

Máster en Ciencia de Datos e Inteligencia Artificial

Trabajo en grupo

Índice

1. Enunciado
2. Asignación de *datasets* y objetivos de predicción



01. Enunciado



1. Enunciado

Fecha de entrega: 8 de febrero 2026

El objetivo del primer trabajo en grupo es realizar el proceso de tratamiento, limpieza, EDA y extracción de *insights* del *dataset* asignado al grupo. Este proceso servirá como fase previa al segundo trabajo en grupo que consistirá en la elaboración de un modelo predictivo. Más concretamente, debes:

- **Desarrollar el proceso de tratamiento y limpieza de información:** se deberá cargar la información desde la fuente asignada y hacer los tratamientos necesarios sobre la misma, de forma que se disponga de un set de datos “limpio” y correctamente preparado. Puedes mezclar los datos con fuentes de datos abiertas a las que tengas acceso y te sirvan para enriquecer los datos de los que dispones.
- **Desarrollar el análisis exploratorio de la información:** para ello puedes realizar los gráficos que consideres. Además, deberás buscar *insights* avanzados dentro de tu información y presentarlo gráficamente de manera clara y sencilla. Puedes emplear las herramientas de visualización estática y dinámicas vistas en clase.

Es importante que el ejercicio sea fácilmente reproducible, es decir, debe de ejecutarse sin necesidad de hacer descargas de datos, instalaciones de paquetes o librerías, etc. En el caso de ser necesarias instalaciones incorporar un fichero explicando que es necesario instalar antes de ejecutarlo.

Atención: los equipos y/o datasets pueden ser modificados para el segundo trabajo en grupo en caso de que la Dirección Académica así lo considere.

Entrega y evaluación

Para la evaluación se deberá entregar:

- Documento o memoria describiendo el proceso realizado.
- Documento o memoria con el análisis exploratorio de información, visualizaciones y detalle de las conclusiones obtenidas durante el proceso realizado.
- Código (aplicaciones desarrolladas)

La evaluación tendrá en cuenta la calidad del código presentado, calidad de las visualizaciones e *insights* obtenidos, así como la exposición en la defensa individual del trabajo.



02. Asignación de *datasets* y objetivos de predicción



2. Asignación de *datasets* y objetivos de predicción

Asignación de *datasets*

#	<i>Dataset</i>	Objetivo de predicción
1	Diabetes Prediction	Clasificador de pacientes con y sin diabetes
2	Productivity Prediction of Garment Employees	Predicción de la productividad de los equipos de trabajo en las fábricas textiles
3	Mobile Price Classification	Clasificador de rango de precios de dispositivos móviles
4	Car Price Prediction	Predicción del precio de venta de un coche

Grupo 1: Juan María Jiménez, Daniel Huarita, David Moya, Iñaki Asúa

Grupo 2: Rodrigo González, Daniel Rodríguez, Rubén Pérez, Inés Heras

Grupo 3: Krishna Soto, Jesús Mesa, Luis Ortiz, Héctor Reyes

Grupo 4: Álvaro Fernández, Adriana Gordillo, Pablo Guerrero, Jaime Cabotá

Máster en Ciencia de Datos e Inteligencia Artificial

Gracias

Afi 