

Sprint 3 - Proyecto

Como analista de datos, tu trabajo consiste en analizar datos para extraer información valiosa y tomar decisiones basadas en ellos. Esto implica diferentes etapas, como la descripción general de los datos y el preprocesamiento.

Descripción del proyecto

En este proyecto, trabajarás con datos reales de transmisión de música online para explorar y procesar información sobre los hábitos de escucha de los usuarios y las usuarias en dos ciudades: Springfield y Shelbyville.

Instrucciones para completar el proyecto

Se te proporciona una plantilla del cuaderno, donde puedes escribir tu código y proporcionar descripciones de tu trabajo. Para completar el proyecto, rellena cada celda de código en la plantilla y edita las celdas Markdown donde la plantilla te pide que expliques tus resultados.

Siempre es una buena práctica incluir una introducción que describa brevemente tus objetivos y una conclusión que resuma tus resultados en forma de celdas Markdown.

El proyecto se divide en dos etapas principales, cada una con objetivos específicos:

Etapas 1: Descripción de los datos.

El cuaderno tiene celdas preparadas con instrucciones sobre qué tipo de código escribir, así como bloques de texto donde puedes anotar tus observaciones.

Etapas 2: Preprocesamiento de datos.

En esta etapa, limpiarás los datos examinando los nombres de las columnas y abordando los valores duplicados y ausentes. Sigue la estructura proporcionada en el cuaderno y asegúrate de anotar tus observaciones al final de esta sección.

Etapas 3: Análisis.

En esta etapa, analizarás si la actividad de los usuarios y las usuarias presenta variaciones según el día de la semana y la ciudad.

Es el núcleo de tu proyecto.

Sigue los pasos de programación necesarios para probar cada declaración y comenta tus resultados en los bloques apropiados.

Finalmente, resume los resultados de todo el proyecto en la sección "Conclusiones".

Diccionario de datos

Los datos se almacenan en el archivo `/datasets/music_project_en.csv`.

También puedes descargarlo [aquí](#).

Descripción de las columnas:

'userID': identifica de forma exclusiva a cada usuario o usuaria;
'Track': título de la canción;
'artist': nombre del artista;
'genre': género musical;
'City': ciudad del usuario o la usuaria;
'time': hora del día en la que se reprodujo la pista (HH:MM:SS);
'Day': día de la semana.