



Análisis Exploratorio del dataset iFood



Empresa iFood

iFood es una empresa líder en **delivery de comida en América Latina**, especialmente en **Brasil**. Es similar a Uber Eats o Rappi y permite a los usuarios pedir comida de restaurantes a través de una aplicación móvil. También ofrece servicios de supermercado y logística de entrega con repartidores propios.



Dataset

El dataset de iFood se usa para analizar el comportamiento de los clientes y contiene información clave como:

- **Income** (*Ingresos anuales del cliente*).
- **Marital_Status** (*Estado civil: Soltero, Casado, etc.*).
- **Kidhome & Teenhome** (*Número de hijos o adolescentes en casa*).
- **Mnt***: Datos de cantidad de productos comprados por categoría

Diccionario de datos (sí, esto es lo que hay)

Feature	Description
AcceptedCmp1	1 if costumer accepted the offer in the 1 st campaign, 0 otherwise
AcceptedCmp2	1 if costumer accepted the offer in the 2 nd campaign, 0 otherwise
AcceptedCmp3	1 if costumer accepted the offer in the 3 rd campaign, 0 otherwise
AcceptedCmp4	1 if costumer accepted the offer in the 4 th campaign, 0 otherwise
AcceptedCmp5	1 if costumer accepted the offer in the 5 th campaign, 0 otherwise
Response (target)	1 if costumer accepted the offer in the last campaign, 0 otherwise
Complain	1 if costumer complained in the last 2 years
DtCustomer	date of customer's enrollment with the company
Education	customer's level of education
Marital	customer's marital status
Kidhome	number of small children in customer's household
Teenhome	number of teenagers in customer's household
Income	customer's yearly household income
MntFishProducts	amount spent on fish products in the last 2 years
MntMeatProducts	amount spent on meat products in the last 2 years
MntFruits	amount spent on fruits in the last 2 years
MntSweetProducts	amount spent on sweet products in the last 2 years
MntWines	amount spent on wines in the last 2 years
MntGoldProds	amount spent on <i>gold</i> products in the last 2 years
NumDealsPurchases	number of purchases made with discount
NumCatalogPurchases	number of purchases made using catalogue
NumStorePurchases	number of purchases made directly in stores
NumWebPurchases	number of purchases made through company's web site
NumWebVisitsMonth	number of visits to company's web site in the last month
Recency	number of days since the last purchase

Table 1: Meta-data table

Objetivo

El objetivo del trabajo es realizar un análisis preliminar del dataset como inicio del Exploratory Data Analysis (EDA).

Este trabajo tiene como finalidad verificar vuestro nivel de programación y capacidad de extraer insights o información relevante a partir de un conjunto de datos pequeño.

Podéis usar el environment de la asignatura ;)

Instrucciones - Grupos

Os podréis juntar en equipos de 3 personas. Os recomiendo entregar el trabajo en Git.

Indicad los compañeros de equipo siempre.

Mi recomendación de entrega es hacer un análisis exploratorio inicial conjunto en un notebook, y añadir 3 notebooks adicionales con cada uno de los 3 ejercicios.

Por lo tanto, os recomiendo que la carpeta ./src se vea así:

1. Análisis inicial. – Conjunto.
2. Ejercicio 1. – Compañero 1.
3. Ejercicio 2. – Compañero 2.
4. Ejercicio 3. – Compañero 3.
5. Ejercicio 4. (adicional quien quiera hacerlo). (Conjunto o compañero 1,2,3... 4 si hay)

Instrucciones – ChatGPT | Copilot

Se puede utilizar para buscar funciones que queramos aplicar. Os recomiendo verificar documentación de las librerías o usar stackoverflow. No quiero que uséis para responder a los ejercicios.

Si entiendo que habéis utilizado Chat para hacer el ejercicio completo sin aplicar vuestros conocimientos, no corregiré ningún otro ejercicio de los compañeros del grupo.

Instrucciones - Uso de Git

Os tendréis que clonar el repositorio con estas instrucciones, el dataset en formato .zip(que tendréis que extraer) y crear el/los notebooks.

Alternativamente, podéis entregar una tarea en Canvas. Si entregáis el repositorio en Github, hacedlo incorporando el enlace de github en la tarea de Canvas.

Ejercicio 1

Los ejercicios no deben ser perfectos. La idea es ver vuestra capacidad de resolver problemas, plantear soluciones y extraer información relevante. Antes de hacer el ejercicio, tendréis que haber procesado ligeramente los datos.

Ejercicio 1.

¿Existe una relación entre los ingresos anuales de los clientes y la cantidad de dinero que gastan en iFood? Visualiza la relación con un gráfico y describe tus conclusiones.

Ejercicio 2

Los ejercicios no deben ser perfectos. La idea es ver vuestra capacidad de resolver problemas, plantear soluciones y extraer información relevante. Antes de hacer el ejercicio, tendréis que haber procesado ligeramente los datos.

Ejercicio 2.

¿Qué grupo de clientes gasta más en iFood: los solteros o los casados? Usa visualizaciones y estadísticas para justificar tu respuesta.

Ejercicio 3

Los ejercicios no deben ser perfectos. La idea es ver vuestra capacidad de resolver problemas, plantear soluciones y extraer información relevante. Antes de hacer el ejercicio, tendréis que haber procesado ligeramente los datos.

Ejercicio 3.

¿Cómo varía el tipo de gasto en comida según los ingresos de los clientes?
Muestra visualmente que tipo de productos son los más consumidos por los clientes según su nivel de ingresos.

Ejercicio 4 (Adicional – Voluntario)

Los ejercicios no deben ser perfectos. La idea es ver vuestra capacidad de resolver problemas, plantear soluciones y extraer información relevante. Antes de hacer el ejercicio, tendréis que haber procesado ligeramente los datos.

Ejercicio 4.

¿Qué tipo de cliente gasta más en productos premium (vino, carne, pescado, productos de lujo)? Analiza si hay un perfil específico según ingresos, estado civil o número de hijos.

Deadline - Entrega

La tarea, entre 3 personas, está pensada para acabarse en las 2 horas de clase. Al menos los primeros 3 ejercicios. Igualmente, las prisas no son buenas. Por ello, tenéis hasta el domingo a las 23:59 para entregar la tarea. El Lunes subiré al repositorio remoto el notebook, que tendréis que actualizar siguiendo los métodos **fetch** y **pull** de git, por si queréis tener mis respuestas a los ejercicios.

Igualmente los corregiré y os daré una nota a vuestro trabajo. La nota es subjetiva, puesto que valoro más que un código ejecutado correctamente.

DOMINGO

23

FEBRERO

23:59 H



(Sabéis más de lo que creéis)