

Caso práctico

1) Objetivo del análisis

Analizar los tópicos y palabras más usadas en 3 medios Ecuatorianos (El Comercio, El Universo y Primicias) al referirse al término de búsqueda Andres Arauz y Rafael Correa

2) Web Scraping

El presente trabajo se basó en la obtención de las noticias de dichos medios a través de web scraping utilizando R.

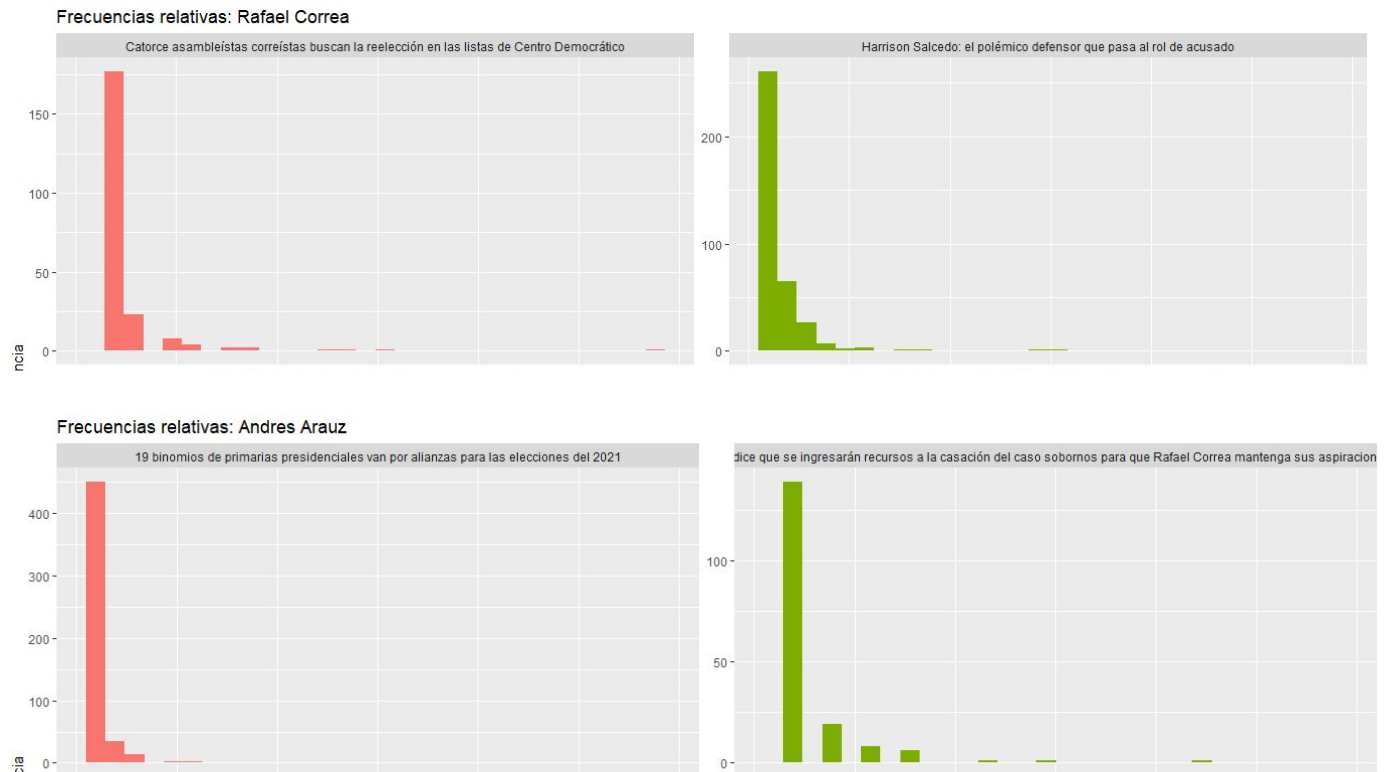
3) Script de R

La estructuración y pre-procesamiento de los datos se encuentra disponible en el código adjunto archivo "Complete.R"

4) Realizar el análisis exploratorio y conclusiones

Se detalla a Continuación

Frecuencia Relativa

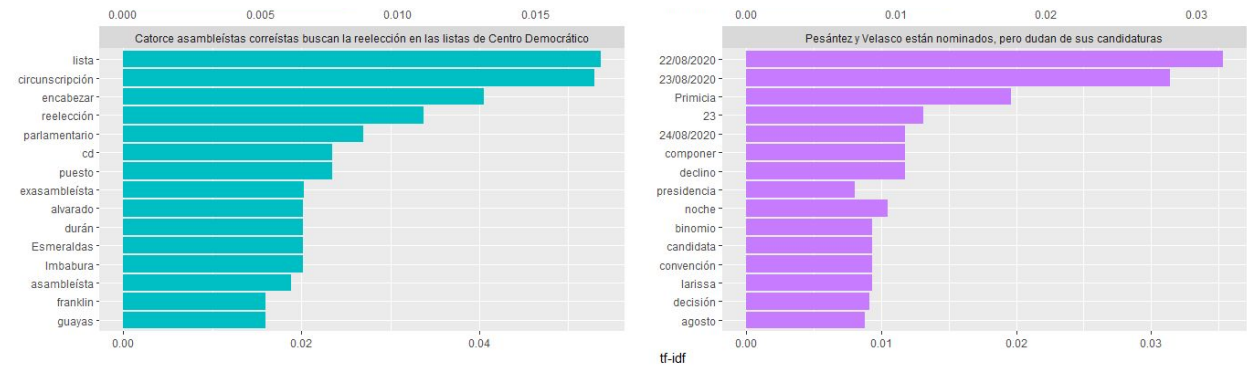


En ambas búsquedas podemos observar el caso donde la frecuencia relativa de las palabras presenta una distribución con colas largas con respecto al total de palabras. Esto indica la existencia de palabras únicas.

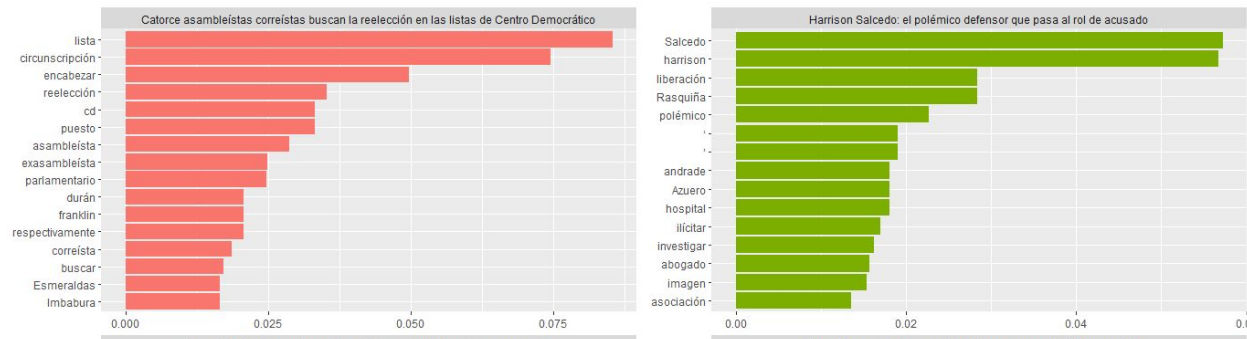
Tf-idf (Busca las palabras importantes en el documento)

Esta medida señala la frecuencia con la que un término específico aparece en un documento. Se puede calcular dividiendo el número de veces que se utiliza la palabra clave en el documento por el número total de palabras de dicho documento.

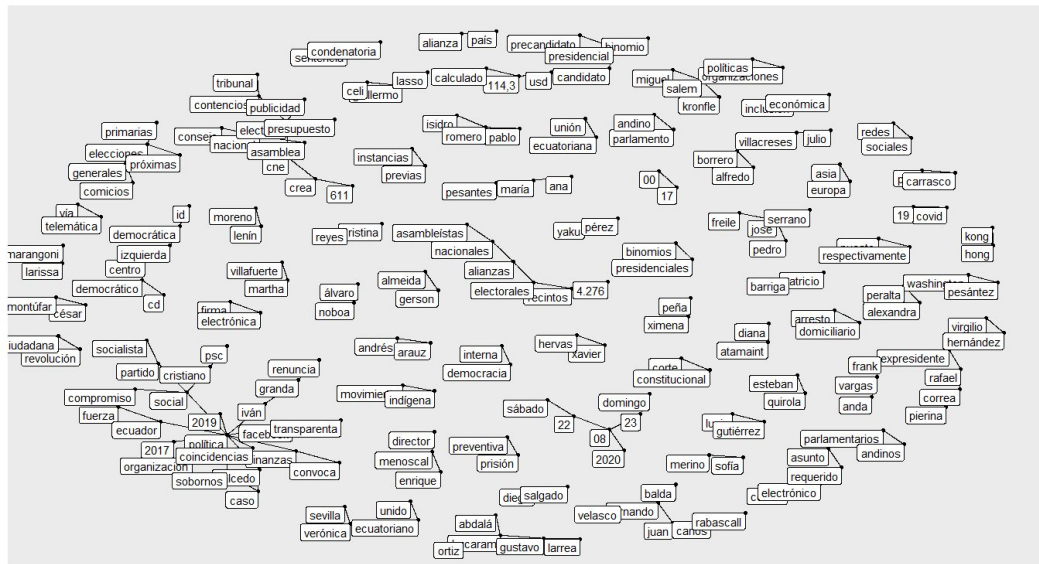
Andres Arauz



Rafael Correa

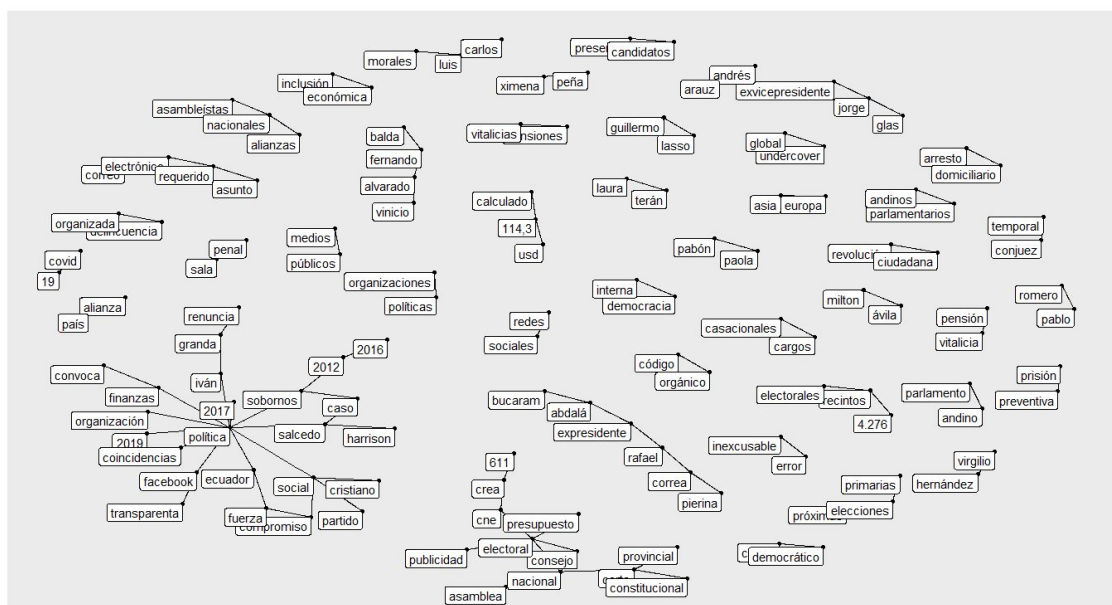


Andres Arauz



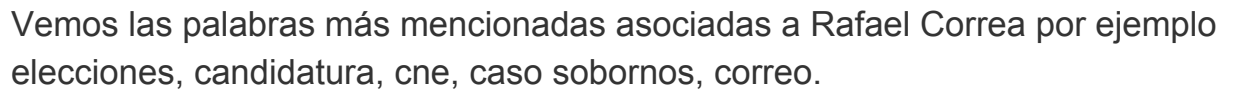
Podemos ver como las noticias giran entorno a la precandidatura de Andres Arauz en su binomio denominado de la esperanza, en un trabajo futuro se analizará este hashtag [#BinomioDeLaEsperanza](#) en twitter para ver los sentimientos asociados y la reacciones sobre el mismo. Vemos un ejemplo Primarias, próximas elecciones, comicios.

Rafael Correa



Nube de palabras Rafael Correa

transparentar pedernales precandidatura ministerio binomio empresario posibilidad postular postulari
inscripcion requerido 2012-2016 consulta provincial septiembre judicial hernandez
temporal matriciacion democratico p provincial septiembre judicial hernandez
previo arduo alvarado presentacion fiscalia presentacion fiscalia presentacion fiscalia
presupuesto



Granda
forma almeida 22/08/2020 gustavo alfredo avanza inscribir sevilla prision insistir octubre correio creaban libertas velasco
homel proceso

Extracción de palabras clave

Rafael Correa y Andres Arauz Keywords respectivamente

	keyword	ngram	freq	rake
1	lista	1	2	8.7567568
2	correa	1	2	8.5961538
3	registro	1	2	8.5000000
4	investigación	1	4	8.4736842
5	tribunal	1	2	8.3200000
6	aceptar	1	2	7.8461538
7	participar	1	2	7.5789474
8	política	1	2	7.4545455
9	asambleísta	1	2	7.3947368
10	abogado	1	4	7.3333333
11	magistrado	1	2	7.3000000
12	exmandatario	1	2	7.2857143
13	caso	1	5	6.8101266
14	juez	1	2	6.7931034
15	justicia	1	2	6.7083333
16	binomio	1	2	6.5833333

	keyword	ngram	freq	rake
1	insistir	1	2	10.20000000
2	lista	1	2	8.69387755
3	fórmula	1	2	8.00000000
4	Vicepresidencia	1	2	7.88000000
5	aceptar	1	4	7.84090909
6	precandidatura	1	3	7.79545455
7	cargo	1	2	7.38461538
8	precandidato	1	3	7.32653061
9	Arauz	1	2	7.10869565
10	movimiento	1	5	7.10769231
11	política	1	2	6.93939394
12	reglamento	1	2	6.86956522
13	binomio	1	5	6.72307692
14	carta	1	3	6.21428571
15	cne	1	3	6.07575758
16	alianza	1	8	6.01754386

Conclusiones

Dada la limitante de google news en cuentas gratuitas no se puede acceder a más información para analizar la evolución de las noticias sobre candidatos políticos, lo que quería observar de estas noticias era la importancia que el peculiar caso de estos candidatos, es decir como los medios de prensa escrita hacen el seguimiento a un juicio político de un candidato que quiere inscribirse para estas elecciones y de su compañero de bancada como estrategia política un personaje sin manchas judiciales pero que queda relegado en importancia por los medios, claramente esta es mi percepción, esta es la parte donde los

resultados deben ser interpretados por personas expertas en temas políticos, sociales y demás, con lo que me doy cuenta que es necesaria la intervención de más profesionales que añadan valor a los resultados expuestos, además de ayudar para que esta herramienta tanto el text mining y el web scraping tengan resultados deseados.

Me ha parecido magnífico poder jugar con estos temas y agradezco a Hugo Porras quien nos mostró el camino a seguir.