

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

Avances

Integrantes

Santiago Suárez Conde

Martha Lucia Cárdenas Soto

Octubre 2018



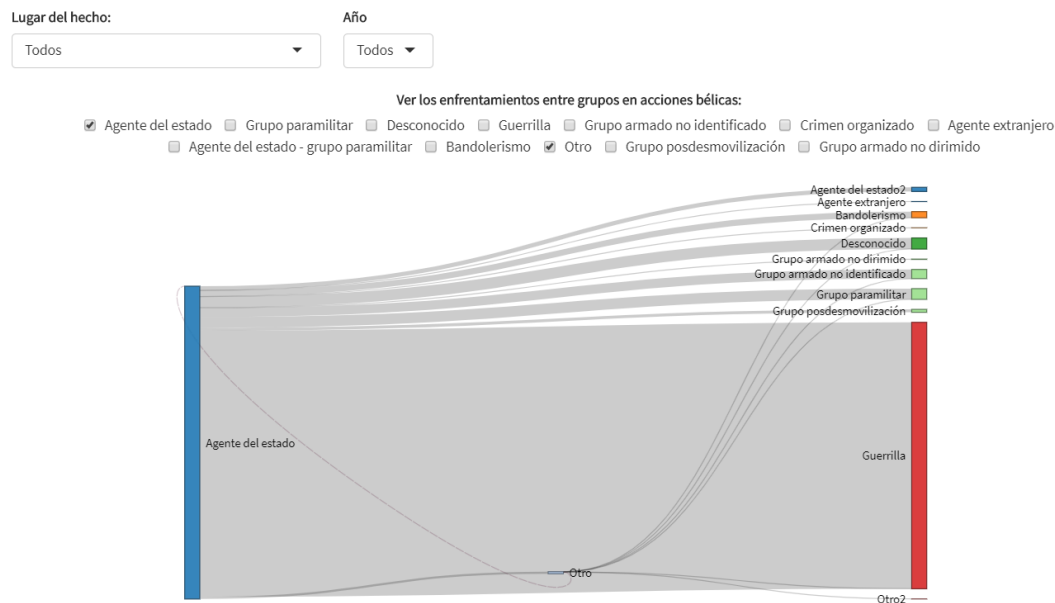
	<p>UNIVERSIDAD DE LOS ANDES</p> <p>Visual Analytics</p> <p>Profesor: John Alexis Guerra Gómez</p>	
---	---	---

TABLA DE CONTENIDO

1. ESTADO DEL ARTE	3
2. AVANCES	4
2.1 Data	4
2.2 Caracterización.....	4
2.2.1 WHAT.....	4
2.2.2 WHY	6
2.2.3 HOW – Método Propuesto	7
2.2.3 Storytelling.....	7
2.2.3.1 Concepto del experto.....	9
2.2.4.1 Concepto del experto	10
2.2.5 Modismo N2 - Tree.....	10
2.2.5.1 Concepto del experto.....	11
2.2.6 Modismo Line Chart	12
2.2.6.1 Concepto del experto.....	12
2.2.7 Modismo Stacked bar chart	13
2.2.7.1 Concepto del experto.....	13
BIBLIOGRAFÍA.....	14

1. ESTADO DEL ARTE

El experto señala que no ha encontrado visualizaciones que muestren algo parecido a la idea planteada, sin embargo, al explorar algunas visualizaciones colombianas, se encuentra la siguiente:



Esta visualización es tomada de <http://centrodememoriahistorica.gov.co/observatorio/bases-de-datos/datos/>, la cual muestra los enfrentamientos entre grupos en acciones bélicas en el país.

Tarea principal:

Explorar las **topologías** generadas por los enfrentamientos entre grupos en acciones bélicas. (En términos de tamara **Explore - Topology**).

Tarea secundaria:

- **Resumir** la **distribución** de los enfrentamientos por cada departamento. (En términos de Tamara **Summarize - Distribution**).
- **Identificar** las cantidades **extremas** de enfrentamientos entre dos grupos (En términos de Tamara **Identify - Extremes**)
- Enjoy

2. AVANCES

2.1 Data

Para la data preliminar entregada se realizaron las siguientes actividades de limpieza y normalización de los datos:

- Las columnas con cifras se ajustaron a valores numéricos, presentaban separados de miles que no eran interpretados correctamente al subir los datos.
- Las columnas numéricas que presentaban la palabra (NA, N/A) por cero.
- Los registros que contenían N/A en la mayoría de las columnas se descartaron.
- Se descargaron la lista de contratos por cada familiar del contratista, con el fin de relacionar a la data original el detalle de los contratos por Entidad Contratante y modalidad de contrato.
- Al realizar la desagregación la cantidad de registros de la data pasó de 75 a 292 registros.

2.2 Caracterización



En relación con la caracterización según el *framework* de Tamara [1], que abstraer a un lenguaje común *What*, *Why* y *How*, se describe cada uno y su correspondiente avance.

2.2.1 WHAT



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

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

Atributo		Tipo
Congresista		Categorico
Votos		Ordenado - Cuantitativo - Secuencial
Contratista		Categorico
Años como contratista		Ordenado - Ordinal
Vínculo		Categorico
Número de contratos		Ordenado - Cuantitativo - Secuencial

	UNIVERSIDAD DE LOS ANDES Visual Analytics Profesor: John Alexis Guerra Gómez	
---	---	---

Valor total contratos		Ordenado - Cuantitativo - Secuencial
Entidad (nuevo)		Categorico
Aportó a la campaña		Categorico
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 (Ajustado)		



	UNIVERSIDAD DE LOS ANDES Visual Analytics Profesor: John Alexis Guerra Gómez	
---	---	---

Nodos	Nombre (Congresita, Familiar, Entidad)	Categórico
Atributos - nodos	Grupo	Categórico
	nro_familiares	Ordenado - Cuantitativo - Secuencial
	nro_contratos_familiares	Ordenado - Cuantitativo - Secuencial
	nro_anios_contratos_familia	Ordenado - Cuantitativo - Secuencial
	monto_contratos_familiares	Ordenado - Cuantitativo - Secuencial
	aporto_al_campana	Categórico
	anio_financiacion	Ordenado - Ordinal
	monto_financiacion	Ordenado - Cuantitativo - Secuencial
Link	Source y target	Enlaces entre los nodos
Atributo - Link	Vinculo familiar_agrupado	Categórico

2.2.2 WHY

La abstracción de las tareas no presentó ajustes y la población objetivo se mantiene. Población objetivo: 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.

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”).

	<p>UNIVERSIDAD DE LOS ANDES</p> <p>Visual Analytics</p> <p>Profesor: John Alexis Guerra Gómez</p>	
---	---	---

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) Derivar** 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”).

2.2.3 HOW – Método Propuesto

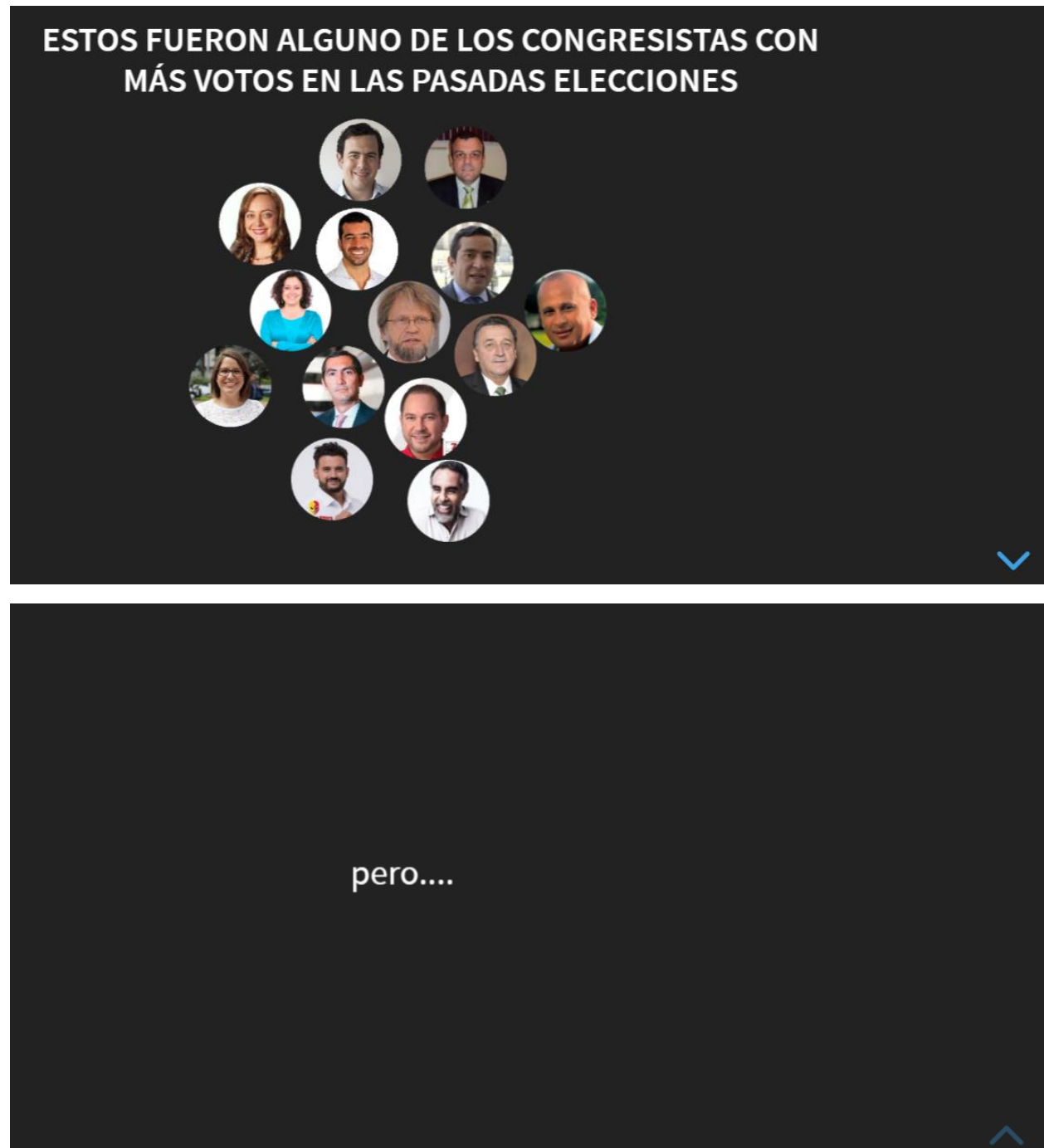
En el proceso que define la visualización teniendo en cuenta los datos, las tareas y los usuarios, se realizó la propuesta del *HOW*, Mediante la utilización de un *StoryTelling* y cuatro modismos.

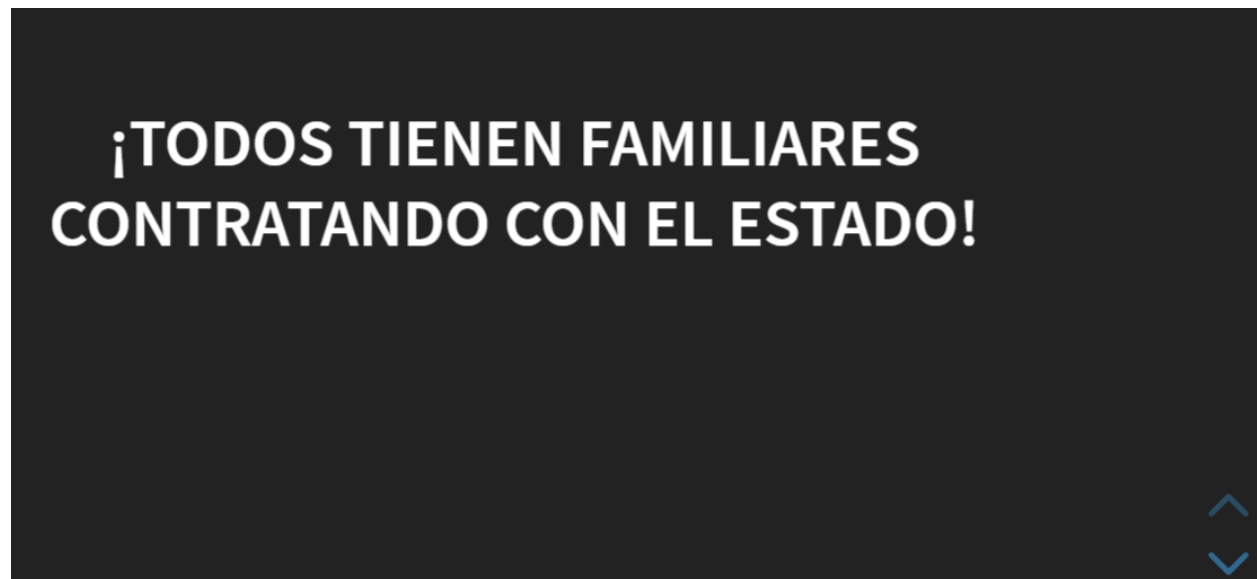
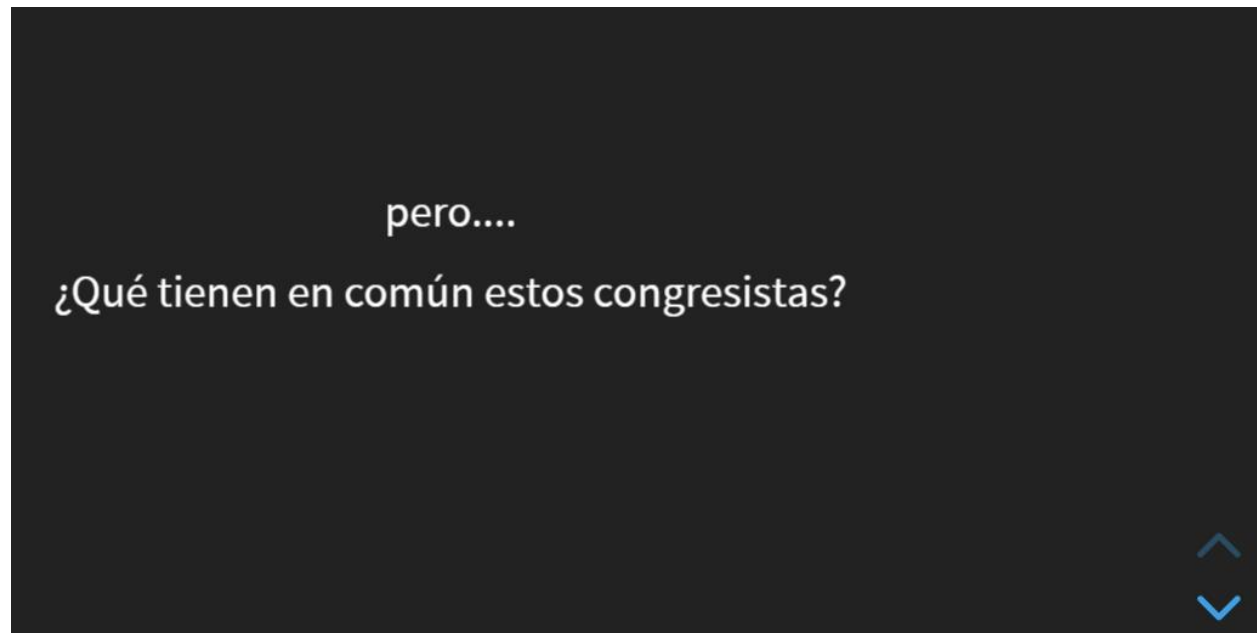
A continuación, se indican los resultados según la validación por parte del experto.

2.2.3 Storytelling

Se implementó una primera versión del Storytelling con la librería revealjs, D3js.V4, html5, Bootstrap y css. Este adelanto se realizó con unos datos provicionales brindados por Cuestión pública. El primer avance fue publicado en Github pages

(<https://santiagoconde0.github.io/familiaresCongresistas/>) en la siguiente imagen se muestra el avance.

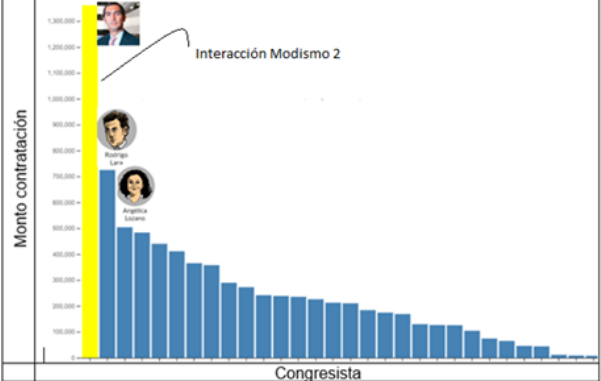




2.2.3.1 Concepto del experto

El uso del Storytelling fue bien recibido por el experto, quién validó y le pareció correcto. Como se está proponiendo transmitir la información de los congresistas, señala la necesidad de retirar los congresistas que no presenta familiares con contratación. Adicionalmente, suministrará algunos textos que se deberían incorporar en el recorrido por las visualizaciones.

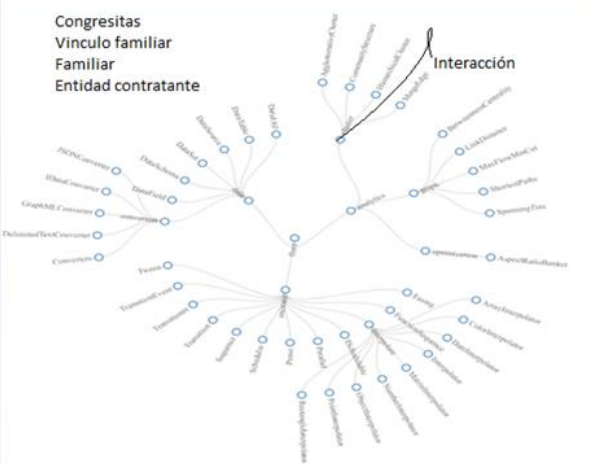
2.2.4 Modismo 1 - Bar Chart

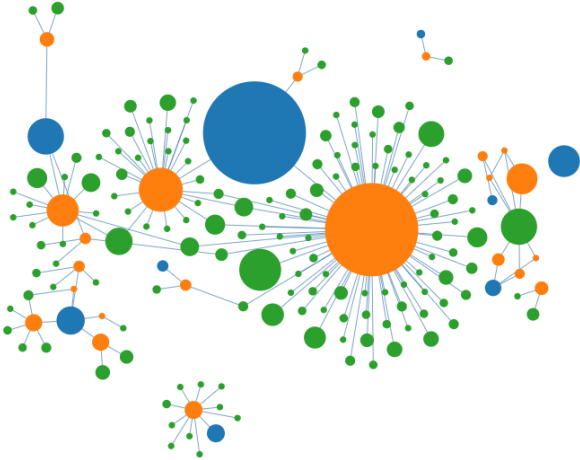
<ul style="list-style-type: none"> • Modismo (Figura N1): Bar Chart • Marca: Línea • Canal: Posición vertical, color (tono) • Arrange : Separar, Ordenar , Alinear • Atributos: <u>Valor total contratos</u> de los familiares por <u>Congresista</u> • Descripción: Representa la distribución del valor contratado (del familiar) por congresista. <p>Aplica para las tareas N1 y N2 descritas en el punto anterior.</p>	
--	--

2.2.4.1 Concepto del experto

El experto concuerda en que el modismo permite visualizar cuál congresista tiene el mayor monto de contratación realizada por sus familiares y sugiere adicionar el atributo del número de familiares que han celebrado contratos con el estado. Igualmente, le parece bien que sea el primero modismo que aparezca en el *Storytelling*.

2.2.5 Modismo N2 - Tree

<ul style="list-style-type: none"> • 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. 	
--	--



<ul style="list-style-type: none"> • Modismo: force-directed • Marca: <ul style="list-style-type: none"> Punto para los nodos (Atributo nombre que contiene el Congresista, Contratista (Familiarl, Entidad). Línea para los enlaces entre los nodos (Source (nombre), Target (nombre)) • Canal: <ul style="list-style-type: none"> Radial axis orientation Encode Arrange Separate y Order Map <u>color</u> Hue atributo grupo, <u>Tamaño</u> nro_contratos_familiares Manipulate: Select 	
--	--

Se realizó una versión de la red utilizando el sistema de fuerzas de d3 como lo muestra la imagen del modismo force-directed. El Dataset usado es el resultado de la derivación realizada de la data provisional brindada por Cuestión Pública.

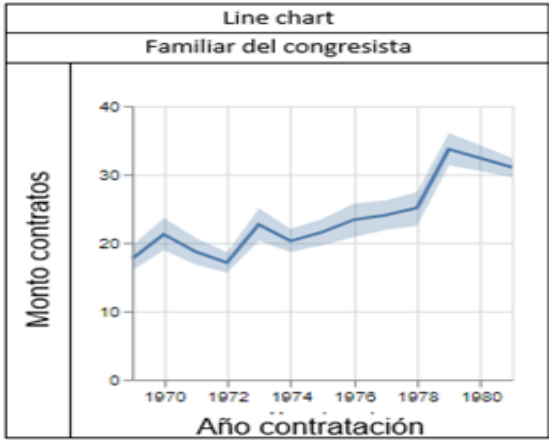
2.2.5.1 Concepto del experto

Según la retroalimentación recibida del experto, el modismo con la utilización de círculos no es el mas acertado para el usuario, dado que ellos han experimentado algo similar en su portal y al evaluar los clicks del sitio, las personas no navegan sobre la visualización. Igualmente, indica que no es muy entendible la información a mostrar.

Por lo anterior, se buscará otro modismo que permita transmitir de manera más efectiva para el tipo de usuario, las relaciones entre el congresista, contratista según su vínculo familiar y la entidad contratante. Por lo anterior, se propone inicialmente la posibilidad de realizar algo parecido al SpaceTree presentado en clase.

	<p>UNIVERSIDAD DE LOS ANDES</p> <p>Visual Analytics</p> <p>Profesor: John Alexis Guerra Gómez</p>	
---	---	---

2.2.6 Modismo Line Chart

<ul style="list-style-type: none"> • 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. <p>Aplica para la tarea número 5.</p>	
--	--

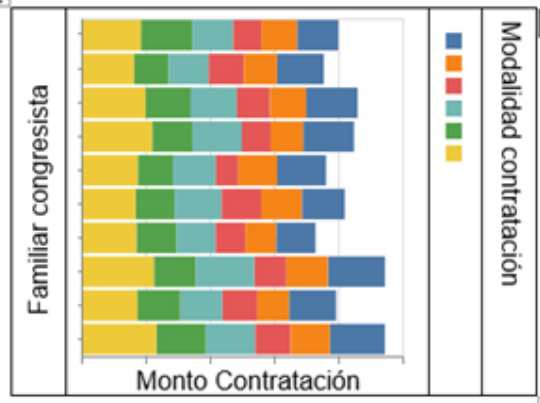
2.2.6.1 Concepto del experto

El experto precisa que el modismo le permite ver en los años como el familiar del congresista ha recibido el monto de contratos, sin embargo, indica que es necesario agregar (adicionar a la data) el año en cada Congresista fue elegido, para visualizar, si el familiar empezó a contratar desde ese mismo año o después, igualmente, si el monto contratado ha aumentado.

Precisa que es necesario mostrar visualmente el cambio cuando el congresista es elegido vs los montos de contratación que ha celebrado su familiar, *“de manera que la gente entienda que aunque la ley no inhabilita que los familiares contraten con el estado, es éticamente reprochable, como por ejemplo: el caso de Angelica Lozano, donde empieza como Congresista el 2014 y su hermano con una empresa de muebles, comienza a contratar en el 2015, en tres años obtiene contratos por 1300 millones, que termina con un contrato de 300 millones repartiendo Fritolay en las cárceles”*, señala el experto.



Por otro lado, el experto expresa la importancia de la desagregación realizada a la data original para establecer las relaciones, la data resultante se envió al experto con el fin de agregar la tipificación por objeto de contrato.

2.2.7 Modismo Stacked bar chart

<ul style="list-style-type: none"> • Modismo: Stacked bar chart • Marca: Line stacked • Canal: Posición horizontal, Longitud, Color (tono) • Arrange: Separar, ordenar y alinear • Atributos: Contratista (Familiar congresista), 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. 	<p>Horizontal Stacked Bar Chart</p> 
---	--

2.2.7.1 Concepto del experto

El experto indica que la modalidad de contratación por sí sola, para el tipo de usuario, no dice mucho y no evidencia el caso de Angelica Lozano expuesto en el punto 2.2.6.1 y el “caso del primo de Vargil que tiene una empresa de infraestructura y celebra contratación directa”, por lo tanto, se adicionará (agregar a la data) una tipificación según el objeto del contrato. Adicionalmente, sugiere mostrar el mismo modismo por vínculo familiar para visualizar la distribución del monto contratado.

	<p>UNIVERSIDAD DE LOS ANDES</p> <p>Visual Analytics</p> <p>Profesor: John Alexis Guerra Gómez</p>	
---	--	---

BIBLIOGRAFÍA

- [1]. T. Munzner. Visualization Analysis and Design. Taylor and Francis Group, 6000 Broken Sound Parkway NW, Suite 300, 1st ed., 2014. doi: ISBN 978-1-4665-0891-0
- [2]. <https://cuestionpublica.com>
- [3]. <https://www.datasketch.co/>
- [4]. 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/