**Sabemos lo que hiciste (Familiares de congresistas que son contratistas del Estado)**

**Propuesta**

**Integrantes**

Santiago Suárez Conde

Martha Lucia Cárdenas Soto

Octubre 2018

**TABLA DE CONTENIDO**

[**INTRODUCCIÓN** 3](#_Toc527324709)

[**1.** **OBJETIVOS** 4](#_Toc527324710)

[**1.1** **Objetivo general** 4](#_Toc527324711)

[**1.2** **Objetivos específicos** 4](#_Toc527324712)

[**2.** **ANTECEDENTES** 5](#_Toc527324713)

[**3.** **DESARROLLO Y DOCUMENTACIÓN DE LA PROPUESTA** 5](#_Toc527324714)

[**3.1** **Tareas** 5](#_Toc527324715)

[**3.2** **Data y Usuarios** 5](#_Toc527324716)

[**3.2.1** **WHAT** 6](#_Toc527324717)

[**3.2.2** **WHY** 7](#_Toc527324718)

[**3.2.3** **HOW** 9](#_Toc527324719)

**[4.](#_Toc527324720)****[PROPUESTA](#_Toc527324720)** [12](#_Toc527324720)

**[CONCLUSIÓN](#_Toc527324721)** [14](#_Toc527324721)

[**BIBLIOGRAFÍA** 15](#_Toc527324722)

# **INTRODUCCIÓN**

Cuestión Pública es un medio digital dedicado a la investigación y al periodismo de datos en el interés público, los cuales han sido destacados por investigar en temas controversiales como lo que los poderosos quieren ocultar, jugadas de corporaciones bancarias, rutas de dineros públicos y arreglos bajo la mesa de contrataciones públicas. Cuestión Pública lanza una serie llamada “Sabemos lo que hiciste” que investiga a los congresistas más votados en Colombia, revelando su patrimonio y sus posibles conflictos de interés.

Actualmente Cuestión Pública adelanta una investigación que quiere mostrarle a los colombianos que 1 de cada 2 congresistas en el país tiene familiares contratistas con el estado. Esta ha sido una ardua investigación, por tratarse de un tema muy sensible que podría ver afectada la reputación de algunos de los congresistas. Una de las grandes dificultades ha sido conseguir y procesar los datos, los cuales no son de fácil acceso y entendimiento. Por esta razón Cuestión Pública ha involucrado la tecnología en su investigación, la cual les ha permitido escarbar a fondo y obtener información importante para mostrarle al país.

Mostrar esta información se convierte en un gran reto, debido a la cantidad y complejidad de los datos involucrados en la investigación, lo que dificulta el entendimiento para un ciudadano común. Por esta razón se quiere utilizar técnicas de análisis visual, para lograr que los ciudadanos puedan hacer *insights* de estos datos y tomar mejores decisiones al momento de elegir sus representantes.

A continuación, se hará un acercamiento al proyecto y se realizará una propuesta de análisis visual, que se ajuste al contexto y que logré generar *insights* sobre los datos recopilados tras la investigación de Cuestión Pública y sus colaboradores a los congresistas que tienen familiares contratistas con el estado.

# **OBJETIVOS**

## **Objetivo general**

Realizar propuesta de visualización para “Sabemos lo que hiciste (Familiares de congresistas que son contratistas del Estado)”.

## **Objetivos específicos**

## 

* Exponer el contexto general
* Caracterizar los datos, tareas y usuarios
* Proponer los modismos que soportaran las tareas caracterizadas

# **ANTECEDENTES**

Cuestión pública lanzó 20 derechos de petición para obtener las declaraciones de bienes y renta de los 10 congresistas más votados en el país. En la segunda etapa de la investigación a estos 10 congresistas se les sacó una lista de familiares, financiadores, laborales, equipo UTL, político, familiar y político, amistades, cercanías y empresas relacionadas. Después de tener la lista, se cruzaron estos datos con los del portal de Servicio de Contratación Pública (SECOP - [www.contratos.gov.co/](http://www.contratos.gov.co/)) con ayuda de un *Scraping* que creo *Datasketch* (<https://www.datasketch.co/>) que es una plataforma digital dedicada al periodismo de investigación y de datos, dicho *Scraping* se dedicaba a hacer match con contratos publicados que tuvieran alguna relación con la lista de los congresistas y de estos resultados se entregó como una sábana de datos.

## **DESARROLLO Y DOCUMENTACIÓN DE LA PROPUESTA**

## **3.1 Tareas**

El pasado jueves 11 de octubre en la reunión con Claudia Báez quien representa a Cuestión Pública, se definieron las tareas y objetivos que debería tener la visualización.

Con la visualización se quiere hacer justicia a la información recopilada en la investigación, con ayuda de algo interactivo que le permita a la gente ver lo escandaloso de que 1 de cada 2 congresistas investigados tenga familiares contratando con el estado. Además, se quiere generar consciencia sobre este tema, para que los colombianos y colombianas tomen mejores decisiones al momento de elegir sus representantes.

Para lograr llamar la atención de los usuarios se quiere que la visualización sea atractiva y que los usuarios puedan hacer veeduría del congresista que eligieron en las pasadas elecciones.

## **3.2 Data y Usuarios**

Para el proceso de caracterización se utiliza el *framework* de Tamara [1], con el propósito de abstraer a un lenguaje común el *Why, What y How .* A continuación, se describe el *What (*abstracción de los datos) y el Why (abstracción de las tareas).

### **3.2.1 WHAT**

Dataset – ***Network*** - ***trees***: Contratación celebrada con el estado por familiares de los diez congresistas seleccionados por los abonados del portal [https://cuestionpublica.com](https://cuestionpublica.com/).

A continuación, se describen los atributos del dataset y sus tipos.

|  |  |  |
| --- | --- | --- |
| **Atributo** |  | **Tipo** |
| Congresista |  | Categórico |
| Votos |  | Ordenado - Cuantitativo -Secuencial |
| Contratista |  | Categórico |
| Años como contratista |  | Ordenado - Ordinal |
| Vínculo |  | Categórico |
| Número de contratos |  | Ordenado - Cuantitativo -Secuencial |
| Valor total contratos |  | Ordenado - Cuantitativo -Secuencial |
| Aportó a la campaña |  | Categórico |
| Año financiación |  | Ordenado - Ordinal |
| Monto financiación |  | Ordenado - Cuantitativo -Secuencial |
| Tipo de contrato | Mod. Cont. Régimen Especial | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Cont. Contratación Directa (Ley 1150 de 2007) | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Cont. Selección Abreviada | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Cont. 27 Otros Servicios | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Cont. Prestación de Servicios | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Cont. Mínima Cuantía | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Cont. Selección Abreviada de menor cuantía. (Ley 1150 de 2007) | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Cont. Licitación Pública | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Cont. Concurso de méritos abierto | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Contr. Directa menor cuantía | Ordenado - Cuantitativo -Secuencial |
|  | Mod. Cont. Subasta | Ordenado - Cuantitativo -Secuencial |
|  | Convenio | Ordenado - Cuantitativo -Secuencial |
| Derivar | Nodes: Congresista, Contratista, Entidad contratante  Link: Vínculo familiar,  Atributos: Número de contratos, Aportó a la campaña, Valor total contratos |  |
|  |  |  |

La versión preliminar del *dataset* anteriormente descrito, es suministrado por Cuestión Pública, que ha consolidado la información de los congresistas y la contratación con el estado realizada por sus familiares. Los contratos fueron extraídos mediante un scraping del portal de contratación de Colombia compra eficiente “Secop” <https://www.contratos.gov.co/consultas/inicioConsulta.do>, vale la pena resaltar que la labor de recolección de información que no es pública, Cuestión Pública, la solicitud mediante derechos de petición a los congresistas.

### **3.2.2 WHY**

Para definir el why, se tiene en cuenta que la población objetivo es el usuario del común, al que se le quiere ofrecer un medio que le permita entender en pocos pasos, cómo el congresista de su elección a manejado su interés público en relación con la contratación estatal y sus vínculos familiares.

El objetivo de la visualización es evidenciar el congresista con más contratación realizada por sus familiares con el estado, así mismo mostrar sus vínculos familiares y como ha sido el comportamiento del monto contratado y su tipo de contratación a través de los años.

1. **Tarea Principal (N1): Resumir** la **distribución** del monto de la contratación por congresista realizada por sus familiares desde el año 2007 al 2018. En términos de Tamara (**Summarize** the **Distribution** of variable “Valor total contratos”).
2. **Tarea Principal (N2): Explorar** de forma **divertida** a los congresistas que tienen familiares contratando con el estado con sus respectivas **características** (Monto contratos, familiar, nombre de empresa, cantidad de contratos, tipos de contratos, etc.) desde el año 2007 al 2018. En términos de Tamara (**Explore** and **enjoy** the **Fetures** of variable “Congresita”).

**Tareas secundarias:**

1. **Tarea (N1): Comparar** la **contratación** por congresista realizada por sus familiares desde el año 2007 al 2018. En términos de Tamara (**Compare** the **features** of variable “Valor total contratos”).
2. **Tarea (N2) Deriva**r los **vínculos** familiares entre los congresistas, los contratistas y entidad contratante. En términos de Tamara (**Derive** the **Topology**).
3. **Tarea (N3):** **Identificar** las **relaciones** (vínculos familiares) entre los congresistas, sus familiares y las entidades contratantes. En términos de Tamara (**Identify** the **Topology** of variables Nodes: Congresista, Contratista, Entidad contratante, Link: Vínculo familiar, Atributos: Número de contratos, Aportó a la campaña, Valor total contratos).
4. **Tarea (N4): Identificar** la **tendencia** del valor contratado por el familiar del congresista desde el año 2007 al 2018. En términos de Tamara (**Identify** then **trend** of variable “Valor total contratos”).
5. **Tarea (N5): Resumir** la **distribución** del total contratatado por los familiares de los congresistas por su modalidad de contratación. En términos de Tamara (**Summarize** the **Distribution** of variables “Tipo de contrato”, Contratista” (Familiar contratista) y Valor total contratos).
6. **Tarea (N6): Explorar** la **distribución** de los familiares de los congresistas que tiene contratos con el estado. En términos de Tamara (**Explore** the **Distribution** of variable “contratista”).

### **3.2.3 HOW**

A continuación, se realiza la abstracción de los modismos que presentan adecuada y efectivamente los atributos del *dataset* suministrado por Cuestión Pública, con el fin de permitir al usuario (usuario del común) que consume la información publicada por el portal, conocer y acercarse al panorama general en relación con la contratación realizada por los familiares de los congresistas y así darles a estos usuarios herramientas para que evalúen sus próximos votos. Las siguientes visualizaciones se mostraran al usuario a modo de ***Storytelling.***

|  |  |
| --- | --- |
| * Modismo (Figura N1): Bar Chart * Marca: Línea * Canal: Posición vertical, color (tono) * Arrange : Separar, Ordenar , Alinear * Atributos: Valor total contratos de los familiares por Congresista * Descripción: Representa la distribución del valor contratado (del familiar) por congresista.   Aplica para las tareas N1 y N2 descritas en el punto anterior. |  |

|  |  |
| --- | --- |
| * Modismo: Radial node-link tree * Marca: link de conexión, punto para los nodos * Canal: radial axis orientation * Atributos: Congresista, Contratista (Familiar congresista), vínculo familiar, entidad contratante, Número de contratos, Valor total contratos * Descripción: Representa las relaciones (vínculo familiar) entre el congresista, contratista y entidad contrante. |  |

|  |  |
| --- | --- |
| * Modismo: Line chart * Marca:Línea/punto * Canal: Posición horizontal, Color (Tono). * Arrange: Express * Atributos: Años como contratista, Valor total contratos * Descripción: Representa el comportamiento del monto contratado por el familiar del congresista a través del tiempo.   Aplica para la tarea número 5. |  |

|  |  |
| --- | --- |
| * Modismo: Stacked bar chart * Marca: Line stacked * Canal: Posición horizontal, Longitud, Color (tono) * Arrange: Separar, ordenar y alinear * Atributos: Contratista (Familiar congresita), Tipo contrato, Valor total contratos * Descripción: Representa del monto total contratado por el familiar del congresista, como se distribuye por el tipo de contrato.   Aplica para la tarea número 6. |  |

**3.3 Usuarios**

La visualización estará enfocada a colombianos y colombianas con cierto interés en la política y el interés público. Los usuarios que se espera estén relacionados con la visualización son ciudadanos del promedio con habilidades computacionales y de análisis promedio, por esta razón la visualización debe ser de uso fácil sin necesidad de contar con grandes habilidades o entrenamiento previo para poder hacer uso de ésta y lograr abstraer los *insights.*

# **PROPUESTA**

Utilizando análisis visual se le quiere mostrar a los colombianos en modo de *Storytelling* que 1 de cada 2 congresistas tiene familiares contratando con el estado. Además, se quiere que los usuarios interactúen con la visualización y puedan hacer veeduría de los 10 congresistas más votados en las pasadas elecciones. Para lograr lo expuesto anteriormente se propone realizar la siguiente visualización, *Mocukup* versión 1 de la visualización propuesta:

**Modismo 1 - Bar Chart**

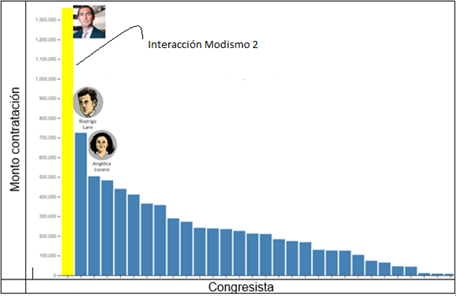


Figura N1.

**Modismo N2 - Tree**

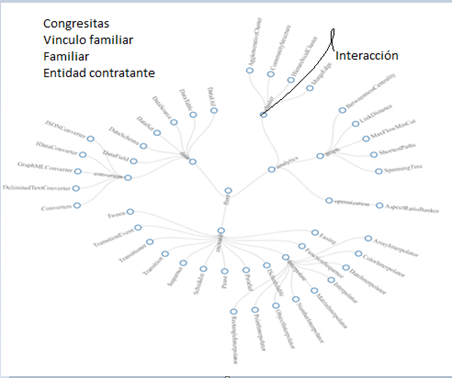


Figura N2.

|  |  |
| --- | --- |
| Modismo N3 | Modismo N4 |
|  |  |

Figura N3.

# **CONCLUSIÓN**

Cuestión pública ha Cuestión Pública ha adelantado sus investigaciones a los congresistas más votados que tienen familiares contratando con el estado para mostrar en su serie sabemos lo que hiciste. Su trabajo está bastante adelantado y se encuentra en la última etapa la cuál es encontrar la forma de exponer los datos recopilados al público, para esto se tomará como apoyo el análisis visual.

La aplicación del framework de Tamara, permitió abstraer los datos, tareas y modismos que permitan cumplir la necesidad de Cuestión Pública de transmitir de manera efectiva sus investigaciones periodísticas en el interés público.

La utilización de la visualización expuesta en el m*ocukup* será evaluada por uno de los representantes del portal Cuestión Pública, con el fin de determinar sí cumple sus expectativas. Así mismo se explorará la factibilidad de realizar con las visualizaciones un *storytelling*.

# **BIBLIOGRAFÍA**

[1]. T. Munzner. Visualization Analysis and Design. Taylor and Francis Group,6000

Broken Sound Parkway NW, SUite 300, 1nd ed., 2014. doi: ISBN 978-1-4665-

0891-0

[2]. <https://cuestionpublica.com>

[3]. <https://www.datasketch.co/>

[4]. [www.contratos.gov.co](http://www.contratos.gov.co)

[5]. <https://cuestionpublica.com/abc-sabemos-lo-que-hiciste/>

[6]. <https://cuestionpublica.com/nosotros/>

[7]. <http://johnguerra.co/classes/visual_analytics_fall_2018/>