

## **Análisis de Ventas y Rentabilidad Superstore**

El objetivo de este proyecto es aplicar técnicas de análisis de datos en un entorno realista utilizando Excel, con el fin de extraer conclusiones clave a partir de una base de datos de ventas. A través de la limpieza, transformación y visualización de los datos, se identifican patrones de comportamiento comercial, rentabilidad por segmento y rendimiento de productos, con una presentación final en formato dashboard interactivo.

## **Instalación y requisitos**

Excel o cualquier otro visor .xls

## **Recursos utilizados**

Link a la tabla original

<https://www.kaggle.com/datasets/vivek468/superstore-dataset-final?resource=download>

## **EDA**

Transformación y limpieza

1. Limpiamos columnas no necesarias o demasiado granulares. ( order # y zip code).
2. Agregamos columna de "Order Month" para trabajar más cómodamente.
3. Cambiamos formato de fechas a inglés (me daba error al intentar pasar los datos a la columna order month).
4. Damos formato al resto de columnas (moneda, cantidades, etc).
5. Creamos una nueva columna para porcentajes ya que el formato actual no lo convierte adecuadamente.
6. La relación entre sales y quantity no tiene mucho sentido, no puede ser que por ejemplo en la primera fila nos de 2 ventas de  $26.196 = 419.136$  de beneficio. No deberíamos de tener más beneficio que el propio importe de la venta.
7. Vamos a quitar las filas que dan datos irreales como por ejemplo un beneficio que es 16 veces superior al precio de la propia venta (tenemos datos sintéticos pero vamos intentar conservar algo de realismo para que no nos salgan resultados

disparatados, y no es que podamos pedirle a nadie un contraste o aclaración sobre los datos...).

Para conseguir esto creamos una nueva columna con la fórmula para verificar si  $\text{sales} * \text{quantity}$  es mayor que profit nos de "verdadero". Después de esto filtramos por verdadero y eliminamos dichas filas.

8. Añadimos un sales total ( $\text{sales} * \text{quantity}$ ) para facilitar el trabajo en las tablas dinámicas.

9. Creamos una columna con un profit % y comprobamos que en algunas ventas perdemos más del 100%, en la mayoría de casos no debería de ser posible o viable para la mayoría de empresas pero...lo dejamos y achacamos a que sean datos sintéticos (podríamos añadir un filtro para eliminar ventas con un % de profit mayor a -100% pero tampoco quiero enrevesar)

### Análisis descriptivo

1. Creamos las distintas táblas dinámicas y sus correspondientes gráficos
2. Evitamos introducir demasiadas tablas y nos apoyamos en la segmentación del dashboard para lograr una visión más completa.

### Dashboard

1. Creo que hay algún conflicto en las fechas porque después de cambiar el formato de fecha varias veces y probar con distintas fórmulas, no he conseguido poner por año).
2. Introducimos las tablas dinámicas.
3. Añadimos segmentadores temporales, de región, envíos y tipos de cliente.

## Resultados y conclusiones

1. Muebles es la peor categoría con diferencia, mostrando un ratio de beneficios minúsculo comparado con el resto de categorías.
2. El segmento Consumer y Corporate tienen un margen de beneficios porcentual inferior al de Home Office, aunque tienen un mayor volumen de ventas y beneficio total
3. El envío en mismo día tiende a tener unos márgenes de beneficios desproporcionadamente mayores en los tipos de cliente Consumer y Home Office pero dan pérdidas en Corporate. Dichas pérdidas están mayormente ligadas con un solo producto.  
Otros tipos de envíos también sugieren diferencias notables, aunque necesitaría un examen más exhaustivo.
4. El top 10 de productos por beneficio representa un porcentaje desproporcionado del total de ganancia, lo cual sugiere una concentración de rentabilidad en pocos ítems
5. La región Central tiene márgenes más bajos en general, las subcategorías de Office Supplies que venden tienen poco margen y la de Muebles es una categoría poco rentable en general, traducándose en un rendimiento peor que el resto de zonas.

Autor

- Qiongxi Liu

- <https://github.com/Legari92/>