

## Indicators of Electoral Systems and Party Systems: R-package: esaps

Nicolás Schmidt<sup>1</sup>

<sup>1</sup>Departamento de Ciencia Política FCS-Udelar,

**Palabras Claves:** datos electorales, análisis estadístico, ciencia política

En ciencia política, el uso de datos electorales para su aplicación a modelos estadísticos implica afrontar una dificultad relativa a la multidimensionalidad de los datos. Esta dificultad proviene de las diferencias entre unidades de análisis y observaciones ya que, entre distintas unidades de análisis (países, estados, municipios, provincias, cantones, etc.), varía el número y el nombre de los partidos políticos. Los indicadores construidos a partir de estos datos, tanto del sistema electoral como del sistema de partidos, se estructuran en función de los votos (dato electoral) que obtiene cada partido en una elección. Es por esta razón que la cantidad de partidos políticos en cada unidad de análisis es lo que termina estructurando los datos.

Por ejemplo, si se desea trabajar con datos electorales en países de América Latina en el período que va desde 1990 hasta 2010, uno de los principales problemas que enfrenta un investigador es que los años de elecciones no coinciden en todos los países y que la cantidad de partidos varía en el tiempo y en el espacio. Esto significa que consolidar toda la información en una única base de datos no resulte ser la estrategia más eficiente ya que pegar esas bases implica aumentar significativamente la cantidad de valores NA. De este modo, es frecuente encontrarse con flujos de trabajo que implican el ingreso de datos por unidad de análisis, utilizando por ejemplo una planilla de cálculo por cada unidad o en un único libro (de excel o calc) con varias hojas y aplicando los indicadores a cada unidad de manera separada. El principal problema de este flujo de trabajo es que se multiplica el trabajo (que es posible automatizar) con los datos conforme aumenta la unidad de análisis y el proceso se vuelve engorroso. Es en esta línea que se trabajó en la elaboración de un paquete de R que, a partir de esta problemática concreta, aporte soluciones para un mejor manejo de los datos. El paquete esaps se estructura a partir de este flujo de trabajo ofreciendo una solución para trabajar de manera sencilla bajo este flujo y permitiendo, al mismo tiempo, trabajar con bases consolidadas o con otros flujos de trabajo. La característica principal que ofrece la función central del paquete (esaps\_object) consiste en que el proceso de cargado y ensamblado de las múltiples bases de datos para el posterior cálculo de distintos indicadores que forman parte del paquete se pueda hacer de manera sencilla, sistemática y eficiente.

Esta función, genera un objeto de la clase esaps que prepara los datos para poder usar cualquiera de las otras funciones del paquete (los distintos indicadores). En el proceso de generación del objeto, se realizan distintos chequeos de estructura y con-

sistencia de los datos, por lo que si hay problemas la función reporta los mismos con bastante detalle para que el usuario no pierda tiempo buscando errores y pueda solucionarlos previamente al uso de los indicadores.

Construir la totalidad del paquete en función de la estructuración de los datos persigue dos objetivos. En primer lugar, si existen problemas en los datos, los mismos serán identificados en el inicio del proceso (creación del objeto de clase `esaps`). En segundo lugar, permite que se pueden incorporar nuevas funciones al paquete siguiendo una determinada estructura de datos de forma sencilla. La elaboración de este paquete además atiende los aspectos más relevantes de la literatura del área. En este sentido, en el estudio de los sistemas de partidos existen dos indicadores importantes y que son los más comúnmente utilizados. Uno de ellos es el número efectivo de partidos (Laakso, Taagapera, 1979), y el otro es la volatilidad electoral (Pedersen, 1979). A su vez, la volatilidad electoral tiene variaciones en sus cálculos dependiendo el aspecto del sistema de partidos sobre el que se desea poner el énfasis: ya sea desde la perspectiva de la volatilidad generada entre los partidos estables en el sistema (Powell, Tucker, 2014; Mainwaring, *et.al.*, 2017) o la generada entre los partidos que se encuentran en equilibrio en el sistema (Torcal, Lago, 2015). Estos indicadores son los que se pueden realizar con el paquete.

Cada uno de estos indicadores tiene particularidades en su construcción que están estrechamente ligados a una estructura de datos determinada para su cálculo. Por ejemplo, la volatilidad electoral es la suma de las diferencias netas entre partidos en dos elecciones sucesivas sobre dos y es común que los partidos, entre elección y elección, desaparezcan, formen coaliciones o aparezcan partidos nuevos. Esto es particularmente un problema propio del objeto de estudio. Este problema tiene una simple solución si el número de casos es pequeño ya que es relativamente fácil identificar divisiones entre partidos o coaliciones. Pero a medida que aumenta el número de casos (países) y el número de observaciones (elecciones) la tarea se vuelve más engorrosa y es necesario modificar con frecuencia la asignación de los votos a los "verdaderos" partidos, aunque tengan nombre distinto. Esto implica que hay que realizar los cálculos de este indicador reiteradas veces en función de la información histórica que se obtiene. Este paquete de R facilita esta problemática al automatizar el proceso de cargado y cálculo de indicadores con independencia de la cantidad de veces que se modifiquen los datos.

En suma, la utilización de R en el proceso de generación, ensamblado, y testeo de bases de datos electorales reporta un beneficio sustantivo. En el caso concreto descrito, para el flujo de trabajo con datos electorales, la utilización de R y la generación de funciones que permitan automatizar procesos implica una ganancia en términos de tiempo y de control sobre el proceso de generación de los indicadores que es sustantiva. El paquete `esaps` es producto de estas dificultades que he enfrentado como investigador en el área de sistemas de partidos.