Proyecto TELECOM X

1. **Descripción**

**Telecom X - Análisis de Evasión de Clientes**

Has sido contratado como asistente de análisis de datos en Telecom X y formarás parte del proyecto **"Churn de Clientes"**. La empresa enfrenta una alta tasa de cancelaciones y necesita comprender los factores que llevan a la pérdida de clientes.

Tu desafío será recopilar, procesar y analizar los datos, utilizando Python y sus principales bibliotecas para extraer información valiosa. A partir de tu análisis, el equipo de Data Science podrá avanzar en modelos predictivos y desarrollar estrategias para reducir la evasión.

**¿Qué vas a practicar?**

✅ **Importar y manipular datos** desde una API de manera eficiente.  
✅ **Aplicar los conceptos de ETL** (Extracción, Transformación y Carga) en la preparación de los datos.  
✅ **Crear visualizaciones** estratégicas para identificar patrones y tendencias.  
✅ **Realizar un Análisis Exploratorio de Datos (EDA)** y generar un informe con insights relevantes.

¡Ahora es tu turno! 🚀 Usa tus conocimientos para transformar datos en información estratégica y ayudar a **Telecom X** a retener más clientes.

1. Extracción de Datos

Para iniciar tu análisis, necesitarás importar los datos de la API de **Telecom X**. Estos datos están disponibles en formato **JSON** y contienen información esencial sobre los clientes, incluyendo datos demográficos, tipo de servicio contratado y estado de evasión.

**¿Qué debes hacer?**

✅ Cargar los datos directamente desde la API utilizando **Python**.  
✅ Convertir los datos a un **DataFrame de Pandas** para facilitar su manipulación.

Este es el primer paso para transformar los datos en información valiosa. ¿Listo para programar? 🚀

1. Conoce el conjunto de Datos

**Descripción**

Ahora que has extraído los datos, es fundamental comprender la estructura del dataset y el significado de sus columnas. Esta etapa te ayudará a identificar qué variables son más relevantes para el análisis de evasión de clientes.

📌 Para facilitar este proceso, hemos creado un **diccionario de datos** con la descripción de cada columna. Aunque no es obligatorio utilizarlo, puede ayudarte a comprender mejor la información disponible.

🔗 **Enlace al diccionario y a la API**

**¿Qué debes hacer?**

✅ Explorar las columnas del dataset y verificar sus tipos de datos.  
✅ Consultar el diccionario para comprender mejor el significado de las variables.  
✅ Identificar las columnas más relevantes para el análisis de evasión.

1. Comprobación de incoherencias

**Descripción**

En este paso, **verifica si hay problemas en los datos** que puedan afectar el análisis. Presta atención a valores ausentes, duplicados, errores de formato e inconsistencias en las categorías. Este proceso es esencial para asegurarte de que los datos estén listos para las siguientes etapas.

1. Manejo de Inconsistencias

Ahora que has identificado las inconsistencias, es momento de aplicar las correcciones necesarias. Ajusta los datos para asegurarte de que estén completos y coherentes, preparándolos para las siguientes etapas del análisis.

📌 **Tips:**

🔗 [Manipulación de strings en pandas: lower, replace, startswith y contains | Alura Cursos Online](https://www.aluracursos.com/blog/manipulacion-de-strings-en-pandas-lower-replace-startswith-y-contains)

1. Columna de Cuentas diarias

Ahora que los datos están limpios, es momento de crear la columna **"Cuentas\_Diarias"**. Utiliza la facturación mensual para calcular el valor diario, proporcionando una visión más detallada del comportamiento de los clientes a lo largo del tiempo.

📌 **Esta columna te ayudará a profundizar en el análisis y a obtener información valiosa para las siguientes etapas.**

1. **Estandarización y transformación de datos (opcional)**

La **estandarización y transformación de datos** es una etapa opcional, pero altamente recomendada, ya que busca hacer que la información sea más **consistente, comprensible y adecuada para el análisis**. Durante esta fase, por ejemplo, puedes convertir valores textuales como **"Sí" y "No"** en valores binarios **(1 y 0)**, lo que facilita el procesamiento matemático y la aplicación de modelos analíticos.

Además, **traducir o renombrar columnas y datos** hace que la información sea más accesible y fácil de entender, especialmente cuando se trabaja con fuentes externas o términos técnicos. Aunque no es un paso obligatorio, puede mejorar significativamente la **claridad y comunicación de los resultados**, facilitando la interpretación y evitando confusiones, especialmente al compartir información con **stakeholders no técnicos**.

Análisis

1. Análisis descriptivo

Para comenzar, realiza un **análisis descriptivo** de los datos, calculando métricas como **media, mediana, desviación estándar** y otras medidas que ayuden a comprender mejor la distribución y el comportamiento de los clientes.

📌 **Consejos:**

🔗 [Documentación de](https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.describe.html) DataFrame.describe()

1. Distribución de Evasión

En este paso, el objetivo es comprender cómo está distribuida la variable **"churn" (evasión)** entre los clientes. Utiliza gráficos para visualizar la proporción de clientes que permanecieron y los que se dieron de baja.

1. Recuento de evasión por variables categóricas

Ahora, exploraremos cómo se distribuye la **evasión** según variables categóricas, como **género, tipo de contrato, método de pago**, entre otras.

Este análisis puede revelar **patrones interesantes**, por ejemplo, si los clientes de ciertos perfiles tienen una mayor tendencia a cancelar el servicio, lo que ayudará a **orientar acciones estratégicas**.

1. Conteo de evasión por variables numéricas.

En este paso, explora cómo las **variables numéricas**, como **"total gastado"** o **"tiempo de contrato"**, se distribuyen entre los clientes que cancelaron (evasión) y los que no cancelaron.

Este análisis ayuda a entender si ciertos valores numéricos están más asociados con la **evasión**, proporcionando **insights** sobre los factores que influyen en el comportamiento de los clientes.

1. Informe Final

Finaliza el desafío elaborando un **informe dentro del mismo notebook** que resuma todo el trabajo realizado. El informe debe incluir:

🔹 **Introducción:** Explica el objetivo del análisis y el problema de **evasión de clientes (Churn)**.

🔹 **Limpieza y Tratamiento de Datos:** Describe los pasos realizados para **importar, limpiar y procesar** los datos.

🔹 **Análisis Exploratorio de Datos:** Presenta los análisis realizados, incluyendo **gráficos y visualizaciones** para identificar patrones.

🔹 **Conclusiones e Insights:** Resume los principales hallazgos y cómo estos datos pueden ayudar a **reducir la evasión**.

🔹 **Recomendaciones:** Ofrece **sugerencias estratégicas** basadas en tu análisis.

Asegúrate de que el informe esté **bien estructurado, claro y respaldado por visualizaciones** que refuercen tus conclusiones. 🚀

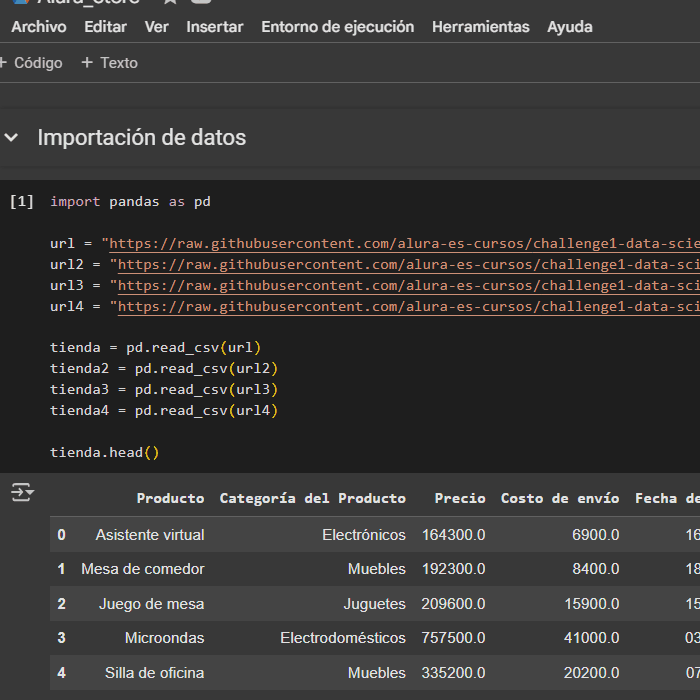
**Git y GitHub**

Git y GitHub son herramientas esenciales para cualquier analista de datos, ya que permiten el versionado y el intercambio eficiente de proyectos.

En este desafío, deberás **subir tu cuaderno de Colab a un repositorio en GitHub**. Esto garantizará que tu progreso esté guardado y accesible desde cualquier lugar.

**Lo que necesitas hacer:**

1. Crea un **nuevo repositorio en GitHub** para almacenar tu proyecto.
2. Exporta tu cuaderno de Colab como un archivo .ipynb.
3. Realiza el **upload** del cuaderno al repositorio.
4. Mantén tu trabajo actualizado utilizando git pull, git add, git commit y git push cuando sea necesario.



Si necesitas repasar conceptos, revisa los siguientes recursos:

* [Git - Configurando Git por primera vez](https://git-scm.com/book/es/v2/Inicio---Sobre-el-Control-de-Versiones-Configurando-Git-por-primera-vez)
* [Cómo subir mi proyecto en GitHub](https://www.youtube.com/watch?v=bhKkeOMysuw)
* [Cómo escribir un README increíble en tu Github](https://www.aluracursos.com/blog/como-escribir-un-readme-increible-en-tu-github)

**README**

El README es esencial para documentar tu proyecto, explicando su propósito, funcionalidades e instrucciones de uso.

Como **desafío adicional**, crea un README.md para tu proyecto **Telecom X**, incluyendo:

1. El propósito del análisis realizado.
2. La estructura del proyecto y organización de los archivos.
3. Ejemplos de gráficos e insights obtenidos.
4. Instrucciones para ejecutar el notebook.

Esto hará que tu proyecto sea más profesional y accesible para cualquier persona que quiera entender tus análisis.