

La base de datos a analizar serán las Circulares de la Corte Electoral disponibles de forma abierta en <https://www.impo.com.uy/basecorteelectoral/>. Las mismas son comunicaciones que realiza la Corporación de forma de dar a conocer cuestiones relevantes referidas a reglamentaciones aprobadas, modificaciones en la atención al público, integración y conformación de agrupaciones de los partidos políticos, entre otros.

Se trabajará con las Circulares de los años 1924 a 2017¹, totalizando casi 100 años.

Realizando búsquedas por cada año considerado en la web previamente citada es posible obtener el intervalo de Circulares por año y así el número de comunicaciones. Esto servirá para definir los vectores que permitan rescatar las Circulares de la web como se verá a continuación. Al definir los intervalos por año es posible realizar la siguiente gráfica.

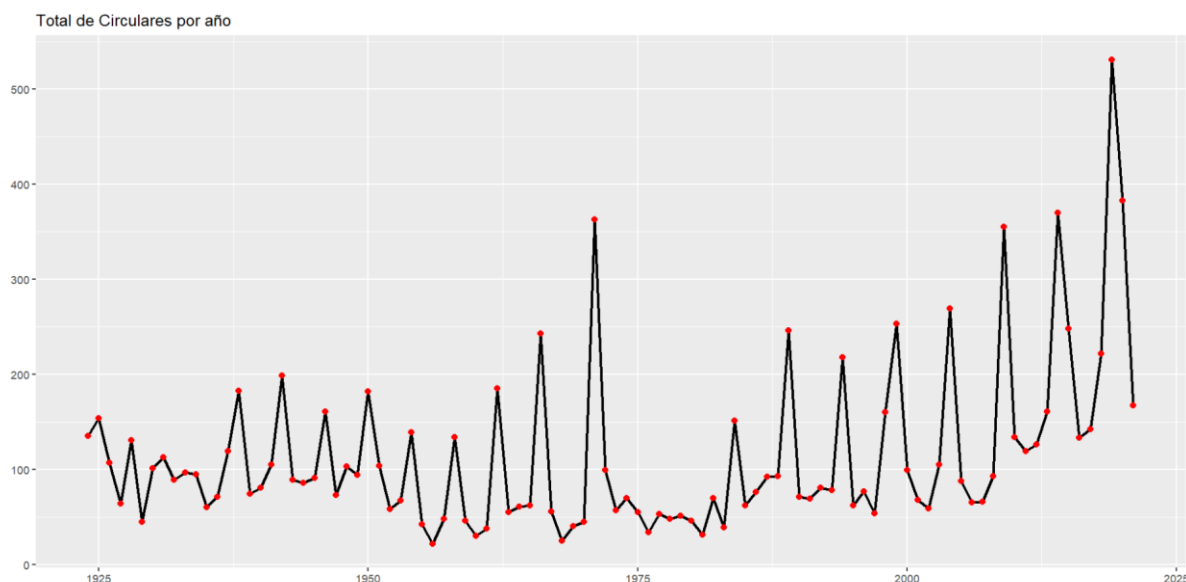


Figura 1: cantidad de Circulares en el período considerado

Esta sola información permitiría en principio realizar un análisis basado en series de tiempo para detectar niveles de medias, ciclos y tendencias de la serie a la vez que predecir valores futuros. Una forma de comprender estas oscilaciones asimismo podría encontrarse al incorporar otras variables en el análisis como el hecho de que se trate de "Año electoral" o "Dictadura". Otra forma de agrupar a las Circulares y por ende, de definir la unidad de análisis, podría ser las distintas integraciones de la Corte Electoral, los distintos regímenes de gobierno (colegiado, unipersonal), los períodos determinados por el gobierno bajo distintos partidos políticos, entre otros.

En el presente trabajo no se explorará esa línea de investigación, sin embargo, se reconoce como posibilidad. Probablemente, no todas estas categorías resulten relevantes para explicar diferencias cuantitativas y/o cualitativas entre períodos pero bien valdría la pena testearlo. Vale la pena comentar

¹ Por razones que se desconocen, el código pudo rescatar hasta las Circulares de 2017. No se cree igualmente que las conclusiones sean significativamente diferentes si se incluyen el resto de los años hasta el presente.

Para obtener una nube de palabras, debe crearse un archivo DFM y a su vez, suprimir aquellas palabras que generen ruido o que no tengan un significado en especial en el contexto que se analiza. Vemos que las palabras más utilizadas en los casi 100 años de comunicaciones tienen que ver con cuestiones electorales como “junta”, “electorales”, “acuerdo”, “partido”, entre otros.



También es posible crear un histograma con las palabras más utilizadas durante este período histórico.

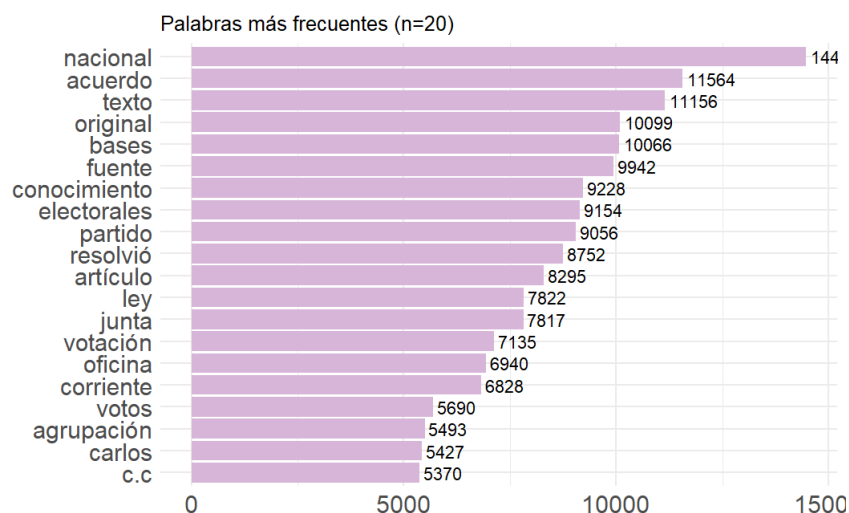


Figura 3: Las 20 palabras más frecuentes en las Circulares

Obviamente, son las mismas, pero resulta más claro percibir cuáles son las que se presentan con mayor frecuencia.

Si por ejemplo tomamos la segunda palabra (“acuerdo”), es posible ver la relación que la misma presenta con otras palabras dentro de todo el corpus de Circulares.

Primeras 15 correlaciones de palabras con “acuerdo”

Palabra	Correlación
resolvió	0.9171623
corriente	0.8665923
presente	0.8600396
texto	0.8415948
fuelle	0.8351532
bases	0.8329756
ppdo	0.8249448
octubre	0.8230356
noviembre	0.8221381
original	0.8214632
próximo	0.8195633
partido	0.8114817
tener	0.7997890
saber	0.7975414
políticos	0.7971855