

# PROYECTO DE AULA

## FUNDAMENTOS DE CIENCIA DE DATOS

### Avances del proyecto – repositorio GitHub

María Bernarda Salazar Sánchez, Ph.D.  
Departamento de Ingeniería de Sistemas  
Universidad de Antioquia, Colombia  
[bernarda.salazar@udea.edu.co](mailto:bernarda.salazar@udea.edu.co)

**Objetivo:** Fomentar la aplicación práctica de los contenidos de las Unidades 2 y 3 del curso, consolidando el inicio del proyecto de aula mediante la creación y gestión de un repositorio en GitHub.

#### Orientaciones para el entregable I

El análisis requerido en este entregable deberá abordar técnicamente al menos:

1. Generación de la base de datos acorde a las variables seleccionadas
2. Análisis exploratorio de los datos
3. Detección y análisis de atípicos

El resultado del Análisis exploratorio **deberá ser realizado en un notebook** y publicado en un **repositorio de GitHub** desarrollado para el proyecto. Recuerde que el repositorio debe incluir un archivo **README** que explique claramente autores, título del proyecto, objetivo y distribución del repositorio. Se debe incluir una carpeta con los datos originales y con los datos que ha seleccionado para este proyecto (mínimo 3 variables continuas y 3 discretas).

Para la creación del notebook del análisis exploratorio se pueden apoyar en los ejemplos sugeridos.

**Evaluación:** esta actividad **tendrá el porcentaje de 15%** del curso.

**Fecha de entrega:** martes siete (07) de octubre de 2025.  
La entrega se realiza por medio del repositorio del proyecto de aula que has creado para este curso.

#### Recomendaciones generales

- Utilicen tablas o esquemas para comparar la información de los artículos seleccionados.
- Utilicen tablas o esquemas para organizar las variables que relacionarán sobre la base de datos seleccionada.

#### Penalizaciones

- La entrega después de la fecha será penalizada con media (0.5) unidad por cada 1h de envío tarde.
- La no generación de un README de acuerdo a las indicaciones será penalizada con media (0.5) unidad.

#### Ejemplos de notebooks con EDA

- <https://www.kaggle.com/code/micheldc55/tutorial-eda-para-data-science>
- <https://www.kaggle.com/code/imoore/intro-to-exploratory-data-analysis-eda-in-python>
- <https://www.kaggle.com/code/agrawaladitya/step-by-step-data-preprocessing-eda>