How does Bogota vote?

By: Alvaro González, Victor Garzón y Dayana Rojas

***Abstract*— How does Bogota vote is a project that with the help of the election observation mission and Diego Laserna was able to explore in a set of extensive data obtained from the registration and deliver from these, graphical analysis and tools for data exploration in order to enable any type of person a portal that allows analyzing the behavior of the parties and the types of voting in time.**

The analysis was divided into three visualizations: localities (spatial visualization), choices (radial visualization) and correlations (visualization of correlations) and their insights and analyzes performed by the work team are documented in this paper.

I. INTRODUCTION

The monitoring of the behavior of the political parties during the time and during the different elections has not been carried out in depth because the main objective after an election is the monitoring of the candidates, but validating the behavior of the parties can show political formulas to obtain voters. The objective of this visualization is to allow the exploration of voting data from 2006 to 2018 pointing to the identification of these behaviors.

II. STATE OF THE ART

Currently, when searching on the web for a solution or technological tool that can perform descriptive or predictive analyzes on voting in Colombia, we find articles from the most important media in the country, and institutions such as the EOM (Electoral Observation Mission) where they make their best efforts to inform and determine the next scenario according to surveys they perform through their website or to describe what happened in past elections through the data thrown by the registry. It can be said that, for the surveys, although it seems to be a good strategy, it is biased by the small sample and the public that can access these means of communication. Therefore the analysis that is made from the data collected, may be somewhat removed from reality. The above, without counting other variables of great importance, which are not collected and which, in turn, can yield conclusions that enrich the predictions.

III. PROPOSED METHOD

To provide the community with a solution to analyze Colombia's election data from the city of Bogotá to perform descriptive and predictive analyzes of voter behavior and correlations between the different voting categories, we proposed a series of visualizations in a technological tool called d3, using the framework for visualizations of Tamara [1].

Within the framework of Tamara, 3 questions to be resolved (WHAT, WHY and HOW) are defined to determine the most effective way to present the information to the end user, in order to find relevant Insights for your business or objective.

As the main input the MOE, provided us with the list of the voting data of Senate, House of Representatives, Presidency, City Hall, Council and Jal for Colombia, during the last 4 periods, corresponding to 2006, 2010, 2014 and 2018 in formats Excel and CSV.

WHAT

Two different datasets were used: 1) A multidimensional table that was generated based on the data of the votes of the Senate, House of Representatives, Presidency, City Hall, Council and consolidated Jal of the last 4 electoral periods in the country, provided by the MOE and 2) Geographical data (geometry - Geojson format) from the UBER application that has as its content the Map of Bogotá [2].

Data Transformation

The data was delivered in Excel and CSV files, but without a standardized format and its understanding was complex, despite having a data dictionary. Another drawback was that the codes used to identify political parties changed from period to period, as well as the

name of the party.

To solve all these problems, a database was implemented in Mongo with several utilities in charge of loading, cleaning and data quality.

Attribute:

Attribute	Description	Type
Year	Electoral period	Order Quantitative Sequential
Zone	Zone of vote	Categorical
Party	Political party	Categorical
Vote	Amount of votes	Order Quantitative Sequential
Type	Category of vote: Senade, House of representatives, Presidency, Town Hall, etc	Categorical
Correlation	Correlation between type of votes (Derived Field)	Order-Quantitative, Diverging

WHY

A. Summarize Trends

Understand the historical evolution of the distribution of votes by parties within the city of Bogotá

B. Identify Dependencies

Identify the vote by party according to the different sectors of the city.

C. Discover and Compare Trends

Discover and Compare the tendencies of the votes within the city of the different categories of votes: Senate, House of Representatives, Presidency, City Hall, Council and Jal.

D. Compare Similarity

Compare the similarity of votes between Senate, House of Representatives, Presidency, City Hall, Council and Jal for the 4 periods by political party

E. Summarize Correlations

Summarize the correlations between the different voting categories of the main political parties during the last electoral periods

F. Identify Patterns of Correlations

Identify patterns of correlations of the votes cast in Bogota during the last electoral periods.

G. Identify, Compare and Explore Correlations

Identify and compare the correlations between votes by different parties

HOW

Task A, B y C

Marks:
- Area: represents the localities
- Lines: represent political parties

Channels
- Spatial region: Location within the Map
- Color Hue: Represents the political party.

Encode
- Use: Location on the map of Bogotá
- Separate: (For the areas) Separate by region the localities.
- Separate, order and align: (For the lines) The parties are separated into lines by colors. They are sorted by number of votes and are aligned on the y axis.

Manipulate:
- Select: You can select a location, and you can see important characteristics of it, such as the name of the party that won and the number of votes
- Select and Change: When selecting a political party, change the alignment of the bars with the x-axis

Facet:
- Juxtapose: The 4 maps of Bogotá are shown on the screen. One for each period.
- Filter: When selecting a location, only the votes of the same are shown.

Idiom
- Map

Task D:

Mark:
- Lines: represent political parties

Channels
- Color Hue: Represents the political party
- Radio: Number of votes

Encode
- Separate: (For areas) Separate political parties
- Express: Shows the percentage of votes in a voting category

- Separate, order and align: Categorize by type of voting

Manipulate:

- Select and change: The year is selected and the radio values are modified

Facet

- SuperImpose: On the radar the results of the most important matches are presented simultaneously

Reduce:

- Filter: When selecting a location, only the votes for this are shown.

Idiom:

- Radar
-

Tareas E, F y G

Marks:

- Points: (Square points) Represent the correlations between the different categories

Channels

- Luminance: Express the correlation between two categories
- Color Hue: Represents the political party.

Encode

- Express: shows the positive and negative correlation variable
- Separate and Align: The correlations are separated from each other and aligned on both the x axis and the y axis

Manipulate:

- Select and change: The political party is selected and the colors of the matrix are modified according to their correlation

Facet:

- Juxtapose: The results of the voting are separated by period and political party.

Idiom:

- Correlation matrix

IV. EVALUATION

The tests were carried out with 2 experts and with people outside the project and seven key aspects of the tasks expected for the visualizations were asked. The questions are listed below and the most representative answers were placed:

1. From 1 to 7, how easy is it to identify how many votes a party had in a type of voting: eg: mayoralty?

Now that I know the trick of putting myself on that option I would say that 7. Without that, I put a 4

2. From 1 to 7, how simple is it to identify a pattern in the voting?

6. Although it seems to be simple, there are a few things to say that it is truly a pattern. For what you see there a similarity can be seen.

3. From 1 to 7, How would you rate the visualization to summarize the votes obtained by each match?

4, it is difficult to see the exact value if the trick is not known.

4. From 1 to 7, how difficult is it to know the exact value just by looking at the radar display?

7, the visualization gives an idea of how good the vote was but knowing the exact or approximate value of the points that are not on the top angle is not possible.

5. From 1 to 7, how intuitive do you think the interaction between the map and the radar is?

Between 3 and 5, most of the users did not identify a relationship and did not think that it could be filtered by the map.

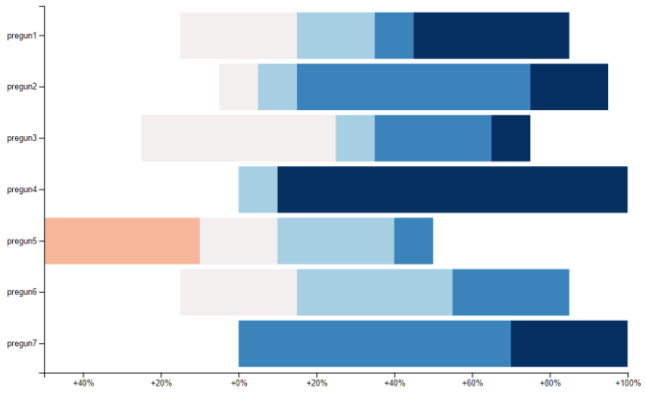
6. From 1 to 7 how intuitive do you think it is to compare the dots on the radar visualization?

4 and 5 given that by explaining to users that this could be done, the behavior was understood and complied with.

7. From 1 to 7 how easy identifies a high correlation between two types of voting

6 since the metric of large colors and rects is clear, but the diagonal rects of correlation equal to 1 are confusing and therefore a 7 is not given.

Below you can see the lickert scale for this tests.

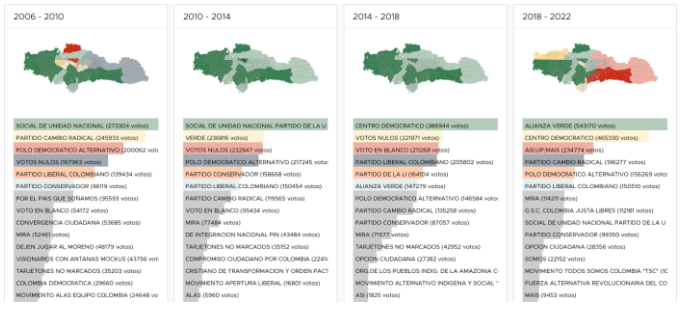


V. CONCLUSIONS

By means of the analysis carried out on the first visualization that shows the data of the votes in each locality the following insights could be obtained:

1. In the presidential elections in the second round, the democratic center did not have as much support in Bogota in 2014 and in 2018 as did its direct rivals, the Santos party and Petro's party respectively, despite the fact that they exchanged their target populations and that the density of voters in the north of the city is higher than in parts of the south such as Usme or San Cristobal.

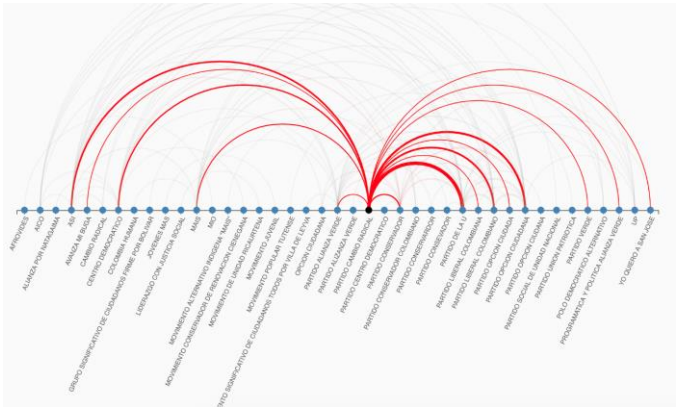
2. The MAIS (Alternative Indigenous and Social Movement), UP (Patriotic Union) and ASI (Independent Social Alliance) parties in the Senate elections were completely outside the political radar, but in the 2018 elections not only did their voters turn up, but they were third parties with more than half of Green Alliance votes and won in the localities of Usme and Ciudad Bolivar.



As for the analysis carried out on the radar and networks that relate matches, the following insights were obtained:

1. The Democratic Center is a party that appeared in 2014 and with great strength in the capital. In 2018, although he has the presidency and a lot of strength in the Senate, it is noticeable that he is losing popularity in the capital. Particularly in 2018 they had a gap of more than 300 thousand votes regarding the next party.

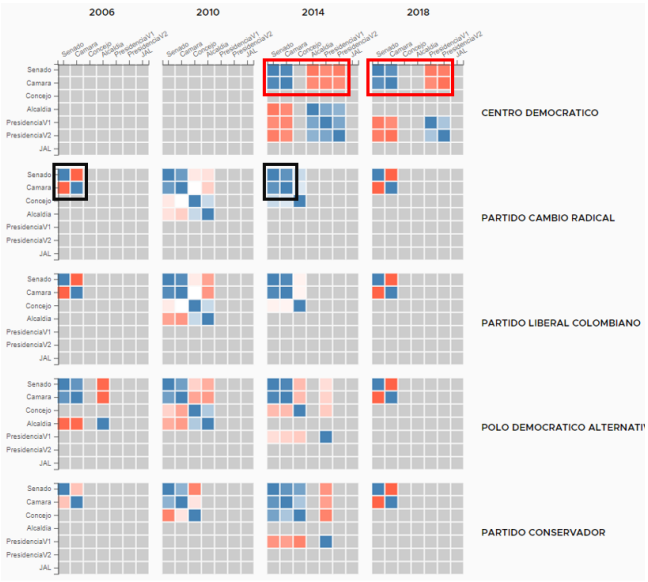
2. There is no sense fighting for political parties. We see that when it comes to getting votes, they all relate to each other. Particularly the radical change party has joined with many others to achieve its electoral goals



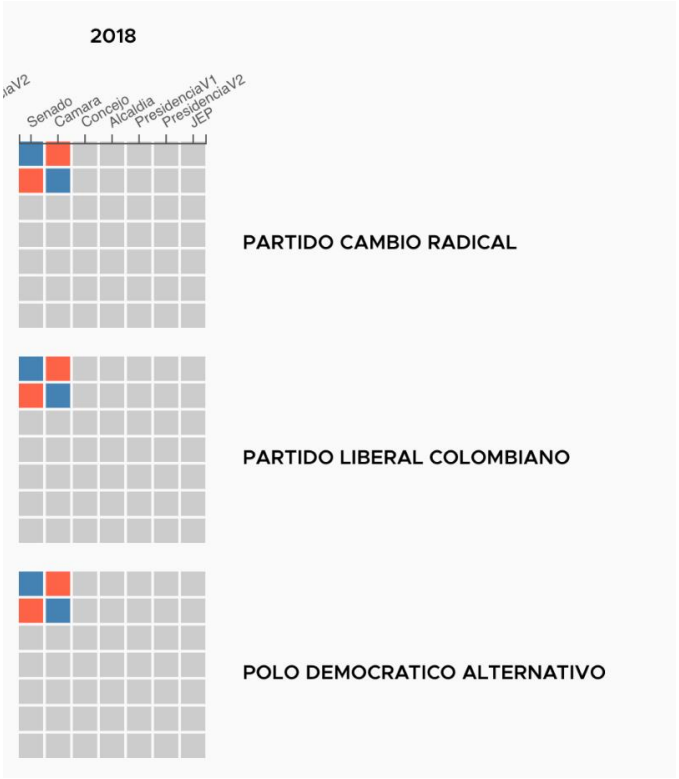
Finally, regarding the analysis of the correlation display, the following insights were identified:

1. There are two predominant patterns in the big parties, which is negative high correlations between the Senate and Chamber, which means that the more votes obtained in one less votes are obtained in the other, or a positive median correlation between the Senate and Chamber, which means that that party did well both in chamber and in senate, but there is a constant pattern in the democratic center and it is a very high positive correlation between senate and camera and a high negative correlation between senate / chamber and presidency / mayoralty. This denotes two things, first is that the lack of coalitions has cost the democratic center in votes in presidency / mayor's office or that the

focus of the democratic center is the legislative power.



2. No pattern of correlations between the types or categories of voting is observed. Each year it changes and in each party behaves differently. Therefore, it can not be said that a type of voting (chamber, senate, mayor, etc.) will increase or decrease the votes of the candidate of the same party that presents itself to the presidency.



REFERENCES

[1] Tamara Muzner, "What's Vis, and Why Do It?," in Visualization, Analysis and Design .New York: Taylor & Francis Group, 2014.

[2] https://movement.uber.com/explore/bogota/travel-times/