



Proyecto 1. Introducción a Python

Uso de elementos básicos de programación

Garduño Hernández Jazmin
Guadalupe

5 de diciembre de 2021



Índice

Introducción	3
Definición del código	4
Solución al Problema	5
Conclusión	11

Introducción

La ciencia de datos es un área en desarrollo y multidisciplinaria que se encarga del análisis y resolución de problemas, así como la optimización de distintos procesos en áreas muy diversas, por medio del uso de datos. Existen diversas herramientas que permiten a los científicos de datos realizar su trabajo, y entre ellas se encuentra el uso de diversos lenguajes de programación. Uno de los lenguajes que han adquirido más popularidad en la ciencia de datos es Python, por lo que en la introducción al mismo es de gran importancia.

El propósito de este proyecto es utilizar los conocimientos adquiridos gracias al curso de introducción al lenguaje de programación Python para realizar el análisis de distintos conjuntos de datos correspondientes a las ventas, inventario y búsquedas realizadas en una tienda de tecnología, con lo que se podría entender el por qué la tienda ha disminuido sus ventas en el último trimestre.

Para lograr realizar este proyecto, no sólo se debe hacer uso de los conocimientos de programación adquiridos en el curso, los cuales son vitales, si no que se tuvo que realizar un análisis detenido de cómo utilizar los datos que se tienen de la tienda, lo cual es parte importante del trabajo de un científico de datos.

A lo largo del proyecto, se hizo la estimación de 3 puntos principales:

1. Productos más vendidos y productos rezagados.
2. Productos por reseña en el servicio.
3. Total de ingresos y ventas promedio mensuales, total anual y meses con más ventas.

A partir de estos puntos, se propondrán los resultados correspondientes a los puntos que fueron planteados por el área de gerencia de la empresa a resolver, los cuales son:

- a) Productos más vendidos y productos rezagados
- b) Productos por reseña
- c) Sugerencia de una estrategia de productos a retirar así como buscar la reducción de inventario

Definición del código

A partir de los primeros tres puntos mencionados en la introducción, se desarrolló el código que está incluido con este proyecto, tomando en cuenta que se importó el archivo `lifestore_file` para hacer uso de las listas en el mismo con toda la información relevante para la solución del problema.

El acceso al repositorio de GitHub es:

https://github.com/JaazGarer/Proyecto1_Emtech.git

Solución al problema

Para dar solución a los puntos planteados por gerencia, y usando los datos obtenidos a través del análisis, se adjuntaran distintos resultados.

Primero, de acuerdo a los resultados obtenidos en cuanto a ventas mensuales totales y ventas en general:

```
Para el mes 11 del año 2019 la venta total fue de 4209, la venta promedio fue de 4209.0 y el número de ventas fue 1
Para el mes 01 del año 2020 la venta total fue de 6372561, la venta promedio fue de 120237.0 y el número de ventas fue 53
Para el mes 02 del año 2020 la venta total fue de 4515699, la venta promedio fue de 110139.0 y el número de ventas fue 41
Para el mes 03 del año 2020 la venta total fue de 8401179, la venta promedio fue de 164729.0 y el número de ventas fue 51
Para el mes 04 del año 2020 la venta total fue de 14497125, la venta promedio fue de 193295.0 y el número de ventas fue 75
Para el mes 05 del año 2020 la venta total fue de 3470184, la venta promedio fue de 96394.0 y el número de ventas fue 36
Para el mes 06 del año 2020 la venta total fue de 406439, la venta promedio fue de 36949.0 y el número de ventas fue 11
Para el mes 07 del año 2020 la venta total fue de 296439, la venta promedio fue de 26949.0 y el número de ventas fue 11
Para el mes 08 del año 2020 la venta total fue de 9231, la venta promedio fue de 3077.0 y el número de ventas fue 3
Para el mes 09 del año 2020 la venta total fue de 4199, la venta promedio fue de 4199.0 y el número de ventas fue 1
Para el mes 11 del año 2020 la venta total fue de 0, la venta promedio fue de 0 y el número de ventas fue 0
```

Se puede observar claramente que los últimos tres meses si existió una baja a los ingresos y ventas de la empresa, pues claramente llega a ceros.

Además encontramos la venta anual de los dos años registrados y los meses con mayor cantidad de ingresos:

```
Para el año 2019 la venta fue de 4209
Para el año 2020 la venta fue de 37977265
[['2019', 4209], ['2020', 37977265]]
Los 5 meses con mejores ventas fueron:
mes 04 año 2020
mes 03 año 2020
mes 01 año 2020
mes 02 año 2020
mes 05 año 2020
```

Así bien, hagamos un análisis de las ventas usando lo pedido por el proyecto:

a) Productos más vendidos y productos rezagados

```
-----Los cinco más vendidos son:
1.-Procesador AMD Ryzen 5 2600, S-AM4, 3.40GHz, Six-Core, 16MB L3 Cache, con Disipador Wraith Stealth
2.-Procesador Intel Core i3-9100F, S-1151, 3.60GHz, Quad-Core, 6MB Cache (9na. Generación - Coffee Lake)
3.-Tarjeta Madre ASRock Micro ATX B450M Steel Legend, S-AM4, AMD B450, HDMI, 64GB DDR4 para AMD
4.-SSD Kingston A400, 120GB, SATA III, 2.5'', 7mm
5.-SSD Adata Ultimate SU800, 256GB, SATA III, 2.5'', 7mm
['Procesador AMD Ryzen 5 2600, S-AM4, 3.40GHz, Six-Core, 16MB L3 Cache, con Disipador Wraith Stealth', 'Procesador Intel Core i3-9100F, S-1151, 3.60GHz, Quad-Core, 6MB Cache (9na. Generación - Coffee Lake)', 'Tarjeta Madre ASRock Micro ATX B450M Steel Legend, S-AM4, AMD B450, HDMI, 64GB DDR4 para AMD', "SSD Kingston A400, 120GB, SATA III, 2.5'', 7mm", "SSD Adata Ultimate SU800, 256GB, SATA III, 2.5'', 7mm"]
```

-----Los diez más buscados son:

- 1.-Procesador AMD Ryzen 5 2600, S-AM4, 3.40GHz, Six-Core, 16MB L3 Cache, con Disipador Wraith Stealth
- 2.-Procesador AMD Ryzen 3 3200G con Gráficos Radeon Vega 8, S-AM4, 3.60GHz, Quad-Core, 4MB L3, con Disipador Wraith Spire
- 3.-Procesador Intel Core i3-9100F, S-1151, 3.60GHz, Quad-Core, 6MB Cache (9na. Generación - Coffee Lake)
- 4.-Procesador Intel Core i3-9100F, S-1151, 3.60GHz, Quad-Core, 6MB Cache (9na. Generación - Coffee Lake)
- 5.-Procesador Intel Core i7-9700K, S-1151, 3.60GHz, 8-Core, 12MB Smart Cache (9na. Generación Coffee Lake)
- 6.-Tarjeta Madre ASUS micro ATX TUF B450M-PLUS GAMING, S-AM4, AMD B450, HDMI, 64GB DDR4 para AMD
- 7.-SSD XPG SX8200 Pro, 256GB, PCI Express, M.2
- 8.-SSD XPG SX8200 Pro, 256GB, PCI Express, M.2
- 9.-SSD Kingston A400, 120GB, SATA III, 2.5'', 7mm
- 10.-SSD Adata Ultimate SU800, 256GB, SATA III, 2.5'', 7mm

-----Los diez menos buscados son:

- 1.-Procesador AMD Ryzen 5 2600, S-AM4, 3.40GHz, Six-Core, 16MB L3 Cache, con Disipador Wraith Stealth
- 2.-Procesador AMD Ryzen 3 3200G con Gráficos Radeon Vega 8, S-AM4, 3.60GHz, Quad-Core, 4MB L3, con Disipador Wraith Spire
- 3.-Procesador Intel Core i3-9100F, S-1151, 3.60GHz, Quad-Core, 6MB Cache (9na. Generación - Coffee Lake)
- 4.-Procesador Intel Core i3-9100F, S-1151, 3.60GHz, Quad-Core, 6MB Cache (9na. Generación - Coffee Lake)
- 5.-Procesador Intel Core i7-9700K, S-1151, 3.60GHz, 8-Core, 12MB Smart Cache (9na. Generación Coffee Lake)
- 6.-Tarjeta Madre ASUS micro ATX TUF B450M-PLUS GAMING, S-AM4, AMD B450, HDMI, 64GB DDR4 para AMD
- 7.-SSD XPG SX8200 Pro, 256GB, PCI Express, M.2
- 8.-SSD XPG SX8200 Pro, 256GB, PCI Express, M.2
- 9.-SSD Kingston A400, 120GB, SATA III, 2.5'', 7mm
- 10.-SSD Adata Ultimate SU800, 256GB, SATA III, 2.5'', 7mm

```

-----Los diez con más reservas son:
1.-Procesador AMD Ryzen 5 3600, S-AM4, 3.60GHz, 32MB L3 Cache, con
Disipador Wraith Stealth
2.-Procesador AMD Ryzen 5 2600, S-AM4, 3.40GHz, Six-Core, 16MB L3 C
ache, con Disipador Wraith Stealth
3.-Procesador AMD Ryzen 3 3200G con Gráficos Radeon Vega 8, S-AM4,
3.60GHz, Quad-Core, 4MB L3, con Disipador Wraith Spire
4.-Tarjeta de Video VisionTek AMD Radeon HD 5450, 1GB DDR3, PCI Exp
ress x16 2.1
5.-Tarjeta Madre ASUS micro ATX Prime H370M-Plus/CSM, S-1151, Intel
H370, HDMI, 64GB DDR4 para Intel
6.-SSD Kingston A400, 120GB, SATA III, 2.5'', 7mm
7.-TCL Smart TV LED 55S425 54.6, 4K Ultra HD, Widescreen, Negro
8.-TV Monitor LED 24TL520S-PU 24, HD, Widescreen, HDMI, Negro
9.-Makena Smart TV LED 40S2 40'', Full HD, Widescreen, Negro
10.-Getttech Audifonos con Micrófono Sonority, Alámbrico, 1.2 Metro
s, 3.5mm, Negro/Rosa

```

A partir de lo observado, notamos que los productos más vendidos NO coinciden en su totalidad con los más buscados. En los elementos con más reservas, encontramos un par de coincidencias los menos buscados. Ahora veamos un ejemplo de la comparación entre ventas y búsquedas considerando la categoría:

Categoría: Procesadores

```

-----Los cinco menos vendidos en la categoria procesadores son:
1.-Procesador AMD Ryzen 3 3300X S-AM4, 3.80GHz, Quad-Core, 16MB L2
Cache
2.-Procesador Intel Core i9-9900K, S-1151, 3.60GHz, 8-Core, 16MB Sm
art Cache (9na. Generación Coffee Lake)
3.-Procesador Intel Core i7-9700K, S-1151, 3.60GHz, 8-Core, 12MB Sm
art Cache (9na. Generación Coffee Lake)
4.-Procesador Intel Core i5-9600K, S-1151, 3.70GHz, Six-Core, 9MB S
mart Cache (9na. Generación - Coffee Lake)
5.-Procesador Intel Core i3-8100, S-1151, 3.60GHz, Quad-Core, 6MB S
mart Cache (8va. Generación - Coffee Lake)

```



```
-----Los diez menos buscados en la categoria procesadores son:
1.-Procesador AMD Ryzen 3 3300X S-AM4, 3.80GHz, Quad-Core, 16MB L2
Cache
2.-Procesador AMD Ryzen 3 3300X S-AM4, 3.80GHz, Quad-Core, 16MB L2
Cache
3.-Procesador AMD Ryzen 5 3600, S-AM4, 3.60GHz, 32MB L3 Cache, con
Disipador Wraith Stealth
4.-Procesador AMD Ryzen 5 2600, S-AM4, 3.40GHz, Six-Core, 16MB L3 C
ache, con Disipador Wraith Stealth
5.-Procesador AMD Ryzen 3 3200G con Gráficos Radeon Vega 8, S-AM4,
3.60GHz, Quad-Core, 4MB L3, con Disipador Wraith Spire
6.-Procesador Intel Core i3-9100F, S-1151, 3.60GHz, Quad-Core, 6MB
Cache (9na. Generación - Coffee Lake)
7.-Procesador Intel Core i9-9900K, S-1151, 3.60GHz, 8-Core, 16MB Sm
art Cache (9na. Generación Coffee Lake)
8.-Procesador Intel Core i9-9900K, S-1151, 3.60GHz, 8-Core, 16MB Sm
art Cache (9na. Generación Coffee Lake)
9.-Procesador Intel Core i7-9700K, S-1151, 3.60GHz, 8-Core, 12MB Sm
art Cache (9na. Generación Coffee Lake)
10.-Procesador Intel Core i5-9600K, S-1151, 3.70GHz, Six-Core, 9MB
Smart Cache (9na. Generación - Coffee Lake)
```

Podemos notar que existe coincidencia entre lo menos vendidos y aquellos que son menos buscados. Hecho que se repite en todas las categorías de los productos. Este hecho es un punto interesante por explorar.

Nota: En este punto hay una falla con el código, en la cual le falta distinguir entre aquellos objetos que están repetidos, lo cual podría dar una estimación aún mejor.

b) Productos por reseña

```
----- Los productos mejor reseñados (5), son:
1.-Procesador AMD Ryzen 3 3300X S-AM4, 3.80GHz, Quad-Core, 16MB L2
Cache
2.-Procesador AMD Ryzen 5 3600, S-AM4, 3.60GHz, 32MB L3 Cache, con
Disipador Wraith Stealth
3.-Procesador AMD Ryzen 5 2600, S-AM4, 3.40GHz, Six-Core, 16MB L3 C
ache, con Disipador Wraith Stealth
4.-Procesador AMD Ryzen 3 3200G con Gráficos Radeon Vega 8, S-AM4,
3.60GHz, Quad-Core, 4MB L3, con Disipador Wraith Spire
5.-Procesador Intel Core i3-9100F, S-1151, 3.60GHz, Quad-Core, 6MB
Cache (9na. Generación - Coffee Lake)
```

```
----- Los productos peor reseñados (1), son:
1.-Tarjeta de Video Gigabyte AMD Radeon R7 370 OC, 2GB 256-bit GDDR
5, PCI Express 3.0
2.-Tarjeta Madre ASUS micro ATX TUF B450M-PLUS GAMING, S-AM4, AMD B
450, HDMI, 64GB DDR4 para AMD
3.-Tarjeta Madre AORUS micro ATX B450 AORUS M (rev. 1.0), S-AM4, AM
D B450, HDMI, 64GB DDR4 para AMD
4.-Tarjeta Madre ASRock ATX H110 Pro BTC+, S-1151, Intel H110, 32GB
DDR4, para Intel
```

Aquí se obtuvieron las dos listas indicadas en el trabajo, para las cuales la mejor reseña se consideró 5 y la peor se consideró 1. Un ejemplo que valdría la pena destacar en esta parte es que si bien, por ejemplo el procesador AMD Ryzen 3 3300X se mantiene entre los menos vendidos entre la categoría de los procesadores y también es de los menos buscados, es uno de los productos que tiene la mejor reseña, por lo que los consumidores que lo adquieren, tienen una buena opinión del mismo.

c) Sugerencia de estrategia

Así pues, basándonos en los resultados que se obtuvieron gracias al código realizado, se podría proponer lo siguiente:

Considerando la coincidencia entre los productos menos buscados y los menos vendidos, se puede tomar en cuenta retirar del mercado aquellos productos que coincidan en ambas listas de cada categoría, pues es claro que es una tendencia que su poca popularidad no les haga venderse.

Si se buscase hacer un análisis más profundo, valdría la pena, además hacer una análisis continuo con el stock de los productos, y mantener limitado el número de stock de aquellos productos que son poco vendidos, y aumentar el stock de aquellos bien recibidos por el público.

El resultado correspondiente a las ventas totales, nos permite ver que en los últimos meses fue a una baja muy apresurada, en ese sentido, sería interesante hacer una análisis de las ventas totales que no sólo incluyan los meses, si no también los productos que fueron vendidos, pues eso permitiría entender y estimar la necesidad de ciertos productos sobre otros.

Conclusión

Si bien se tiene una idea clara de qué tipo de productos se podrían quitar de inventario, es importante considerar que aún queda un análisis completo por realizar, pues hay aspectos que aún no fueron tomados en cuenta y podrían cambiar los resultados del proyecto. También valdría la pena hacer un análisis más completo categorizando por producto: sus ventas, meses con mayores ventas, búsquedas y su comparación con respecto a otros productos de su misma categoría.

Cosidero que la idea del proyecto se cumplió satisfactoriamente, pues se dio a gran manera el uso de los elementos básicos de programación en Python, además fue también de gran utilidad al identificar todas las posibilidades y escenarios que un conjunto de datos nos brinda. Lo más

interesante fue reconocer que los datos nos pueden brindar tanta información como sepamos obtener de ellos.