European Parties Explorer - Visual Analytics project

Stefano Bonetti, 1764624

Abstract

In this ever changing world, the European Union and, more generally, the European continent have been facing several challenges during the last twenty-five years: some of the most important ones are the 2008 economic crisis and Russian invasion of Georgia, the migrant crisis, episodes of Islamic terrorism, EU cohesion and Brexit, climate change, the Covid pandemic, the rise of authoritarian governments, the Russian invasion of Ukraine and European defense. All of this inevitably affects European citizens and countries, and politics is becoming increasingly more present in everyday life; it is probably safe to assume that basically all European political parties have evolved in one way or another, and many more were born during the last years. For these reasons, this paper will present a visual analytics application aimed at showing the evolution of European parties and their beliefs.

1 Introduction

European Parties Explorer is a visual analytics system, developed with the D3.js library, suited for tasks like exploring how parties from most EU countries and the UK have evolved in the last years, how their views over the most important topics have changed, comparing parties in one or many countries, viewing how polarized nations or political factions are on certain topics, and possibly discovering further similar insights. The political parties are represented in different charts and can be filtered according to the year, the country they come from and their faction (left, right, etc.), but it is also possible to make selections directly on the available charts.

European Parties Explorer is based on the datasets provided by the Chapel Hill Expert Surveys, a set of surveys conducted in 1999, 2002, 2006, 2010, 2014, 2019 and 2024, which were compiled by many political scientists so that they could evaluate the most important parties in many European countries during the years. To fully understand this visual analytics application, we first need to analyze and understand the provided data.

2 Dataset

2.1 Dataset description

The data used by this project was built from two files available on the CHES website:

- the *1999-2019 Chapel Hill Expert Survey trend file*,[1] a single file containing the data from all surveys conducted between 1999 and 2019;
- the *2024 Chapel Hill expert survey*.[2]

The two csv files have a similar structure, with the former having a total of 1196 entries from 1999 to 2019, the latter 279 entries for the year 2024. Each row represents a single European party analyzed during one of the years in which the surveys were conducted; the columns contain either data about the party (e.g. its country, unique identifier, name abbreviation, political family) or a value, contained in a certain range (in most cases 0-10), assigned by the political scientists expressing in general how much the party agrees or disagrees with a certain topic or policy in the given year.

Considering that we have two distinct csv files and that the many surveys were often conducted in different ways during the years (in fact most questions were asked only in some of the various years; this means that the topics they are bound to were not evaluated during the remaining years, which implies there are a lot of missing values in the dataset), some form of data preprocessing is needed in order to handle the missing data and create a single coherent data file.

2.2 Data preprocessing

As we saw, the many surveys were brought out in distinct ways during the years, and they were always composed by different questions from one another. Since the 1999-2019 trend file gathers all the surveys up to 2019, it contains a lot of columns (around eighty) with most of them storing data only for some years; the first step in data preprocessing is therefore

---

[1] Jolly, Seth, Ryan Bakker, Liesbet Hooghe, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen, and Milada Anna Vachudova. 2022. "Chapel Hill Expert Survey Trend File, 1999-2019." Electoral Studies. https://doi.org/10.1016/j.electstud.2021.102420

[2] Rovny, Jan, Ryan Bakker, Liesbet Hooghe, Seth Jolly, Gary Marks, Jonathan Polk, Marco Steenbergen, and Milada Vachudova. "25 Years of Political Party Positions in Europe: The Chapel Hill Expert Survey, 1999-2024," working paper.
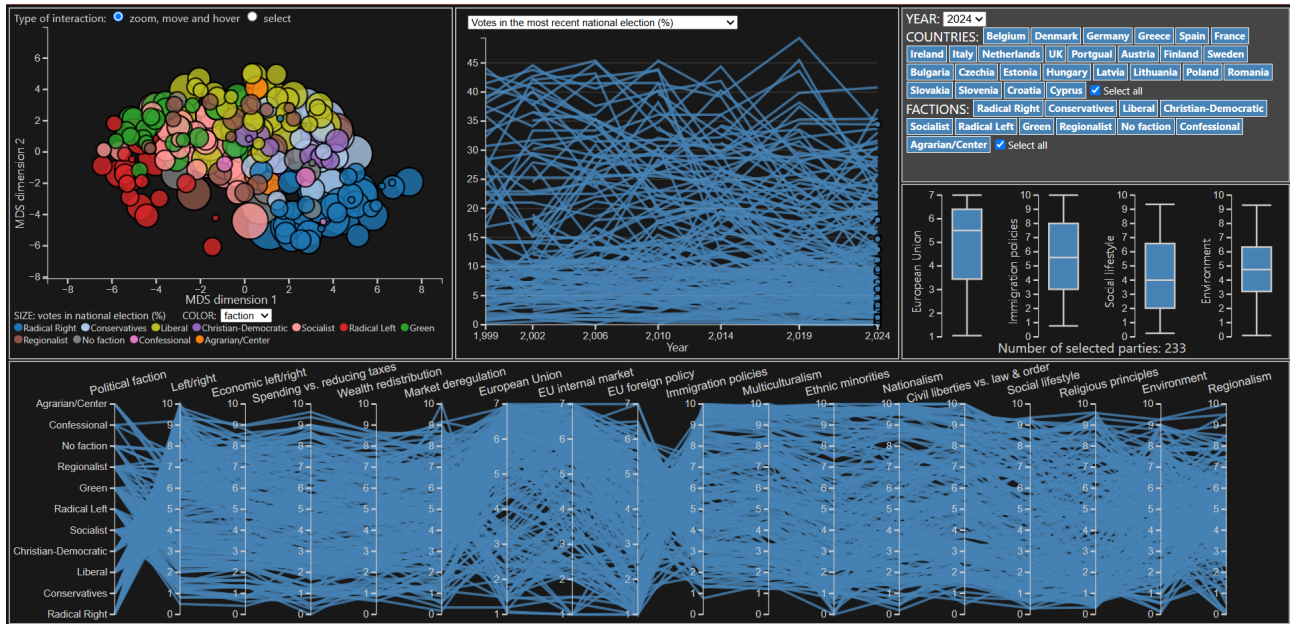
Figure 1: *Preview of the system*

deciding only what columns (basically the topics on which the parties were evaluated) we want to keep and analyze. Considering that the project wants to visualize the evolution and comparison of political parties, the columns we need are:

- the ones storing parameters essential for the application's functioning (like a party's unique identifier and country);
- columns regarding important and non-redundant topics (e.g. economy, environment, immigration);
- columns about topics evaluated during at least a good amount of consecutive years.

Here you can find the list of attributes that in the end were kept from the 1999-2019 file and used for this project and how they are mostly described on the CHES site (check their codebook for more details).

- **country**: unique (numerical) identifier for each country (where the party comes from);
- **year**: year for which party experts were asked to evaluate (1999, 2002, 2006, 2010, 2014, 2019);
- **party_id**: unique identifier for each party;
- **party**: party abbreviation;
- **vote**: vote percentage received by the party in the national election most prior to the given year;
- **seat**: seat share of the party in the national election most prior to the given year;
- **epvote**: vote percentage received by the party in the European Parliament election most prior to the given year;
- **family**: the party's political faction (1 = Radical Right, 2 = Conservatives, 3 = Liberal, 4 = Christian-Democratic, 5 = Socialist, 6 = Radical

Left, 7 = Green, 8 = Regionalist, 9 = No family, 10 = Confessional, 11 = Agrarian/Center);
- **eu_position**: overall orientation of the party leadership towards European integration in the given year (1 = strongly opposed, 2 = opposed, 3 = somewhat opposed, 4 = neutral, 5 = somewhat in favor, 6 = in favor, 7 = strongly in favor);
- **eu_intmark**: position of the party leadership in the given year on the internal market (i.e. free movement of goods, services, capital and labor; 1 = strongly opposed, ... 7 = strongly in favor);
- **eu_foreign**: position of the party leadership in the given year on EU foreign and security policy (1 = strongly opposed, ... 7 = strongly in favor);
- **lrgen**: position of the party in the given year in terms of its overall ideological stance (0 = extreme left, 5 = center, 10 = extreme right);
- **lrecon**: position of the party in the given year in terms of its ideological stance on economic issues; 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 = center, 10 = extreme right);
- **spendvtax**: position on improving public services vs. reducing taxes (0 = strongly favors improving public services, 10 = strongly favors 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

favors 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);
- **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 favors a liberal policy on immigration, 10 = strongly favors a restrictive policy on immigration);
- **multiculturalism**: position on integration of immigrants and asylum seekers (0 = strongly favors multiculturalism, 10 = strongly favors assimilation);
- **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 decentralization to regions/localities (0 = strongly favors political decentralization, 10 = strongly opposes political decentralization);
- **ethnic_minorities**: position towards ethnic minorities (0 = strongly supports more rights for ethnic minorities, 10 = strongly opposes more rights for ethnic minorities);
- **nationalism**: position towards cosmopolitanism vs. nationalism (0 = strongly promotes cosmopolitan conceptions of society, 10 = strongly promotes nationalist conceptions of society).

At this point, processing of the first file is done, since I only needed to decide which data to use. The 2024 expert survey file has some differences with respect to the 1999-2019 file, therefore a little more work on the former was necessary to make it coherent to the latter:

- the 2024 has no *year* column (since all the data refers to the same year), thus it was added storing of course "2024" as value;
- the 2024 file has no *social lifestyle* column, but it has two new attributes called **womens_rights** (position on policies supporting women's rights, e.g. equal pay, family leave, reproductive health) and **lgbtq_rights** (position on policies supporting LGBTQ+ rights, e.g. marriage equality, adoption rights, transgender rights); given that they

are important topics but that the first dataset only has the *social lifestyle* column, I decided to join them into a new single column called again **sociallifestyle** by computing the average between the two values;
- of all the columns in the dataset, I again kept only the ones we already saw for the first file.

Having now decided which data is needed for the project, the next step is merging the two csv files in one single coherent dataset; then I made the following operations on this new file.

- Deleting the data regarding Türkiye, Norway, Switzerland and Iceland (because they are present only in 2024 and have no evaluation for the European Union questions), Malta (since there is too much missing data for 2024 and Malta only has two parties with, at least from 2014 to 2024, basically always the same electoral results), and Luxembourg (it has too many empty values and no data for 2024): removing whole countries is not ideal, but since these ones have a good amount of missing data and that the choice was between adding them but removing the evaluations for the EU from everyone, or not including them and having a more accurate analysis of the other relevant nations, I preferred removing them;
- handling the missing data: even in the remaining countries, it happens that some parties have missing values for some columns; for simplicity and for not distorting the data, I decided to just delete the rows of parties who have no data for one or more topics which should have been evaluated in the given year (luckily, we're talking mostly about parties with a small vote share, between 0 and 2%). The computed coordinates were then added in the dataset as two new columns.

2.3 Dimensionality reduction

The last data preprocessing operation is dimensionality reduction, so that in the application it will be possible to visualize on a scatter plot how similar or not the many parties are. The algorithm that was used is MDS, which was applied by building a dissimilarity matrix for each year based on Euclidean distance between parties on the columns regarding the topics on which they were classified (*eu_position, eu_intmark, eu_foreign, lrgen, lrecon, spendvtax, deregulation, redistribution, civlib_laworder, sociallifestyle, religious_principles, immigrate_policy, multiculturalism, environment, regions, ethnic_minorities, nationalism*). The computed coordinates were added in the dataset as two new columns.

3 Views

We can now start analyzing the project itself and its various components.

3.1 Filters

Figure 2: *Filters view with 2024 as selected year, Germany, France and Italy as selected countries, and all political factions*

The filters section allows us to choose for what year we want to conduct the analysis, what countries, and what political groups; remember that the available years are the ones in which the expert surveys were conducted (1999, 2002, 2006, 2010, 2014, 2019 and 2024). Of course these filters affect all the other views we are about to see, which will display data only for the selected year, countries and political families.

3.2 Parallel coordinates
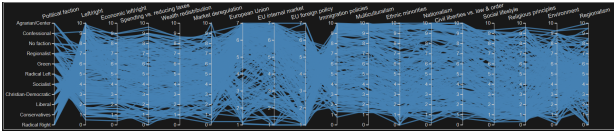
Figure 3: *Parallel coordinates plot (year 2024, all countries and factions)*
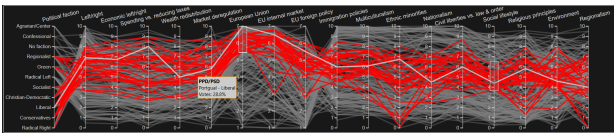

Figure 4: *Parallel coordinates plot with brushing and hovering*

The parallel coordinates visualize many attributes all at the same time, with a single piece of data being represented by one line connecting its associated values for all attributes. In this case, each line encodes a single party, and the represented attributes are the same used in the dimensionality reduction plus the political family. It is possible to do brushing with your mouse on the axes (that is, selecting subsets of parties representing some user-defined constraints) with the selected parties being highlighted, and when

hovering a single line you get information about that party's abbreviation, country, family and share of votes all in the year selected in the filters view. Notice that the attributes were reordered so that axes bound to one general topic are close together, which makes it easier to do comparisons and notice trends inside that topic. In particular, we have in order: three axes about political families (with the last one regarding the economy too), then three axes strictly about economic issues, three axes regarding the EU, four axes roughly about immigration and foreign cultures, then five axes on other topics. Also, since as we saw in the preprocessing section, many topics were evaluated by the experts only in more recent years, the parallel coordinates chart is updated depending on the year chosen in the filters view, visualizing only the axes about topics of that year.

Finally, it is possible to hover over the axes' legends in order to obtain a brief description of the attribute they encode and what the minimum and maximum values mean.
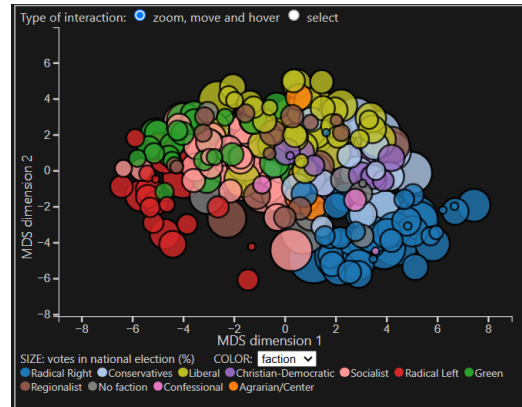
3.3 Scatter plot

Figure 5: *Scatter plot (year 2024, all countries and factions, color by factions)*
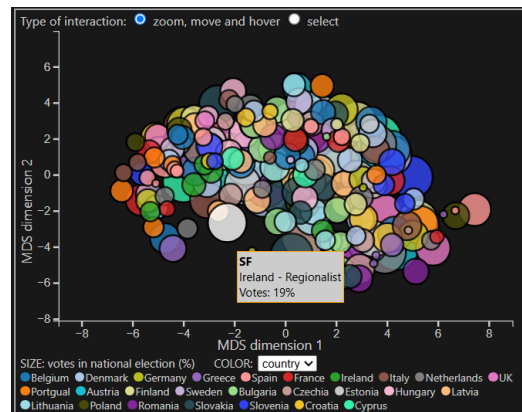

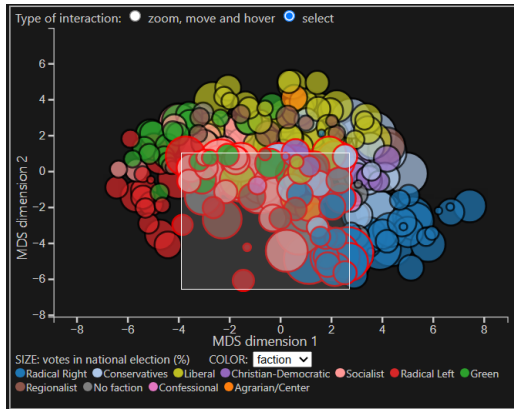**Figure 6:** *Scatter plot with hovering (color by countries)*

Figure 7: *Scatter plot with brushing*

Figure 9: *Line chart on a topic evaluated only from 2006 onwards, with brushing and hovering*

In the scatter plot, parties are encoded as circles positioned according to the coordinates computed with MDS so that it is possible to have an idea of which parties are similar, and their size depends on the votes received (parties with a higher vote share are bigger). The points are colored according to either the parties' political family or the nation they come from, at the user's choice. The colors were taken from D3's `d3.schemeCategory20` categorical color scheme, and in the case of coloring by political group, colors typically associated with each family were mostly respected (red for left-wing parties, green for greens, yellow for liberals, blue for right-wing parties). Below the chart there is a color legend that updates according to the selected coloring mode.

The scatter plot con be interacted with in two different ways: the user can either freely zoom in and out, move the points and hover over them (with a tooltip showing information about the hovered party, just like in the parallel coordinates), or they can make brushing, thus highlighting the chosen circles.
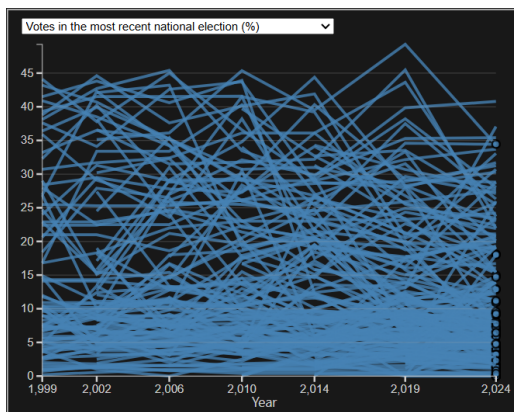
3.4 Line chart

Figure 8: *Line chart (year 2024, all countries and factions)*

While the charts we have seen allow the user to witness the "state" of European political parties in-

side one single year (they are basically a "picture" of what parties look like in the selected year), the line chart allows to have a quick overview of the parties' opinion over the years regarding one chosen topic, or their electoral results. As usual, the displayed parties are the ones existing in the year selected in the filters section and respecting the other filters. The political parties are encoded as lines spanning over the years in which they existed and in which the selected attribute was computed, unless we have a party that existed only in the filtered year; in this case, it is represented with a point.

Thanks to a drop-down menu, it is possible to choose what attribute we are interested in. These attributes are the same found in the parallel coordinates plot except the political family, and with the addition of:

- votes share in the national election most prior to the given year;
- seat share in the national election most prior to the given year;
- votes share in the European Parliament election most prior to the given year.

As for all the above plots, the user can hover over a line or point to observe a tooltip giving information about the hovered party in the filtered year. There is a difference with brushing which can not be carried out on the line chart itself, but the lines are highlighted according to the brushing made on the parallel coordinates and the scatter plot.

3.5 Box plots

The last element remaining in this project is the box plots area. It is an additional analytics part constituted of four box plots, each one showing how parties are distributed over the values associated to four attributes (*eu_position*, *immigrate_policy*, *sociallifestyle*, *environment*). I chose these attributes as I feel like
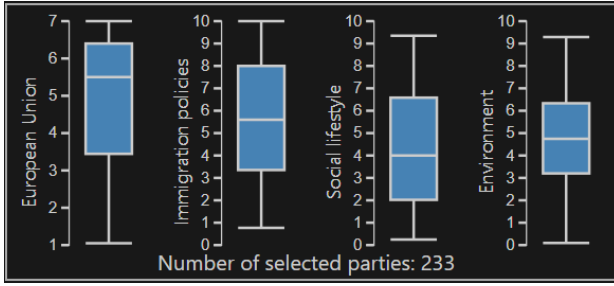
Figure 10: *Boxplots (year 2024, all countries and factions)*

they represent well enough "macro-topics" we mentioned in the parallel coordinates section (European Union, immigration, and two important topics in the remaining ones).

The box plots allow for a quick overview of how parties are distributed about the four attributes, simplifying the study of how much they agree or are polarized for instance among specific political factions or countries (thanks to the filters). The four box plots update according to the global filters and the brushing made on the scatter plot and the parallel coordinates, and in order to know how "valuable" they are, there is also a counter displaying how many parties the box plots are taking into account (that is, the number of brushed parties or all parties if instead there are no selections).

3.6 Interactions between views

To sum up, we have filters (year, countries and factions), scatter plot, line chart, parallel coordinates plot, and four box plots. The parties considered when drawing the various charts are only the ones existing in the selected year and belonging to the filtered countries and factions. It is possible to hover the data on the scatter plot, line chart and parallel coordinates, displaying information about that party in the given year; when you hover a party in one of these charts, the same party is highlighted in the other two. Brushing is possible on the scatter plot and parallel coordinates, highlighting the selected parties in the other view and in the line chart. The box plots visualize the distribution of all parties in the chosen year, countries and political groups if no brushing was made, otherwise they do the same thing but only for the brushed parties.

4 Related work

Now that we fully know the many parts of this project and how they work, we can analyze similar systems and compare them to European Parties Explorer, finding similarities and differences.

When looking for systems similar to this project I re-

alized I wasn't having much success, in my opinion because the CHES dataset can be considered pretty niche. When talking about political parties, the only visualizations that can easily be found are the ones regarding electoral results and similar, therefore there isn't really something around about the evolution of parties (rather than their results at the elections or polls). Still, I managed to find a couple of systems using the CHES data itself.

4.1 CHES Interactive

The CHES website provides an interactive section with three visualizations using their own data:

- a **bar chart**, where you can choose a nation, a year and one topic, and see how that country's parties are evaluated on the chosen topic thanks to vertical bars; hovering a party gives its ranking on the selected attribute;
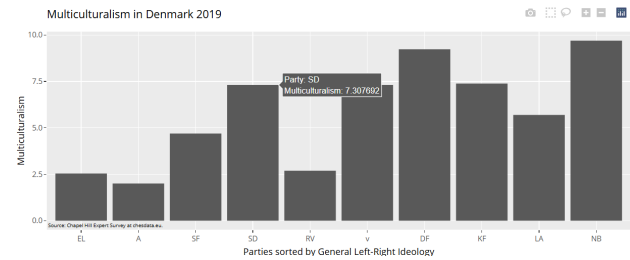

Figure 11: *CHES Interactive's bar chart*

- a **scatter plot on one country**, where the user can choose again country and year but also two attributes for the chart's axes, and the parties are placed accordingly; their size reflects the vote share, and hovering gives some information;
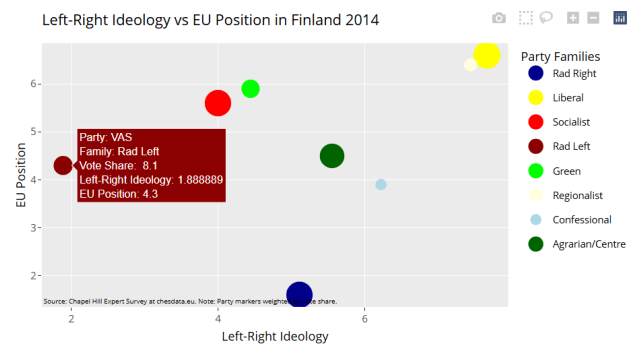

Figure 12: *CHES Interactive's one country scatter plot*

- a **scatter plot on all countries** that works exactly like the above scatter plot, except that of course it displays all nations.

CHES Interactive is an excellent tool for quickly comparing political parties, especially within a single country, but it's missing many features that this project is trying to implement:
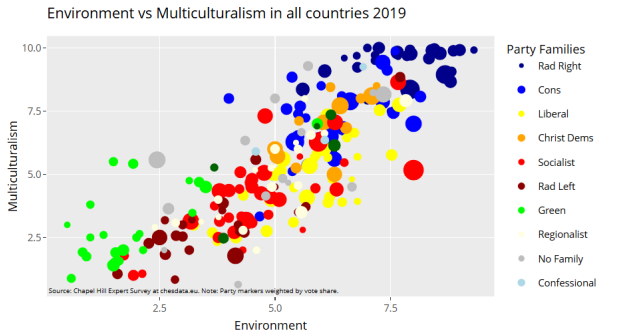
Figure 13: *CHES Interactive's all countries scatter plot*

- the only filters on the parties are year and country (one at a time), and the year 2024 is missing for now;
- the comparison between parties is very effective, but it happens only regarding one or two topics at a time; with European Parties Explorer, you can quickly compare all available attributes thanks to the parallel coordinates, and instead of having a "picture" of one year at a time, you can also see how parties evolved on one topic with the line chart;
- CHES Interactive is very good at showing trends and correlations among attributes, and this is true also in European Parties Explorer (for instance there is a clear correlation in figure 13, but it's easy to notice it also in the parallel coordinates by comparing the appropriate axes);
- parties are placed on the scatter plots using two attributes, but using dimensionality reduction you can observe a more "universal" positioning discovering what parties are similar on a general basis;
- there are no explanations about the many topics and their minimum and maximum values, but this is understandable since we are using this system on the CHES website itself.

4.2 Foundations of European Politics

Foundations of European Politics - A Comparative Approach[3] is a textbook for students about research on European Politics. The book's website describes many political datasets including the CHES one, and presents some related visualizations.

In the section dedicated to the CHES data we can find, similarly to CHES interactive, a scatter plot that positions parties according to two selected variables. The interaction with this chart is very limited, there are no filters and the user can choose among a small number of topics, but this is understandable since the plot was made as part of an exercise from the

---

[3] Catherine E. De Vries, Sara B. Hobolt, Sven-Oliver Proksch, Jonathan B. Slapin

textbook. A very interesting feature is the addition of a line explicitly highlighting the potential correlation between the two variables.
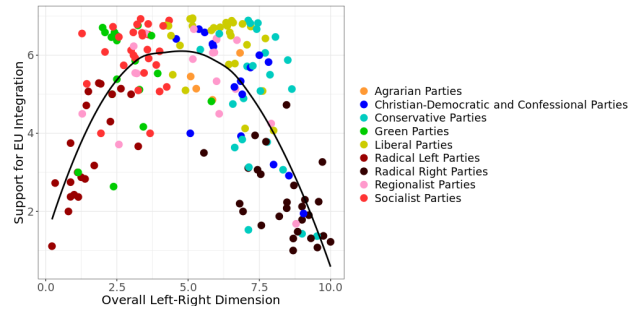

Figure 14: *Foundations of European Politics' scatter plot*

4.3 Election results charts

When talking about political parties, the first thing that comes to mind is electoral results. For this reason, visualizations about elections are very important, and were among the first results when looking for reference material during the first stages of creating this project. Great inspiration came from *Politico*'s visualization of European election results, even though elections are not European Parties Explorer's main focus.

Not only does Politico's visualization display the distribution of seats in the European parliament, it is also very interactive. The user can compare the 2024 results with those from 2019 and can check which and how many seats belong to each single country; it is also possible to see exact election results in one nation. This system's functionalities are pretty far from the ones of this project, but it still was a good reference and source of inspiration.
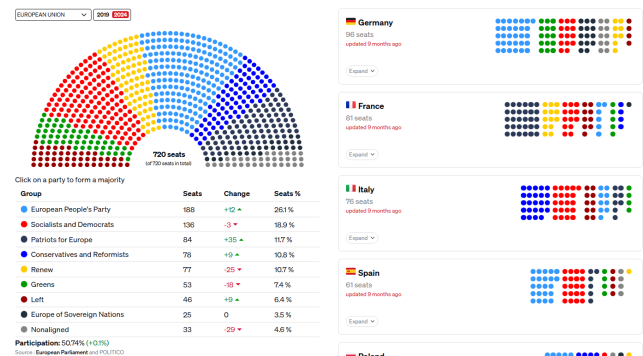

Figure 15: *2024 European elections results on politico.eu*

5 Possible users

6 Insights