Recap primera sesión: EDA inicial

Entendimiento general del conjunto de datos y de las columnas.

En esta sesión de trabajo hemos cargado los datos y el sistema de carpetas en un repositorio de GitHub. Hemos duplicado el Excel con el conjunto de datos para poder trabajar las transformaciones que sean necesarias en la copia.

DESCRIPCIÓN DE LAS COLUMNAS:

Customer ID: Identificador único de cada cliente.  
✅ Habría que comprobar si los valores son únicos.

Churn Label: Etiqueta que indica si el cliente se fue (Yes) o no (No).  
✅ Estas dos columnas nos dan la misma información, la diferencia es que en Churn Label los valores son Yes/No y Churned son 0/1. Habría que comprobar que son exactamente iguales y eliminar una de ellas.

Churned: Variable binaria (0 o 1) que indica si el cliente abandonó el servicio (1) o se mantuvo (0).

Account Length (in months): Tiempo que el cliente ha estado suscrito al servicio, en meses.  
✅ Hay que comprobar la diferencia entre esta columna y Tenure (in months).

Customer Service Calls: Número de llamadas realizadas al servicio de atención al cliente.

Avg Monthly GB Download: Promedio mensual de datos descargados por el cliente en GB.

Unlimited Data Plan: Indica si el cliente tiene un plan de datos ilimitado (1: Sí, 0: No).

Extra Data Charges: Cargos adicionales por uso de datos más allá del límite del plan.

State: Estado de residencia del cliente.

Phone Number: Número de teléfono del cliente.

Gender: Género del cliente.

Age: Edad del cliente.

Under 30: Indica si el cliente es menor de 30 años (1: Sí, 0: No).

Senior: Indica si el cliente es mayor de 65 años (1: Sí, 0: No).

Group: Indica si el cliente tiene contratado un plan familiar.

Number of Customers in Group: Número de clientes que pertenecen al mismo plan familiar.

Device Protection & Online Backup: Indica si el cliente tiene protección de dispositivo y respaldo en línea.

Contract Type: Tipo de contrato (mensual, anual, etc.).

Payment Method: Método de pago utilizado por el cliente.

Monthly Charge: Cargo mensual que el cliente paga por el servicio.

Total Charges: Cargos totales acumulados durante la suscripción.

Churn Category: Categoría que indica el motivo de la deserción (churn).

Churn Reason: Razón específica por la que el cliente abandonó el servicio.

Contact Date: Fecha del último contacto con el cliente.

Last Transaction Date: Fecha de la última transacción realizada por el cliente.

Tenure (in months): Tiempo total de suscripción del cliente.

Customer Segment: Clasificación del cliente según su nivel de uso o importancia (Alto, Medio, Bajo).

Average Monthly Expenses: Gasto promedio mensual del cliente.

Number of Complaints or Support: Número de quejas o interacciones con el servicio de soporte.

Preferred Contact Method: Método preferido por el cliente para ser contactado (email, teléfono, etc.).

Applied Discount: Indica si se ha aplicado algún descuento al cliente.

Hobby: Pasatiempos del cliente.

Favorite TV Show: Programa de televisión favorito del cliente.

Internal Notes: Notas internas relacionadas con el cliente.

COLUMNAS RELEVANTES PARA NUESTRO ANÁLISIS:

✅ Churn label: Ojo Silvia que esta es importante.

✅ Duración de la cuenta.

✅ Frecuencia de llamadas a atención al cliente.

✅ Número de quejas o soporte.

✅ Método de pago utilizado.

✅ Tipo de contrato.

✅ Segmento de cliente y su relación con el abandono.

✅ Cargos adicionales por datos.

✅ Descuentos aplicados.

✅ Promedio mensual de gastos.