

Data∞p

Business Case

Advanced Analytics

Advanced Analytics

En este ejercicio, se proporciona un conjunto de datos en formato CSV (llamado “Variables_entrevista.csv”) que contiene información relacionada a el comportamiento crediticio de una cartera de clientes de una auto financiera. Cada registro en el conjunto de datos incluye un ID de crédito, modelo de auto, monto del préstamo, entidad federativa, actividad económica, bandera de empleo, ingreso mensual, propiedad casa y la variable objetivo Default Crediticio.

ID_Credito:	ID único.
Modelo:	valor (string), modelos de unidades automotrices para las que se solicita el credito.
Monto_prestamo:	valor (integer), es el monto que solicitaron para la compra de cada unidad.
Entidad_Federativa:	valor (string), son las Entidades Federativas dentro de México donde se hizo la compra de la unidad
Actividad_Economica:	valor (string), en esta columna se incluyen las actividades económicas que realiza cada cliente
FLG_EMPLEO:	valor (integer), variable bandera que indica si el cliente tiene empleo o no, el indicador es 1 si tiene empleo y 0 si no lo tiene
Ingreso_mensual:	valor(integer), es el promedio del ingreso mensual de cada cliente de la institución
Propiedad_casa:	valor(string), es una variable que indica de quien es la casa en donde habita el cliente
Default_Crediticio:	valor(integer), variable objetivo la cual indica si el cliente deajo de pagar el crédito (default), donde 1 es que lo deajo de pagar y 0 es que no tiene mora.

El objetivo de este business case es: poder **predecir la probabilidad de impago** por cliente y presentar cuales son las variables que explican principalmente el impago, es decir cuáles son las variables y los valores que hacen que un cliente sea más propenso a caer en impago.

La presentación deberá incluir la justificación del modelo utilizado, los resultados del modelo y las variables que explican el modelo.

En este ejercicio, se requiere principalmente el uso de R, **Python o SAS** como lenguaje de programación para el desarrollo. Además, de que la presentación debe de ser ejecutiva y visual en donde se respeten los siguientes colores:

- Color principal: #4866F7
- Color complementario: #4BBBF7

Además de las habilidades técnicas, valoramos mucho los siguientes aspectos:

1. **Comunicación clara de los resultados:** Se espera que presentes los hallazgos de manera concisa y comprensible, destacando los puntos clave del análisis exploratorio y los resultados del modelo de pronóstico.
2. **Accionables definidos:** Es importante que identifiques acciones o recomendaciones específicas basadas en los resultados obtenidos. Estos deben ser prácticos y orientados a la toma de decisiones con el fin de maximizar el impacto empresarial.
3. **Buena presentación:** Se valora la capacidad de presentar la información de manera visualmente atractiva y organizada. Utiliza gráficos, tablas y otros recursos visuales para resumir y comunicar eficazmente los *insights* y los resultados del análisis.

Resumen de los entregables finales:

- 1. **Presentación en PDF:** Incluirá insights obtenidos del análisis exploratorio de datos, y del modelo final, así como un resumen del análisis de los resultados del modelo. Se destacarán los insights tanto a nivel total como por grupo de vendedores. Recomendamos de 4 a 6 diapositivas de contenido (i.e. excluyendo portada(s) y/o agenda).
- 2. **Dataset con el output final en CSV:** Proporcionará una proyección de las ventas por día, por grupo y en total desde el 1 de enero de 2023 hasta el 1 de julio de 2023. Los datos estarán en un formato fácil de entender y utilizar.
- 3. **Jupyter Notebook o Script en Python (*.ipynb/*.py):** Incluirá todas las funciones y clases utilizadas en el ejercicio. Esto permitirá a otros revisar y reproducir el análisis de manera eficiente.

Esperamos que esta clarificación sea de ayuda. Si tienes alguna pregunta adicional, no dudes en preguntar.

¡Éxito!