anti-elite rhetoric, immigration, redistribution, decentralization, and environmental policy.

The questions of these surveys were then answered by the many political experts, who estimated the political parties' views on different topics; more specifically, according to the so-called **Codebook** [11], a document available on the CHES-Europe website, acting as a guide for the dataset:

- in the 1999 survey, 116 experts estimated the positioning of 143 parties;
- in the 2002 survey, 250 experts estimated the positioning of 171 parties;
- in the 2006 survey, 235 experts estimated the positioning of 227 parties;
- in the 2010 survey, 343 experts estimated the positioning of 237 parties;
- in the 2014 survey, 337 experts estimated the positioning of 268 parties;
- in the 2019 survey, 421 experts estimated the positioning of 277 parties;
- in the 2024 survey, 609 experts estimated the positioning of 279 parties.

To be more clear and illustrate a part of the data collection process, two questions taken from the 2019 survey for UK parties, available as a sample questionnaire again on the CHES-Europe website, are presented as examples in Figures 3.1 and 3.2.
For each question in the surveys, the experts were asked to assign a certain score to every party they were able to evaluate, and the endpoints of the numerical scales represent opposing ideological stances (e.g. "strongly opposed" vs. "strongly in favor"). The final value assigned to a party for a specific question is the mean of expert responses assigned to that party, and these aggregated values were stored in the so-called **1999-2024 CHES trend file**, a file containing the data obtained from all surveys. This file was then used as a basis for *The 2024 Chapel Hill*

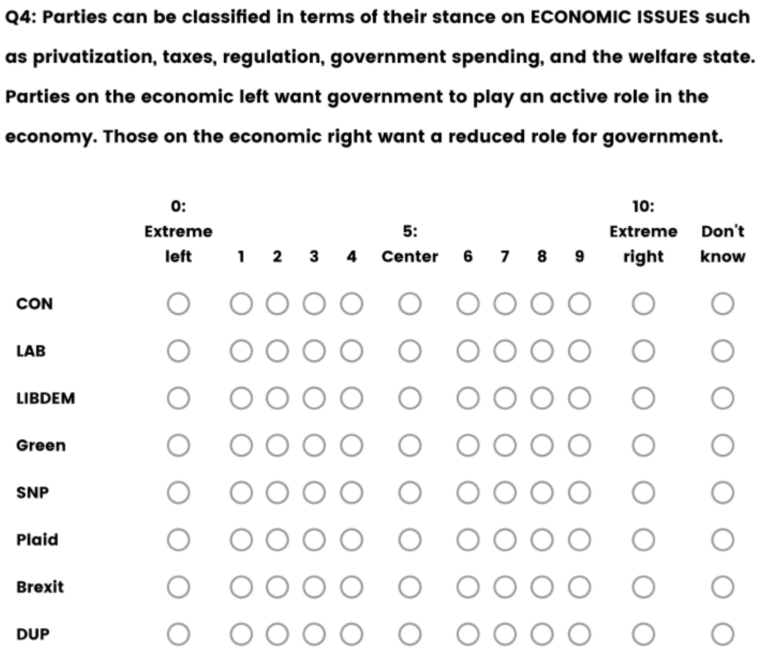


General Questions on European Integration

Q1: How would you describe the GENERAL POSITION ON EUROPEAN INTEGRATION that the party leadership took during 2019?

	1: Strongly opposed	Opposed	Somewhat opposed	Neutral	Somewhat in favor	In favor	7: Strongly in favor
CON							
LAB							
LIBDEM							
Green							
SNP							
Plaid							
Brexit							
DUP							

Figure 3.1. Example of question on European integration from the 2019 expert survey for UK parties. The scale ranges from 1 (strongly opposed) to 7 (strongly in favour). Source: CHES-Europe website.



Q4: Parties can be classified in terms of their stance on ECONOMIC ISSUES such as privatization, taxes, regulation, government spending, and the welfare state. Parties on the economic left want government to play an active role in the economy. Those on the economic right want a reduced role for government.

	0: Extreme left	1	2	3	4	5: Center	6	7	8	9	10: Extreme right	Don't know
CON												
LAB												
LIBDEM												
Green												
SNP												
Plaid												
Brexit												
DUP												

Figure 3.2. Example of question on economic issues from the 2019 expert survey for UK parties. The experts were asked to position parties on a left-right spectrum (using a scale from 0 to 10) regarding issues like spending and taxes. Source: CHES-Europe website.

Expert Survey on political party positioning in Europe: Twenty-five years of party positional data [6], the latest article from the same authors of the CHES-Europe project studying the evolution of European parties, already mentioned in Chapter 2. The CHES trend file is available on the CHES-Europe website as both Stata and csv file, and was of course used as dataset for this thesis' application.

3.1.2 Relevant topics

The various surveys were not structured in the same way each year; in fact, when considered singularly, they have many differences one another. While the CHES' core mission and philosophy has remained the same, with some questions that were asked in all or most years, there are also some subjects that were considered only in, for instance, one or two years, depending on the political landscape of the time. The earlier surveys were also pretty simple and short, thus containing just a small number of questions and gathering less information and data.
In order to make the political analysis more significant, especially in the case of analysis across different years, it was decided to take into consideration for this thesis' project only data regarding the most important political topics that were present in a good amount of consecutive expert surveys (at least three surveys), while also filtering out subjects that are too redundant. Furthermore, the 1999-2024 trend file contains a great number of dimensions (90 columns in the csv file), therefore reducing the amount of used data also makes the final application more manageable for the user.

In short, what we will be analysing is:

- the vote percentage received by the parties in the national election and European election most prior to the given survey year;
- the parties' political families;
- the parties' ideological stances (left or right-wing);
- the parties' opinions on the following subjects:
  - European Union;
  - economic issues (e.g. deregulation of markets, redistribution of wealth from the rich to the poor);
  - "social" issues, like LGBT rights and gender equality;
  - the role of religious principles in politics;
  - immigration policies and ethnic minorities;
  - nationalism;
  - environmental sustainability;
  - regionalism.
- the parties' "metadata" (basic information like name and country), essential for the application's functioning.

3.2 Reliability of the Chapel Hill Expert Survey

Since this thesis project relies on expert judgements rather than objective measurements, one might rightfully question the reliability of a dataset based on "opinions"; however, the Chapel Hill Expert Survey's reliability has been tested and confirmed many times in political science literature.

3.2.1 Comparison with other sources

A possible method to test the accuracy of expert surveys is to compare them with other existing ways to evaluate party positionings, for instance academic projects analysing party manifestos or aggregated public opinions. Steenbergen and Marks [21] established the framework for this evaluation, demonstrating that expert surveys possess "high *convergent validity* with other measures of party positions": this means that the CHES data strongly correlates with these external sources, confirming that the experts provide evaluations of political phenomena in the same way that other measurement instruments do (as mentioned, analysis of official party documents and voter opinions). More recently, in their analysis of the 1999-2010 CHES trend file, Bakker et al. [4] found strong consistency between positions deriving from the CHES data and those coming from mass surveys, confirming that expert evaluations are a reliable measurement for party ideologies.

3.2.2 Inter-expert reliability

A potential criticism of expert surveys is the subjectivity of respondents. However, Hooghe et al. [5] showed that the cohesion of expert scores in the CHES data is pretty high ("*inter-expert reliability*"): by analysing the standard deviation of expert placements, they found high consensus among experts for most parties. Furthermore, the authors observed that where there are higher standard deviations, they generally reflect actual internal dissent within a party, rather than measurement error (it is not "noise", but actual "signal").

3.2.3 Advantages over other methods

Lastly, expert surveys can offer some advantages over methods like the already mentioned text analysis of election manifestos. As stated again by Steenbergen and Marks [21], manifestos are static documents published only once per election cycle and may be intentionally vague on controversial topics, to avoid the loss of voters. Political experts, instead, can effectively evaluate a party's ideology on the basis of many different parameters, like media rhetoric, and day-to-day political actions. This makes CHES suited for estimating party positions on complex and always evolving issues, especially when a party's actual behaviour might differ from what is stated on its official program.

3.3 Related work: tools for the CHES dataset

As hinted in Chapter 2, there exist a few visual tools for using the CHES data; let us describe them here, now that we have a better understanding of the dataset.

3.3.1 CHES Interactive

The Chapel Hill Expert Survey website provides an interactive section named *CHES Interactive*[6], showing three simple charts:

- a bar chart (Figure 3.3) where the user can choose a nation, a year and one topic, and see how that country's parties, ordered by left-right ideology, are evaluated on the selected topic thanks to vertical bars; hovering a party gives its name and score on the selected attribute;
- a scatter plot on one country (Figure 3.4), where the user can again choose country and year but also both the axes' variables; the parties are placed accordingly (bringing out possible correlations among the attributes), they are colored according to their political family, their size reflects the vote share, and hovering gives some information;
- a scatter plot on all countries (Figure 3.5) that works exactly like the above one, except that of course it displays all nations simultaneously.

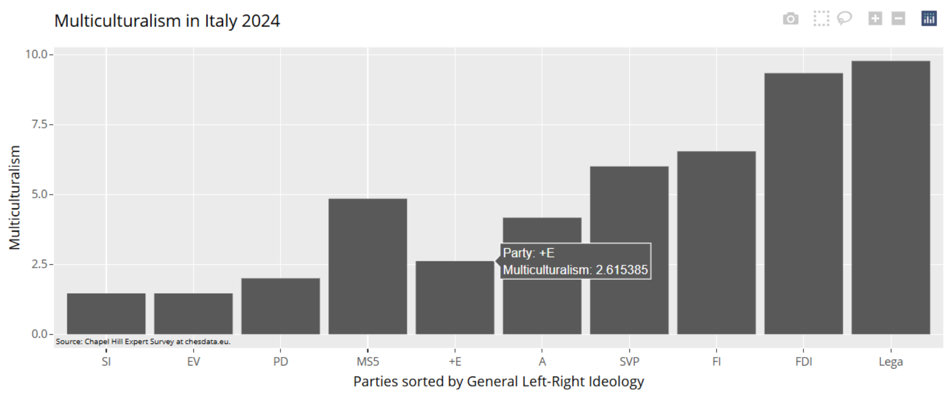


Figure 3.3. CHES Interactive's bar chart, displaying Italian parties in 2024 (ordered by left-right ideology) and their positioning with respect to multiculturalism (a tall bar means the party is against multiculturalism, favoring assimilation).

In all three charts there is also a simple zooming mechanism.
CHES Interactive is an excellent tool for quickly comparing political parties, especially within a single country, and for showing trends and correlations among different dimensions, but there are inevitably some limitations:

- there are only two filters on the parties (year and country);
- the comparison between parties is pretty effective, but it is possible only with one or two topics at a time;

[6] https://chesdata.shinyapps.io/Shiny-CHES/

- making comparisons across the years is not too well supported, since the user needs to continuously change the year and wait for the loading of data;
- parties are placed on the scatter plots using two attributes, but we will see in Section 5.1.3 how to implement a more "universal" positioning with *dimensionality reduction*, showing similarities among parties on a general basis;
- there are no explanations about the many selectable topics and their minimum and maximum values, but this is understandable since we are using this system on the CHES website itself.

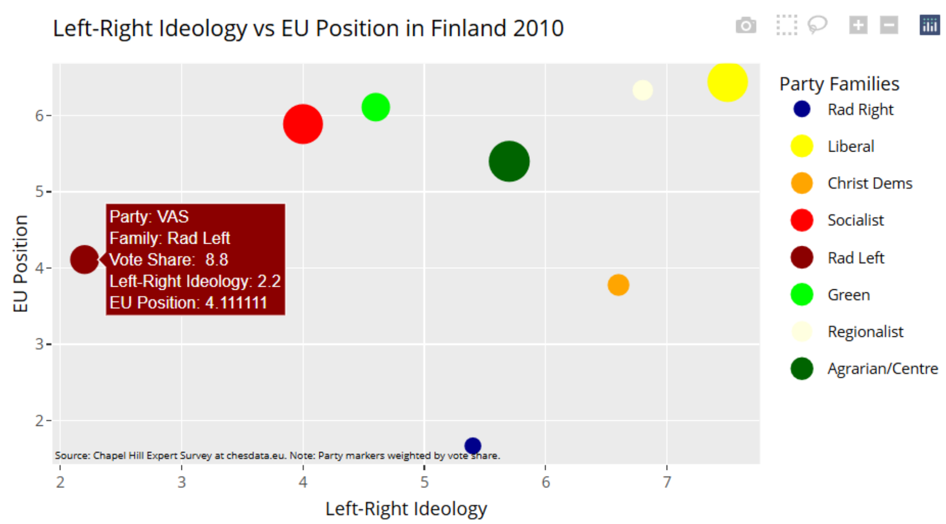


Figure 3.4. CHES Interactive's one country scatter plot, showing Finnish parties in 2010 placed according to their left-right ideology on the x axis and their position regarding the EU on the y axis (a greater value means "Europeanist").

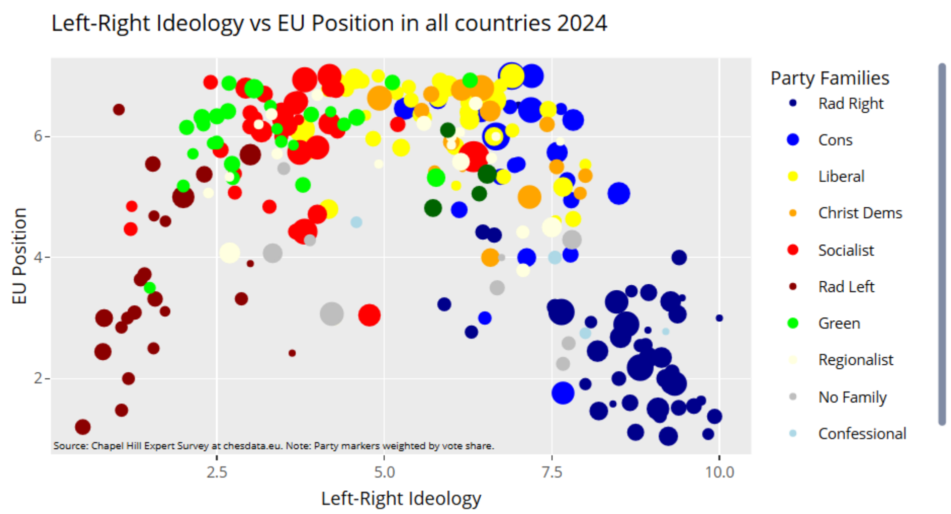


Figure 3.5. CHES Interactive's scatter plot for all countries, displaying parties of all nations tracked in 2024; the axes are the same of the one country scatter plot.

3.3.2 Data exercises from Foundations of European Politics

Another visual tool can be found on the website of *Foundations of European Politics - A Comparative Approach*[7], a textbook for students about research on European politics, by De Vries et al. [12]
The website presents some "data exercises" to teach political concepts utilizing political datasets, including one exercise[8] using of course the CHES data. Similarly to CHES Interactive, this section presents a scatter plot that positions parties according to two selected variables (Figure 3.6) and colors them depending on their faction. Interaction with this chart is very limited, there are no filters, the user can choose only a small number of topics and there is no support for open-ended exploration of data, however this is understandable since this tool was designed as a guided tutorial with educational purposes, guiding students to discover predetermined relationships. Nonetheless, this system introduces a very interesting feature, that is a line explicitly highlighting the potential correlation between the two variables.

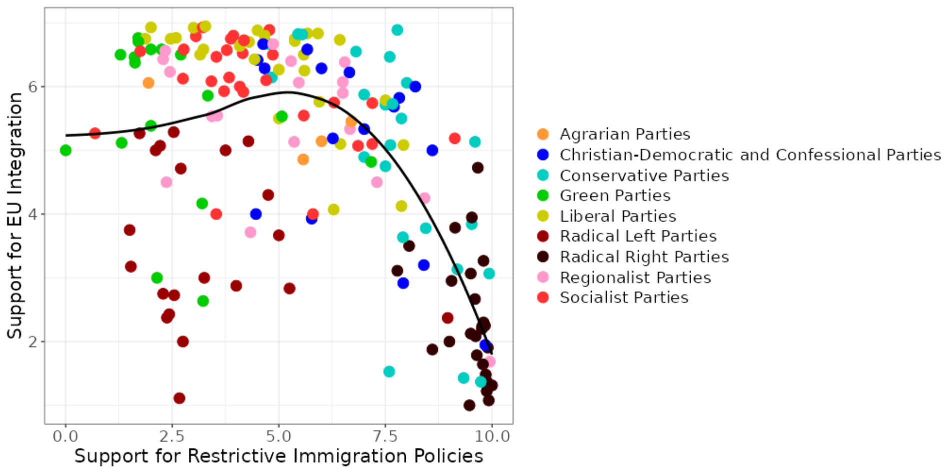


Figure 3.6. The scatter plot tied to Foundations of European Politics' exercise on the CHES data, showing the correlation between support for restrictive immigration policies and support for the EU.

European Parties Explorer will aim to support these tools' features, such as year and country selection, comparison among parties of one country, inspection of single parties, finding relationships between topics, and clear display of parties and distinction of families. However, it will also try to overcome their limitations, for instance with better support to comparison between years, the introduction of more filters thanks to the brushing technique, the analysis of multiple attributes at the same time, clear topic explanations, and allowing the political analysis to start not only from a single country, but also from a faction or a specific topic.

[7]https://www.foundationsofeuropeanpolitics.com/
[8]https://www.foundationsofeuropeanpolitics.com/project/chapter8_shiny/

Chapter 4

User analysis and system requirements

We have learnt about the Chapel Hill Expert Survey and how it gathers data about European parties. We also had a brief introduction on what topics of interest may arise from the CHES dataset; these subjects could be used to accomplish some form of political analysis on the various parties at our disposal or groups of parties.
However, data alone is of little value without a clear understanding of who will use it and for what purpose. In the field of Visual Analytics and Human-Computer Interaction, a user-centred design approach is essential to create an effective system. Therefore, let us now take the initial steps in designing this thesis' project by first reasoning about its users and their objectives.

4.1 Possible users

As anticipated, a good starting point when creating an application is of course trying to imagine who could be interested in utilising it. This is particularly relevant in the case of European Parties Explorer, because depending on who its potential users are we can better hypothesise what types of analysis can be carried out on the basis of the CHES dataset and how to support them; also, different kinds of users may possess different levels of domain knowledge and technical skills, therefore a "casual" user might want a more or less quick answer to a simple question, while a domain expert may need to perform a complex and open-ended analysis.
Let us then list the possible categories of users and what they might want to accomplish with a system built upon the Chapel Hill Expert Survey data.

4.1.1 Voters

The first category represents the casual or novice users: voters. Given that the CHES data is basically a collection of "scores" representing in general how much a certain party agrees on a specific issue, a first task, for now conceptually simple, could be "studying" the parties of one's own nation.
In the current information landscape, voters often have to deal with polarised debates and lots of rhetoric, and it may not be easy to move beyond political slogans.

We should then remember that the objective of the surveys' expert responses is to estimate the positioning of political parties on many subjects, and the scores assigned by the experts are well suited for understanding the actual opinions of a party, as demonstrated in Section 3.2. All of this brings us to the first possible category of users, which is citizens that are about to vote; indeed, the CHES data could be useful to identify parties that are closest to one's own ideologies.
In short, European Parties Explorer could be utilised by voters to:

- understand the ideas of parties running for elections and what political families they belong to;
- find which party or parties are more in line with their ideas on some or all available political topics; as an example, a voter prioritising measures against climate change might want to filter parties to find those with the highest scores on the **Environment protection** dimension within their country;
- check how coherent parties have been across the years with respect to their ideas, or in general understand if there has been a change of direction in a given party's leadership and ideology; for instance, a potential voter could be wondering if a certain party has always been Eurosceptic or if this was a recent populist turn. All of this is made possible by comparing the expert evaluations among different survey years.

4.1.2 Journalists

The potential analysis carried out by voters, as we just saw, is mainly a study made in a single year for a single country; however, it is also possible to achieve a type of analysis focused on changes over many years. Furthermore, we could also think about analysing different groupings of parties, rather than only parties considered individually.
Other possible users could therefore be journalists and data reporters, who often require a broader perspective, especially considering the case of data journalism with its articles aiming at extracting useful information from data. Their goal is not personal decision-making, but rather analysing, communicating and explaining; these users need to extract facts, verify trends and find evidence to support their articles.
In short, journalists may want to use the application to:

- write articles about parties based on historical data, using visualisations to tell how they have changed during the years;
- report on the political situation in one or many nations, observing for example how divided their parties are on a subject or which ones share similar ideologies, or identifying if a national trend is an isolated case or a phenomenon shared among countries;
- verify claims made by politicians regarding political stances, comparing them with the objective and aggregated consensus of the academic experts.

4.1.3 Political scientists

The last class of possible users is of course the "domain experts": researchers, professors or even students in political science, specifically the ones conducting studies on European politics.
In this case, the users may possibly already be familiar with the CHES dataset, but they need an efficient way to navigate it and extract information and insights. Consequently, the application should support the most complex types of analysis, and when designing the system we should not only understand how to effectively support these kinds of analysis, but also try to figure out questions to which these expert users may want to find the answer, or what insights they are trying to discover.
Here is then an example list of potential tasks political scientists might be interested in:

- analysing ideological evolutions over time of national parties, countries, political families or other kinds of party groupings;
- finding "trends" and defining characteristics of political groups;
- observing the composition of the European political landscape, finding which groups are the biggest ones whether in all of Europe or just a subset of countries;
- searching for "outlier" parties that somehow deviate from their own faction or nation;
- finding similar parties from different countries;
- focusing on a specific topic and studying how opinions on it have changed during the years;
- checking how polarised countries or factions are on a chosen subject, or even arbitrary subsets of parties;
- trying to figure out if opinions on a topic may have influenced the outcome of an election.

4.2 Task complexity and the need for Visual Analytics

At a first glance, many of the user tasks listed in the above sections might appear simple, at least on a conceptual level. Questions like "How has this party's stance on the European Union changed?", or "Which parties are more restrictive on immigration?" are clearly defined and easy to understand.
The problem is that however, a simple question does not necessarily imply a likewise trivial way to find the answer, or a simple answer at all. This is true especially if we consider a dataset as rich as the Chapel Hill Expert Survey. The difficulty of the user tasks derives from the nature of the data itself and its intrinsic complexity, and the cognitive load that may be required to make a political analysis starting from this amount of data. Therefore, we can now try to explicitly identify what

factors are contributing to this complexity, which derives mainly from the analysis being multidimensional, the temporal factor, and making comparisons within a high quantity of data.

4.2.1 High dimensionality

As explained in Chapter 3, the CHES dataset provides a collection of parties and political ideologies. A great amount of data is stored for every single party, which implies that the parties are defined by a multidimensional vector of attributes storing opinions on the European Union, economic stances, policy issues like environmentalism or nationalism, and specific party data like its name and political family; in Section 3.1.2 we even noticed how there was the need to reduce the amount of dimensions used.
In order to perform some sort of analysis, the user is required to observe these multiple dimensions simultaneously and make deductions based on them. As an example, determining if one selected party is an outlier with respect to its faction means comparing all its dimensions (its scores) with the faction's trends on all those same topics. It is easy to understand how this is a challenging task, because we would be asked to compare many numerical attributes with the ones of many other parties; high dimensionality is also why Visual Analytics needs multiple coordinated views in the first place, and in Chapter 2 we noticed how even researchers use many charts to show their findings, even though they may be conceptually simple. Without a visual aid, reasoning effectively on a high-dimensional dataset and finding complex patterns is very difficult, if not impossible.

4.2.2 Temporal factor

Not only does the CHES dataset store lots of information about parties coming from many countries; there are also different years to take into consideration (seven years, as we saw: 1999, 2002, 2006, 2010, 2014, 2019 and 2024), with many dimensions repeated in the various years and most parties appearing in more than one year only.
Time adds a new layer of complexity in the data, but it is fundamental for understanding the evolution of parties and trends across the years. In the case of a user interested in the history of a party, they are not looking for a single piece of data, but rather for many arrays and comparing them. In some cases, the different years don't even store the same party attributes, and there could also be parties in the dataset appearing in multiple years that are not contiguous. A party present in 1999 may not exist in 2024, or it may have changed its name or political family; we also definitely need to handle outright missing data in a given year. A Visual Analytics application is of great help when tracking temporal trends and variations in the expert evaluations.

4.2.3 The scale of data

Finally, the scale itself of the dataset adds an additional layer of complexity that cannot be ignored: after all, we have observed that representing great amounts of data in an understandable and explorable way is the objective of Visual Analytics itself. The tracked national parties come from more than twenty countries, and there are seven different years to consider. Finding insights in data of this magnitude is a very hard task, and without an application providing interactive, coordinated visualisations and filters there is no way of making this data useful; even if we need to study only a subset of parties, a visual system would be very helpful thanks to its ability to make selections on the data.

In short, while the user goals might be conceptually simple, extracting insights from the raw CHES dataset is certainly not a trivial task. Objectives like comparing groups of parties, finding trends and studying opinions on a subject are reasonably easy to understand, but accomplishing them is a complex task that justifies the need of a Visual Analytics application complete with data preprocessing, coordinated views, filters and high interactivity like this thesis' application.

4.3 Tasks and analysis patterns

While presenting the types of users, we talked about their tasks (which we have stated need visual tools to be accomplished) at a fairly high level of abstraction, meaning that we mentioned them only as potential objectives the political experts (or the other categories of users) may want to achieve in their political analysis; however, we did not say anything about how to accomplish such goals and the possible intermediate passages that are required. Therefore, in the following sections we will go deeper in the user tasks, attempting to translate these abstract goals into a number of more concrete **analysis patterns**: these patterns can be understood as structured workflows, or a number of "procedures", that the users will be able to follow in the final application to discover insights, of course depending on the type of analysis that is being made too. By explicitly listing these analysis patterns, we will be able to have a better understanding of the users' needs and what the Visual Analytics system must do, ensuring that this project allows the user to accomplish their tasks and supports their whole reasoning process.

4.3.1 Basic functionalities: visualising party information

Before dealing with the "actual" analysis patterns, let us discuss a preliminary and basic form of analysis first.
The simplest, yet most fundamental kind of analysis is the one where the user just wants to gather information on one single party or even several parties but considered individually, one independent to each other. In this case, we are mainly talking about basic functionalities, essentially atomic actions that should be available for the user, who just wants to "read" what parties are in the dataset without the need to find particular insights; from a Visual Analytics point of view, it is mostly a matter of displaying information, especially in a way that is immediately readable and contextualised (for instance giving reference frames for numerical scales). The cognitive effort required for this first type of analysis is relatively low, as it does not involve comparisons between many elements, and also it is understood that these

elementary features will be the foundation of the more complex classes of analysis. The Visual Analytics system should therefore support the following operations:

- give the ability to select one single party or a subset or parties, highlighting them and if needed dimming the other elements; of course, highlighting one or more elements on one visualisation should make the same elements evident in the other visualisations, in a coordinated fashion;
- show basic information like the name of the party, its country and its political faction;
- use visualisations to encode the expert evaluations of the CHES data, for example using axes to map the used scales;
- display how the expert evaluations of a specific party have changed over time, connecting data points across the many available years;
- display electoral results, allowing the user to distinguish between fringe parties and more mainstream ones and to visualise the possible electoral successes and failures.

As mentioned, these are just basic operations that it is reasonable to expect from a project like European Parties Explorer, and they are going to be useful for every kind of analysis imaginable.

4.3.2 Single parties analysis

Let us now put the elementary features just described to use, and combine them in order to construct more complex analysis patterns.
There are many models of analysis achievable with the CHES dataset, but for the sake of clarity they can essentially be grouped into two macro-cases: the case in which the user wants to analyse parties taken singularly, and the case in which the focus is not on the single piece of data but rather in groupings of parties (for instance when searching for general trends). Let us then observe what can be achieved with an analysis of the former case.

Comparing and filtering parties

The single parties analysis is a type of study particularly well-suited for the voter user persona, who as we discussed might want to understand the ideologies of their own country's parties.
In this context, the user is not satisfied with just obtaining basic information about individual parties; instead, the core task becomes comparison. The user needs to compare parties one against the other to find differences, and filter out a number of them to find the ones they might be of their liking (in the best case, finding the "ideal" party). Let us then describe a series of steps that should be achievable in the Visual Analytics system and which could be carried out by citizens about to vote:

- filter the dataset to select parties from the user's nation (or of course, in general any nation);

- understand the parties' stances on the available subjects (through the scores from the expert evaluations);
- understand which parties are ideologically similar;
- use filters so that the user can find parties closest to their ideology, choosing ideal scores for topics selected by the user;
- check how the parties' ideas have changed over time.

4.3.3 Aggregate analysis

In order to discover broader and more complex insights, it is likely necessary to deal with more sophisticated types of analysis. This leads us to what we can call "aggregate" analysis, where the unit of analysis is no longer a single row in the dataset, but rather a group of elements treated mostly as a single entity.
In this case, the analysis patterns do not concern single parties, rather they mostly deal with generic groups of data. For instance, a political scientist may want to discover outliers, trends or how much a faction is polarised on a subject; these tasks all require to study the behaviour of entire groups of parties, and then in case analyse the individual parties as added knowledge.
We can identify three main approaches to the aggregated type of analysis: by faction, by country, and by topic.

Analysis of factions

As the name suggests, the analysis of factions is an analysis pattern in which the user wants to find insights regarding one or more political families, like temporal trends or their opinions on a number of subjects.
The steps we can imagine when carrying out such analysis are the following, not necessarily in the presented order:

- select a political faction to isolate it from the rest of the dataset;
- find out how big this family is and what parties it contains;
- see the geographical distribution of the faction, meaning which countries have no parties belonging to it or where they are in greater numbers;
- observe its general electoral trend over the years, determining if its parties are rising or declining in popularity across Europe;
- analyse its ideas thanks to the expert evaluations, how divided or cohesive it is on certain or all topics (e.g. a faction could be united on economics but split on social rights), or maybe find out that the chosen family presents no trends on the available subjects;
- make a historical analysis, studying the temporal evolution of its opinions (for instance, a family moving more to the right or becoming more pro-EU);
- find out if there is an identity topic that the political group cares a lot about;

- look for “outlier” parties with respect to the whole family, meaning that they formally belong to it but have some significantly different views;
- easily make comparisons with other factions to understand the ideological differences.

Analysis of countries

As in the case of political groups, it should be possible to also gather insights about one or more selected countries. The steps to follow in this form of analysis are not too dissimilar from the ones in the analysis by factions, but of course the focus is on a (very) different grouping of parties, especially considering that the parties within a country are in direct competition:

- select a country;
- check what parties are in it and how many;
- see what political families are represented in the chosen nation and which ones are the biggest;
- study its electoral history, observing which parties rose or fell and possible political shifts for instance to the left or to the right;
- analyse the opinions of the parties in the country and how polarised it may be in one year;
- carry out a historical analysis on the many topics available in the application, extracting the country’s opinions over the years;
- compare with the political landscape of other nations.

Analysis of topics

We have seen the analysis by families and by countries; in both cases, what we do is focus on a subset of parties, find insights about it, and maybe compare it to other chosen subsets. However, the CHES data can also be used to obtain knowledge about a specific attribute: this is the case of the analysis by topic, where we concentrate on one topic and find out what all parties think about it, and only later we can also decide to observe the behaviour of a selection of parties.
If we think more in terms of the raw dataset file, where each row represents one party, and the party data and the expert evaluations are stored in the columns, then reasoning by faction and by country in practice means “cutting” the data vertically, leading to a form of analysis based on the data of a subset of rows. Reasoning on one or more topics implies instead cutting the dataset horizontally, paying attention to one or some columns specifically, and then drawing conclusions by comparing the data on the selected dimensions.
Going back specifically to the support for the analysis of topics, here is the list of actions the user may want to undertake in the Visual Analytics system:

- select one topic, using it as starting point in the reasoning process;

- observe how parties are distributed on the selected attribute, for instance most of them may be in favour, or maybe parties are very polarised;
- study its temporal evolution, determining the general consensus over the course of many years;
- find out if it has some form of correlation with other dimensions (e.g. negative correlation);
- focus only on some values of the expert evaluations (for instance low scores) to discover the behaviour of the parties associated to those values and what factions are more in line with such scores;
- intersect this analysis pattern with the other patterns, thus discovering the opinions of specific nations or factions on the selected topic, whether during one single year or across the decades.

4.4 Functional requirements

We have defined the above analysis patterns (basic functionalities, single parties analysis, and three types of aggregate analysis) in order to make the user tasks more explicit and to understand the operations the users need to carry out for such tasks, independently of the actual system implementation; making sure that the established analysis patterns are well supported implies that the users themselves are well supported and that the system is indeed tailored for them. Therefore, it was deemed appropriate to make an additional step before designing the system and define some high level **functional requirements** for European Parties Explorer, because respecting them means that the final system effectively supports the analysis patterns (the tasks), and thus the actual users. The functional requirements will then also guide us during the design process of this thesis' application.
When observing the analysis patterns, we can notice that some features are in common: for instance, all patterns mention in one way or another comparing multiple attributes, or the possibility to switch to an analysis over time. Thus here is the list of functional attributes, based on these features, that if respected by the system guarantee the support of the user tasks.

- **FR1 - High-dimensional analysis**: the system must allow the user to observe, correlate and filter a large number of attributes simultaneously.
- **FR2 - Temporal analysis**: the system must allow the user to see detailed information for each year, enabling both the inspection of single years ("snapshots") and temporal trends.
- **FR3 - Multi-granularity comparisons**: the system must support different levels of granularity (e.g. single party vs. whole faction), allowing the definition and comparison of arbitrary subsets of data.
- **FR4 - Coordinated exploration**: every time data is manipulated, all views immediately reflect this change (multiple coordinated views).

Now, we can finally start talking about the actual design choices of European Parties Explorer under the guide of the defined functional requirements.

Chapter 5

System implementation

We have studied everything we need to know to finally talk about the effective implementation and coding of European Parties Explorer: we have learnt how the Chapel Hill Expert Survey data is gathered, the application's possible users with the tasks they can accomplish with a political analysis based on the CHES dataset, and we have explained some analysis patterns to describe sets of actions the users can follow to carry out the aforementioned tasks, which helped us define a set of functional requirements.
Let us then proceed with an accurate description of the system's design choices and implementation, starting from the first fundamental step in the Visual Analytics cycle: the operations on the data, which were implemented with Python. European Parties Explorer was instead coded in JavaScript as a Node.js application, using the D3.js library for the visualisations.

5.1 Data preprocessing

In Chapter 3 we already talked about the main building block of European Parties Explorer, which is the Chapel Hill Expert Survey data. However, we presented it at a mostly abstract level, explaining only what the CHES project is, how it gathers data (that is by storing the average for each party of the responses given by the political experts when answering the questionnaires), the main topics that can be extracted from it. Let us therefore explain the real structure of the dataset file and the many operations made on it.

5.1.1 Dataset file structure

As shown in the chapter about the dataset, the data for this project was taken from the **1999-2024 Chapel Hill Expert Survey trend file**, which can be found on the CHES-Europe[1] website. This file is a unique collection of the data gathered with the many expert surveys conducted, as a reminder, in 1999, 2002, 2006, 2010, 2014, 2019 and 2024, and is available in csv and Stata format.
For this project, the csv version of the trend file was used. This means of course that the data is stored as a table:

[1]https://www.chesdata.eu/ches-europe

- the trend file contains 1441 rows, one for every tracked party in each survey year, storing party data and expert evaluations for that party in that survey year;
- there are 90 columns, most of them storing the expert evaluation scores of the parties in the rows, but there are also columns containing essential information such as the survey year and the parties' countries.

In practice, what we have is a multidimensional dataset where each piece of data (each party) is constituted by a large amount of attributes (the columns).

The information about the CHES data is either extracted directly from the dataset itself, or taken from the Codebook [11], which we remind is available on the CHES-Europe site and acts as description for the 1999-2024 CHES dataset. This document provides especially a thorough description of the data attributes: most of the columns are associated to expert evaluations, and the Codebook explains the questions the evaluations were answering to and the meaning of the scales used (e.g. 0 = strongly opposed, 10 = strongly in favour).

5.1.2 Extracting and cleaning the data

As we observed in Chapter 2, in Visual Analytics applications it is necessary to first of all carry out operations on the dataset used. Not only do we need to deal with missing data, there may also be errors in the data or just data that for some reason we are not interested in using. We will see that actually, we are going to have the need to compute new, additional data too.

Columns filtering

As we already explained in Section 3.1.2, we are not going to use all the available attributes of the dataset; let us see more in detail the reasons for this decision.

- The expert surveys are all structured in different ways: in each survey year, there was always a different set of questions. There are some questions that were asked in all surveys (for instance the parties' opinion on the EU), but most questions were asked only in a small number of years from 2006 onward. There are even some questions asked only in one year (e.g. the position towards US power in world affairs, where 0 = the party strongly opposes strong US leadership in world affairs, 10 = the party favours strong US leadership in world affairs, was asked only in 2006), and questions asked in many consecutive years but that were "skipped" during one year in the middle. These differences among questionnaires force us to use only a subset of columns in order to make an application that is coherent, does not use sparse data, and allows the user to make temporal analysis thanks to attributes used in multiple consecutive years (**FR2**).
- There are lots of attributes regarding redundant or less important questions. We mentioned a question about the opinion on the European Union (the expert evaluations are stored in the `eu_position` column), but in the trend file

there is also another attribute (`eu_salience`) about the "relative salience of European integration in the party's public stance" (0 = European integration is of no importance, never mentioned; 10 = European integration is the most important issue); basically, in some cases the dataset is allowing us to know not only what the parties thinks about a certain topic, but also "how important" it is to them. This is certainly not a useless piece of information, since when we think about a political party we know that they give priority to some topics at the expense of other ones, but still what we mostly care about is their general ideology and what they think of all possible matters.

- The amount itself of attributes must not be underestimated. Having lots of rows in the dataset means representing more data (more lines, more points, etc.), but having more columns implies instead having to represent more dimensions in the final visualisations. Representing multidimensional data in an effective and clear way is not an easy task, and adding new dimensions makes the charts more complex, increasing the amount of reasoning the user has to do and increasing the amount of graphical elements in the final system, such as for example axes and menus. **FR1** imposes us to support multidimensional analysis, but this certainly should not overwhelm the user.

In the end, of all the available dimensions, it was decided to utilise only the ones with the following properties:

- columns that store parameters essential for the application's functioning (like a party's unique numerical identifier and its country);
- columns regarding important and non-redundant topics (e.g. economy, environment, immigration);
- columns with data present not only in a good amount of consecutive survey years, but whose most recent year is 2024 (so that we can use the most recent data). It was decided therefore to utilise attributes evaluated in at least the last three survey years (2014, 2019 and 2024), because for the few ones tracked only from 2019 there is unfortunately too much missing data.

Finally, here is the complete list of attributes that were used in European Parties Explorer, with their names as they appear in the dataset's columns and their descriptions taken from the CHES trend file's Codebook. These will be the dimensions represented in the project's visualisations and they reflect the topics that were already identified in an abstract way in Section 3.1.2.

- `year`: year for which party experts were asked to evaluate (1999, 2002, 2006, 2010, 2014, 2019, 2024).
- `country`: unique identifier for each country (where the party comes from, see Table 5.1).
- `party_id`: unique (numerical) identifier for each party.
- `party`: party abbreviation.

Table 5.1. Mapping of country identifiers used in the dataset.

ID	Country	ID	Country
1	Belgium	21	Czechia
2	Denmark	22	Estonia
3	Germany	23	Hungary
4	Greece	24	Latvia
5	Spain	25	Lithuania
6	France	26	Poland
7	Ireland	27	Romania
8	Italy	28	Slovakia
10	Netherlands	29	Slovenia
11	United Kingdom	31	Croatia
12	Portugal	37	Malta
14	Finland	38	Luxembourg
16	Sweden	40	Cyprus
20	Bulgaria		

- `vote`: vote percentage received by the party in the national election most prior to `year`.
- `epvote`: vote percentage received by the party in the European Parliament election most prior to `year`.
- `family`: numerical identifier for the party's political family (Table 5.2).

Table 5.2. Mapping of faction identifiers used in the dataset.

ID	Country	ID	Country
1	Radical Right	7	Green
2	Conservatives	8	Regionalist
3	Liberal	9	No family
4	Christian-Democratic	10	Confessional
5	Socialist	11	Agrarian/Centre
6	Radical Left		

- `lrgen`: position of the party in `year` in terms of its overall ideological stance (0 = extreme left, 5 = centre, 10 = extreme right).
- `lrecon`: position of the party in `year` in terms of its ideological stance on economic issues. Parties can be classified in terms of their stance on economic issues such as privatisation, taxes, regulation, government spending, and the welfare state. Parties on the economic left want government to play an active role in the economy. Parties on the economic right want a reduced role for government (0 = extreme left, 5 = centre, 10 = extreme right).

- `eu_position`: overall orientation of the party leadership towards European integration in `year` (1 = strongly opposed, 2 = opposed, 3 = somewhat opposed, 4 = neutral, 5 = somewhat in favour, 6 = in favour, 7 = strongly in favour).
- `spendvtax`: position on improving public services vs. reducing taxes (0 = strongly favours improving public services, 10 = strongly favours reducing taxes).
- `deregulation`: position on deregulation of markets (0 = strongly opposes deregulation of markets, 10 = strongly supports deregulation of markets).
- `redistribution`: position on redistribution of wealth from the rich to the poor (0 = strongly favours redistribution, 10 = strongly opposes redistribution).
- `civlib_laworder`: position on civil liberties vs. law and order (0 = strongly promotes civil liberties, 10 = strongly supports tough measures to fight crime).
- `sociallifestyle`: position on social lifestyle (e.g. rights for homosexuals, gender equality; 0 = strongly supports liberal policies, 10 = strongly opposes liberal policies).
- `womens_rights`: position on policies supporting women's rights (e.g. equal pay, family leave, reproductive health; 0 = strongly supports women's rights policies, 10 = strongly opposes women's rights policies).
- `lgbtq_rights`: position on policies supporting LGBTQ+ rights (e.g. marriage equality, adoption rights, transgender rights; 0 = strongly supports lgbtq+ right, 10 = strongly opposes lgbtq+ rights).
- `religious_principles`: position on role of religious principles in politics (0 = strongly opposes religious principles in politics, 10 = strongly supports religious principles in politics).
- `immigrate_policy`: position on immigration policy (0 = strongly favours a liberal policy on immigration, 10 = strongly favours a restrictive policy on immigration).
- `multiculturalism`: position on integration of immigrants and asylum seekers (multiculturalism vs. assimilation; 0 = strongly favours multiculturalism, 10 = strongly favours assimilation).
- `nationalism`: position towards cosmopolitanism vs. nationalism (0 = strongly promotes cosmopolitan conceptions of society, 10 = strongly promotes nationalist conceptions of society).
- `ethnic_minorities`: position towards ethnic minorities (0 = strongly supports more rights for ethnic minorities, 10 = strongly opposes more rights for ethnic minorities).

- `environment`: position towards environmental sustainability (0 = strongly supports environmental protection even at the cost of economic growth, 10 = strongly supports economic growth even at the cost of environmental protection).

- `regions`: position on political decentralisation to regions/localities (0 = strongly favours political decentralisation, 10 = strongly opposes political decentralisation).

- `eu_intmark`: position of the party leadership in `year` on the internal market (i.e. free movement of goods, services, capital and labor; 1 = strongly opposed, 2 = opposed, 3 = somewhat opposed, 4 = neutral, 5 = somewhat in favour, 6 = in favour, 7 = strongly in favour).

- `eu_foreign`: position of the party leadership in `year` on EU foreign and security policy (1 = strongly opposed, 2 = opposed, 3 = somewhat opposed, 4 = neutral, 5 = somewhat in favour, 6 = in favour, 7 = strongly in favour).

Countries filtering

In this first step of data preprocessing we filtered many columns leaving only the essential ones, going from 90 dimensions to just 26.
There was actually the need to make some filtering on the rows too because of missing data; unfortunately, all parties from Malta and Luxembourg had to be removed from the dataset.

- Malta's parties were evaluated in the surveys of 2014, 2019 and 2024, but for the 2024 data there are just too many missing values; this means that we cannot make an accurate analysis, especially considering that we would like to have the most recent data. Luckily we are not losing too much data with respect to the whole CHES dataset, since Malta's data contains only two parties (one left-wing, one right-wing).

- Luxembourg's parties were evaluated only in 2014 and 2019. Not only does this push us to exclude Luxembourg because we do not have the 2024 data, but most expert evaluations are also missing from the two tracked years.

Being a small nation unfortunately implies that it is hard to find a consistent number of political experts capable of fully evaluating these countries' parties, therefore Malta and Luxembourg had to be excluded from the data used by European Parties Explorer.

Missing data

Even the bigger countries tracked for most years happen to have missing data. There are some parties for instance that were evaluated on most topics, but which have one or two attributes with no values. It is understood that we can not include these parties in the final application, since we should treat them as special cases and it is not clear what we should do with them in the final visualisations (for instance, in the parallel coordinates plot we are going to see in Section 5.3.2, how are we supposed to draw a line connecting all the axes, if one axis is missing its value?). Luckily for us, almost all the deleted rows contain information about small parties (vote share less than 2%), so we still retain data about the most relevant ones. The amount of removed data is also not too high, given that (considering the removal of Malta and Luxembourg too) from the 1441 rows of the original dataset we end up with 1397 rows.

Adjusting some scales

Now that we have removed the too small countries and all the parties with missing data, there are only some "adjustments" left to do to the dataset, mainly to improve the user experience and to better support the research of new insights.

- First, it was decided to invert some scales used for the expert evaluations. Let us take two attributes as example: `deregulation` and `redistribution`. As described earlier, in the case of `deregulation`, 0 means "strongly **opposes** deregulation of markets", and 10 means "strongly **supports** deregulation of markets"; for `redistribution` of wealth from the rich to the poor, 0 means "strongly **favours** redistribution", whereas 10 means "strongly **opposes** redistribution".
  What we get is basically an attribute where 10 = "in favour", and another one where 10 = "against". When representing all the expert evaluations and their scales, it may become confusing to have some scales in which the maximum value means one thing, but the same value means the opposite thing in other scales. When testing earlier versions of European Parties Explorer, having scale extremes with different meanings only made interacting with the application and finding insights needlessly harder, since the user always has to check what the minimum and maximum scores mean for each available range. For this reason, it was decided to "invert" some dimensions (`spendvtax`, `redistribution`, `sociallifestyle`, `womens_rights`, `lgbtq_rights`, `multiculturalism`, `ethnic_minorities`, `environment`, `regions`), so that in all scales 0 means "strongly opposes the specific topic", and 10 means "strongly favours that same topic"; for instance, a party that had an evaluation of "6" in `redistribution` will now have "4". This way, it is always clear for the user how to interpret the scales without having to check the meaning of each attribute, making the multidimensional analysis requested by **FR1** simpler and less mentally demanding.

- We talked mainly about scales going from 0 to 10, but there are three attributes (`eu_position`, `eu_intmark` and `eu_foreign`) whose scales were defined as follows: 1 = strongly opposed, 2 = opposed, 3 = somewhat opposed, 4 = neutral, 5 = somewhat in favour, 6 = in favour, 7 = strongly in favour. These ranges all respect the rule established above such that the maximum score means "strongly favours", but to make the expert evaluations consistent it was decided to recompute the data and move the minimum value to 0 and the maximum value to 10.

- The last special case is the one of `lrgen` and `lrecon`, where 0 = extreme left, 5 = center, 10 = extreme right. This seems coherent with the other 0-10 scales, but these two attributes are slightly different: while the other scales encode contrary or favourable opinions on a topic, in this case we are talking about two scales representing the parties' actual political alignment (left vs. right). Thus, to remind the user that these scales should not be interpreted as "opposes" vs. "favours", their data was recomputed as well so that the range goes from a minimum of -5 to a maximum of 5. This way, the "centre" value is also placed at 0.

Computing new attributes

Finally, in order to support the future visualisations and the discovery of further insights, we can also compute some new useful dimensions ourselves.

- First, we need to deal with the `sociallifestyle` attribute, defined as "position on social lifestyle (e.g. rights for homosexuals, gender equality)". This topic was evaluated during the surveys of 2006, 2010, 2014 and 2019, but not in 2024; instead, in the last survey (and only in that year) the experts were asked to evaluate the `womens_rights` and `lgbtq_rights` attributes.
  Since we are talking about a relevant topic and we want to allow analysis across many years (**FR2**), including especially the most recent one, it was established to add in the dataset, for each party, the values for `sociallifestyle` in 2024, computing them as the average of the specific party values for `womens_rights` and `lgbtq_rights`; these last two dimensions were then discarded from the dataset and will not be used in the application. Tests in later versions of European Parties Explorer proved that this approach generated values for `sociallifestyle` in 2024 coherent to the ones already stored in the previous years.

- In Section 4.3.3 we talked about the aggregate analysis patterns, dividing them in analysis of factions, of countries and of topics. Rather than just analysing by country, it is also interesting to find differences among European regions, for instance Western Europe and Eastern Europe, given their different history (and this is also something that has been done in the CHES articles). This would introduce a new type of analysis pattern (although very similar to the one by countries, at least from the point of view of the steps to take), that is the analysis of European regions.
  To this aim, we can take the `country` identifier and use it to assign to each party the European region it belongs to in a new `region` attribute: **West**, **North**, **South** or **East**. In this project we adopted the *UN M49 standard* [28], also known as *United Nations geoscheme*, a system that divides the countries of the world into many subregions for statistical purposes.

  - **Western Europe**: Austria, Belgium, France, Germany, Netherlands.
  - **Northern Europe**: Denmark, Estonia, Finland, Ireland, Latvia, Lithuania, Sweden, United Kingdom.

  - **Southern Europe**: Croatia, Cyprus[2], Greece, Italy, Portugal, Slovenia, Spain.
  - **Eastern Europe**: Bulgaria, Czechia, Hungary, Poland, Romania, Slovakia.

- Finally, we need to compute new attributes for the left-right alignment. The reason for this will be clear in Section 5.3.4 when we discuss about a particular visualisation called RadViz; for now, it is enough to know that we need to "split" the `lrgen` scale into two new scales, `leftgen` and `rightgen`, according to the following philosophy:

  - if a party has a value less or equal than 0 in `lrgen` (left-wing), then store that value changing its sign in `leftgen` and store 0 in `rightgen`;
  - if instead a party has value greater or equal than 0 in `lrgen` (right-wing), store that value in `rightgen` and 0 in `leftgen`;

  The same rules have to be applied to `lrecon`, creating the `leftecon` and `rightecon` dimensions.
  Essentially, the `leftgen` and `leftecon` scales (from 0 to 5) tell how far to the left a party is, storing 0 if it is right-wing; similarly, `rightgen` and `rightecon` tell how far to the right a party is, storing 0 if it is left-wing. Again, the reasons for these computations will be clear once we deal with RadViz.

To summarise, from the original CHES dataset we selected only a subset of columns to use, we removed the parties and countries with missing data, we adjusted the scales so that they all go from 0 to 10 and the maximum value means "strongly in favour" (except for `lrgen` and `lrecon` going from -5 to 5), we computed the `sociallifestyle` data for 2024 and created the new attributes `region`, `leftgen`, `rightgen`, `leftecon` and `rightecon`.

5.1.3 Dimensionality reduction

As it is clear by now, the CHES dataset is inherently high-dimensional, containing more than 20 attributes even after the columns filtering we described above; this means that our data is in a multidimensional R^n space. Unfortunately, human perception is limited to visualising data in two or at most three dimensions, and while it is possible to create 2D charts representing multidimensional data (like the aforementioned parallel coordinates plot we will see in Section 5.3.2), they inevitably present some defects and are more complex than a "simple" Cartesian chart. Most importantly, in these types of views it is hard to visualise data in a way that reveals similarities between data objects.
That is the context in which **dimensionality reduction** techniques come into play. As defined by van der Maaten et al. [20], dimensionality reduction is the process of mapping data from a high-dimensional space to a lower-dimensional one (typically 2D) while retaining the geometry of the data as much as possible. Thanks

[2] According to the United Nations geoscheme, Cyprus is classified as belonging to Western Asia; however, for the purpose of this thesis' project and its focus on European politics, it was manually reclassified as Southern Europe.

to dimensionality reduction algorithms, it will be possible for us to place parties in a simple bidimensional space (according to their many original attributes) and easily find which ones are similar; this will not allow us to directly compare multiple attributes, but it will still contribute to the multidimensional analysis (**FR1**) by making it simple to find similarities in the data.

Dimensionality reduction with MDS

Among the various dimensionality reduction algorithms, **multidimensional scaling** (**MDS**), originally defined by Torgerson [26], was the one chosen for this project. In short, MDS works as follows:

- the input is data in an R^n space;
- the data is used to compute a *dissimilarity matrix*, that is a matrix where each value represents how "distant" two elements are (0 means they are equal);
- the data points are randomly plotted in an R^2 space;
- a new matrix is computed, storing the Euclidean distance between point pairs in the new R^2 space;
- the two matrices are compared, evaluating a stress function that tells how different they are;
- the points' positions are adjusted lowering the stress function; the last three steps are repeated.

In our case, MDS was not applied to all parties at once, but only between parties present in the same year; it was therefore executed seven times, giving us one point arrangement per year. Also, not all attributes available up to this point were used to compute the dissimilarity matrices: the parties should be compared only according to the "real" expert evaluations, and it makes no sense to use for instance their country or vote percentage, nor metadata like their identifier. These dimensions were thus excluded: `year`, `country`, `party_id`, `party`, `vote`, `epvote`, `family` and the newly computed `region`, `leftgen`, `rightgen`, `leftecon`, `rightecon`.
In Figure 5.1 we can see how the points encoding parties in 2024 were placed thanks to MDS; the points of each year will be displayed in an appropriate visualisation (a scatter plot described in Section 5.3.1). It is important to notice that the axes and points orientation have no real meaning, the only thing that matters is the distance between points. The points' MDS coordinates were stored in two new attributes added to the dataset, named `mds1` and `mds2`.

Dimensionality reduction concludes the data preprocessing phase of this project. More specifically, referring again to the Visual Analytics process presented in Section 2.1.1, by now we have completed two of its phases: the preliminary operations on the data, and the creation of a data model (the data stored in the application, plus the 2D representation given by the dimensionality reduction). In later sections, we will of course explain the views and examples of how to obtain knowledge, completing the full cycle of Visual Analytics.

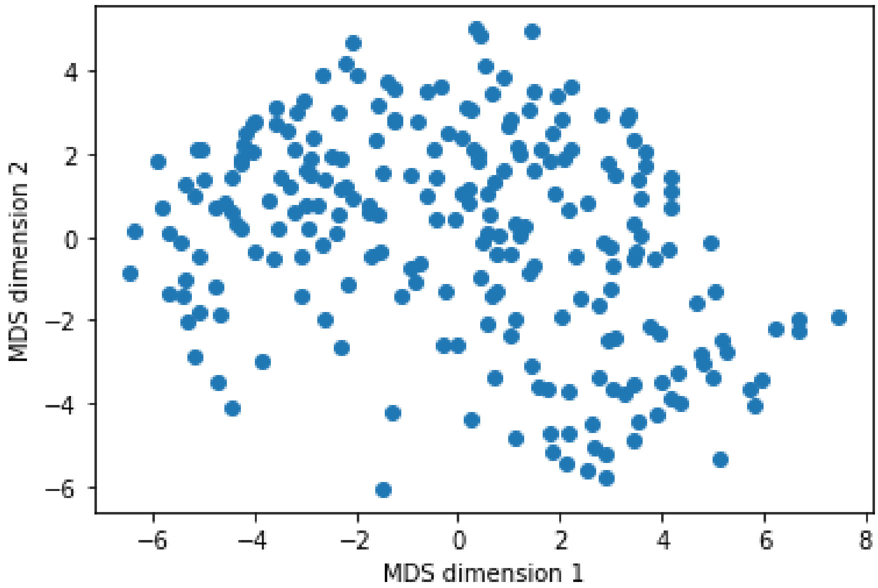


Figure 5.1. How MDS places parties of 2024 according to their attributes (one point represents one party). The only meaningful piece of information is the distance between points, with parties being closer the more similar they are.

5.1.4 Final dataset

Having seen all the operations carried out as part of the data preprocessing, let us have a quick overview of the final state of the dataset. This is the data that is going to be effectively used by European Parties Explorer.
From the original dataset of 1441 rows and 90 columns, storing party data for each survey year, we extracted a dataset of 1397 rows and 31 columns, with some dimensions created during preprocessing. Given that the number of tracked nations increased during the years, as explained when introducing the Chapel Hill Expert Survey project, what follows is what countries we are going to find in the final dataset and from which year their parties were evaluated by the experts:

- countries tracked from 1999: Belgium, Denmark, Germany, Greece, Spain, France, Ireland, Italy, Netherlands, United Kingdom, Portugal, Austria, Finland, Sweden;
- countries tracked from 2002: Bulgaria, Czechia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia;
- countries tracked from 2006: Estonia;
- countries tracked from 2014: Croatia, Cyprus.

Here is instead how the attributes are distributed, reflecting the increase of questions during the years in the expert surveys:

- attributes tracked from 1999: `country`, `year`, `party_id`, `party`, `vote`, `epvote`, `family`, `lrgen`, `lrecon`, `eu_position`, `eu_foreign`, `region`, `leftgen`, `rightgen`, `leftecon`, `rightecon`, `mds1`, `mds2`;

- attributes tracked from 2002: `eu_intmark`;
- attributes tracked from 2006: `spendvtax`, `deregulation`, `redistribution`, `civlib_laworder`, `sociallifestyle`, `religious_principles`, `immigrate_policy`, `multiculturalism`, `ethnic_minorities`, `regions`;
- attributes tracked from 2010: `environment`;
- attributes tracked from 2014: `nationalism`.

The attributes contain either "metadata" (party id, country etc.), electoral results or expert evaluations; the expert evaluations were made all coherent, with scales in the 0-10 range where 10 means "the party strongly agrees with the given topic", except of course for `lrgen` and `lrecon` going from -5 to 5 and the four attributes we derived from them. The `leftgen`, `rightgen`, `leftecon` and `rightecon` dimensions will be used for RadViz and the peculiar way it represents data, whereas `mds1` and `mds2` will be utilised to position the parties when displaying them thanks to the dimensionality reduction.

Attributes renaming

Some of the visualisations used in European Parties Explorer that we are going to see in a few sections have to inherently display what data attributes they are showing; the problem is that the raw dimensions provided by the CHES dataset are identified by technical names, and viewing these names without explanations does not make much sense. To ensure a user-friendly experience in the interface, these attributes were renamed in the views to more meaningful and readable labels, while also making more explicit what a high value in the attributes means. Nonetheless, there will be ways in the system to read clear and extensive explanations of the many attributes.
Table 5.3 presents the mapping between the original dataset identifiers and the labels displayed to the user in European Parties Explorer.

5.2 The model-view-controller architecture

We have seen all the preliminary operations carried out on the data, thus we can start talking about the project's actual structure and implementation.
To ensure the modularity, maintainability, and scalability of European Parties Explorer, the system architecture follows the **Model-View-Controller** (**MVC**) design pattern, originally formulated by Trygve Reenskaug in 1979 [22][23] and later popularised by Glenn E. Krasner and Stephen T. Pope [19].
MVC is a popular architectural pattern in software design that allows to structure code such that there is a clear separation between the internal representation of data and the way information is presented to the user. This separation is particularly critical in Visual Analytics applications where not only there are multiple complex visualisations (the views) representing the internal data and responding to the underlying logic (the model), but they must also coordinate themselves and adapt to the user's interaction (satisfying **FR4**).

Table 5.3. Mapping between original CHES attribute names and labels displayed in the user interface.

Original Code	Display Label
`family`	**Political faction**
`country`	**Country**
`region`	**European region**
`vote`	**Votes in the most recent national election (%)**
`epvote`	**Votes in the most recent European election (%)**
`lrgen`	**Left/right**
`lrecon`	**Economic left/right**
`leftgen`	**Left**
`rightgen`	**Right**
`leftecon`	**Economic left**
`rightecon`	**Economic right**
`eu_position`	**European Union**
`eu_intmark`	**EU internal market**
`eu_foreign`	**EU foreign policy**
`nationalism`	**Nationalism**
`immigrate_policy`	**Immigration restriction**
`civlib_laworder`	**Law & order**
`religious_principles`	**Religion in politics**
`deregulation`	**Market deregulation**
`spendvtax`	**Public spending**
`redistribution`	**Wealth redistribution**
`environment`	**Environment protection**
`sociallifestyle`	**Social lifestyle**
`multiculturalism`	**Multiculturalism**
`ehtnic_minorities`	**Ethnic minorities**
`regions`	**Regionalism**

In our JavaScript implementation, the three components interact as described below.

5.2.1 The model

The model is the part of code that manages the data and the logic of the application. The model's job is to store the data, expose it to the application and update it when needed, all independently of the actual user interface; the model does not even "know" anything about the rest of the system, including the specific visualisations used to represent the data.
In the case of European Parties Explorer, this is how the model works:

- it is represented by an object of the class `Parties`, and as soon as the program starts it takes all the rows (the parties) of the loaded csv file and stores them internally after "converting" them in objects containing the data of all attributes, including the ones computed during preprocessing (such as the MDS coordinates);

- the model acts as the single source of truth for the application state; this means that it is the only piece of code where current “global” variables can be retrieved, such as the currently selected year (e.g. if the user selects 2019, than the application should display the party evaluations of 2019) and possible active filters.

As mentioned, the model does not possess any knowledge of the DOM or how the data is displayed; it simply exposes methods to retrieve the data and updates it according to the user interaction (via the controller, see below).

5.2.2 The view

The view is the part responsible for the visual representation of the information. In European Parties Explorer, the views correspond to the interactive charts described in the following sections (parallel coordinates, box plots, scatter plot, RadViz, line chart). Implemented using the D3.js library, each view is an independent component that:

- receives the data from the model (through the controller), both initially and whenever there is an update in the data;
- renders the data through visualisation elements (SVG shapes, axes, labels) on the DOM.
- captures user interactions (like mouse clicks, hovering or brushing events) and forwards them to the controller.

Just like the model, the views are not fully “aware” of the rest of the system: they do not make decisions on how to filter data or update the state; instead, they strictly focus on rendering the current state provided to them by the controller, and notify the controller of current user interactions.

5.2.3 The controller

The missing piece in the MVC architecture is the controller. Given that the model and the view are unaware of each other’s existence and that they do not directly communicate with each other, the controller’s job is to act as the intermediary between them: it handles the application flow and the user’s input.

- When the application starts, the controller triggers the data loading process in the model and initialises the views;
- when a user interacts with a view (e.g. changing the year to display, or hovering a graphical element), the view notifies the controller; the controller then interprets this action and updates the model’s state accordingly (for instance setting its `selectedYear` variable as the newly selected year, or updating the piece of data corresponding to the just hovered element);

- finally, once the model's state has changed, the controller triggers an update function that retrieves the new data from the model and delivers it to all views, making sure that the visualisations are all coordinated and reflect the latest interal state.

This system architecture ensures that the multiple coordinated views described in the next section are always synchronised, as they all immediately react to the changes propagated by the controller.

5.3 Visual Analytics environment

It is clear that European Parties Explorer is designed according to the multiple coordinated views paradigm explained in Section 2.1.2. The system employs five types of visualisations, as visible in Figure 5.2, each optimised for a specific type of data analysis and working together to satisfy the functional requirements of the application: four of them (scatter plot, RadViz, parallel coordinates plot and box plots) are dedicated to the analysis of the dataset in one year selected by the user (essentially, a "snapshot" of the data), whereas the fifth one (line chart) is more suited to support temporal analysis across the years; exactly the kinds of analysis that **FR2** asks us to support, together with the ability to switch from one to the other. This also means that the interface includes a global year selector that allows the user to choose the year to display, updating all linked views simultaneously.
In the following sections, we will describe how these visualisation techniques work, how they were adapted for this project, and how they support the user tasks identified in Chapter 4.3,

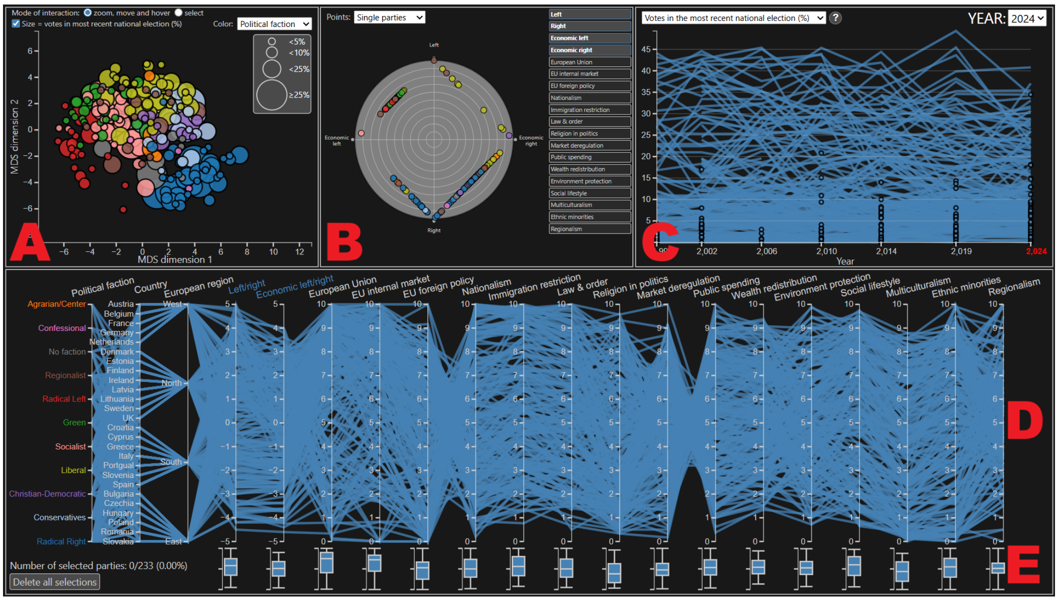


Figure 5.2. European Parties Explorer in its initial state, before any user interaction. We can see scatter plot (A), RadViz (B), line chart (C), parallel coordinates plot (D) and box plots (E).

5.3.1 Scatter plot

The scatter plot (A in Figure 5.2) is one of the most common types of charts, being the conventional way to represent data with two dimensions using Cartesian coordinates. In the case of European Parties Explorer, each circle in the scatter plot represents a party placed according to the coordinates computed by MDS and stored in the `mds1` and `mds2` attributes (compare the positions in Figure 5.1 and Figure 5.3). As explained in Section 5.1.3, MDS was applied only among parties belonging to the same year, thus there are seven possible dispositions depending on the year selected globally. Most importantly, similar parties are placed next to each other, with this similarity computed starting from the aforementioned dissimilarity matrix built upon the expert evaluations.

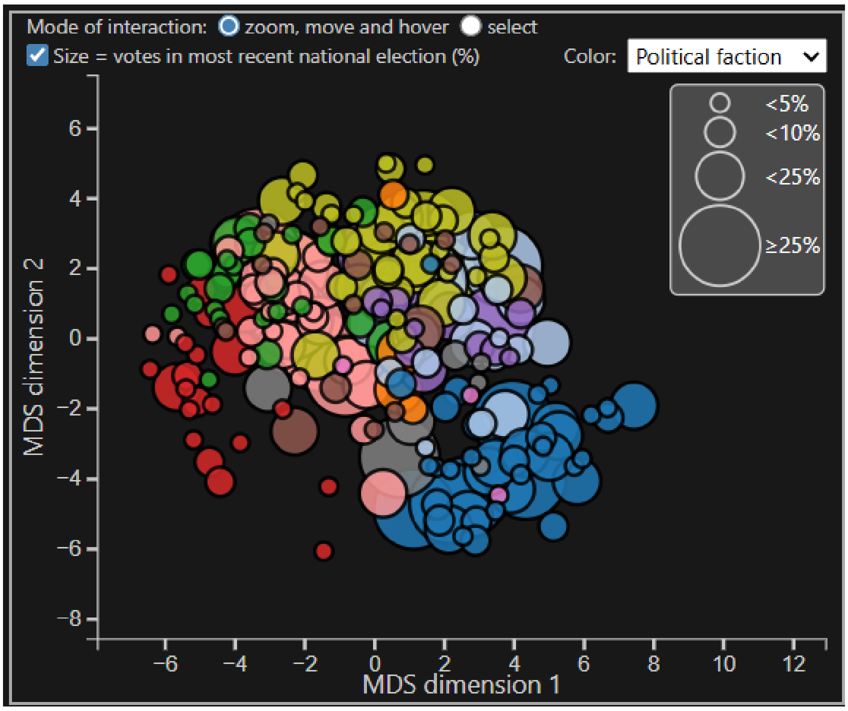


Figure 5.3. Scatter plot for the year 2024. Each circle represents a party, and parties with similar attributes are placed close together thanks to MDS. Colours encode political factions, sizes encode the amount of votes.

Parties are coloured according to their political family (Figure 5.3) or European region (Figure 5.4), at the user's choice. The colours were chosen from D3's categorical scale `schemeCategory20`, and the way they were assigned to factions tries to reflect the way political groups are "traditionally" represented in politics (e.g. blue for right-wing, red for left-wing, green for greens, yellow for liberals; see for instance Politico's visualisation in Figure 2.2). The colour legend is incorporated in the parallel coordinates, as shown in the next Section 5.3.2.
The circles' size depends on the votes received in the national election most prior to the selected year (see Figure 5.3.1), with a legend showing the four possible sizes: the smallest one for parties with less than 5% of votes, then one for less than 10%, one for 25%, and the biggest one for at least 25% of vote share. A "continuous"

range of sizes (reflecting the share of votes 1:1) was tested, instead of using these four predetermined intervals: this just made the scatter plot too confusing, introducing a lot of visual noise. Anyway, the user can also set all the circles at the same size as in Figure 5.4, if they prefer so. In this case, the legend disappears.
Lastly, it is essential to remember that with MDS, the axes and the actual position of points have no real meaning: we must reason only on the basis of the distances between the points, their colour and their radius. One proof of this is the fact that a faction placed, for instance, on the top right in one year, may end up positioned on the left in another year.

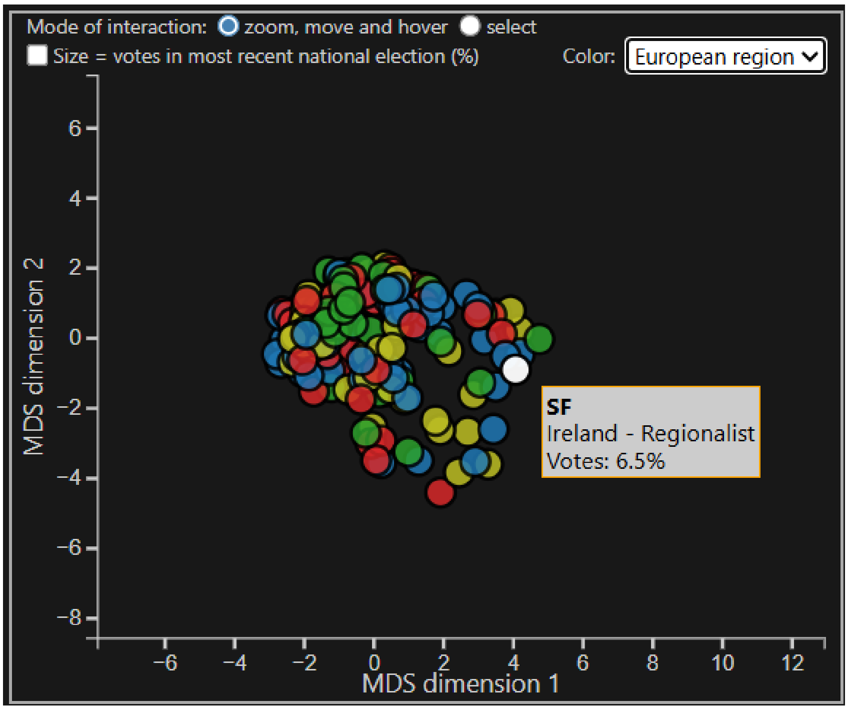


Figure 5.4. Scatter plot for the year 2002. In this case MDS had to operate with a small set of attributes, so the points are placed closer with respect to later years. The different sizes are disabled, the colours encode European regions. The **SF** party from Ireland is mouse hovered, displaying a tooltip with its information of 2002.

Interacting with the visualisation

The user can interact with the scatter plot in different ways. Aside from the aforementioned disabling of the size and the possibility to choose between colouring by faction or by European region, the user can of course hover the parties to see their basic data (name, country, region and share of votes, see Figure 5.4). A hovered circle becomes white to highlight it and it is brought to the foreground, on top of all the other circles.
The user can zoom in and out on the scatter plot (Figure 5.5). This operation is useful when many parties are too close to each other, and this happens especially in the years 1999 and 2002 (check for example the difference between 2024 and 2002 in Figures 5.3 and 5.4 respectively): MDS uses the expert evaluations to position

the points, and we should remember that the expert surveys had less questions in the first years. This implies that the dimensionality reduction algorithm has less available attributes to separate the parties (e.g. in 1999 it can use only **Left/right**, **Economic left/right**, **European Union** and **EU foreign policy**) placing them closer in the early years. Of course the user can also move the points around, exploring new areas.
Finally, it is clear that brushing is implemented too. In figure 5.5 we can see a rectangular brush highlighting the circles inside it, and it can be resized and moved around to brush new parties.

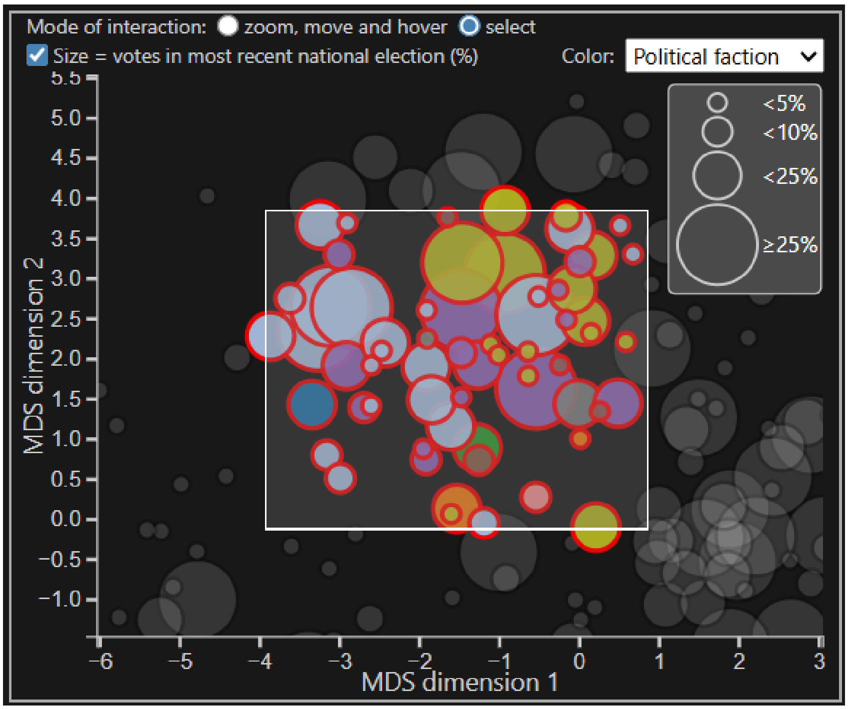


Figure 5.5. Scatter plot for the year 2014. Here we can see it was zoomed in and brushed.

Functional requirements and tasks

As explained in Section 4.4, the visualisations should be designed following the defined functional requirements (**high-dimensional analysis**, **temporal analysis**, **multi-granularity comparisons**, **coordinated exploration**), since this guarantees us the full support of the user tasks ("formalised" through the analysis patterns).
European Parties Explorer's scatter plot mainly satisfies **FR1**: while it does not allow the user to directly compare the numerical values of the many dimensions, it still uses them to position the parties, revealing similarities. Analysing multidimensional data projected on a 2D space is a typical way of carrying out high-dimensional analysis, especially considering that the views will be coordinated for **FR4**: as such, hovering one party on the scatter plot will allow us to observe its actual attributes on other charts. In any case, some dimensions can easily be compared on the scat-

ter plot, also thanks to the tooltip (used in all views): political family, country, European region and votes.
The scatter plot was presented as a “snapshot” chart, displaying the data only of the selected year, thus it does not directly support the temporal analysis of **FR2**. However, just like all the other snapshot views, it is possible to change year and observe the differences among them, so it is still possible to do some form of temporal analysis. As for the multi-granular comparisons required by **FR3**, in the scatter plot we can mainly compare single parties, but we will see that it is possible to make considerations among arbitrary groups of data by using brushes on the other coordinated visualisations.
As for specific subtasks feasible in the scatter plot, some examples are: finding ideologically similar parties, seeing how big a faction is and how much it was voted, distinguishing between small and big parties, finding ouliers with respect to their political family, observing how “unified” a group of parties is.

5.3.2 Parallel coordinates

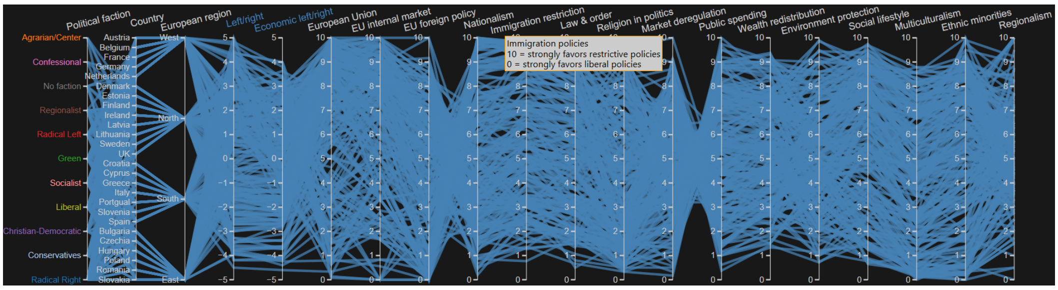


Figure 5.6. Parallel coordinates for the year 2024: each line represents a party, connecting its values for each dimension. The **Immigration restriction** label is mouse hovered, displaying a short explanation of the attribute. The **Political faction** tick labels are coloured as legend for the scatter plot. Some dimension labels are coloured in blue to quickly identify what attributes are being used in RadViz.

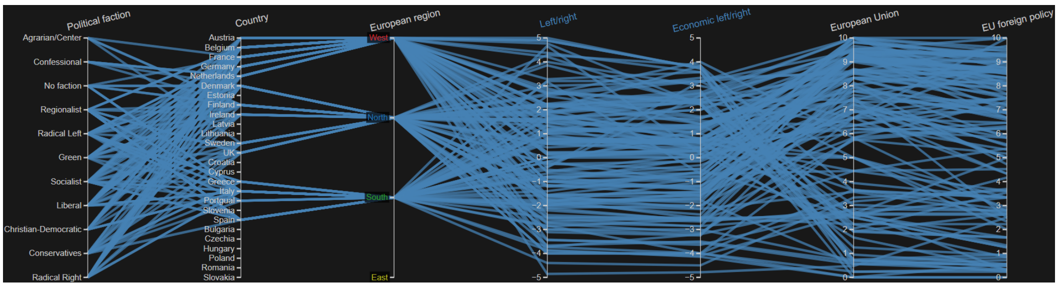


Figure 5.7. Parallel coordinates for the year 1999: we can clearly see the lower amount of dimensions available in the earlier years. In this case, the **European region** tick labels act as legend because of the scatter plot being coloured according to that attribute.

Parallel coordinates (D in Figure 5.2) is a visualisation technique popularised by Alfred Inselberg [17], specifically designed to represent high-dimensional data. Unlike Cartesian coordinates which use orthogonal axes, this view places axes parallel

to each other: each dimension of the dataset corresponds to a vertical axis, and a data point is represented as a line connecting its values across all axes. This layout allows for the simultaneous visualisation of many attributes, making it possible to spot correlations: for example, parallel lines between two axes indicate a positive correlation, while crossing lines indicate a negative correlation.
In European Parties Explorer, each party is encoded by a line, while the axes correspond to both categorical (**Political faction**, **Country**, **European Region**) and numerical attributes (the expert evaluations). The arrangement of the axes in parallel coordinates significantly impacts the readability of the chart, since patterns are more easily visible only between adjacent axes. As demonstrated by Ankerst et al. [2], the best way to order the axes is to place similar or correlated attributes next to each other, which reduces visual clutter and reveals hidden relationships. In European Parties Explorer however, we decided to also adopt a semantic approach, grouping together as much as possible dimensions belonging to similar topics (e.g. **Public spending** is near **Wealth redistribution**, **Nationalism** is near **Immigration restriction**) and facilitating the user's exploration of the data. The final arrangement is structured as follows:

- categorical axes (which will be fundamental for brushing);
- axes about left and right;
- axes about the European Union (they are similar, generally following the same patterns);
- a set of correlated axes, going in Figure 5.6 from **Nationalism** to **Market deregulation** (in general, parties with high values in one of them have high values in all the others, and vice-versa for low values; they are mostly next to semantically similar axes);
- another set of correlated axes, with values generally inverted with respect to the previous set, starting from **Public spending** (and again internally semantically ordered).

As we will observe in Section 6.3, ordering the axes in this way allows the user to immediately spot trends in the political families and easily compare them.
As usual, in earlier survey years there are less attributes, so the parallel coordinates plot adapts and displays only the available dimensions, as in Figure 5.7. In order to not hide information from the user, hovering the axes names shows a tooltip (Figure 5.6) with the *Codebook* explanation of the attribute and the extreme values.
Lastly, the visualisation contains the colour legend for the scatter plot, colouring the faction ticks accordingly (Figure 5.6); Figure 5.7 shows instead the updated legend when the scatter plot is coloured by European region. We can also see some blue axis labels: these are the attributes selected in RadViz so that the user can quickly find them in the parallel coordinates, as we will see in Section 5.3.4.

Interacting with the visualisation

All the axes in the parallel coordinates can be brushed, allowing the user to specify a certain range of values they are interested in. The brushed lines are highlighted

in red, while the other ones are greyed out and put in the background; of course, when there are more brushes, only the lines passing through all the selected ranges are highlighted (Figure 5.8).
It is possible to hover single lines, revealing information about that party with the same tooltip of the scatter plot; however, when brushing is active, the filtered out lines are disabled and do not display information on hover, since in that moment the focus is on the brushed parties. Also, allowing hovering on the deselected lines would be confusing, because it would not be immediately clear if a hovered party was also brushed or not (since the line is coloured in white, not red or grey), unless the user checks if the line crosses all the existing brushes.

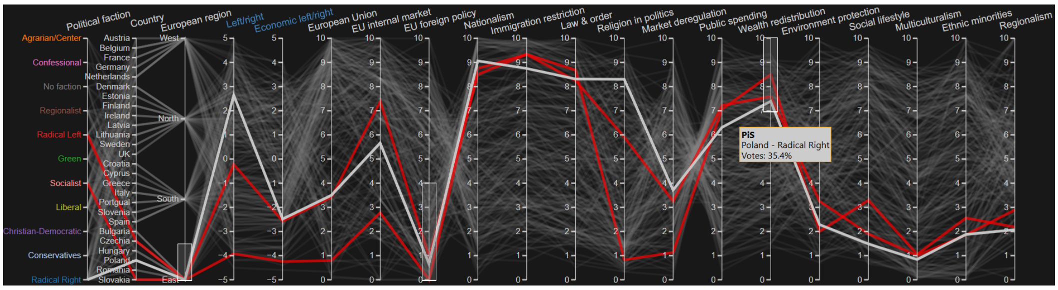


Figure 5.8. Parallel coordinates for the year 2024, with the parties remaining from the three brushes on **European region**, **EU foreign policy** and **Wealth redistribution**. **PiS** from Poland is mouse hovered.

Functional requirements and tasks

The parallel coordinates plot is probably the most "complete" view and the starting point of most analyses. It obviously supports **FR1**, since it is a visualisation that was created for high-dimensional analysis, and allows us to observe the values of almost all dimensions of the dataset and compare them. Correlations can be found at a first sight only between adjacent axes, but brushing can reveal further relationships (e.g., when selecting high values for **Nationalism** and sliding towards lower values, we can see that **Multiculturalism**'s values follow the opposite trend, thus being negatively correlated). In one way or another, parallel coordinates support the same actions that were possible in the related visualisations seen in Section 3.3, like discovering correlations among attributes or observing how parties of a country are ordered on one specific attribute.
Parallel coordinates are the best instrument for **FR3**, supporting comparisons at different levels. It is possible to compare single parties on their dimensions, but by brushing we can define arbitrary sets of parties and compare them. For instance, brushing one family in the **Political faction** axis reveals its ideology, and by sliding the brush we can compare it to the ones of the other factions: in this case, the comparison is not between single parties, and the faction becomes the "unitary" element; the same considerations can be made for **Country** and **European region**. Brushes are a very powerful instrument in parallel coordinates, and combining more of them allows the user to define any kind of subset of parties.
Temporal analysis is not possible with parallel coordinates on their own, but by

brushing a set of parties and selecting a different year we can see how that set's behaviour changed.

5.3.3 Box plots

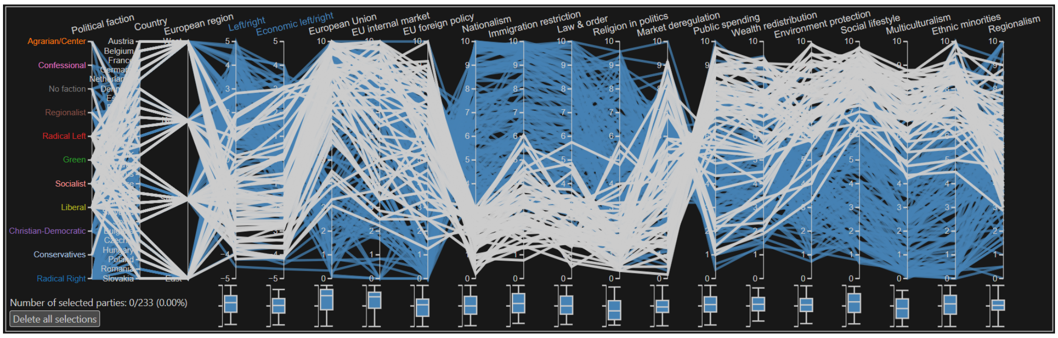


Figure 5.9. Parallel coordinates with relative box plots of 2024, displaying how parties are distributed on each dimension. The **Nationalism** box plot is hovered on its lower whisker, showing in the parallel coordinates the parties belonging to that percentile.

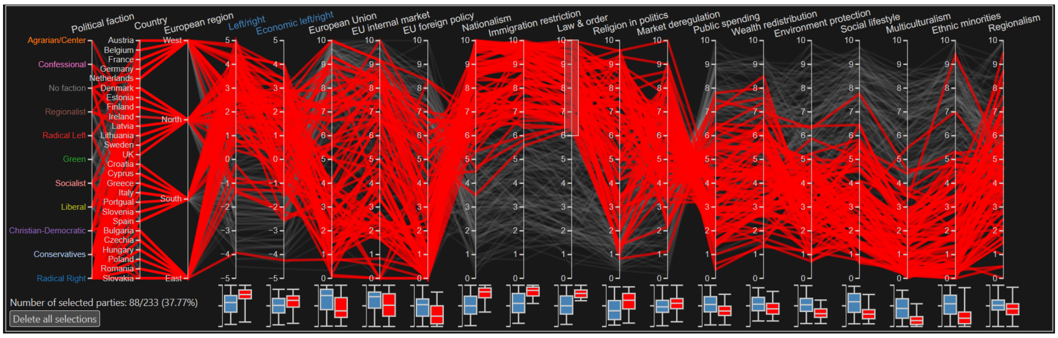


Figure 5.10. The **Law & Order** axis is brushed on its values from 6 to 10, which renders additional red box plots showing the distribution of the brushed data.

While parallel coordinates are excellent for analysing individual parties and finding trends of arbitrary groupings, they can suffer from "overplotting" when displaying tens or hundreds of lines, making it difficult to understand the statistical distribution of the data; to address this, box plots were integrated into the visualisation (E in Figure 5.2).
Box plots are a popular tool for displaying the distribution of data [27]. They work as follows:

- the central box represents the "middle" 50% of the data, that is the data going from the first quartile (25th percentile) to the third quartile (75th percentile);
- the central horizontal line inside the box marks the median;
- lines (called *whiskers*) extend from the box to display the range of the remaining data.

Box plots are therefore great for showing how parties are distributed on a certain topic: for instance, a "short" box plot indicates a common line of thought, whereas a "tall" one means polarisation; a high median tells us instead that most parties agree with the given topic.
In European Parties Explorer, box plots are rendered directly below the parallel coordinates axes, associating one box plot to one axis and how data is distributed on it (Figure 5.9). Since the parallel coordinates' axes can change on the basis of the selected year, the box plots are dynamically updated according to the existing ones. Given the nature of box plots, it makes no sense to use them with the categorical axes, so they are defined only for the numerical dimensions (the expert evaluations). If brushing is used on any visualisation, then additional box plots are displayed in red, visualising the distribution of the brushed parties (Figure 5.10).

Interacting with the visualisation

There is only one way to interact directly with the box plots, which is by hovering its regions. Whenever the user hovers one of the box plots' areas, the corresponding parties are highlighted in the other views, as in Figure 5.10. Of course, this works for both kinds of box plots.

Functional requirements and tasks

The inclusion of box plots is crucial for the aggregate analysis patterns, as they allow users to observe the cohesion of a group, or to understand the general opinion on a chosen topic. Since they allow the comparison among arbitrary groups of parties (mainly countries and factions), specifically their distribution on all numerical dimensions, they help in supporting the high-dimensional analysis required by **FR1**, especially when combined with the parallel coordinates plot. Furthermore, comparing box plots from different years shows how opinions on the topics have changed, particularly when brushing a political family or a nation, which means that the combination of box plots and parallel coordinates can somehow support temporal analysis (**FR2**).

5.3.4 RadViz

The last "snapshot" view is RadViz (Radial Coordinate Visualisation, B in Figure 5.2), a non-linear visualisation technique originally introduced by Hoffman et al. [16] It places *dimensional anchors* (representing attributes) around the circumference of a circle, and data points are placed inside the circle based on a physical "spring" model: a point with a high value for a specific attribute will be "pulled" closer to that attribute's anchor, as if there were a spring between the point and the anchor, stronger than the ones of the other anchors.
RadViz is a great instrument for analysing the relationships between attributes, but the spring mechanism can sometimes deceive the user: let us make some examples with three fictitious attributes called **A**, **B** and **C**, with values going from 0 to 10.

- If a point is close to the **A** anchor, the user might think that it has a high value for the **A** dimension. In reality, since the point is placed where the metaphorical springs are at the equilibrium, it just means that the value for **A** is higher than **B** and **C**, and no considerations can be made on the absolute values. A point with **A**=0.5, **B**=0 and **C**=0 will be positioned on top of the **A** anchor, even though 0.5 is a very low value.

- Points with different data can end up in the same spot. A point with the three attributes all equal to 1, and another point with the attributes equal to 8 will both be placed at the centre of Radviz on top of each other, because in both cases their springs apply the same force relative to each other.
- For this last example, let us introduce a fourth dimension **D**, and let us suppose that the anchors are arranged in the following clockwise order: **A**, **B**, **C** and **D**. In some cases, the points may appear as being "mispositioned": according to the previous example, a point at the centre can imply the same value for all attributes, but a point with **A**=10, **B**=0, **C**=10 and **D**=0 will elso end up at the centre of RadViz, because the contributions of **A** and **C** cancel each other out.

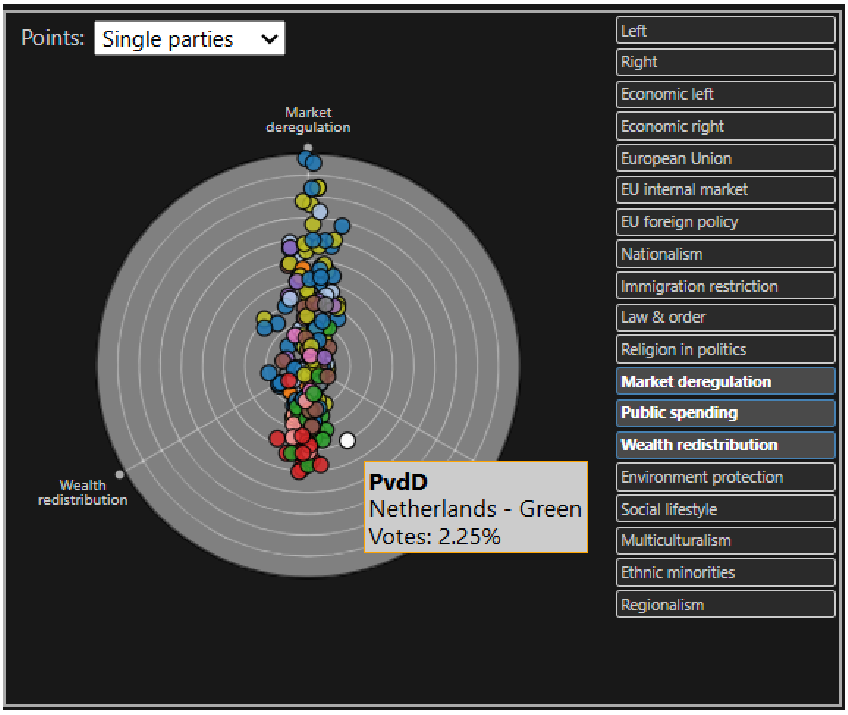


Figure 5.11. RadViz showing how three economical attributes are correlated in 2024, with parties coloured by faction as in the scatter plot. With respect to **Market deregulation**, **Public spending** and **Wealth redistribution** basically balance each other out, presenting proportional values whether they are low or high. The party **PvdD** from the Netherlands is hovered.

In practice, what we should not do in RadViz is reasoning in terms of absolute values, but rather on the **proportions between dimensions**. Also, as demonstrated in the third example, the main reason why RadViz can deceive users is the anchors arrangement, because some orderings can decrease the quality of the visualisation,

leading to wrong insights: if the example's anchors were ordered as **A**, **C**, **B** and **D**, the first two would pull and place the point close to them, succesfully showing that they are the dominating dimensions. In the case of European Parties Explorer, the library implementing RadViz uses a heuristic that "optimises" the arrangement of anchors [1], leading to a visualisation that can indeed reveal insights.
In the context of European Parties Explorer, RadViz encodes parties as points, and it can be used as a tool for analysing a subset of attributes: it is great for understanding patterns among many different dimensions and observing how parties are distributed on the chosen topics (see Figure 5.11). It is clear now why we created `leftgen`, `rightgen`, `leftecon` and `rightecon` during preprocessing: RadViz needs positive (high) values to pull the points towards the anchors, and so the `lrgen` and `lrecon` scales, where -5 means left and 5 means right (or, originally, 0 and 10), are not appropriate to position the parties and equally represent both political ideologies. We can see the new four dimensions in Figure 5.12.

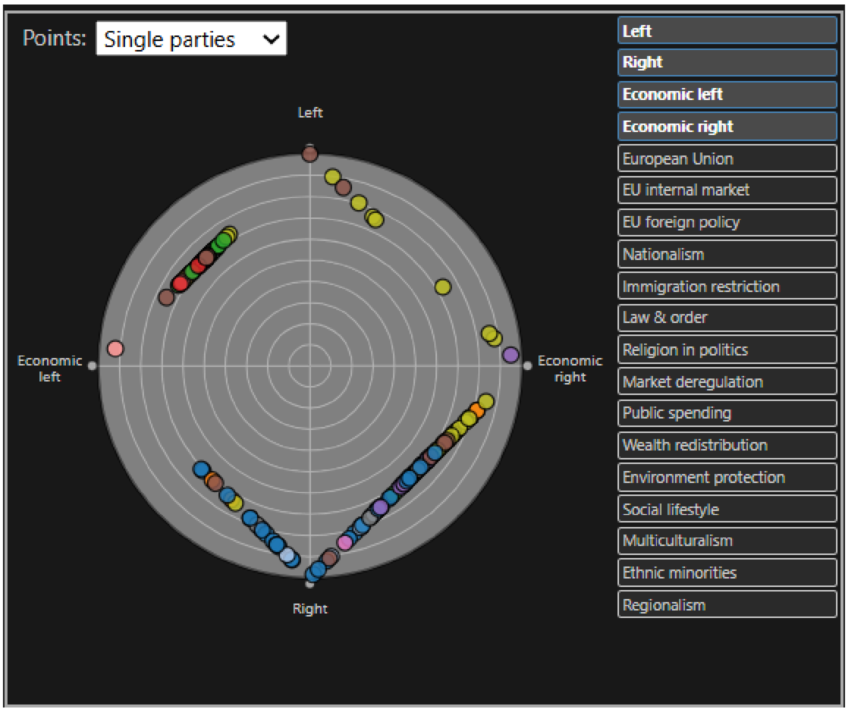


Figure 5.12. RadViz in its initial state, using the new attributes computed during data preprocessing. When using them together, they show which parties are left-wing for both scales, right-wing for both scales, or left in one scale and right in the other.

Interacting with the visualisation

In our application, the user can choose which attributes to use as dimensional anchors, revealing possible correlations. Only a maximum of four anchors are allowed, because even with the optimisation heuristic the quality of the visualisation can drastically decrease with too many dimensions. As usual, which attributes are available depends on the current selected year.
With RadViz we must focus on the proportions among dimensions, but by hovering

the points we can also reason on the absolute values, thanks to the coordinated parallel coordinates (remember **FR4**). To facilitate this, as we saw earlier, the parallel coordinates' axis labels corresponding to the dimensions selected in RadViz are coloured in blue.
Lastly, we integrated an alternative way of displaying parties. Thanks to a drop-down menu, the user can decide do represent on RadViz entire factions rather than single parties, utilising the factions' average values for the selected attributes (see Figure 5.13). This way, we can more clearly see the relationships among dimensions in the context of entire political families, and hovering one of these aggregated points highlights in the other views all parties belonging to the corresponding faction.

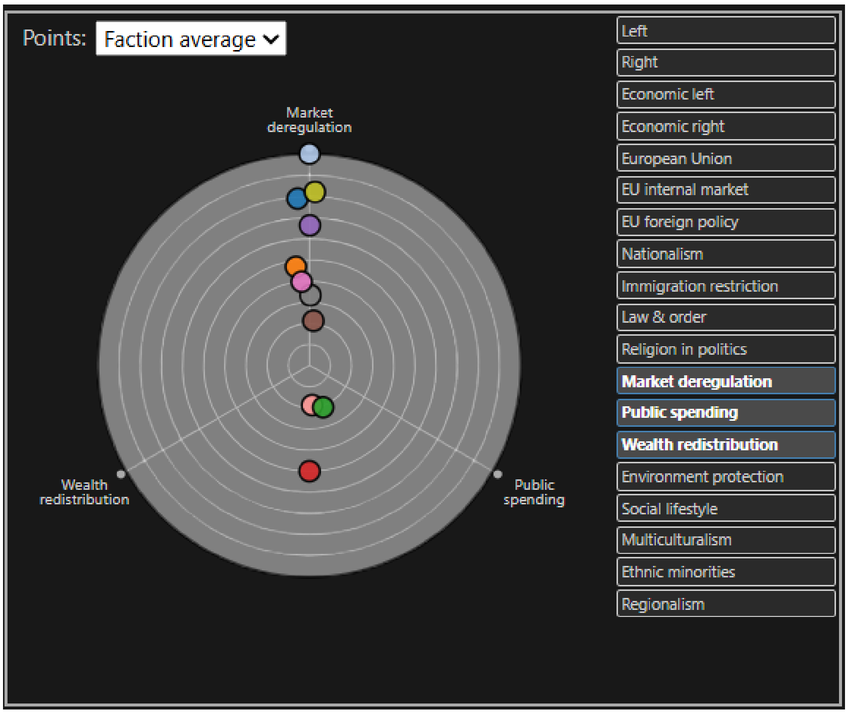


Figure 5.13. RadViz in 2024 displaying again the three economical topics of Figure 5.11, but using the faction average for each dimension instead of single parties.

Functional requirements and tasks

RadViz is particularly powerful for the analysis of topics, and as usual it respects **FR1** by allowing the comparison of many dimensions and to see how they relate to each other. It also can support multi-granular analysis (**FR3**), because when brushing different countries or factions (or other groups of data) we can clearly observe their differences on the selected attributes, and even without brushing it is possible to analyse single parties while also being able to quickly see how all parties are distributed. Most importantly, the drop-down menu allows to quickly switch from single parties to factions, easily changing the granularity of the analysis.

5.3.5 Line chart

Lastly, while the previous views show a "snapshot" of the selected year, the line chart (C in Figure 5.2) in European Parties Explorer is the main tool for the temporal analysis. It contains lines encoding parties that were present multiple times in the CHES dataset, showing the evolution of expert scores during the years, but differently from the other views it can also display the history of electoral results, as in Figure 5.14. It also uses points, for parties appearing only in one year.

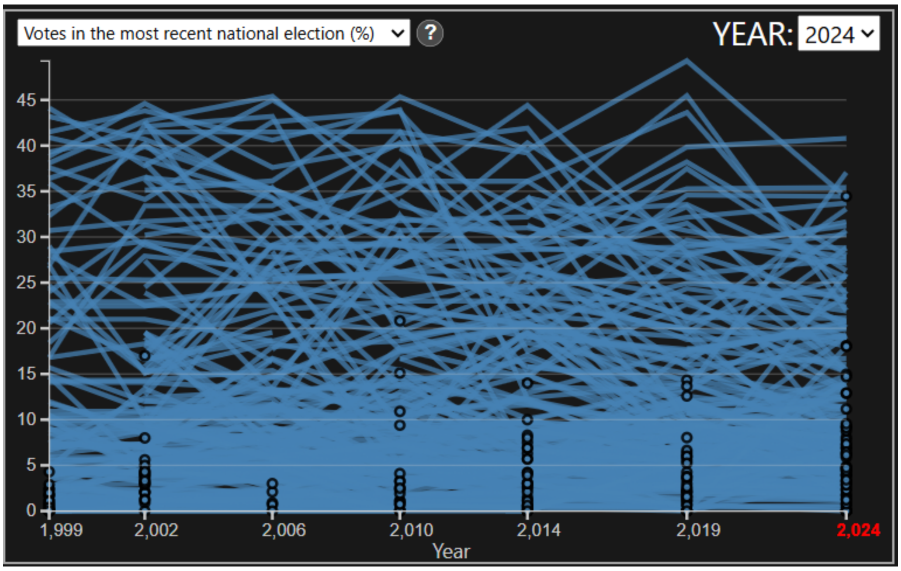


Figure 5.14. Line chart showing the amount of votes for each party during the years. Points are used to encode parties appearing only once.

The attribute to display in the line chart can be selected thanks to a drop-down menu, and a nearby question mark describes the attribute just like when hovering the axes labels in the parallel coordinates plot. The currently selected year is highlighted in red on the horizontal axis, to remind the user that even though the line chart displays the history of parties, the focus is still on the current year: parties can drastically change over time, so when brushing for instance liberal parties in 2024, the line chart will highlight only the parties that belong to the liberal family **in 2024**, even if in the past they were from another faction, and ignoring parties that were liberal in previous years but not in 2024. This approach is necessary to make the snapshot visualisations and this "temporal" visualisation coherent and to not confuse the user.
As demonstrated in Figure 5.15, the line chart succesfully displays attributes encoding the many expert evaluations that are available only from later years.

Interacting with the visualisation

Aside from choosing which attribute to visualise, of course the line chart supports mouse hovering on both lines and points. However, since this view focuses on temporal analysis, it would not make much sense to use the same tooltip shared by the other visualisations, displaying party information only for the currently selected year. Therefore, the line chart employs a different tooltip that visualises the hovered party's name, family and share of votes for each year in which it has data. This way, it is clear what the history of the hovered party is and how it has changed, as shown for example in Figure 5.15.

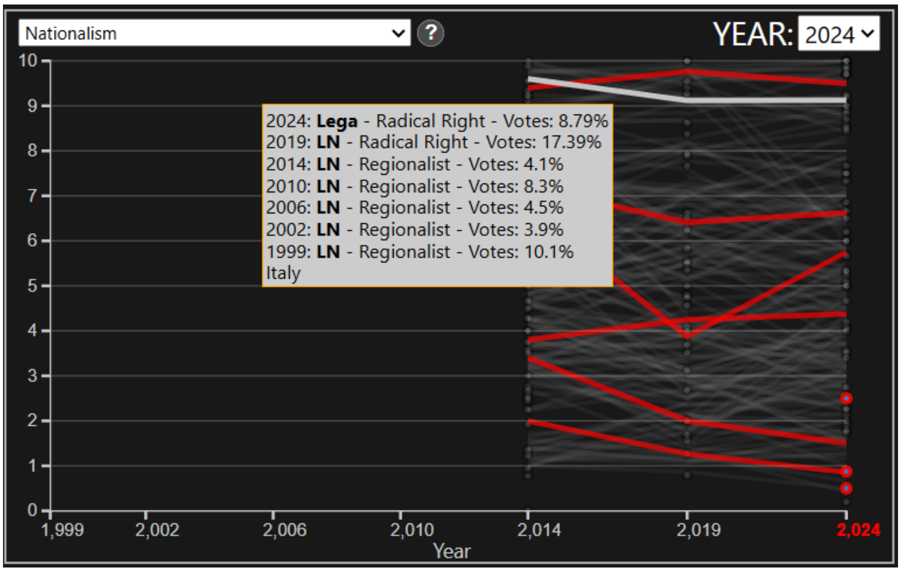


Figure 5.15. Line chart showing the evolution of **Nationalism** for the Italian parties tracked in 2024; the expert scores for this attribute are available only from 2014. We can see how the tooltip is different in the line chart, showing the party's history.

Functional requirements and tasks

It is clear that the line chart is the best visualisation to support temporal analysis (**FR2**), being very useful when carrying out the temporal comparison steps in the defined analysis patterns. The user can choose to display only one topic at a time, so it does not really support the comparison among different dimension, if not for party name, family and votes: the focus is on comparing many values of the same attribute in different years.

Regarding **FR3**, the line chart alone allows the analysis of single parties, but through brushing it is possible to analyse the evolution and temporal trends of many different groups of parties.

Chapter 6

Case studies

We now have a complete knowledge of European Parties Explorer, from the dataset that inspired this work and how it was processed, to the design choices that were made, the supported tasks, and the practical implementation of views respecting the functional requirements.
It is therefore time to get our hands on the application and validate it by putting it to the test. In this chapter, we will demonstrate how the system works in practice by walking again through the analysis patterns described in Section 4.3 and by simulating some real-world usage scenarios. Our objective is to verify that European Parties Explorer effectively supports the user personas we defined and to discover meaningful insights.

6.1 Basic functionalities

First of all, we should ensure that the fundamental features we described in Section 4.3.1 are respected and well-supported. Let us remind very briefly what these features were:

- highlight one or more parties, with coordination among views;
- show party name, country and faction
- encode the expert evaluations of the CHES dataset;
- visualise how the expert evaluations changed over time;
- display electoral results.

As we stated in that chapter, these functionalities are the building blocks of the whole interaction with the application and of the other more complex analysis patterns, and for now we are only dealing with correctly and quickly displaying information.
It is not hard to verify that indeed our design choices and the visualisations that were implemented support the correct displaying of fundamental information; here is why:

- hovering an element in any of the available visualisations allows the user to see the hovered party's basic information in the tooltip, that is to say its name, country, political family and the electoral result in the most recent national election;
- the parties' country and family are also available as data in the parallel coordinates, and the faction is represented in the scatter plot and RadViz too thanks to their color encodings; the vote share can be found again in the scatter plot with the circles' areas, even though only four predetermiend areas are used to avoid cluttering;
- the expert evaluations are clearly encoded and displayed thanks to the parallel coordinates' axes, and a single line is enough to identify one specific party with all its expert evaluations; also, instead of reasoning in absolute values, RadViz allows us to explore the proportions and relationships between different attributes;
- the line chart quickly shows how the expert evaluations of one selected dimension have changed over time, and it is possible to choose which attribute to visualise; we can also just compare how all scores of one party have evolved by selecting all years in sequence and observing the changes in the parallel coordinates and RadViz;
- the electoral results are available not only in the tooltip and in the scatter plot, but rather the whole electoral history can be found in the line chart for both the national and European elections;
- coordination among views exists when hovering a single party, but it works also when brushing multiple parties.

The above functionalities are shown all together in Figure 6.1, with scatter plot, RadViz, line chart and parallel coordinates encoding expert evaluations, party data and electoral results in many different ways; two brushes (one in the scatter plot with a rectangular area, one in the parallel coordinates selecting **Socialist**, **Green** and **Radical Left** in the **Political faction** axis) highlighting in all charts the parties resulting from the brushes intersection; coordination among views on mouse hover, with the tooltip showing party name, country, factions and the vote share of the most recent national election. Finally, a temporal study can be conducted both on the line chart and the other views by changing year.

6.2 Single parties analysis

As previously observed in Section 4.3.2, the first analysis pattern (focusing on the comparison among parties considered singularly) is the one more bound to the voter persona. This user is typically interested in a targeted search, filtering out parties they dislike and finding the ones more aligned to their personal ideas; furthermore, they may also want to identify parties that are ideologically similar and observe the parties' expert evaluations over the years.
Let us therefore imagine, as an example, an Austrian user who wants to identify

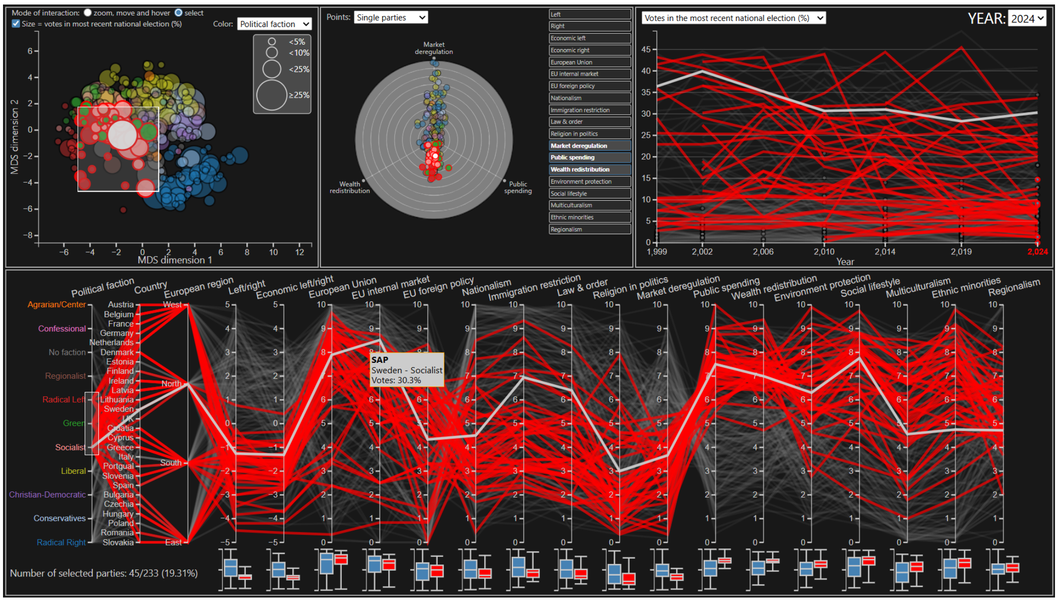


Figure 6.1. Screenshot of European Parties Explorer showcasing the application's basic functionalities; the two brushes select the parties deriving from their intersection.

a party that matches their pro-EU and environmentalist views, and see what set of actions they can follow in European Parties Explorer (not necessarily all in the order below):

- the user starts by interacting with the parallel coordinates plot, and by utilising the **Country** axis they apply a brush to isolate Austria; this highlights in the parallel coordinates all tracked Austrian parties, and immediately all other views are updated as well;

- the user can utilise the parallel coordinates to have a first overall view of the Austrian parties, their data and ideologies;

- thanks to the scatter plot, the user can quickly observe the proximity of the parties (if two parties are close together in the chart, then they share a similar ideological profile, regardless of their actual political family); also, the size of the bubbles gives a rough idea of the electoral results in the national election most prior to the last survey year;

- at this point, the user can move to the specific policy axes of the parallel coordinates and brush the top range of the **European Union** axis and the top range of the **Environment protection** axis; this way, the system filters out all parties not meeting these criteria, highlighting in all views only Europeanist and environmentalist parties existing in the last survey year;

- in case no parties are left because of too strict filters, the user can "relax" the brushes on the axes to find the next best options, that is parties still relatively close to the user's desires;

- as an additional option, the user can take advtange of the line chart to carry out a temporal analysis of all parties or only the ones remaining from the brushes, discovering in case if a party is coherent on a certain policy or if it drastically changed its mind;

The Austrian user example just described has been simulated in Figure 6.2. As a result of the selections made with the brushes, two parties to choose from are left: **NEOS** (Liberal) and **Grune** (Green).

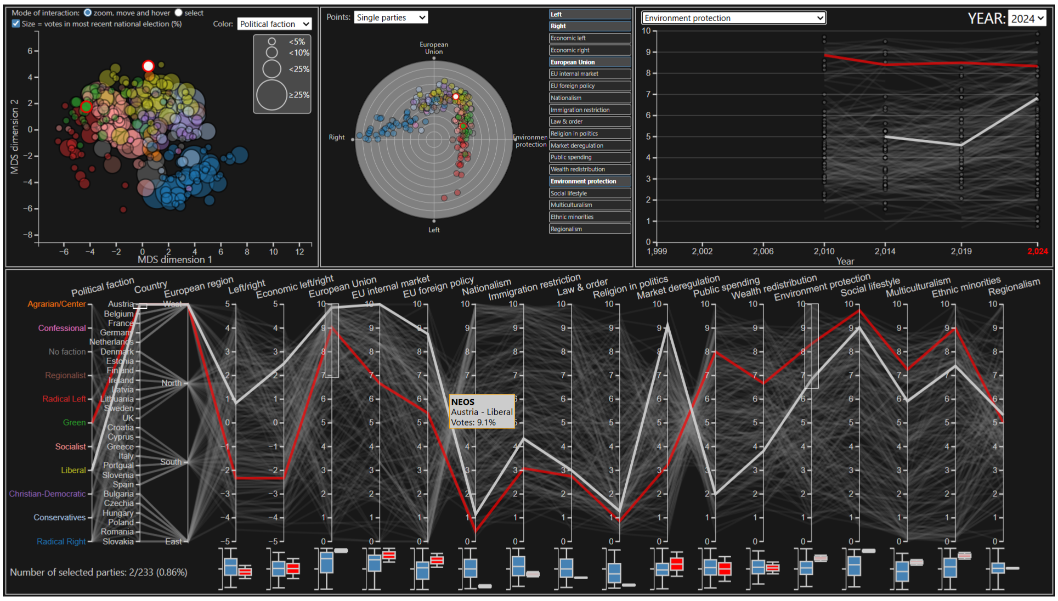


Figure 6.2. Screenshot of European Parties Explorer where the user is trying to find Austrian parties that are pro-EU and environmentalist; one of the two remaining parties is hovered (**NEOS**). For a deeper search, the user could also have brushed the **EU internal market** and **EU foreign policy** axes.

6.2.1 Case study: outlier parties

Thanks to the scatter plot (or while doing brushing), it is possible to see parties belonging to a family but being "far" from it; here are some examples.

- **BSW** (Germany) and **KSCM** (Czechia, Figure 6.3) in 2024: both Radical Left parties somewhat close to the Radical Right too, because of their conservative policies; the same happens also for some Socialist parties, like **SMER-SD** (Slovakia), even though they are not as far from their faction as the Radical Left ones.

- **SDS** (Bulgaria): a Radical Right party that is very close to Liberal and Conservative parties. This is due to a switch to the Radical Right family in its most recent election, since in the previous years SDS was actually Conservative. With respect to 2014, the party's policies have changed a little, but not so dramatically to resemble most Radical Right parties.

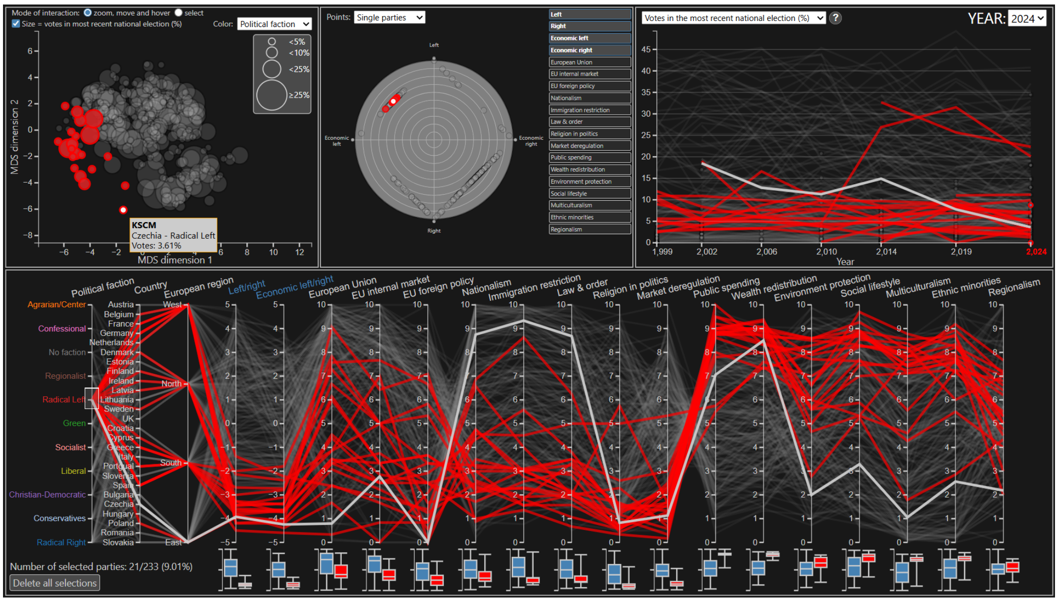


Figure 6.3. KSCM, an outlier radical left party: the scatter plot shows how it is far away from the other red circles, and indeed its line in the parallel coordinates follows a different pattern from the general one.

6.3 Analysis of factions

Let us now see the different classes of the second type of analysis, the aggregate ones. These patterns are more suited for the domain experts who want to discover complex insights, such as political scientists and journalists, as it was explained in earlier chapters. We will start with the analysis focused on factions.
Let us remind in short the main steps of the analysis by factions exposed in Section 4.3.3:

- select a political family and see its parties, its size and its geographical distribution in the European continent;
- study its electoral trend;
- analyse its ideas and trends, both in one year only and in multiple years, and if there are polarising subjects;
- search for outlier parties;
- compare with other groups.

Considering the steps just described and the visualisations that are used in European Parties Explorer, it is possible to notice that we can actually find some further "subclasses" of analysis, depending on which subtasks the user is carrying out. Indeed, while following the steps described above, we are fundamentally doing one of three things:

- analysing only one faction in one year, its patterns, its parties (first, third and fourth subtask);

- making a historical analysis of the faction over the course of many years (first, second and third subtask);

- comparing with other factions (first and fifth subtask).

Of course these are not strict rules, and we can and should use the application intertwining the above subtypes of analysis. For instance, nothing prevents us from looking at the single parties of one chosen family in one year, then compare the family's trends with those of other groups either in the same year or in a longer lapse of years. The presented subclasses of analysis patterns (and even the original "macro"-classes) are just a means to formalise and identify all possible steps to follow while using European Parties Explorer, and to verify that it supports all users and user tasks.

6.3.1 Analysis of one faction in one year

The goal here is to understand the internal composition and homogeneity of a political family, for instance Conservatives in 2014.

- The user starts the analysis by selecting the year 2014 from the drop-down menu and by brushing the Conservatives tick on the **Political faction** axis of the parallel coordinates plot;

- the user can observe how many parties belong to the political group thanks to the counter on the bottom left, and the **Country** axis on the parallel coordinates shows in which countries there are no Conservative parties (Austria, Germany, Netherlands, Finland and Romania);

- the parallel coordinates plot also displays the Conservatives' ideology and trends, for example it is possible to see they are right-leaning, have medium-high evaluations on the axes going from **Nationalism** to **Market deregulation** and medium-low evaluations on all the other axes on the right;

- the red box plots show the internal cohesion of the faction; when inspecting them, we can see for instance very tall box plots under the **European Union**, **EU internal market** and **EU foreign policy** axes, meaning in 2014 Conservative parties were pretty split on the EU, even though most of them were still Europeanist;

- the user can watch on the line chart and the scatter plot how many votes the parties got, and by encoding the scatter plot's bubble color as European region the user can also use the bubbles' size to see in which European regions the most voted conservative parties are;

- furthermore, the scatter plot displays how similar the parties are one another (the faction may be cohesive, or its parties may be distant from each other) and possible outliers; hovering the outlier parties on the scatter plot shows on the parallel coordinates and RadViz why they are classified as such;

- finally, the user can make tests on RadViz by selecting various attributes in order to discover a topic where the faction is “unbalanced” or just how it is distributed between the selected dimensions.

The example is visible in Figure 6.4.

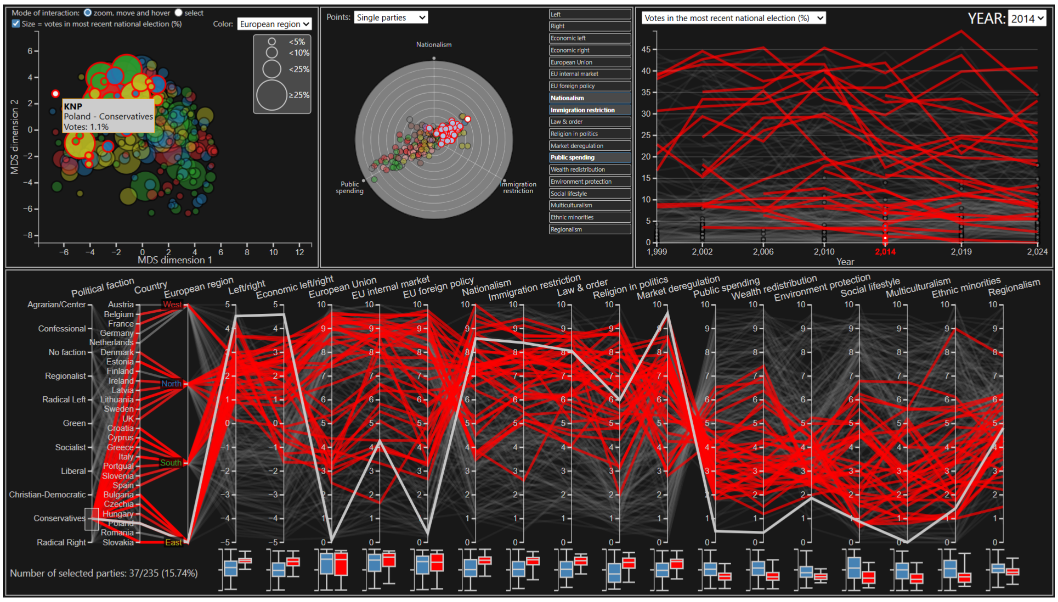


Figure 6.4. European Parties Explorer showing the state of Conservative parties in 2014; the focus here is on the whole faction in that year, not the single parties. Hovering the outlier party **KNP** from Poland shows in the other charts its positioning with respect to all Conservatives.

6.3.2 Temporal analysis of one faction

This subclass of analysis is useful to understand how a political family has evolved over time. For example, the user may be wondering: "Have Radical Right parties moved more to the right? Are they more voted?".

- The user can observe on the line chart the electoral trends of the chosen faction, but they can also choose to see the evolution over the years of one, some or all attributes;
- another way to carry out the temporal analysis is by selecting all years and look for changes in other visualisations; for instance, some box plots may shrink or have some quartiles move in a certain direction;
- other than the box plots, the user can compare the other views, for example finding possible differences in the parallel coordinates patterns;
- RadViz in various years can show how the relationships among chosen attributes have changed;
- while changing years, a useful action is always checking the party counter, which tells the user if the political family has grown across the years.

A comparison between the Radical Right in 2010 and in 2024 can be found in Figures 6.5 and 6.6. Some notable differences are the increased percentage of Radical Right parties, bigger circles in the scatter plot, and more parties with high scores in the **Economic left/right** attribute.

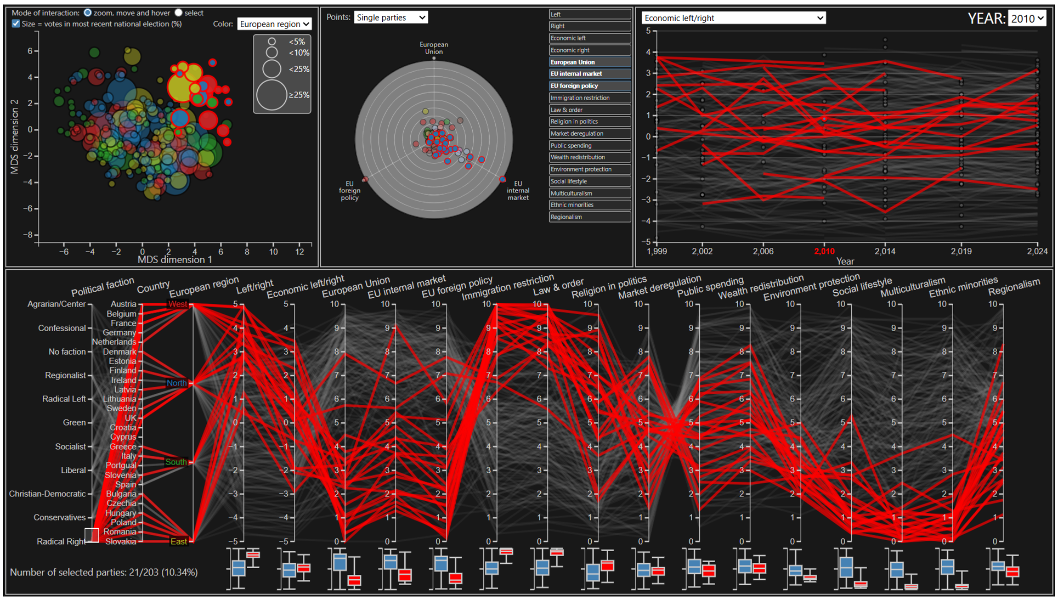


Figure 6.5. European Parties Explorer displaying the state of Radical Right parties in 2010; this is just an example year for when carrying out a temporal analysis, comparing different years.

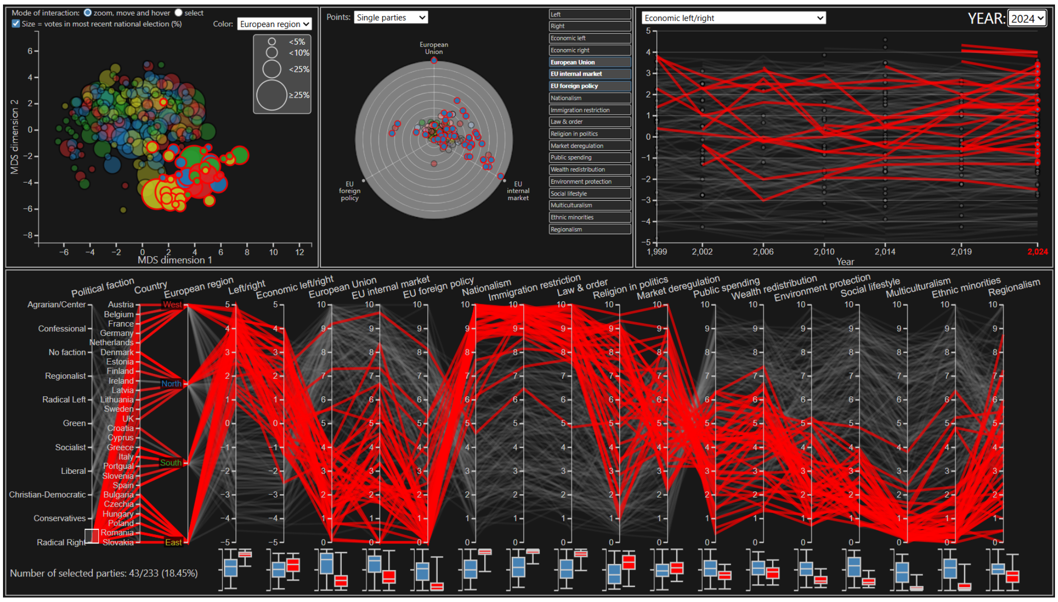


Figure 6.6. European Parties Explorer showing again the Radical Right, but this time in 2024. This setting can be compared with the one of 2010: more parties, different box plots and distribution in RadViz, parties more economically right-leaning.

6.3.3 Comparing more factions

The analysis by faction imposes us to focus not on the single parties, but rather to use a selected family as a starting point in the analysis and treat it mostly as a single entity. However, it is clear that we are not forced to stick with just one faction: useful insights can be obtained also by the comparison among different groups.
Supposing the user has already studied one faction, here is what they can do to compare it to other ones.

- The user can move the brush on the **Political faction** axis of the parallel coordinates plot to switch between groups and observe the diverse trends displayed on the lines, e.g. the Radical Right has very low scores in **Multiculturalism** and **Ethnic minorities**, whereas the Radical Left's scores are very high;
- further insight on how the groups are distributed on the axes is given by the red box plots, changing together with the brush;
- even without brush (but using it helps in isolating the parties of interest), the scatter plot shows which political families are far away from each other, on the basis of all attributes; RadViz instead, displays the differences on the basis of a subset of topics.
- one last important piece of information is the number of parties in each faction, thanks to the counter showing what groups are more populous.

As already explained earlier, these subtypes of analysis should not be considered as unchangeable and independent one another. Rather, they can and should be intertwined and used together, combining the many steps that were described above; for example, comparing more political families is something that comes to naturally carry out in the same selected year, but nothing is preventing us to change the year and make comparisons in the context of a temporal analysis. We are not trying to limit the user, instead we are just trying to formalise how they can interact with European Parties Explorer and ensure that all tasks are supported. The only limits are the application's functionalities and the user's ability to find insights.

6.3.4 Case study: differences and trends among political groups

Let us see what trends we can find in each political family, understanding their ideologies; of course these are general considerations and clearly there are outlier parties.

- Radical Right: by analyzing the parallel coordinates in each available year (Figures 6.5 and 6.6), we can observe that generally, Radical Right parties are (obviously) right-leaning regarding economic topics, mostly Eurosceptic (low scores), and highly against immigration, progressive policies and environmentalism (high scores). They have values ranging from low to high in religious principles and regionalism, so there isn't really a common policy or line of thought on these two topics. When looking at differences between years, it's possible to notice that Radical Right parties have become increasingly more right-leaning on economic topics, especially from 2014 on.

- Conservatives: Conservative parties are right-leaning in economy, similarly to Radical Right ones; in many cases they are even slightly more right-leaning, but still they do not reach the extreme values that some Radical Right parties have. They are generally Europeanist and have medium-high scores on all other topics (immigration, progressive policies, religious principles, environment, regionalism); they are after all "conservative", but they are not as against as the Radical Right which has extremely high values on these attributes.

- Liberals: Liberal parties are centre/centre-right economically, are extremely Europeanist (in fact they are the most Europeanist faction and have been since the first expert survey, as proven by the parallel coordinates and the boxplot), and have average to low values on all other attributes. This is not surprising, since it means they favour generally liberal policies.

- Christian-Democratics: Christian-Democratic parties are pretty similar to Liberal ones, being economically centre/centre-right and Europeanist; the differences are on the remaining topics, given that they are more conservative on immigration and the other attributes, as shown by the average-high values. Christian-Democratics' scores have been very stable throughout the years, with basically no outliers.

- Socialists: when talking about the economy, Socialist parties seem to have convincingly shared policies; their scores on the economical axes show of course that they are left-leaning, but the interesting fact is that while other parties have slightly scattered scores on these axes, Socialists' evaluations in the surveys are instead way narrower. About other attributes, Socialist parties are very similar to Liberal ones, being mostly Europeanist and favouring liberal policies.

- Greens: Green parties are very similar to Socialist parties regarding the economy, but when talking about liberal policies they have more in common with the Radical Left. The most interesting topic to analyse is Europeanism, because it is easy to see that Greens have become increasingly more Europeanist during the years: Green parties' scores on EU attributes have in general steadily increased, and while in 1999 there were both Eurosceptical and moderately Europeanist parties, nowadays almost all of them agree with the EU.

- Radical Left: Radical Left parties share the same narrowness of Socialists on the economical axes, but obviously they are even more left-leaning. They aren't Europeanist, but they are not totally Eurosceptic: their scores go from low to medium-high. They agree on liberal policies, with scores fairly lower than Socialists.

- Regionalist, No faction, Confessional and Agrarian/Centre parties: these groups are discussed all together since they belong to factions that do not necessarily reflect the left/right split; for this reason, there isn't really a common line of thought among these families. The only notable things are that the few Confessional parties are basically conservative and neutral towards the EU, in general the whole four groups are fine with the EU or at least neutral, and unsurprisingly Regionalists favor political decentralisation. The "most regionalist" parties are Spanish.

6.3.5 Case study: the rise of the radical right

A common belief, especially after the latest European Parliament election in 2024, is that the far right has been rising during the last years. Is this confirmed by the data? Indeed, this project proves that it is true. A quick proof is provided by observing the scatter plot during the years, with the dark blue points becoming more and bigger. In fact, by using the counter in the box plot area, we can see that the number of Radical Right parties was pretty stable between 1999 and 2014, then it increased from 20 parties in 2014 to 31 in 2019 and 43 in 2024 (the most numerous family). This is because of new far right parties being born and Conservative parties turning more radical, in fact the number of Conservative parties has decreased from 38 in 2014 to 22 in 2024. The line chart proves too that the Radical Right is nowadays more voted than ever.

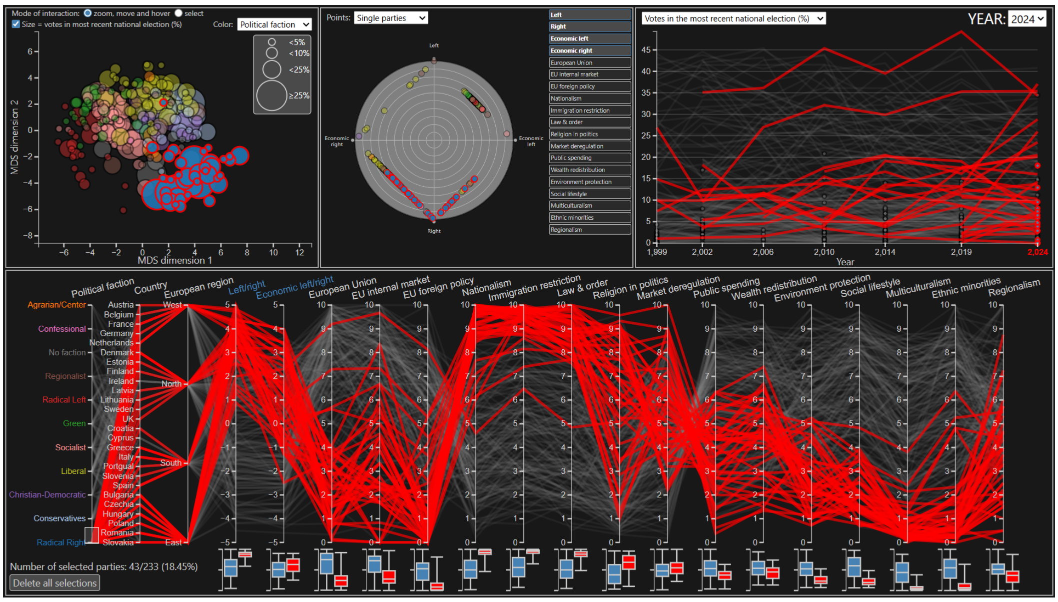


Figure 6.7. The rise of the radical right, shown by the increase in votes share in the line chart, bigger blue circles in the scatter plot, and a greater number in the counter with respect to older years.

6.4 Analysis of countries and European regions

Having seen all ways to conduct a faction-based analysis, we can now discuss about the second type of aggregate analysis, that is the one focused on countries.
Again, let us very shortly list as a reminder the steps of the analysis by country described in Section 4.3.3:

- select a country to observe its parties, the factions that are present and in what percentage of votes;
- study its electoral history;
- observe the parties' patterns and scores and how compact they are on the many topics, whether in one year or more years;
- compare with other nations.

As already explained in the just mentioned chapter, the analysis of countries is pretty similar to the one by factions, at least from the point of view of the actions taken; the real differences are the focus and starting point of the analysis (country vs. political family), and the parties themselves constituting the group that will result from the brushing operation.
There is also another consideration to make. When originally describing the analysis pattern, it made sense to concentrate on countries as the "unitary" element. However, given that now we know how European Parties Explorer is structured and the views it uses, we can actually start the aggregate analysis from an additional, very similar attribute that was not present in the original CHES dataset: the European regions (Northern, Southern, Eastern and Western Europe). Given that in this new case the steps in the analysis are exactly the same (instead of using single countries we are using further groupings of countries), from now on, when describing the analysis pattern and referring to countries, it is implicit that we will be talking about single countries or European regions.

Going back to the analysis pattern by countries, given that it is comparable to the one by factions, it is clear that we can subdivide it in the same three subclasses of analysis described in the previous section (one year, more years, comparison). Even though the steps to undertake are mostly the same, there are still some small differences, thus for the sake of completeness we are listing them again here.

6.4.1 Analysis of one country/region in one year

In this case, the objective is studying the political landscape of a selected nation, e.g. Greece in 2024.

- The user starts the analysis by brushing the Greece tick on the parallel coordinates' **Country** axis while displaying the data for the year 2024;
- the user can see how many tracked parties from Greece there are (eight), as always because of the party counter (and in most cases, since in this type of analysis we are considering only one country instead of a transnational

political faction, the number of parties will be less than the one we had during the analysis by faction);

- thanks to the **Political faction** axis in the parallel coordinates plot and the color encoding from the scatter plot and RadViz, it is possible to notice that only five families are represented in Greece (Radical Right, Conservatives, Socialists, Radical Left, Confessional parties);
- now there is a significant difference with respect to the analysis by faction, because in that case, the brushed lines on the parallel coordinates plot were mostly showing a common pattern (the faction's ideology); in the countries case, the parties are in competition and from different families, therefore there is not in general a common trend, instead the user can inspect the few parties' ideologies;
- the user can observe the red box plots (or even the few lines in the parallel coordinates plot) to understand how divided Greece is on the available subjects;
- as usual, it is possible to check for similar parties in the scatter plot, but as mentioned it is very likely that the bubbles tied to the selected country are pretty scattered around; the scatter plot is probably more helpful in quickly understanding which parties were the most voted in the most recent national election;
- finally, RadViz could show particular patterns with respect to the selected attributes.

A screenshot about this example can be found in Figure 6.8.

6.4.2 Temporal analysis of one country/region

Now the goal is determining the political evolution of a nation or European region over the course of many years.

- The user can utilise the line chart to observe both electoral results and changes in expert evaluations over time;
- just like in the analysis of factions, the user can change the year and look for changes in the many visualisations;
- in this type of analysis, the party counter is just a tool to tell us how many parties were tracked in the selected country in the current year.

In Figure 6.9 we have the political landscape of Greece in 2010. There are some clear differences with respect to 2024: the **Economic left/right** red box plot is shorter (meaning that economically, most parties were center-left leaning); the **Immigration restriction**, **Law & order** and **Religion in politics** red box plots are way higher up in 2024, in particular because of four parties having high scores, which interestingly received actually 0% votes except for one party with a percentage of 4.45; finally, while **ND** (Conservative) was the most voted party in 2024 with

40.79% of votes, in 2010 it was in second place (33.48%), behind only **PASOK** (Socialist, 43.92% of votes).

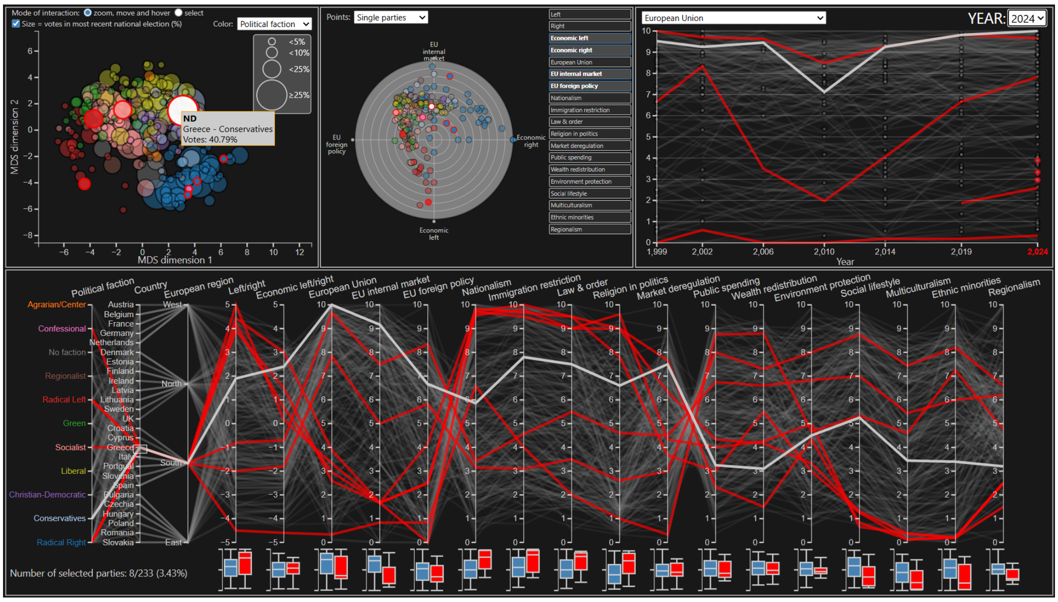


Figure 6.8. Screenshot of European Parties Explorer showcasing the political landscape of Greece in 2024. Contrary to the factions case, the lines on the parallel coordinates are very diverse and the scatter plot's and Radviz' brushed points are all scattered around. The most voted party was **ND** (Conservative).

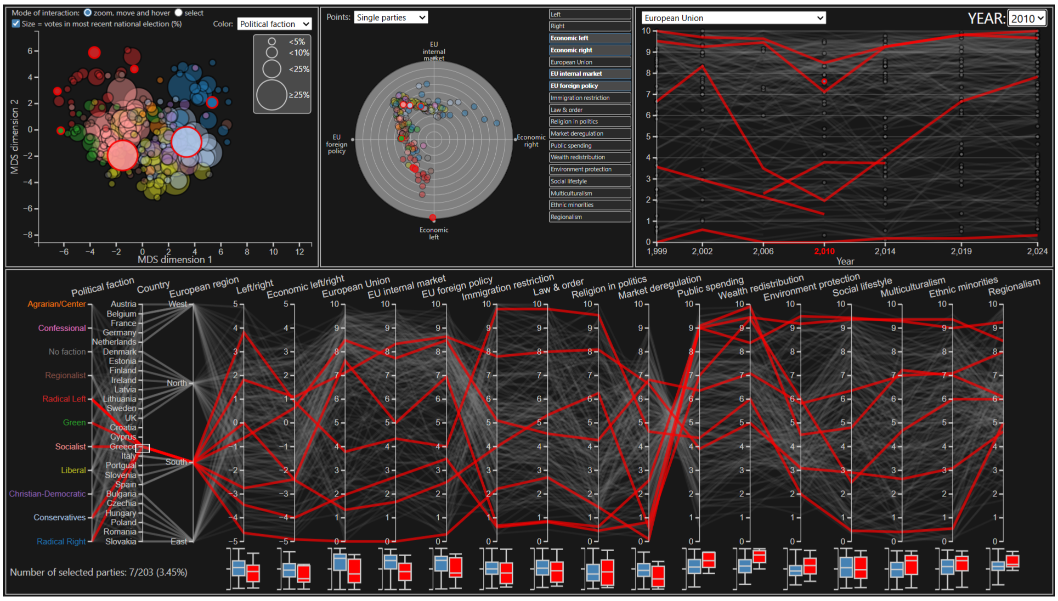


Figure 6.9. European Parties Explorer displaying the Greek situation in 2010, for a comparison with 2024.

6.4.3 Comparing more countries/regions

To conclude the analysis of countries, of course there is also the need to compare with other nations.

- In order to explore different countries, the user can move the brush on the **Country** axis of the parallel coordinates plot; this way, the lines will change and allow the comparison between the two (or more) political landscapes, and to observe what parties are in each nation;
- the box plots, as usual, can demonstrate which country's parties are more divided on certain topics, or if they are shifted towards one of the two axis' extremes;
- the scatter plot and the line charts allow for a quick overview of the latest electoral results, showing if there are some factions in common dominating the elections.

As always, it is a good idea to combine all the above kinds of analysis in order to discover complex insights.

6.4.4 Case study: Brexit and Europeanism

The EU has faced in the past many challenges about its cohesion, with Euroscepticism even leading to Brexit. We saw that actually most political families are in favor of the European Union, with the Greens increasingly more Europeanist, and the general level of Europeanism seems stable enough; we could argue that, if there really is more Euroscepticism, that's because the Radical Right parties have had their share of votes increased.
Speaking of Brexit (2016), let us see the level of Europeanism in the UK. In 2014, most parties were in favor of the EU, except for the Conservatives. In 2019, the two main parties (Conservatives and Labourists) were more Eurosceptic, and both had increased their share of votes (especially the former). In 2024, their opinion about the EU remained stable, and the elections were won by the Labour party thanks to the Conservatives having their votes halved in favor of Reform UK, a Radical Right, highly Eurosceptic party.
In short, while there are many Europeanist parties in the UK, the most voted ones are those indefferent to the EU or completely against it, even though the party that led Brexit had its votes halved form the previous election (Figure 6.10).

6.5 Analysis of topics

Finally, the user may wish to abstract away from parties and countries to focus on the topics themselves.

- The user examines the parallel coordinates. They can look for the relationship between two adjacent axes:
  - parallel lines indicate a positive correlation;.

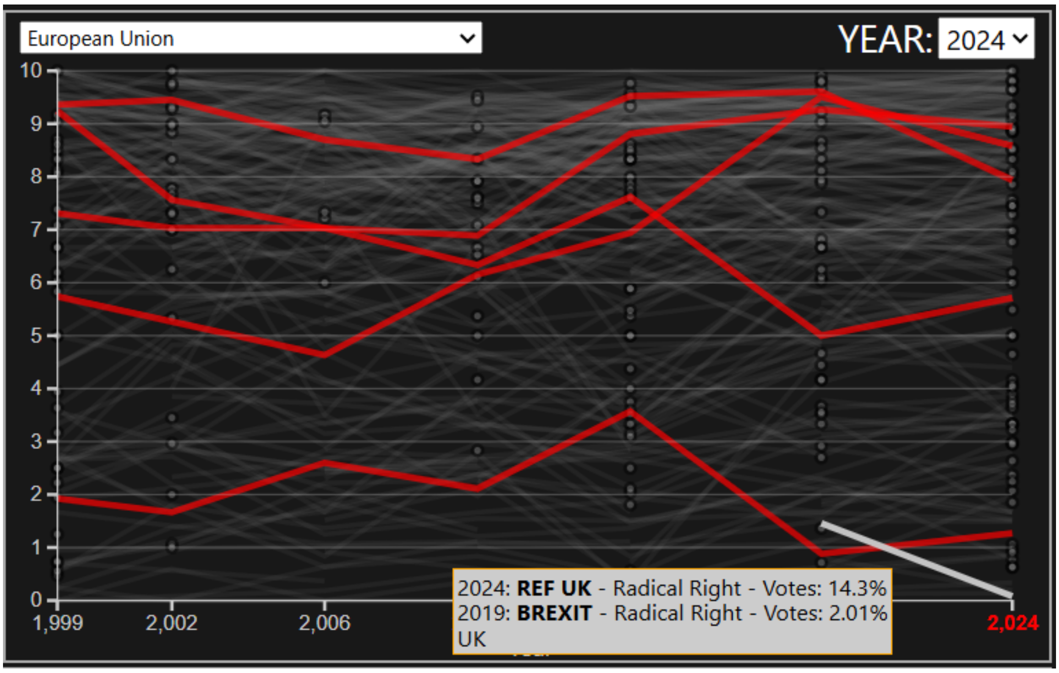


Figure 6.10. Evolution of the level of Europeanism in the UK, with the most Eurosceptic party going from 2% of votes in 2019 to 14% in 2024.

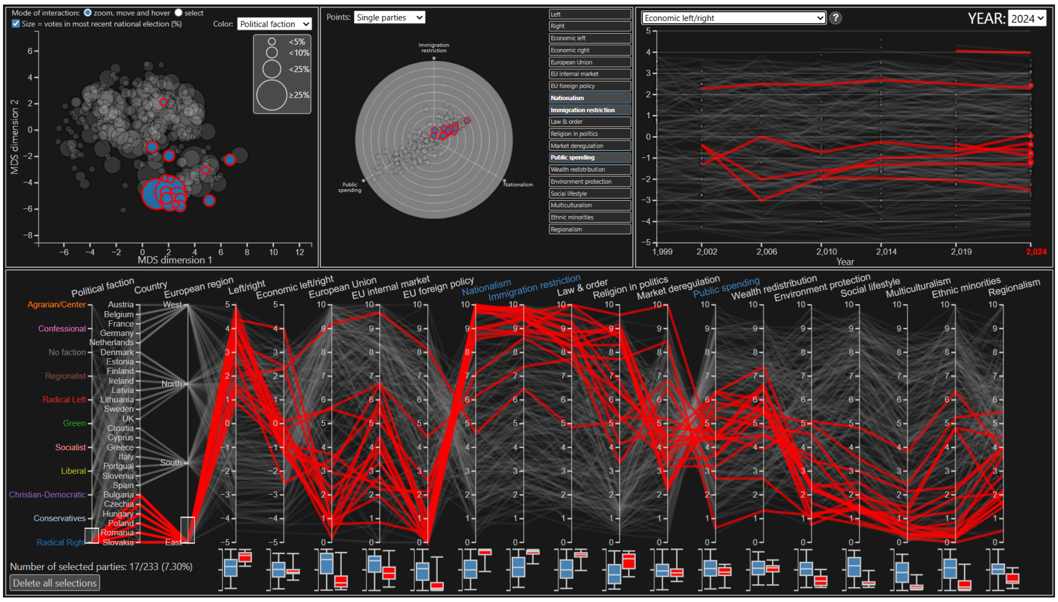


Figure 6.11. The radical right in Eastern Europe, with the views demonstrating how its parties are almost left-wing economically.

- crossing lines indicate a negative correlation.
- scattered lines indicate no correlation.

Furthermore, by moving the brush on one axis it is possible to see how the other axes behave, depending on lower or higher values in the brush.

- The RadViz view is particularly powerful here. The user observes which anchors (variables) are pulling the parties most strongly. If most parties are clustered near the center, it implies that no single topic is polarising the dataset. If parties are pushed to the edges near specific anchors, it identifies those topics as the primary "dividing lines" of European politics. Also, the implementation of "aggregated" Radviz allows to see for instance if a faction has an identity topic distinguishing it from the other groups.

- Using the line chart, the user can see the global trend of a specific issue (e.g., European Union) across all Europe. Rising lines would indicate a continent-wide trend towards Euro-optimism, independent of specific national dynamics.

6.5.1 Case study: economy and radical right in Eastern Europe

One of the most interesting findings from the CHES papers [6] is that the radical right has been moving more and more to the left when dealing with the economy, differently from Western Europe. As shown in Figure 6.11 we can see in fact that, when selecting radical right parties from Eastern Europe, the values in **Economic left/right** are mostly around the axis' centre point, meaning that they are in the middle between left and right. The line chart too shows new radical right parties (the points) being born around middle values.

Chapter 7

Conclusions

The primary objective of this thesis was to design and develop a Visual Analytics system capable of navigating the complexity of the European political landscape through the Chapel Hill Expert Survey. Starting from the observation that high-dimensional political data is often inaccessible to non-experts, we followed a rigorous user-centred design process to build European Parties Explorer.

The development process began with an in-depth analysis of the domain and the identification of three user personas: voters, journalists, and political scientists. By translating their analytical needs into functional requirements, we addressed the challenges of high dimensionality, temporal evolution, and multi-level comparison. The implementation used advanced visualisation techniques:

- parallel coordinates and box plots proved effective in handling the multidimensional nature of party ideologies, allowing the simultaneous comparison of economic, social, and cultural attributes;
- dimensionality reduction (MDS and RadViz) provided users with an intuitive spatial map of political similarity, simplifying the identification of clusters and outliers;
- temporal views enabled the tracking of ideological shifts over a 25-year period.

As demonstrated by the usage scenarios, the system allows users to successfully replicate academic findings, validating its utility as both an educational tool and an instrument for insight discovery.

Despite the successful implementation of the core features, the current system presents some limitations that should be acknowledged.

- The CHES dataset is updated approximately every four years. While sufficient for analysing long-term structural changes, this low frequency prevents the analysis of short-term political events or rapid reactions to crises, limiting the tool's usage for "breaking news" journalism.
- While techniques like parallel coordinates and RadViz are powerful for experts, they have a steep learning curve. Evaluation with novice users (e.g., voters) might reveal difficulties in interpreting these charts correctly.

- The high number of parties and dimensions can lead to visual clutter, particularly on smaller screens. While interaction techniques like brushing and filtering mitigate this, the system is currently optimized primarily for desktop usage.

To further enhance the capabilities of European Parties Explorer, several directions for future development can be identified, focusing on automation, user guidance, and data enrichment.
Currently, the discovery of insights relies entirely on the user's manual exploration. Future versions could implement automated insight discovery algorithms. For instance, the system could automatically calculate and highlight statistical anomalies, such as the most polarised country in a given year. By pre-calculating these metrics server-side, the system could proactively suggest interesting starting points for the analysis, helping the user.
A great improvement would be a guidance engine. This could take the form of guided analytics or interactive tours, where the system walks the user through a pre-defined analysis path. This step-by-step approach would help novice users (such as voters) understand how to read the visualisations before allowing them to explore freely.

In conclusion, this project highlights the potential of Visual Analytics to spread access to complex political data. By providing the tools to explore, compare, and verify positions, we hope to contribute to a more transparent understanding of the forces shaping Europe.

Bibliography

[1] Marco Angelini et al. "Effectiveness Error: Measuring and Improving RadViz Visual Effectiveness". In: *IEEE Transactions on Visualization and Computer Graphics* 28.12 (Dec. 2022), pp. 4770–4786. DOI: 10.1109/TVCG.2021.3104879.

[2] Mihael Ankerst, Stefan Berchtold and Daniel Keim. "Similarity Clustering of Dimensions for an Enhanced Visualization of Multidimensional Data". In: *Proceedings of the International Conference on Information Visualization*. 1998, pp. 52–60. ISBN: 0-8186-9093-3. DOI: 10.1109/INFVIS.1998.729559.

[3] Ryan Bakker et al. "Chapel Hill Expert Survey trend file, 1999–2019". In: *Electoral Studies* 75 (Feb. 2022). DOI: 10.1016/j.electstud.2021.102420. URL: https://www.sciencedirect.com/science/article/pii/S0261379421001323.

[4] Ryan Bakker et al. "Measuring party positions in Europe: The Chapel Hill expert survey trend file 1999-2010". In: *Party Politics* 21.1 (Jan. 2015), pp. 143–152. DOI: 10.1177/1354068812462931.

[5] Ryan Bakker et al. "Reliability and validity of the 2002 and 2006 Chapel Hill expert surveys on party positioning". In: *European Journal of Political Research* 49.5 (Aug. 2010), pp. 687–703. DOI: 10.1111/j.1475-6765.2009.01912.x.

[6] Ryan Bakker et al. "The 2024 Chapel Hill Expert Survey on political party positioning in Europe: Twenty-five years of party positional data". In: *Electoral Studies* 97 (Oct. 2025). DOI: 10.1016/j.electstud.2025.102981. URL: https://www.sciencedirect.com/science/article/pii/S0261379425000873.

[7] Richard A. Becker and William S. Cleveland. "Brushing Scatterplots". In: *Technometrics* 29.2 (May 1987), pp. 127–142. DOI: 10.1080/00401706.1987.10488204. URL: https://www.tandfonline.com/doi/abs/10.1080/00401706.1987.10488204.

[8] Liliana Bounegru, Lucy Chambers and Jonathan Gray, eds. *The Data Journalism Handbook*. 2012.

[9] Luca Carrieri, Nicolò Conti and Marco Morini. "Electoral mobilization, party support and EU issue voting: an analysis of five countries". In: *Italian Political Science Review/Rivista Italiana di Scienza Politica* 53.2 (2023), pp. 220–238. DOI: 10.1017/ipo.2023.1.

[10] CHES Project. *Chapel Hill Expert Survey*. 2025. URL: https://www.chesdata.eu (visited on 31/12/2025).

[11] CHES Project. *Chapel Hill Expert Survey Trend File Codebook (1999-2024)*. Version 2025.1. 2025. URL: https://www.chesdata.eu/s/1999-2024_CHES_codebook.pdf (visited on 31/12/2025).

[12] Catherine E. De Vries et al. *Foundations of European Politics - A comparative approach*. Oxford University Press, 2021.

[13] Marina Dunbar. *Political poll news site 538 to close amid larger shuttering across ABC and Disney*. The Guardian. 5th Mar. 2025. URL: https://www.theguardian.com/us-news/2025/mar/05/abc-news-538-shut-down (visited on 15/01/2026).

[14] FiveThirtyEight. *FiveThirtyEight (Archived)*. Archived via Wayback Machine. 2023. URL: https://web.archive.org/web/20230130224343/http://fivethirtyeight.com/ (visited on 15/01/2026).

[15] Sara B. Hobolt and Catherine De Vries. “Turning against the Union? The impact of the crisis on the Eurosceptic vote in the 2014 European Parliament elections”. In: *Electoral Studies* 44 (Dec. 2016), pp. 504–514. DOI: 10.1016/j.electstud.2016.05.006. URL: https://www.sciencedirect.com/science/article/pii/S0261379415300603.

[16] Patrick Hoffman, Georges Grinstein and David Pinkney. “Dimensional Anchors: A Graphic Primitive for Multidimensional Multivariate Information Visualizations”. In: *Proceedings of the 1999 Workshop on New Paradigms in Information Visualization and Manipulation (NPIVM ’99)*. ACM, 1999, pp. 9–16. DOI: 10.1145/331770.331775.

[17] Alfred Inselberg. “The plane with parallel coordinates”. In: *The Visual Computer* 1.2 (Aug. 1985), pp. 69–91. DOI: 10.1007/BF01898350. URL: https://doi.org/10.1007/BF01898350.

[18] Daniel Keim et al., eds. *Mastering the Information Age: Solving Problems with Visual Analytics*. Eurographics Association, 2010. ISBN: 978-3-905673-77-7.

[19] Glenn E. Krasner and Stephen T. Pope. “A cookbook for using the model-view controller user interface paradigm in Smalltalk-80”. In: *Journal of Object-Oriented Programming* 1.3 (Oct. 1988), pp. 26–49.

[20] Laurens van der Maaten, Eric Postma and Jaap van den Herik. *Dimensionality reduction: a comparative review*. Technical Report TiCC TR 2009-005. Tilburg University, 2009.

[21] Gary Marks and Marco Steenbergen. “Evaluating expert judgments”. In: *European Journal of Political Research* 46.3 (May 2007), pp. 347–366. DOI: 10.1111/j.1475-6765.2006.00694.x.

[22] Trygve Reenskaug. *Models - Views - Controllers*. Technical Note. Xerox PARC, 1979. URL: http://heim.ifi.uio.no/~trygver/1979/mvc-2/1979-12-MVC.pdf.

[23] Trygve Reenskaug. *The Model-View-Controller (MVC): Its Past and Present*. 2003. URL: https://folk.universitetetioslo.no/trygver/themes/mvc/mvc-index.html (visited on 08/01/2026).

[24] Nate Silver. *The Signal and the Noise: Why So Many Predictions Fail - but Some Don't*. Penguin Press, 2012.

[25] James J. Thomas and Kristin A. Cook, eds. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society Press, 2005.

[26] Warren S. Torgerson. "Multidimensional scaling: I. Theory and method". In: *Psychometrika* 17.4 (Dec. 1952), pp. 401–419. DOI: 10.1007/BF02288916. URL: https://doi.org/10.1007/BF02288916.

[27] John W. Tukey. *Exploratory Data Analysis*. Addison-Wesley, 1977.

[28] United Nations Statistics Division. *Standard Country or Area Codes for Statistical Use (M49)*. 2026. URL: https://unstats.un.org/unsd/methodology/m49/ (visited on 07/01/2026).

[29] Andreas Veglis and Charalampos Bratsas. "Reporters in the age of data journalism". In: *Journal of Applied Journalism & Media Studies* 6.2 (June 2017), pp. 225–244. DOI: 10.1386/ajms.6.2.225_1.

[30] Michelle Q. Wang Baldonado, Allison Woodruff and Allan Kuchinsky. "Guidelines for using multiple views in information visualization". In: *Proceedings of the Working Conference on Advanced Visual Interfaces*. ACM, 2000, pp. 110–119. ISBN: 1581132522. DOI: 10.1145/345513.345271. URL: https://doi.org/10.1145/345513.345271.