



**EMTECH**

**DATA SCIENCE: GRUPO 1**

**PROYECTO\_1**

*Introducción a Python*

César Romero Canul

Septiembre 2020

# Contents

0.1	Introducción . . . . .	1
0.2	Definición del código . . . . .	2
0.2.1	Comprensión de los datos . . . . .	2
0.2.2	Limpieza . . . . .	2
0.2.3	Preprocesamiento . . . . .	2
0.2.4	Simulación simple . . . . .	4
0.3	Solución al problema . . . . .	6
0.4	Conclusión . . . . .	7

## 0.1 Introducción

Como primera parte de curso, se ha desarrollado un primer análisis sobre un caso particular donde se han puesto en práctica métodos principales sobre programación en python.

LifeStore es una tienda virtual que maneja una amplia gama de artículos y requiere un análisis basado en sus ventas y búsquedas para evitar acumulación de inventario.

Este reporte tiene como finalidad mostrar el proceso de análisis así como los resultados obtenidos y las consideraciones tomadas para dar solución al problema. Para el código se usó anaconda y el entorno de programación Jupyter notebook

## 0.2 Definición del código

### 0.2.1 Comprensión de los datos

En primera instancia, se visualizó el archivo "lifestore\_file.py" proporcionado previamente y el cual contiene 3 listas con datos sobre la tienda.

- lifestore-products. Donde cada elemento contiene los sig. valores: id del producto, nombre, precio, categoría y stock respectivamente.
- lifestore-sales. Donde cada elemento contiene los sig. valores: id de venta, id del producto, calificación (1 a 5), fecha y reembolso (1 si ocurrió o 0 si no)]
- lifestore-searches. Donde cada elemento contiene los sig. valores: id de búsqueda y id del producto.

### 0.2.2 Limpieza

El archivo "lifestore\_file.py" fue modificado ya que se encontró algunas inconsistencias en los datos. Las modificaciones fueron mínimas y se realizaron a consideración de obtener resultados más claros.

Los cambios fueron hechos en la lista lifestore-sales para las fechas de las ventas con id 119 y 219 respectivamente.

```
[119, 17, 1, '05/09/2020', 1]
```

```
[219, 54, 4, '04/05/2002', 0]
```

Para la venta 119, suponiendo que el fin de las ventas registradas fue en Agosto, esta fecha del 05 de Septiembre se tomó como realizado en el año 2019, quedando como: [119, 17, 1, 05/09/2019, 1].

Para la venta 219, el error ortográfico del año 2002 en la fecha, fue corregido a 2020, quedando como: [219, 54, 4, 04/04/2020, 0].

### 0.2.3 Preprocesamiento

Como paso inicial se usó el módulo "import" para importar el archivo "lifestore\_file.py" en el archivo principal del proyecto.

Ya importado el archivo, se renombraron las listas (lifestore-sales, lifestore-searches, lifestore-products) a sales, searches y product respectivamente para un mejor manejo de las mismas.

## 1. Productos más vendidos y productos rezagados

Para obtener el listado de los *productos más vendidos*, primero se tuvo que contar todas las ventas, para posteriormente obtener la cantidad de veces un producto fue vendido. En el código, los diccionarios `sales_by_index` y `sales_by_names` contienen el índice y los nombres, respectivamente, de los productos y la cantidad de veces registradas que ese producto se vendió de esos datos, se obtuvo la lista `top_sales` la cual contiene el nombre de los productos ordenados según su cantidad de ventas.

- \* Se hizo uso de diccionarios para un mejor manejo de la información, primeramente con los índices de los productos ya que los nombres son bastante largos.
- \* El ordenamiento para este y todos los demás casos fueron realizados con la función lambda `x: x[1]` como key donde se ordenan los nombres de los productos tomando en cuenta sus valores.

Para obtener los *productos rezagados*, se tomaron en cuenta los productos cuyo id no aparecieran en la lista de ventas, a partir de ello, se obtuvo la lista `lagging-products` la cual contiene el nombre de los productos sin ventas y la cantidad de stock disponible que se tiene de ellos.

Para obtener los *productos más buscados* y *productos menos buscados*, se tomaron en cuenta los productos que aparecieron en la lista `searches` y la cantidad de veces que fueron buscados, finalmente, se obtuvo la lista ordenada `top_searches` la cual contiene los nombres de los productos y número de veces que fueron buscados en la por los clientes. De esta lista, se tomaron los primeros 25 productos como los más buscados y los últimos 25 como los de menor búsqueda.

Para obtener las *ventas por categoría* y *búsquedas por categoría*, se tomaron en cuenta las categorías dentro de la lista `prodcuts`, luego para cada categoría se tomaron los productos vendidos en la lista `sales` y los productos buscados de la lista `searches` y se contó, por categoría y productos dentro de esta, la cantidad de ventas y búsquedas que los productos tuvieron. Esto se muestra contenido en el código como las listas `top_categories_sales` y `top_categories_searches` respectivamente.

- \* Se registraron un total de 8 categorías: procesadores, tarjetas de video, tarjetas madre, discos duros, memorias usb, pantallas, bocinas y audífonos.

## 2. Productos por reseña en el servicio

En esta sección, como reseña del servicio se tomó en cuenta la calificación dada a los productos y si tuvo reembolso o no.

Para obtener los *productos mejor calificados* y *productos peor calificados* se tomaron en cuenta los productos vendidos de la lista `sales`, donde se tomaron las calificaciones que los usuarios le asignaron y se generó el diccionario

to.best\_rated el cual contiene como claves los valores del 1 al 5 (calificaciones) y cada uno de ellos contiene una lista de los productos con esa calificación y las veces que la recibieron. Por lo tanto, el top 20 de los productos con mejor reseña muestra el nombre de los productos con calificaciones más altas y por otro lado, los productos con peor reseña muestra los productos con calificaciones más bajas.

\* Adicionalmente, se obtuvieron también los productos que fueron reembolsados y la cantidad de veces que esto ocurrió. (lista products\_refunded)

### 3. Total de ingresos y ventas promedio mensuales

La cantidad total (en pesos) de *ingresos* que la tienda registró en sus ventas fue de 760,177. Esto se obtuvo de sumar el precio de todos los productos vendidos de la lista sales.

Por otra parte, los *ingresos anuales* se obtuvieron tomando en cuenta los años de las fechas de los productos vendidos registrados en la lista sales, donde únicamente se obtuvieron 2 años: 2019 y 2020 con ingresos de 8,408 y 751,769 pesos respectivamente.

Para 2019 se registraron un total de 2 ventas en 2 meses, y un promedio de 1 venta por mes. Para 2020 se registraron un total de 281 ventas en 8 meses consecutivos, lo cual nos da un promedio de 35 ventas mensuales para este año. En el código, la lista montly\_sales muestra el total de *ingresos mensuales*, esto se obtuvo tomando en cuenta las ventas por mes y el precio de cada producto, datos tomados de la lista sales; finalmente se obtuvo que el mes con mayores ingresos fue Abril con un total de 193,295 pesos.

#### 0.2.4 Simulación simple

La forma en la que se presenta la vista para el usuario es a través de un simple menú interactivo donde la persona interesada podrá acceder en uno de dos modos: Administrador o usuario.

```

*****/ Bienvenido a LifeStore \*****
Ingresar como:
1.Administrador
2.Usuario
Número de elección: 
```

Para los administradores, las opciones a las cuales se puede acceder son las siguientes considerando una vista con la información de toda la tienda.

```
\\\\\\\\\\\\\\\\ Menú \\\\\\\\\\\\\\\\\\\  
1: Productos más vendidos  
2: Productos rezagados  
3: Ventas por categoría  
4: Productos más buscados  
5: Productos menos buscados  
6: Búsquedas por categoría  
7: Productos mejor calificados  
8: Productos peor calificados  
9: Reporte de ingresos  
10:Salir
```

Mientras que para los usuarios las opciones se limitan a las siguientes considerando los intereses de los usuarios y la información que es adecuada mostrarles.

```
\\\\\\\\\\\\\\\\ Menú \\\\\\\\\\\\\\\\\\\  
1: Productos más vendidos  
2: Más vendidos por categoría  
3: Productos más buscados  
4: Más búscados por categoría  
5: Productos mejor calificados  
6:Salir
```

### 0.3 Solución al problema

here. Considerando los resultados obtenidos, la tienda lifestore debería considerar lo siguiente:

- Para los productos que se han vendido y actualmente mantienen nada o muy poca cantidad en stock, surtir la tienda con esos productos lo antes posible tomando en cuenta el top de mejores ventas en cuanto a productos.
- Si el producto se encuentra en el top de ventas, búsquedas y reseñas, considerar subir el precio un 5% de su valor.
- Para los productos rezagados(no vendidos) considerando que han pasado casi un año sin ventas, se propone lo siguiente: 1. Si el producto cuenta con un stock acumulado en la tienda menor( $i$ ) a 50 pzs, hacer promoción reduciendo el precio un 20%  
2. Si el producto cuenta con un stock acumulado en la tienda mayor( $i$ ) a 50 pzs pero menor( $i$ ) a 100 pzs, hacer promoción reduciendo el precio un 25%  
3. Si el producto cuenta con un stock acumulado en la tienda mayor( $i$ ) a 100 pzs, hacer promoción reduciendo el precio un 50%  
Esto considerando solo recuperar la inversión hecha en esos productos para luego invertir en otros.
- En la tienda, ya sea física o en línea, por secciones poner a primera vista los productos más vendidos y más buscados.
- Considerando los meses con mayores ventas (Marzo y Abril), suirtr productos frecuentemente y realizar promociones con productos afines.

## 0.4 Conclusión

Como conclusión del proyecto, se obtiene que la tienda cuenta, de 96 productos que tiene, con ventas de solo 42(43.75%) de ellos y con otros 54(56.25%) que no se han vendido en un largo tiempo.

Esto ha causado un descenso en sus ingresos a partir del mes de Mayo hasta Agosto de un 95.15% aproximadamente (ya que en agosto solo hubo de ingresos de \$3,077 de los \$63,348 promedio mensuales que se esperaban).

Por lo tanto, considerando los puntos antes mencionados sobre como dar solución al problema, se espera que la tienda aumente para los próximos 2 mes al menos un 25% sus ingresos.