

# Práctica 1 (35% nota final) Presentación

## 1. Contexto.

La información recolectada es sobre los sondeos de las elecciones generales con la estimación porcentual de cada partido, en el cual aparecen sondeos de diversas fuentes y medios gracias al sitio web electocracia que ofrece esta información porque el equipo que hay detrás persigue un ideal difícil de observar en política como es la objetividad, la integridad, la independencia y la ecuanimidad. En dicha página no solo hay información electoral y política, también nos brindan información empresarial (indicadores, publicidad, consultoría de marketing...)

## 2. Definir un título para el data set.

Evolución histórica y estimaciones porcentuales de los principales partidos políticos en las elecciones generales de España según los sondeos (2017-2019).

## 3. Descripción de la data set.

En el data set nos encontramos con todos los sondeos de los 5 partidos políticos con mayor influencia de España con sus respectivas estimaciones porcentuales desde el más reciente (11/04/19) hasta el primero del 2017 (08/01/27) y ofreciendo entre otros datos el medio de comunicación, su fuente y el tamaño.

## 4. Representación gráfica.

La idea de la representación gráfica es mostrar en base a la fecha de publicación todas las estimaciones porcentuales de los 5 partidos principales (es decir un gráfico de líneas con múltiples columnas). En otras palabras, en el eje horizontal se encontraría la fecha de publicación y en el eje vertical el rango de estimaciones porcentuales y cada línea representa un partido.

(nos hemos retrasado con el gráfico, pero en breves estará hecho en el código no está como se pretende que esté)

## 5. Contenido.

Se recogen las siguientes características:

- MEDIO: el medio de comunicación por el que se ha informado del sondeo
- FUENTE: la fuente que ha realizado el sondeo
- PUBLICACIÓN: fecha de publicación del sondeo
- FIN CAMPO: última fecha en la que se cogieron encuestas

- TAMAÑO: viene refiriéndose al tamaño de cada sondeo
- PSOE: Partido Socialista Obrero Español
- PP: Partido Popular
- Cs: Ciudadanos
- UP: Unidos Podemos
- VOX: partido político que no son siglas su nombre, es abreviatura en latín 'voz'.
- Otros/Blancos: resto de partidos políticos o en blanco
- MUESTRA:
- FECHA:
- FIABILIDAD: fiabilidad según los autores del sitio web del sondeo
- VALOR (1-5): valor que se le acredita al sondeo

Los datos han sido recolectados mediante técnicas de web scraping desarrolladas en el fichero code para coger los datos de la página web HTML.

## 6. Agradecimientos.

Agradecemos al equipo de Electocracia (José Miguel Silva, José Manuel San Millán, Rubén Rodríguez y Miguel Ángel Ramos) por ofrecer información política de calidad e independiente.

## 7. Inspiración.

Primero de todo por ser un tema de extrema actualidad (el último sondeo es del 11/04/2019), también por el interés mutuo por la política actual, sin mencionar que el data set da puede dar mucho de que hablar y analizar; por ejemplo, se puede establecer una comparación de este data set con los resultados reales de las elecciones para ver la credibilidad y grado de acierto de los sondeos. No deja de ser útil porque posee el data set sondeos desde 2017 hasta ahora, lo cual da para contestar preguntas de evolución e históricas, como observar la evolución decreciente del PP en los últimos años. Y con 226 sondeos lo que significa que el data set posee información relevante.

## 8. Licencia.

Hemos escogido la licencia CC BY-SA 4.0 por las siguientes razones:

- Consideramos importante reconocer el trabajo de electocracia y con esta licencia se le reconoce el trabajo original y si se han hecho aportaciones o cambios.
- Se puede hacer un uso comercial, así es más probable que se de a conocer más este sitio web que cede datos de los sondeos de forma gratuita e independiente.

- Se debe compartir igual, lo que significa que si se cambia o transforma el proyecto deberá seguirse la misma licencia, lo que promueve este trabajo del autor original y darse más a conocer.

**Código.** Se ha realizado con Python

## 9. Data set. Presentar el data set en formato CSV

El CSV se encuentra en la carpeta de Git Hub.

## Recursos

Los siguientes recursos son de utilidad por la realización de la PEC:

- Subirats, L., Calvo, M. (2018). Web Scraping. Editorial UOC.
- Masip, D. El lenguaje Python. Editorial UOC.
- Lawson, R. (2015). Web Scraping with Python. Packt Publishing Ltd. Chapter 2. Scraping the Data.