**“Caso de Estudio"**

Se entrega el siguiente set de datos de una empresa del sector retail, **BigMartSales.csv**, el cual es un dataset de ventas de 2013 para 1559 productos en 10 tiendas en diferentes ciudades. Además, se han definido ciertos atributos de cada producto y tienda. Debido a la nueva coyuntura que estamos viviendo COVID-19, han observado que algunos productos que normalmente tenían alta rotación se han estado quedando en el almacén generando un grave problema de sobre stock y quitándoles espacio para otros productos por ende le han encargado al área de analítica de la empresa poder ayudarlos con la manera óptima de realizar el inventario y pedidos a los proveedores pues tienen metas comerciales muy altas en el Q1-2021.

Las variables que se disponibilizan son:



Item\_Identifier ID de producto único.

Item\_Weight Peso del producto.

Item\_Fat\_Content El producto es bajo en grasa o no?.

El% del área de visualización total de todos los productos

Item\_Visibility

en una tienda asignada a un producto en particular.

Item\_Type La categoría a la que pertenece el producto.

Item\_MRP Precio minorista máximo (precio de lista) del producto. Outlet\_Identifier ID de tienda única.

Outlet\_Establishment\_Ye El año en que se estableció la tienda.

Outlet\_Size El tamaño de la tienda en términos de superficie cubierta Outlet\_Location\_Type El tipo de ciudad en la que se encuentra la tienda.

Si la tienda es solo una tienda de comestibles o algún tipo

Outlet\_Type

de supermercado



**Consignas:**

1.- Definir el problema de la naturaleza que se tiene a continuación, además de los objetivos de negocio bien definidos.

Problema: Sobre stock de productos a causa de la pandemia que modificaron los hábitos de consumo de los consumidores.

Objetivos:

Poder optimizar el espacio del depósito.

Determinar compras a proveedores que sean acordes a los hábitos de consumo.

2.- ¿Qué tipo de variables se utilizan en el problema de negocio?

Item\_Identifier

ID de producto único.

Variable categórica nominal.

Item\_Weight

Peso del producto.

Variable cuantitativa continua

Item\_Fat\_Content

El producto es bajo en grasa o no?.

Variable Categórica ordinal

Item\_Visibility

El% del área de visualización total de todos los productos en una tienda asignada a un producto en particular.

Variable cuantitativa continua

Item\_Type

La categoría a la que pertenece el producto.

Variable Categórica nominal

Item\_MRP

Precio minorista máximo (precio de lista) del producto.

Variable cuantitativa continua

Outlet\_Identifier

ID de tienda única.

Variable categórica nominal.

Outlet\_Establishment\_Ye El año en que se estableció la tienda.

Variable cuantitativa discreta

Outlet\_Size

El tamaño de la tienda en términos de superficie cubierta

Variable Categórica ordinal

Outlet\_Location\_Type

El tipo de ciudad en la que se encuentra la tienda.Si la tienda es solo una tienda de comestibles o algún tipo

Variable Categórica ordinal

Outlet\_Type de supermercado

Variable Categórica ordinal

Item\_Outlet\_Sales

Variable cuantitativa continua

3.- ¿Cómo podríamos resolver este problema de negocio y cumplir con los objetivos planteados a través de la ciencia o analítica de datos?

No supervisado, cluster, según la rotación de las ventas por tienda.