


# Parte I: selección del conjunto de datos

Esta actividad, primera parte de la práctica final, consiste en la selección por parte del estudiante de un conjunto de datos de su interés que será usado en el proyecto de creación de la visualización de datos, de acuerdo con unos criterios establecidos. Básicamente, la temática es libre, pero se valorarán los aspectos siguientes:

1. [10%] Justificad brevemente vuestra selección, sea por motivos personales o profesionales.
2. [10%] La relevancia del conjunto de datos en su contexto. ¿Son datos actuales? ¿Tratan un tema importante por algún colectivo concreto? ¿Se ha tenido en cuenta la perspectiva de género?
3. [25%] La complejidad (medida, variables disponibles, tipos de datos, etc.). Debe tener del orden de miles de registros mínimo. Y debe tener un mínimo del orden de decenas de variables. ¿Combina datos categóricos y cuantitativos? ¿Incluye otros tipos de datos? La riqueza en tipología de variables os puede ayudar a realizar un trabajo más brillante: valores discretos, continuos, fecha u hora, lógicos, cartográficos.
4. [25%] La originalidad. Se valora **no repetir los conjuntos de datos clásicos o muy trabajados**  ([https://medium.com/@dhruval\\_/50-popular-datasets-for-your-next-data-science-project-11550daf8d2%20](https://medium.com/@dhruval_/50-popular-datasets-for-your-next-data-science-project-11550daf8d2%20)) ni temas ya muy tratados (p. ej. Covid-19, tráfico, criminalidad...) Podéis combinar o mejorar el conjunto de datos. En el primer caso, enriquecer el conjunto de datos con otros diferentes para dar un enfoque nuevo. En el segundo caso, generando nuevas métricas o indicadores con las variables existentes mediante transformaciones. ¿Hay otras visualizaciones basadas en este conjunto de datos? ¿Es una evolución o actualización de un conjunto anterior? ¿Habéis enriquecido un conjunto de datos ya existente?
5. [30%] Las cuestiones que responderéis con la visualización de datos, ¿Tienen en cuenta los puntos anteriores? ¿Han sido planteadas en otras visualizaciones u otros proyectos? ¿Son adecuadas para el conjunto de datos elegido? En este

punto, elaborar un diccionario de las variables, su significado y si es un hecho a estudiar o una dimensión que lo mide, os puede ayudar.

Debéis entregar un documento PDF que contenga los puntos mencionados anteriormente.

NOTA: Si se considera que esta parte no se ha resuelto de forma adecuada, el profesor os puede proponer otro conjunto de datos. Una C+ en esta actividad es un indicador de que podréis tener problemas después en la hora de crear la visualización de datos en la segunda parte de la práctica. Si obtenéis una A o una B podéis considerar que vuestra propuesta de datos ha sido aceptada para utilizarla a la segunda parte de la práctica.

## Sobre plagio y uso de IA

**IMPORTANTE:** En esta actividad sólo se permite un uso limitado de herramientas de inteligencia artificial. Concretamente en el contexto de las artes y el diseño, las herramientas de IA, tanto textuales como visuales, pueden servir para generar y esbozar ideas, contextualizar conceptos, plantear prototipos o combinar elementos de manera innovadora o alternativa. En caso de utilizar estas herramientas, es necesario explicitar y citar correctamente su uso en la elaboración del trabajo. Esto implica indicar en cada apartado de la actividad entregada qué herramientas se han utilizado, los objetivos, los prompts que se han empleado, la respuesta obtenida y el proceso que se ha seguido para revisar y editar la salida. Para hacerlo correctamente, consulta la guía [¿Cómo citar la IA en los trabajos? \(https://openaccess.uoc.edu/bitstream/10609/148823/2/U2\\_17\\_GuiaCitarIA\\_ESP.pdf\)](https://openaccess.uoc.edu/bitstream/10609/148823/2/U2_17_GuiaCitarIA_ESP.pdf). En el plan docente y en la [Web sobre integridad académica y plagio de la UOC \(https://campus.uoc.edu/estudiant/microsites/plagi/es/index.html\)](https://campus.uoc.edu/estudiant/microsites/plagi/es/index.html) encontrarás información sobre qué se considera conducta irregular y las consecuencias que puede tener. Recuerda también que en caso de usar herramientas de IA

generativa, no debe proporcionarse información personal, confidencial o protegida por propiedad intelectual. Ten en cuenta que un uso inadecuado de la IA generativa, como la falta de revisión de la respuesta obtenida o la falta de citación de las herramientas utilizadas, será considerada una conducta irregular en la evaluación. En caso de duda, consulta con el profesorado colaborador.