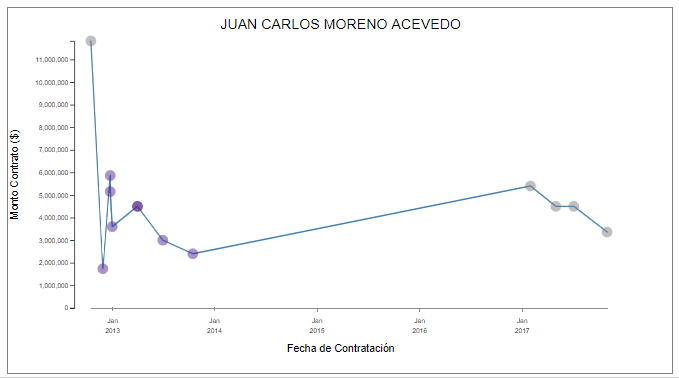
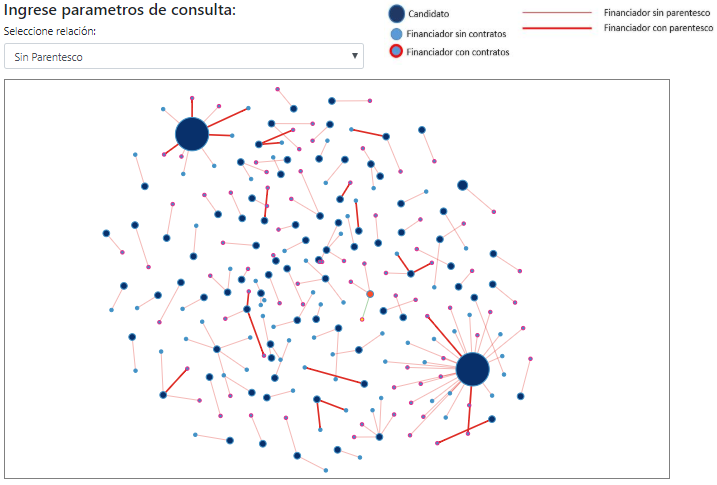
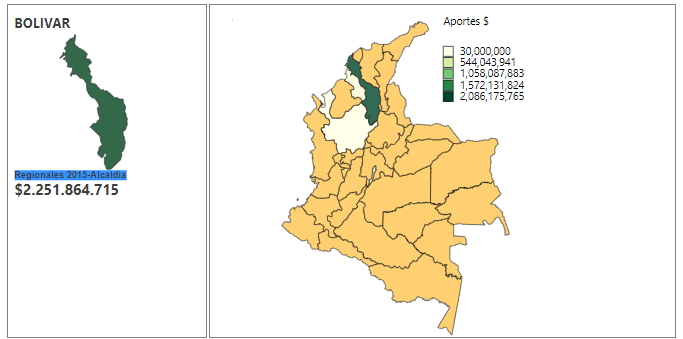
Redes de Contratación Pública y Financiación de Campañas

Iván Romero Flórez 201717315, Gerardo Pérez Clavijo 201519403,

Stanley Melo Fontalvo 201723966



**Abstract**— Datasketch is a digital platform specialized in journalistic research and data. It allows journalists and citizens to learn and consult data visualizations about different conjuctural subjects. Its main objective is the democratization of knowledge. Nowdays, it’s interested in investigations that allow them to evidentiate the relationship between public procurements and the financing of campaings in Colombia. As part of this project, Datasketch along with the Citizen Monitor, developed the portal “*Elecciones y Contratos*” (<http://elecciones-contratos.monitorciudadano.co>, which provides a tech solution that allows the data of popular election candidates and their lenders, to be known. However, they express their interest in extending this analysis, in order to include departamental behaviours and relationships between the financing and celebration of public biddings. As a result of the process of data and tasks extraction, 3 interactive visualizations are presented in accordance with the needs identfied in Datasketch, based on the concept of the data visualization mantra design (Overview first, summand filter and details on demand).

**Index Terms:** candidates, financing, public procurement, irregularities, Colombia, visualization.

INTRODUCTION

Datasketch is a digital platform specialized in journalistic research and data. It provides a portal that allows journalists, data scientists, social scientist and citizens in general, to learn and consult about data visualization, tools, software and in-depth research about several conjunctural subjects. It has free data tools and different projects in order to create a link between the data and the citizenship, facilitating the democratization of knowledge and a critical review about social realities based on information contrasts.

This document shows a methodological framework that, based on a common language framed in the visualization developments, allows to make data and task abstractions, proper of the business problem Datasketch provides, and that also helps to select the best idioms in order to generate insights to the tool users about the context of the problem on its own.

**OBJECTIVE**

Provide visualization tools that let the possible irregularities in the financing of political campaigns for Congress and territorial aggregates at a departmental level to be explored, and their influence on public procurements.

# COMPONENTS OF THE PROJECT

## Users

* **Juliana Galvis** E-mail: [jgalvisn@datasketch.co](mailto:jgalvisn@datasketch.co)

Politologist and candidate for a master’s in Digital Humanities. She’s currently leading the development of the "Who Is" database, as well as supporting journalistic researches and databases creations.

* **Camila Achuri** E-mail: [camila@randommonkey.io](mailto:camila@randommonkey.io)

Statistician and expert in programming R language. She has developed various applications of data visualization in mobility and open data topics

## Tasks

### Main Tasks

* Summarize the contributions made to political campaigns nationwide per year.
* Identify each candidate’s campaign funders, as well as their relationships and contracts with the State.
* Identify the relationship between the financing consideration and the quantity of the contracts awarded to the funder.

### Secondary Tasks

* Obtain a list of contributions made to political campaigns at a department level
* Identify each candidate’s campaign funders and their contracts with the State.
* Show the relationships between candidates and campaign funders.
* Identify the correlation between the financing consideration and the quantity of contracts assigned to the funders.

## Available Data

Datasketch has provided the following data set to carry out the project:

* A table of candidates
* A table of political campaign donors
* A table of public contracts related to those donors / candidates.

# STATE OF THE ART

Currently the Citizen Monitor ([www.monitorciudadano.co](http://www.monitorciudadano.co)), based on the work that has been doing to provide to the citizens technological tools, that allow them to interact them with corruption and its actors date, has developed the portal “*Elecciones y Contratos*” (<http://elecciones-contratos.monitorciudadano.co>), which seek to offer a technological solution that allows to get to know data about the candidates for popular elections and their funders. What’s mentioned above, contributes to social control and also, facilitates the realization of information crossings between official data of financed campaigns supplied by data sources as: *Cuentas Claras* and official procurement contracts provided by *SECOP*

# WHAT

## Dataset

The data set is a table type and it’s descripted as following:

* Contratos\_financiadores\_data: this table contains 149,070 records and its purpose is to obtain public procurement information from state entities, with regard to the corresponding service providers. (2010 -2017) funders that have had contracts with the state.

## Types of attributes:

The table Contratos\_financiadores\_data originally has 79 attributes from which 17 were picked. They are shown in the following table:

|  |  |
| --- | --- |
| Attribute | Type |
| Corporation or Categorical Position | Categorical |
| Electoral Circumscription | Categorical |
| Department | Categorical |
| Candidate Name | Categorical |
| Funder | Categorical |
| Consideration | Quantitative |
| Department of Categorical Contributions | Categorical |
| Relationship | Categorical |
| Registr. of Sequential Movement Date | Sequential |
| Campaign | Categorical |
| Contract Number | Categorical |
| Candidate | Categorical |
| Start Date of Contract Execution | Sequential |
| Contract Amount | Quantitative |
| Object | Categorical |
| Group | Categorical |
| Entity | Categorical |

## Availability

A Scheme was generated in which used datasets are published in OneDrive and GitHub

# WHY

## Main Task 1.

Summarize the contributions made to political campaigns nationwide per year.

* ACTION (Summarize) (department: Categorical)
* TARGETS – (Features) (campaign: Categorical, total contributions: quantitative calculated)

### Secondary Task 1:

Obtain a list of contributions made to political campaigns at a department level.

* ACTION (Summarize) (department: Categorical)
* TARGETS – (Features) (campaign: Categorical, total contributions: quantitative calculated, department of origin of the contribution: Categorical)

### Secondary Task 2:

Identify departments with greater and lesser contributions to campaigns.

* ACTION (Summarize) (department of origin of the contribution: Categorical)
* TARGETS – (Features) (campaign): Categorical, total contributions: quantitative calculated, department: Categorical)

### Secondary Task 3:

Identify the distribution of contributions to a campaign by income of the department.

* ACTION (Discover - Locate- Identify) (department of origin of the contribution: Categorical)
* TARGETS– (Distribution) (Consideration: quantitative)

## Main Task 2:

Identify each candidate’s campaign funders, as well as their relationships and contracts with the State.

* ACTION (Present – Explorer- Identify) (Funder: Categorical)
* TARGETS– (Feature) (campaign: Categorical, Contract Number: Categorical, consideration: quantitative, relationship: Categorical)

### Secondary Task 1**:**

Present candidates' relationships with campaign funders.

* ACTION (Present – Explore - Identify) (candidate: Categorical)
* TARGETS– (Feature) (funder: Categorical, relationship: Categorical, consideration: quantitative)

### Secondary Task 2**:**

Identify funders of multiple campaigns

* ACTION (Present – Explore - Identify) (Funder: Categorical)
* TARGETS– (Feature) (Candidate: Categorical, consideration: quantitative)

### Secondary Task 3**:**

Summarize the total contributions per candidate

* ACTION (Summarize) (candidate: Categorical)
* TARGETS – (Features) (Funder: Categorical, total contributions: quantitative calculated)

### Secondary Task 4**:**

Identify the candidates with the most campaign financing.

* ACTION (Summarize) (candidate: Categorical)
* TARGETS – (Features) (Funder: Categorical, total contributions: quantitative calculated)

## Main Task 3:

Identify the relationship between the amount of financing and the number of contracts awarded to the funder.

* ACTION (Present – Locate - Identify) (Funder: Categorical)
* TARGETS - (Features) (contract: Categorical, consideration: quantitative y contract date: sequential)

### Secondary Task 1**:**

Know the contracting behavior for the funder before and after a contribution.

* ACTION (Present – Locate - Identify) (funder: Categorical)
* TARGETS- (Features) (campaign: Categorical, total of contributions: quantitative calculated, consideration: quantitative, contract date: sequential)

### Secondary Task 2**:**

Present the number of contracts assigned to a funder in a period of time.

* ACTION (Present – Locate - Identify) (Funder: Categorical)
* TARGETS – (Features) (contracts: Categorical, start date: sequential)

# HOW

## Graphic Areas

This visualization consists in choosing departments, through a Colombian map, which based on a color scale emphasizes the relationship between the selected department and the financing departments of the political campaigns. This color scale will have as main objective, to attract the perception of the user, pointing out those contributions that are in the high, medium and low scale.

On the other hand, when selecting the department, by a network chart, it will be observed, the relationship between candidates of the campaign and the selected departments, and its link with funders who have concluded contracts with the stare, and that belong to the same department or to different departments.

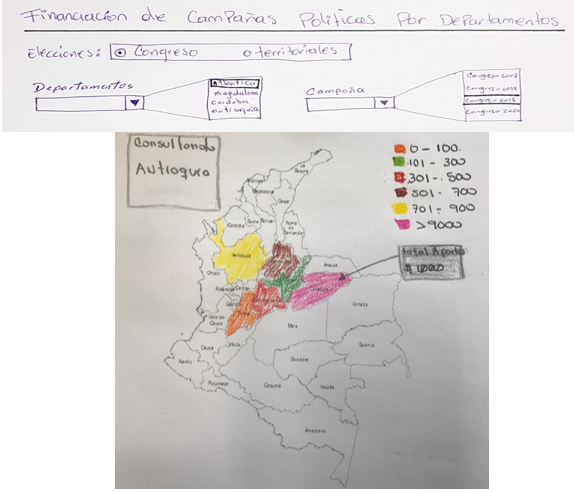


Image1. Display Prototype No.1

|  |  |
| --- | --- |
| HOW | Description |
| Marks | * AREAS (Department) |
| Channels | * SATURATION (Attribute quantitative consideration) |
| Encode | * ARRANGE * EXPRESS (Attribute quantitative consideration) |
| Manipulate | * NAVIGATE * SELECT (Department) |
| Facet | * JUXTAPOSE * MULTIFORM * OVERVIEW /DETAIL: * AGGREGATE (Consideration quantitative) |
| Reduce | * FILTER: he proposed idiom allows the application of a filter (Categorical department) |

## Network Chart

With this visualization the user gets to see, the candidates of political campaigns, the state contractors that are registered as funders of its own campaigns. It consists in identifying the direct link with one or several candidates of the campaign and its categorization of the contribution in the high, medium and low rank, in addition to show or not, the relationship to the user’s judgement.

This visualization presents interactivity in the following way:

- Allows the redistribution of campaign financiers according to their relationship or not.

- Allows to redistribute the visualization of campaign financiers, according to the amount of the contribution.

- Allows, through a click event, to look at the time line for the conclusion of contracts with the state, for each selected funder.

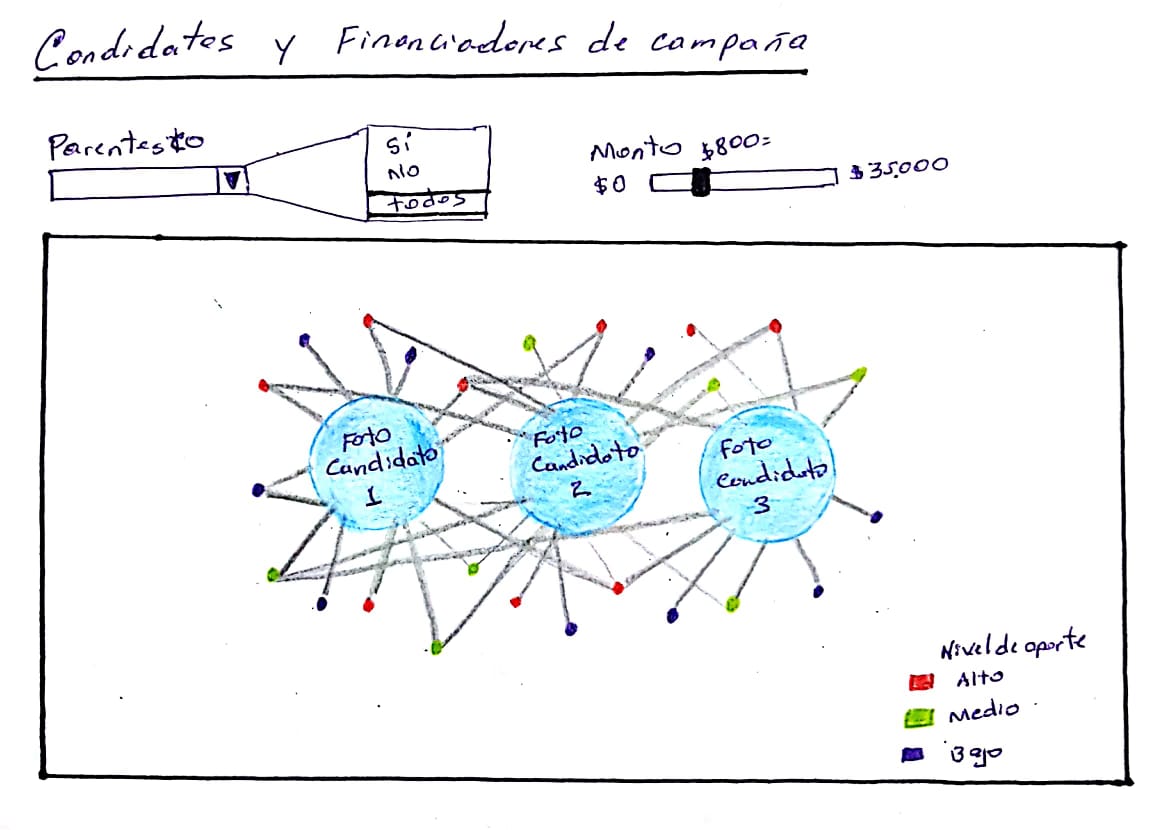


Image 2. Display Prototype No.2

|  |  |
| --- | --- |
| HOW | Description |
| Marks | * Points (Candidate - Funder) * Lines (Relationship Candidate - Funder) |
| Channels | * HUE (Attribute Categorical Candidate - Funder) * SIZE (size of Candidate - Consideration) |
| Encode | * ARRANGE * EXPRESS   (Consideration quantitative) |
| Manipulate | * NAVIGATE * SELECT (Relationship) |
| Facet | * JUXTAPOSE * MULTIFORM * OVERVIEW /DETAIL: * AGGREGATE (Consideration quantitative) |
| Reduce | * FILTRO: the proposed idiom allows the application of a filter (categorical Relationship) |

## Timeline Chart

It allows to observe timeline corresponding to the conclusion of contracts of a selected campaign funder. Each point of the chart shows in which campaigns a funder has been registered.

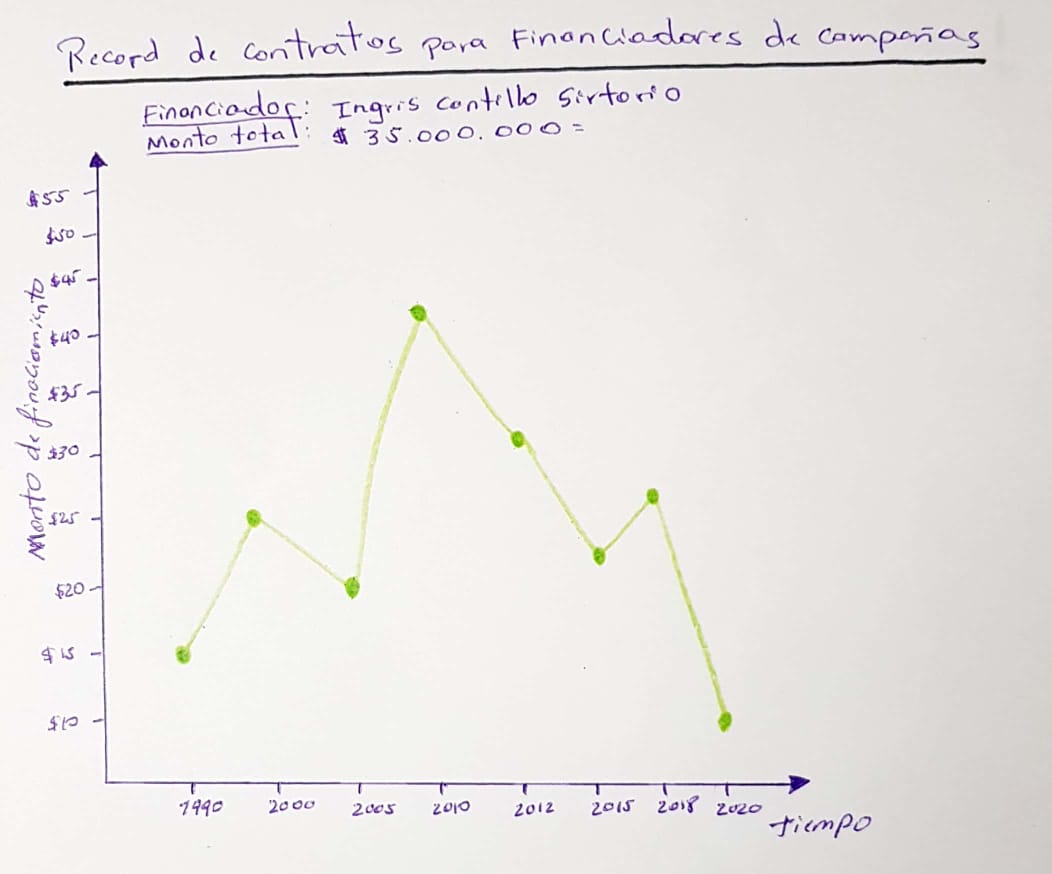


Image 3. Visualization prototype No.3

|  |  |  |
| --- | --- | --- |
| HOW | Description | |
| Marks | * Lines (Contract Amount) * Points (Consideration, Campaign / Contract Number) |
| Channels | * Position X quantitative (Start Date) * Position Y quantitative (Contract Amount) |
| Encode | * ARRANGE * EXPRESS (Consideration quantitative) |
| Manipulate | * N/A |
| Facet | * JUXTAPOSE * MULTIFORM, * OVERVIEW /DETAIL: * AGGREGATE (**Contract Amount**) |
| Reduce | * N/A |

# TECHNOLOGICAL BASE

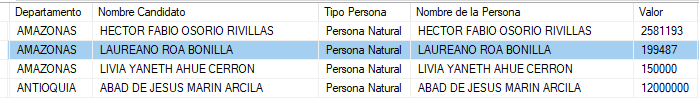
For the development of this project it was agreed along with the client to use the following technologies

* HTML
* D3 version 4 (visualization)
* GitHub (repository and deployment)
* File format .csv

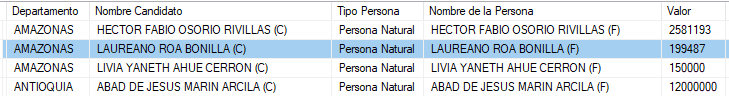
# DATA PREPARATION

As a result of the construction process of the functional prototype, quality needs were evidenced on the data. Final attributes necessary to be used in the idioms were identified as well.

Likewise, it was necessary to debug records that didn’t contemplate the necessary quality conditions of completeness. Also, as part of the necessary tasks on the data, the modification was made in the cases in which the "Name of the Person" was equal to the "Name of the candidate". To do that, all the records were marked with a flag that allowed its effective distinction. (see screenshot)



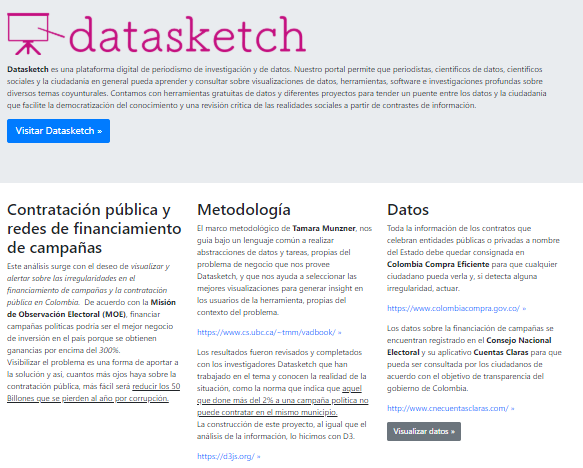
As a result of this process, an example is observed below with the inclusion of the flag for its differentiation:



These tasks allowed the previously selected idioms to be implemented in the expected manner.

# FUNCTIONAL PROTOTYPE

For the development of the functional prototype, the techniques learned in class have been used, typical of Tamara's framework, which allowed us to perform data abstractions and tasks on a common language that orients our work to the selection of the best idioms, and the consequent generation of insight for the user. According to the indications provided by the client, the GitHub platform (https://github.com/ ) was used as repository of the project and there, using the technologies mentioned in point 6, the deployment was carried out in the following url: <https://63anp3ca.github.io/VA/>



Three sections are identified in the prototype: (1) visualization section Map of Colombia, (2) visualization section of nets (see screenshots) and (3) visualization section of line chart of campaign finance contracts.

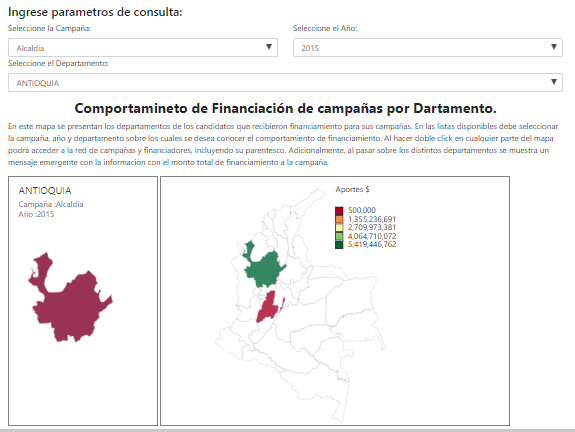


Image 4. Screenshot display No.1

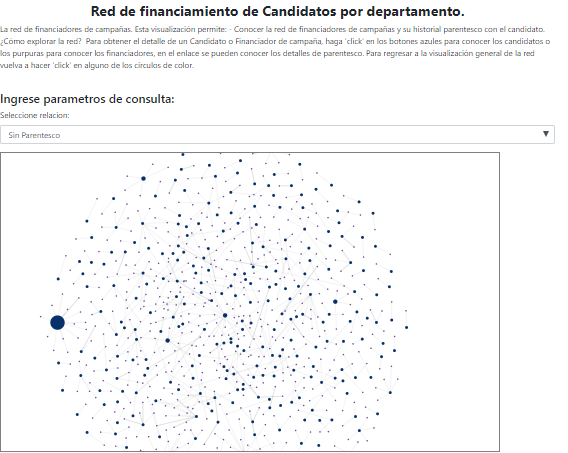


Image 5. Screenshot display No.2

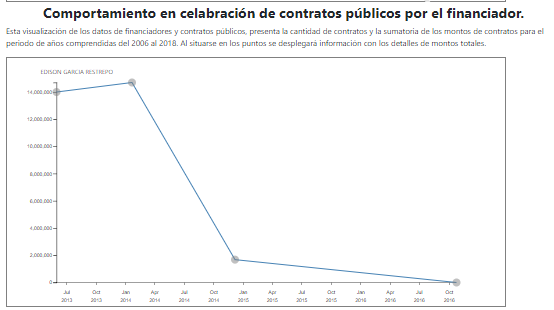


Image 6. Screenshot display No.3

# USABILITY EVALUATION

In order to validate the selected idioms based on the results of the abstraction of data and tasks, in addition to the application of the principles of expressiveness and effectiveness, we perform an evaluation test of the usability of the tool, taking into account the following characteristics:

* Datasketch has extensive knowledge about the application of visualization techniques.
* A first controlled experiment was developed, so that clients and external users could interact with the visualizations and draw conclusions tending to evaluate the usability of the tool.
* Each one was asked about the result of their interaction, if they consider that they were useful or not and the proposed improvements.

## Usability Assessment Test Criteria

|  |  |
| --- | --- |
| CATEGORY | DESCRIPTION |
| Identity | Identify the corporate image of the institution that owns the site |
| Contents | Determine the quality assigned to the contents and whether the way they are presented allows the user to get a concrete idea of the information that is being delivered through the website. |
| Navigation | Evaluate how appropriate is the organization of the information of the website, according to the experience, knowledge and expectations that the user has. |
| Web Graphics | Evaluate how meaningful the graphs or visualizations offered to the user are, and their perception of the speed of deployment. |
| Search | Establish if the search system meets the needs of the user. |
| Usefulness | General summary of the experience of navigating the site. |

## Profile Sociodemographic

|  |  |
| --- | --- |
| **User characteristics** | **Description** |
| Level of study | Professional |
| Degree of  knowledge | Full knowledge of the context of the problem |
| Office Domain | E-mail address.  Word processors.  Browsers  Spreadsheets and statistics. |
| Age | Range: 21- 60 years old |

## Steps for Application of The User Test

Before the test

* There is a list of 6 questions, two for each member of the group, taking into account the areas of: identity, content, navigation, web graphics, search, feedback and usefulness.
* Schedule a meeting with the client to express opinions on the prototype of the visualization.
* In agreement with the members of the group, we choose the person who will expose the prototype to the users.

During the test

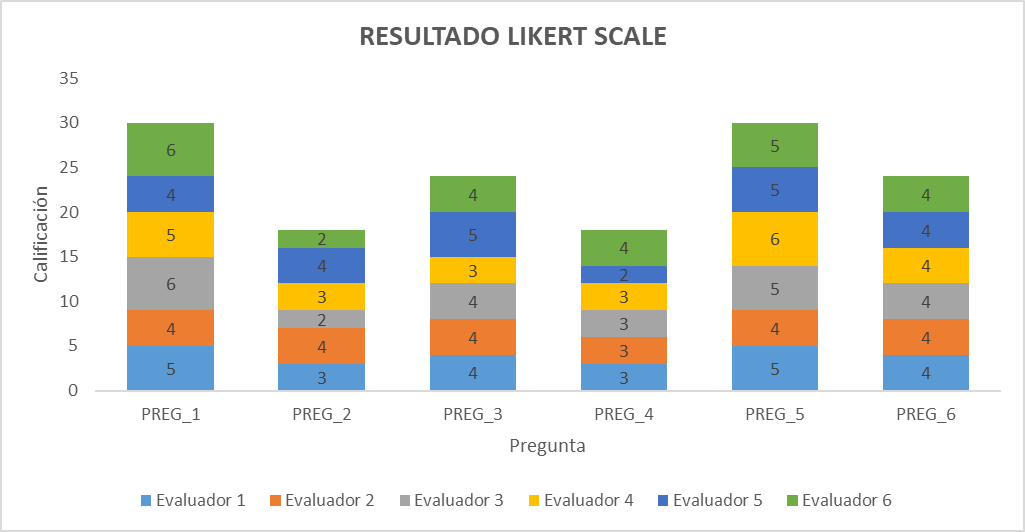
* A member of the group will explain the main idea in the development of this test and its role.
* The group members will take careful note of each user's opinions and reactions.

After the test

* With the user opinions, we will discuss the changes to be applied to the prototype according to their priority.
* Each of the recommended changes will be made.

Below is a table and visualization that summarizes the question scores per respondent, total and average respondent.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| COD | LIKERT SCALE | | | | | | Total | Average |
| Eval.1 | Eval.2 | Eval.3 | Eval.4 | Eval.5 | Eval.6 |
| PREG\_1 | 5 | 4 | 6 | 5 | 4 | 6 | 30 | 5 |
| PREG\_2 | 3 | 4 | 2 | 3 | 4 | 2 | 18 | 3 |
| PREG\_3 | 4 | 4 | 4 | 3 | 5 | 4 | 24 | 4 |
| PREG\_4 | 3 | 3 | 3 | 3 | 2 | 4 | 18 | 3 |
| PREG\_5 | 5 | 4 | 5 | 6 | 5 | 5 | 30 | 5 |
| PREG\_6 | 4 | 4 | 4 | 4 | 4 | 4 | 24 | 4 |



# APPROVED PROTOTYPE.

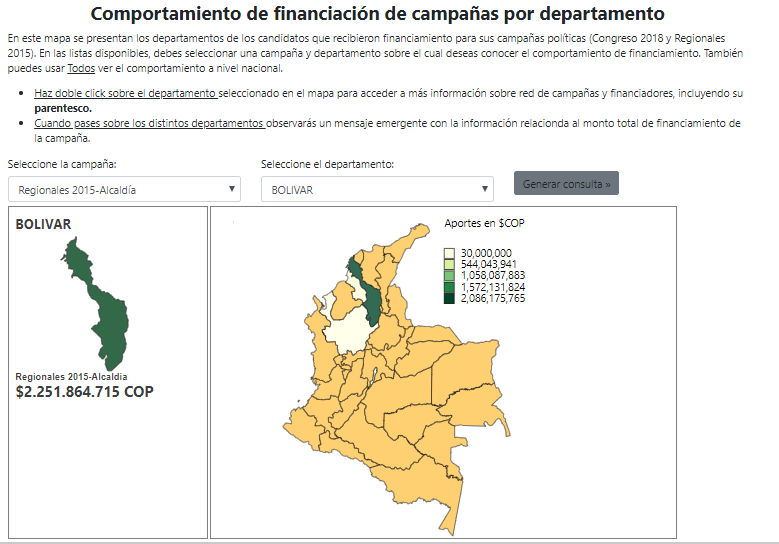


Image 7. Capture Prototype visualization No.1 Approved

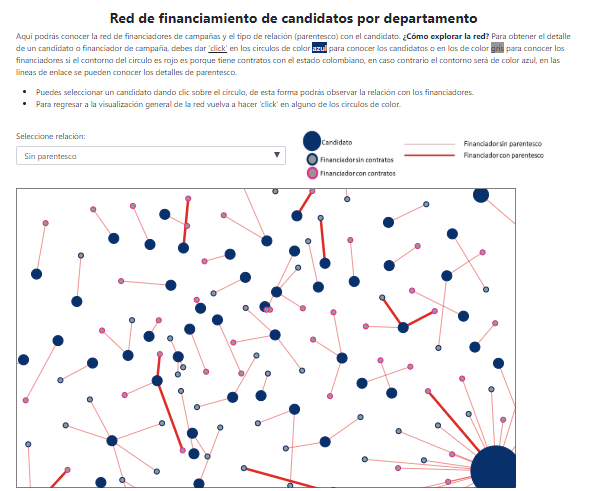


Image 8. Capture Prototype visualization No.2 Approved

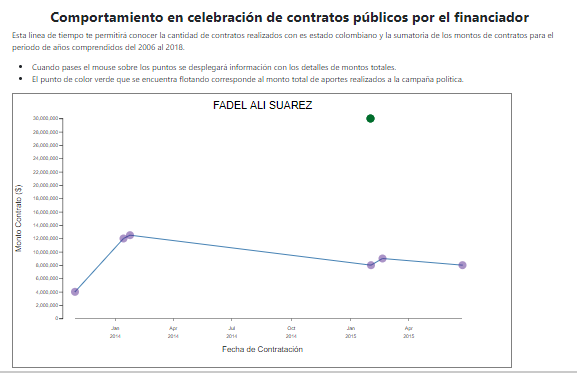


Image 9. Capture Prototype visualization No.3 Approved

# INSIGHTS

In the preliminary analysis we obtained thefollowing findings:

* For the 2015 campaign of the Mayor's Office in Bogota, funding is identified from the departments of Antioquia and Santander.
* For the 2015 mayoral campaign in Bolívar, funding is identified from the departments of Antioquia and Bogotá.
* For the 2015 Mayor's campaign in Cundinamarca, funding is identified from the departments of Tolima, Bogotá and Magdalena.

These findings allow us to appreciate that there is a correlation of closeness between the departments. In most cases, it is observed that the border department is part of the financiers of the candidates. It is also identified that there are departments where the main cities are located, which finance other departments regardless of their territorial proximity. (for example: Antioquia and Bogotá)

# PRIVACY

As it is information that is available as open data from the National Electoral Council (CNE) and its portal of Cuentas Claras (<http://www.cnecuentasclaras.com>). The information is published in compliance with the principles of transparency and publicity as stipulated in Article III of Resolution 1044 issued by the National Electoral Council (CNE) on August 25, 2011.

In addition, this tool allows the CNE to comply with some of the provisions contained in the Law on Transparency and Access to Public Information (Law 1712 of 2014), which reiterates that all information of public interest -such as that related to campaign financing- must be timely, objective, truthful, complete, reusable and available in formats accessible to all.

# BLOG

|  |  |
| --- | --- |
| **Number** | **1** |
| Date | 09/10/2018 |
| Participants | Juliana Galvis |
| Format | Email Address |
| Details:   * Sent mail for customer contact and coordinate appointments | |
| Commitments:   * Datasketch will send project data set | |

|  |  |
| --- | --- |
| **Number** | **2** |
| Date | 11/10/2018 |
| Participants | * Juliana Galvis * Camila Achuri |
| Format | Email Address |
| Details:   * Data set delivery by Datasketch:   + Contratos\_candidatos\_data   + Contratos\_financiadores\_data   + Cuentas\_claras\_data | |
| Commitments:   * Exploration of the dataset by project students. | |

|  |  |
| --- | --- |
| **Number** | **3** |
| Date | 12/10/2018 |
| Participants | Juliana Galvis  Camila Achuri |
| Format | Skype |
| Details:   * - Presentation of the client and working group * - To know Expectations and needs of the client * - Clarification of the data set * - Definition of technology and end user | |
| Commitments:   * Present proposals with scopes, tasks and characterization of the data | |

|  |  |
| --- | --- |
| **Number** | **4** |
| Date | 16/10/2018 |
| Participants | Juliana Galvis  Camila Achuri |
| Format | Email Address |
| Details:   * Proposal sent to Datasketch | |
| Commitments:   * Coordinate Skype call to validate the proposal | |

|  |  |
| --- | --- |
| **Number** | **5** |
| Date | 16/10/2018 |
| Participants | Juliana Galvis  Camila Achuri |
| Format | Skype |
| Details:   * Revision of the proposal * Customer Feedback * Approval of the scope | |
| Commitments:   * Include adjustments and improve the proposal | |

|  |  |
| --- | --- |
| **Number** | **6** |
| Date | 28/10/2018 |
| Participants | Juliana Galvis  Camila Achuri |
| Format | Email Address |
|  |  |
| Details | Proposal sent to Datasketch |
| Commitments: | Coordinate Skype call to validate the proposal |

|  |  |
| --- | --- |
| **Number** | **7** |
| Date | 30/10/2018 |
| Participants | Juliana Galvis  Camila Achuri |
| Format | Skype |
|  |  |
| Details | Validation of the proposal to Datasketch |
| Commitments: | Include adjustments and improve the proposal |

|  |  |
| --- | --- |
| **Number** | **8** |
| Date | 07/11/2018 |
| Participants | Juliana Galvis  Camila Achuri |
| Format | Skype |
|  |  |
| Details | Validation of the proposal to Datasketch |
| Commitments: | * Include adjustments * Usability evaluation * Improve the proposal |

|  |  |
| --- | --- |
| **Number** | **9** |
| Date | 14/11/2018 |
| Participants | Juliana Galvis  Camila Achuri |
| Format | Skype |
|  |  |
| Details | Approval of the final proposal of visualizations. |
| Commitments: | Insights Validation |

# CONCLUSIONS

After carrying out this work and applying Tamara's own framework concepts, we can see the great value it brings in the realization of visualizations that generate a significant impact on the work of the users, reflected in the manifestation of insight for the business. Taking into account the design concepts that improve the user usability indexes, we made use of the visualization mantra technique, to guide the user from the general to the particular that would allow him to explore and identify significant events in his research.

# BIBLIOGRAPHY

* <http://www.datasketch.co/>
* <http://elecciones-contratos.monitorciudadano.co/>
* <https://www.cs.ubc.ca/~tmm/talks.html#minicourse14>
* [www.monitorciudadano.co](http://www.monitorciudadano.co)