

Grado en Ciencia e Ingeniería de Datos Visualización de Datos

Práctica 1: Análisis Elecciones a Rector/a UAM

Comienzo	G-7312: 18 de septiembre de 2025 / G-7311: 19 de septiembre de 2025
Entrega	G-7312: 16 de octubre de 2025 (Hora: 09:55) / G-7311: 17 de octubre de 2025 (Hora: 09:55)

Introducción

En esta práctica trabajaremos con datos reales de las elecciones a Rector/a de la Universidad Autónoma de Madrid (UAM), celebradas desde el año 2017. El objetivo es utilizar estos datos como base para aprender a **visualizar información con las librerías matplotlib y seaborn**, dos de las más utilizadas en ciencia de datos para la generación de gráficos.

Contexto del problema

Las elecciones a Rector/a en la UAM se celebran cada cuatro años y desde 2024 cada seis años. Cada año de elecciones pueden constar de una o dos vueltas (v1, v2), dependiendo de si hay más de dos candidatos/as y un/a candidato/a no obtiene mayoría absoluta en la primera votación. Además, en algunos años se celebraron dos procesos electorales diferentes (por ejemplo, en 2025a y 2025b). El objetivo de esta práctica será explorar la participación y los resultados de voto en cada elección, representarlos gráficamente y compararlos entre diferentes años y vueltas.

Datos disponibles

Para cada proceso electoral disponemos de dos ficheros CSV, cuyo nombre sigue la siguiente estructura:

```
[AÑO]_v[VUELTA]_participation_data.csv  
[AÑO] v[VUELTA] votes data.csv
```

Donde:

- **AÑO**: año de la elección (ejemplo: 2017, 2021, 2025a, 2025b).
- **VUELTA**: vuelta de la elección (1 = primera vuelta, 2 = segunda vuelta).

Estructura de los ficheros

Los datos incluidos en cada fichero son los siguientes:

Ficheros de participación (*_participation_data.csv)

- **Center**: Centro o facultad de la UAM donde se emitieron los votos.
- **Category**: Estamento al que pertenecen los votantes (Profesorado permanente, Profesorado no permanente, PDIF, Estudiantes, PTGAS).
- **Censo**: Número total de personas con derecho a voto en ese estamento y centro.
- **Votos**: Número de personas que acudieron a votar.

Ficheros de votos (*_votes_data.csv)

- **Center**: Centro o facultad de la UAM donde se emitieron los votos.
- **Candidate**: Nombre del candidato/a.
- **Category**: Estamento al que pertenecen los votantes.

- **Votos:** Número de votos recibidos por ese candidato/a (o en blanco/nulos) en el estamento y centro correspondiente.

Ponderación de votos por estamento

En estas elecciones, los votos no tienen el mismo peso para todos los colectivos, sino que se aplica una ponderación:

- Profesorado permanente: 55%
- Profesorado no permanente: 5%
- PDIF: 4%
- Estudiantes: 27%
- PTGAS: 9%

Para calcular el porcentaje final de apoyo a cada candidatura se debe:

1. Calcular el porcentaje de votos que recibe cada candidatura en cada estamento.
2. Multiplicar ese porcentaje por el peso del estamento.
3. Sumar los resultados de todos los estamentos.

Ejemplo simplificado:

Supongamos que un candidato obtiene los siguientes resultados para el estamento de Profesor permanente:

- 60% de los votos $\rightarrow 60 \times 0.55 = 33\%$ **de los votos ponderados.**

De este modo, el resultado oficial no depende únicamente del número bruto de votos, sino de la distribución por estamentos.

EJERCICIOS

Parte 1: Gráficos básicos con matplotlib

En esta primera parte se debe realizar varias visualizaciones a partir de los datos de las elecciones a Rectorado de la UAM proporcionados. El objetivo no es solo generar los gráficos, sino también justificar la elección del tipo de representación (barras, líneas, circular, etc.), en función de lo que mejor se adapte a la información que se quiere transmitir.

Todos los gráficos deben realizarse con la librería matplotlib. No se especifica el tipo de gráfico que debes usar: deberás decidir qué tipo de representación es más adecuada en cada caso (barras, líneas, etc.). En el notebook, tras cada gráfico deberás escribir una breve justificación del tipo de gráfico que has elegido, además de responder a las preguntas de cada apartado.

Apartado 1: Participación por estamento (Elecciones 2021)

Representa los porcentajes de participación por estamento (Profesorado permanente, Profesorado no permanente, PDIF, Estudiantes y PTGAS) en el conjunto de la UAM, para la primera y la segunda vuelta de las elecciones de 2021.

Cada una de las vueltas deberán aparecer en un gráfico separados. Una vez realizado, responde a las siguientes preguntas:

- ¿En qué vuelta se observa una mayor participación por cada estamento?
- ¿En cuál de los estamentos, esta diferencia es mayor?

Apartado 2: Votos ponderados por candidatura (Elecciones 2021)

Representa el porcentaje de votos ponderados obtenidos por todos los candidatos en el conjunto de la UAM por cada uno de los estamentos, tanto en la primera como en la segunda vuelta de las elecciones de 2021. También el porcentaje total de los votos ponderados por todos los candidatos, es decir, sin diferenciar los estamentos de la UAM.

Una vez realizado, ve a la página oficial de la UAM donde se muestran los resultados de las elecciones. Compara los gráficos que usan para mostrar esos mismos resultados y explica brevemente en qué características tus gráficos mejoran a los de la página oficial de la UAM.

Página web elecciones UAM (Elecciones 2017):
https://sede.uam.es/UAM_Formularios2/secretariaGeneral/idActasElectoresRectorado/informeVotacionRectorado.jsf?fechaElecciones=04-05-2017

Apartado 3: Histórico de participación por estamento (2017–2025)

Representa la evolución de la participación de cada estamento en todas las elecciones disponibles (2017, 2021, 2025a, 2025b), incluyendo las vueltas correspondientes. El gráfico debe permitir comparar la tendencia de participación de cada estamento a lo largo del tiempo. Una vez realizado, responde a las siguientes preguntas:

- ¿En qué elecciones ha habido más participación en cada uno de los estamentos?

Parte 2: Múltiples gráficos con matplotlib (subplots)

En esta segunda parte vamos a repetir los mismos ejercicios de la Parte 1, pero ahora desglosando los resultados por centro o facultad de la UAM. El objetivo es aprender a trabajar con múltiples gráficos en una misma figura (subplots) y a mejorar su presentación visual.

Todos los gráficos deben realizarse con la librería matplotlib. No se especifica la disposición de los subplots ni las modificaciones de estilo que debes aplicar: deberás decidir qué organización es más adecuada en cada caso (rejilla de subplots, disposición por filas o columnas, etc.). En el notebook, tras cada figura deberás escribir una breve justificación explicando por qué has elegido esa representación con subplots y, en caso de haber aplicado modificaciones de estilo (como compartir ejes, unificar la leyenda, modificar colores o tipografías, ajustar espaciados, etc.), deberás justificar también esas decisiones de diseño.

Apartado 1: Participación por estamento y centro (2021)

Representa los porcentajes de participación por estamento, desglosados por centro, para la primera y la segunda vuelta de las elecciones de 2021. Cada subplot debe mostrar los datos de un centro distinto. Una vez realizado, responde a la siguiente pregunta:

- ¿Qué centro y qué estamento tuvo la máxima participación en cada una de las vueltas de las elecciones de 2021?

Apartado 2: Votos ponderados por candidatura y centro (2021)

Representa los porcentajes de votos ponderados por candidato, desglosados por centro, tanto en la primera como en la segunda vuelta de las elecciones de 2021. Cada subplot debe corresponder a un centro distinto.

- ¿En qué centro obtuvo el/la candidato/a ganador/a el mayor porcentaje de votos ponderados?

Apartado 3: Histórico de participación por estamento y centro (2017–2025)

Representa la evolución de la participación de cada estamento en todas las elecciones, pero desglosada por centro. Cada subplot debe mostrar los resultados históricos de un centro. Responde a la siguiente pregunta:

- ¿En qué año, centro y estamento se obtuvo la mínima y la máxima participación?

Apartado 4: Exportación de las gráficas

Exporta las gráficas de los Apartados 1-3 en ficheros PDF. Las gráficas deben tener la mejor calidad de imagen posible.

Parte 3: Gráficos avanzados con seaborn

En esta tercera parte vamos a utilizar la librería seaborn para generar visualizaciones avanzadas que permitan analizar la distribución de los datos, identificar patrones y comparar resultados de manera más clara. El objetivo es aprender a emplear tipos de gráficos que en matplotlib requieren mayor complejidad de implementación, como diagramas de caja o violín, mapas de calor o diagramas de dispersión estilizados.

Todos los gráficos deben realizarse con la librería seaborn. No se especifica el tipo exacto de gráfico que debes usar: deberás decidir qué representación es más adecuada en cada caso (por ejemplo, boxplot o violinplot para distribuciones, heatmap para valores tabulados, scatterplot para analizar relaciones, etc.). En el notebook, tras cada figura deberás incluir una breve justificación explicando por qué has elegido ese tipo de gráfico y, en caso de haber aplicado modificaciones de estilo (como añadir paletas de color, modificar escalas, añadir títulos o etiquetas personalizadas, etc.), deberás justificar también esas decisiones de diseño.

Apartado 1: Distribución de participación por centro y estamento

Representa la distribución de la participación en las elecciones de 2021, diferenciando por estamento y centro. El gráfico debe mostrar cómo varía la participación entre los distintos centros dentro de cada estamento, de manera que se puedan identificar centros con niveles de participación más altos o más bajos respecto al resto.

Apartado 2: Mapa de calor de participación histórica

Construye una tabla con la participación media por estamento en cada proceso electoral (2017, 2021, 2025a y 2025b). Representa estos datos en forma de heatmap, de manera que se visualicen las diferencias de participación según el estamento y su evolución a lo largo del tiempo.

Apartado 3. Relación entre tamaño del censo y participación

Representa en un diagrama de dispersión (scatterplot) la relación entre el tamaño del censo de un centro (eje X) y la participación porcentual (eje Y). Diferencia los puntos por estamento mediante colores o marcadores distintos. El objetivo es analizar si existe relación entre el tamaño del colectivo y su nivel de participación.

Instrucciones de la Entrega

La práctica se realizará en parejas. El estudiante deberá entregar un único notebook de Jupyter que contenga:

- El código utilizado para la resolución de todos los ejercicios.
- Las gráficas generadas, que deberán visualizarse correctamente al ejecutar el notebook.
- Las justificaciones y respuestas a las preguntas planteadas en cada apartado.

En la primera celda del notebook deberá aparecer claramente:

- El número del grupo (G-7311 o G-7312).
- El número de la pareja.
- Los nombres de los miembros de la pareja.

Además, se deberá incluir una carpeta con todos los gráficos exportados en formato PDF, organizados de forma clara.

La entrega se realizará en un fichero comprimido **ZIP** cuyo nombre seguirá el formato VD_G<GRUPO>_P<PAREJA>.zip, donde GRUPO será el grupo al que se pertenece (7311 o 7312) y PAREJA el número identificativo de la pareja. Un ejemplo de nombre sería: VD_G7311_P01.zip. **Las entregas deberán realizarse en la plataforma Moodle, en la entrega especificada para la práctica.**

Criterios de evaluación

Cada apartado de la práctica tendrá una valoración de **1 punto**, y se evaluará:

- El **uso correcto de la librería** correspondiente (matplotlib o seaborn).
- La **calidad del código** empleado (claridad, organización, buenas prácticas).
- La **claridad y legibilidad de las gráficas** (etiquetas, leyendas, escalas, etc.).
- La **justificación de la elección de la visualización** utilizada en cada caso.
- La **adecuación de las modificaciones de estilo** aplicadas para mejorar la interpretación de los resultados.

Además, se podrán realizar penalizaciones en los siguientes supuestos:

- Se podrá penalizar hasta un 10% de la nota total de la práctica si la entrega no se ajusta a las normas especificadas en el apartado Instrucciones de la Entrega.
- Se podrá penalizar con un 10% de la nota por cada día de retraso en la entrega de la práctica a contar desde la fecha y hora que se ha especificado en las instrucciones de la práctica.
- Los apartados que no se puedan ejecutar durante la corrección serán calificados con un 0.