

# Tools For Enjoy Election Data From Colombia Since 1958.

Camilo Andrés Escobar<sup>1</sup>, Felipe Mate<sup>2</sup>, David Ricardo Mayorga<sup>3</sup>, Mónica Pachón Buitrago<sup>4</sup>, John Alexis Guerra Gómez<sup>5</sup>, and José Tiberio Hernández Peñaloza<sup>6</sup>

<sup>1</sup>ca.escobar2434@uniandes.edu.co

<sup>2</sup>f.mate10@uniandes.edu.co

<sup>3</sup>dr.mayorga20@uniandes.edu.co

<sup>4</sup>Universidad del Rosario

<sup>1,2,3,5,6</sup>Ingeniería de Sistemas y Computación, Universidad de los Andes

December 6, 2016

## Abstract

The knowing of politics is mostly an unknown problem on a large scale in Colombia. Here is proposed a web visualization project, in which the historical information of the votes and the elected representatives are presented in an entertaining and inclusive way, in order to generate a feeling of empathy or politician relevance in the spectator creating the assumption that there's a familiar relationship between the user's and the politicians elected with the same las

## 1 Introduction

The lack of interest in the politician development of the country is a serious problem because, to achieve better development and better political representation. It is important that we observe our past and learn from it to build a better country and not repeat the same mistakes.

This lack of interest can be observed how a large part of citizens abstain from voting,(Figure 1) how young people and adults do not know beyond the last two presidents, are not aware of politic parties and their trends. How the issue of politics is never treated in people's homes, except for important moments such as the election of a president or a plebiscite for peace.

There are several causes that generate this problem, the lack of communication in certain regions of Colombia, personal resentment on the subject, the little education that was provided in schools on these topics, among others.

In order to generate interest and trying to threath this problem, it was decided to take the politician history of the last 50 years in Colombia and present it to people in a more striking, sympathetic and without the may have the

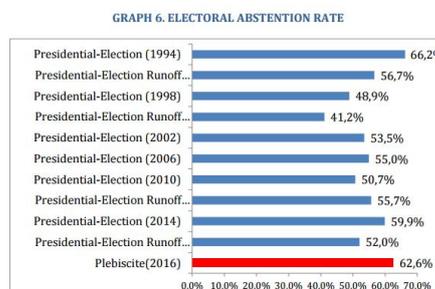


Figure 1: Colombian Electoral Abstention Rate. [1]

theme. It is known that the surnames of a person are the family legacy which is transmitted from generation to generation and There is a tiny probability about find a family relationship between two people with a common surname. However it is important to clarify that we are talking about two unknown persons with a fictitious family relationship. Taking this premise the reconstruction of a genealogical chain between the spectator and the thousands of political positions take place for the last 60 years that we would call "political family". With this idea of "family" the team will show the spatial distribution among Colombia, the different political offices and a comparison with the most political outstanding "families".

## 2 State of Art

Before the team could design something, it is necessary to focus on how had been the problem boarded on the same politician topic or on others but wit the same intention: make social aware trough interactive visualizations. On then one hand are the news media which have been

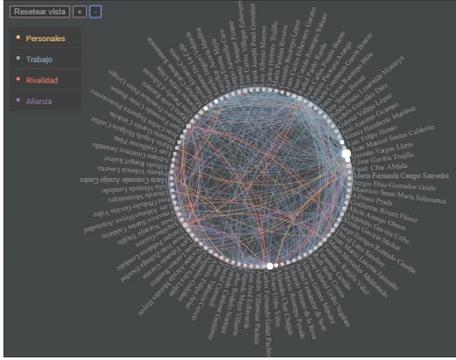
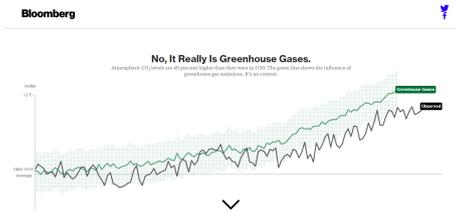


Figure 2: Who is who? by La Silla Vacía. [2]



created politician visualizations about the team has focus on to learn about them and apply to the proposed method. In this case stay afloat the Colombian portal media La silla Vacía, a recognized information media which has become that prestige because of their controversy politician news either the innovate multimedia content it offers. That last required because their viewers are mostly social network's users. Therefore that interactive media they got to approach ... In 2 is shown one of the last visualizations which shows the interpersonal relations between the most important politics and public figures in Colombia.

On the other hand we could find some visualization which have to make aware people of their issues, situations which will be affect in the future. The better example is the actual global warming problem. However a visualization which just shows the actual global warming situation *printed on the screen* is not enough. The team is focusing on an insight which offers the same data not just by an enjoyable, but intriguing way. The Bloomberg portal shows an entertaining vis to find out what caused the global warming. [4] Using the scrolling function the global warming trend is being compared with potential natural i.e. ozone pollution, deforestation or natural causes like the sun's temperature or earth's orbit. At the end of the scroll and a lot of interest accumulated the vis conclude contrasting very well the temperature with the greenhouse gases. Figure 3

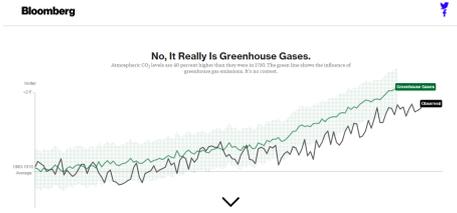


Figure 3: Global warming visualization by Bloomberg . [4]

### 3 Proposed Method

Something like the scrolling vis from Bloomberg is a visualization the team has proposed to create. One which does not show just data visualized printed on the screen but one which increases the viewer curiosity through the interaction he has with it. At the end of this he had to made aware of the importance of politics on his country. Otherwise he have at least realized in the fact that politics are not people so different of him, . Contrariwise he could had had some relatives occupying a public office a few past years ago. To realize all this project the team was based on the proposed method by T. Munzner to visualize data. This method treats the 3 questions (What? Why? and How? )to generate not just visualization but enjoy-insights for the viewers: [3].

All the data about the elections in Colombia for the last 60 years have been collected by the Universidad del Rosario. The next visualizations are based on that dataset proposing activities which supposes there is an family relationship between the user and the politics who have the same last name. This assumption was established just for the enjoyable development of the activity. It is not true information neither have been confirmed: it has been planned with the only proposal about the joy of the user.

The original dataset was mainly a couple of tables grouped by election year and politic office. To make the visualizations it had to pass through an ETL process. Originally the team had become a 2GB group of tables in dta format. Thus they used Stata and a Java script to convert it in CSV files. This new format let threat the tables as SQL sentences to insert them in a SQL database. However the team found some difficults on the database loading step caused because incoherent or null values or because there were larger than the processing team's computers. In this case the tables have to been divided on smaller files. Finally every relational table added on MySQL Workbench had all the values about the quantity of votes for a candidate for a specific politic office, town, department and

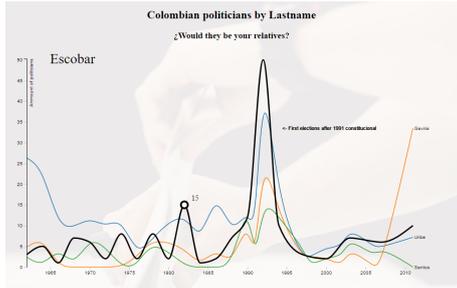


Figure 4: Visualization 1

year including if he or she was or not elected. Once the visualizations were created using d3.js v4 it has been complemented with the framework Node.js letting the user consume the REST services which the database offered. Lastly the web page server was uploaded using Heroku.

## 4 Results and Evaluation

### 4.1 Vis 1: Recognized political last names vs user's last names

In the first visualization we wanted to show how powerful is the user's last name generally. Thus all the politicians with a common last name who have been elected were aggregated and grouped by the election year no matter which office they have won. Creating a line chart amount vs. year. (figure fig:VisWebPage1.png) the user could compare his political trend through the years with other last names which have been frequently heard on the politician context. In this case the users last name was compared with very renowned families: Santos, Uribe and Gaviria.

### 4.2 Vis 2: Distribution of people with a specific last name throughout the country

This visualization consisted of a map of Colombia divided by departments, where each department has a specific color based on a value calculated as a quotient. There were 2 possible options given a last name: Firstly was the amount of people that was elected for that department having that last name over the amount of people that was a candidate for that department with that last name, both for a specific year; Secondly was the amount of people that was elected for that department having that last name over the total of people that was candidate for that department no matter which one was their last name. Also, the visualization has a scroll bar down for the different years, allowing the user to move this bar, and see how the departments

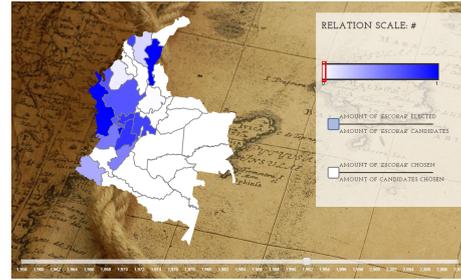


Figure 5: Visualization 2

change their color, representing the movement of that specific "family" over the years through Colombia. Finally, the visualization has a space with a gradient color bar, that once the user click a department the bar shows a marker over the color that has the department, allowing in that way the user to compare different departments by clicking them.

### 4.3 Vis 3: Different political positions of my family through the years

This visualization shows the number of people with a given last name, who performed as politician in the Colombia history. In the x-axis the time step is growing while the graph shows several figures that are increasing or decreasing. Each color represents a different political position and the area shows the quantity of people with that last name that exerted such position in such a period of time. From the graph you lose the facility of being able to see exactly how many people there were by charge, but it allows the observer to see more naturally the flow that is in the passage of time, in addition to highlighting easily and simply when there is a large difference in size. From the entire data set, only the candidate's surnames, the type of charge and the number of candidates per position that is calculated are used. The data that are shown in the graph are those that are in the database. From this visualization one could find patterns of how a "family" of a certain last name tends to occupy certain political positions.

### 4.4 Evaluation

It's clear that the team could have sacrifice some expressiveness terms in order to make more enriched the user's experience. In the first and last visualization the team has made the assumption that the elections years were a continuous data, untrue fact. It was better to use a bar chart to represent more effective the data more but the priority task for those two visualizations treated

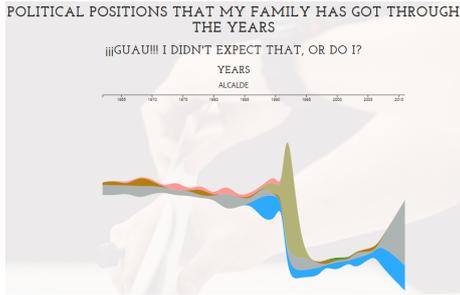


Figure 6: Visualization 3

about comparing trends and find easily patterns on political offices. Those trends are easier to compare in a continuous domain than in separated bars.

## 5 Discussion and Analysis

The large data set rarely collected in Colombia before shows some interesting results which could not have been discussed until today when they can be seen at a glance. The fact which mostly stands out is the higher amount of politicians in 1992 no matter the last name or political office. This can be explained through the Colombian political history during 1991 when a new constitution, the governing document were written. Therefore one of the new actions it demanded was to make new elections for political offices which were not elected popularly i.e. department governors.

Although the project were proposed to make aware the viewer about politics the team can find some. After the construction of the 3 visualizations new insights emerge:

- Which are the most effective families throughout history?
- What are the most remarkable years in the displacement of a "family"?
- In what year was the political awakening for certain "families"?
- There are families that share a similar story?
- Which are the last names that have exerted positions with greater power?

## Conclusions

- Here we can find an effective channel to present the data political elections which have been rarely collected for the last 60 years.

- The creation of interest in citizens by their country politics is a serious topic for the development of the country.
- In addition to having a large data set to continue implementing new visualizations, whether for a deeper analysis or for the expansion of the game.

## References

- [1] Voting for peace: understanding the victory of "no", October 2016. [Online; posted 20-October-2016].
- [2] LaSillaVacía. Quién es quién - las conexiones del poder, 2014.
- [3] T. Munzner. *Visualization Analysis and Design*. CRC Press, 2014.
- [4] E. Roston. What's really warming the world?, June 2014.