

EMTECH

DATA SCIENCE

## **REPORTE: PROYECTO 1**

*Gil Estéfano Rodríguez Rivera*

*ge.rodriguezrivera@ugto.mx*

31 de agosto de 2021

## Índice

1. Introducción	3
2. Definición del código	3
3. Solución al problema	4
4. Conclusión	5

(**META-COMENTARIO:** PARA EL REPORTE Y EL CÓDIGO, SE HACE LA SUPOSICIÓN DE QUE LA FECHA DEL ANÁLISIS ES LA MISMA QUE LA QUE CORRE AL ENVIAR ESTE REPORTE. ESTO ES PORQUE TAL SUPOSICIÓN OFRECE MÁS POSIBILIDADES PARA EL ANÁLISIS DE SOLUCIONES, POR EJEMPLO: ¿CUÁL PUEDE SER EL MOTIVO DETRÁS DE ALGUNOS MESES SIN VENTAS EN ABSOLUTO?.)

## 1. Introducción

El caso que se nos ha encomendado es sobre una tienda virtual: **LifeStore**. La Gerencia de la tienda solicitó un análisis de su inventario. Específicamente, sobre los artículos más comprados y buscados en la tienda así como los menos comprados y consultados. También será importante conocer las reseñas de los productos. El motivo del análisis es comprender la problemática para proponer una estrategia que permita reducir la acumulación de inventario así como para aumentar las ventas mensuales.

El código del análisis así como el de las bases de datos y este mismo reporte los puede encontrar en el siguiente repositorio de **GitHub**: [https://github.com/gilesitorr/EmTech\\_Proyectos/tree/main/Proyecto\\_1](https://github.com/gilesitorr/EmTech_Proyectos/tree/main/Proyecto_1).

## 2. Definición del código

El código está estructurado de tal modo que primero importa las bases de datos de la tienda (la de las ventas, la de las búsquedas y la de los productos del catálogo). Luego, se definen las funciones que se usan regularmente dentro del programa (aquellas que se usan para ordenar o filtrar las listas, por ejemplo). Después, sigue la sección donde se hace el procesamiento de datos para llevar a cabo todas las tareas deseadas. Al final, está la sección donde se imprime la información procesada.

Al inicio de la sección de procesamiento del código se implementó un sistema de inicio de sesión (sólo permite acceso al usuario **admin** con contraseña **12345**). Ese inicio de sesión funciona con un bucle while, del que se puede salir sólo saliendo del programa (finalizándolo introduciendo el valor “s” cuando es apropiado) o iniciando sesión correctamente. Una vez que se inicia sesión, se procesa y despliega toda la información solicitada por el cliente. Todo ese flujo (si se ejecuta o no el procesamiento), se controla por medio de la variable **salir\_del\_programa**: si su valor es “s”, no se ejecuta el procesamiento ni la visualización; de lo contrario, sí se ejecuta.

Lo primero en procesarse son los 50 productos con mayores ventas así como los 100 productos más buscados. En segundo lugar, se procesaron los productos con menores ventas así como los que tienen menores búsquedas (en esta ocasión, el ranqueo se hizo dependiendo de las categorías de los artículos: es decir, se rankearon puros teclados, luego se rankearon puros mouse, etc.).

La siguiente parte del procesamiento se centró en generar dos listas de productos: una con los productos que tienen las mejores reseñas y otros con los peores. Los objetos sin reseñas no se consideraron en estas listas, pues el hecho de que no las tengan implica necesariamente que no se han vendido unidades de dicho artículo.

La última parte del procesamiento estudió las ganancias del 2020 en distintas

variantes, así como la cantidad de ventas hechas en el mismo año. Además, se procesaron los meses del 2020 donde hubo más ganancias. Note que en esta parte del análisis se encontraron datos anómalos de ventas pertenecientes a los años 2002 y al 2019. Esos datos no fueron tomados en cuenta en el estudio.

### 3. Solución al problema

Toda la información que se usará en el diagnóstico del problema se encuentra en el código. Sólo se requiere de iniciar sesión para tener acceso a la misma información que se cita.

La primera cuestión que surge al ver que no hubo ventas en los últimos meses del año es si la base de datos de la tienda se ha actualizado. En caso de ser ese el problema, la recomendación es revisar la infraestructura de la tienda virtual. Puede que se requiera mantenimiento en el proceso de captura de datos y el manejo de las bases de datos. Esto es porque parece improbable que las ganancias de los últimos meses sean absolutamente cero. De hecho, desde Agosto del 2020 debe de haber un problema, pues en ese mes hubo una ganancia de \$3,077, valor que no coincide ni siquiera con el orden de magnitud del mes que le precede en el ranking, que es Julio, en el cual se obtuvo una ganancia de \$26,949. Sin embargo, desde Junio hubo una disminución en las ventas con respecto a los primeros meses del año.

En el resto de la sección, suponga que no se trata de un problema en el sistema de captura de datos o el tratamiento de las bases de datos. Entonces la segunda mitad del año está en ceros completos. Note que hay similitudes entre las listas de los más buscados con las listas de los más vendidos. Sin embargo, la verdadera similitud yace en los rankings de los menos buscados y la de los menos vendidos. Además, también hay una relación entre las ventas y las reseñas de los productos (aunque esa relación es menos fuerte que la de las búsquedas y las ventas). Pero la relación es más notoria cuando se trata de los objetos más vendidos: pues los objetos menos vendidos son menos propensos a tener, siquiera, una reseña. Lo mencionado puede mostrarse con las siguientes relaciones:

Más Buscados  $\iff$  Más Vendidos  $\implies$  Mejores reseñas

Menos Buscados  $\iff$  Menos Vendidos  $\nRightarrow$  Peores reseñas

Como se ve en las relaciones de arriba: a más exposición, es posible que haya más ventas. Una estrategia posible es implementar una rotación de los artículos recomendados cada cierto tiempo, para evitar que los artículos se rezaguen. Aunque para evitar recomendar puros artículos *desconocidos* en las recomendaciones, debería

poder determinarse la proporción apropiada de artículos más buscados y artículos rezagados que se promueven y recomiendan.

Lo anterior es algo que se puede implementar en todo el año, para evitar rezagos. Sin embargo, puede verse que hay un problema en cierta época del año: la segunda mitad. Esta diferencia de ganancias entre mitad y mitad del año puede deberse a la recesión económica que se produjo por la pandemia de COVID-19 (es decir, en vistas de que la situación económica iba “*para largo*”, la clientela dejó de comprar artículos no primordiales). Incluso si es el caso, el hecho de que no haya habido ventas ni siquiera en temporada navideña, indica que puede que los precios o las promociones de la página no hayan podido competir con las de otras páginas que ofrecen el mismo catálogo de productos o, incluso, que no haya habido suficiente publicidad sobre la página en esas temporadas, haciendo que los clientes optasen por páginas más conocidas. Suponga que el tráfico de la página ha sido el problema y que la razón por la que no ha habido ganancias ha sido que el precio de los productos no es competitivo (de todos modos, se recomienda revisar las estrategias del equipo de Marketing).

La estrategia en este caso empieza por revisar los precios medios de los productos del catálogo en otras páginas web así como sus fluctuaciones a lo largo del año. En base a eso, se puede determinar si el precio es competitivo o no. También se recomienda revisar si las páginas competidoras ofrecen promociones para la segunda mitad del año (que seguramente es el caso, sobretodo en temporada de Black Friday y de Navidad). En base a eso, se propone variar los precios en función de la temporada. Además, en la segunda mitad del año, se recomienda lanzar promociones de equipos. Por ejemplo, en la compra de un artículo -que ya se vende mucho-, te puedes llevar este artículo -que no se vende mucho- por una pequeña diferencia de precio. O si se trata de un artículo que ya se pretendía discontinuar, se puede dar de regalo en la compra de cualquier otro artículo.

Esto es lo que se propone realizar, dado que sólo se tienen datos sobre un año completo (2020). Sin embargo, con los datos del primer trimestre del 2021 se podría determinar si el problema son las fluctuaciones de los precios por las temporadas y las promociones de los competidores o si el problema es otro (como que el catálogo esté desactualizado o que la publicidad no sea apropiada).

## **4. Conclusión**

A pesar de las soluciones y estrategias que se dieron, se recomienda fuertemente primero revisar que el sistema de captura de datos funcione correctamente. Casi todas las demás estrategias discutidas no tienen sentido si ese sistema no funciona correctamente (y que las ganancias sean absolutamente cero no alimenta sino la

hipótesis de que hay un error en el sistema). Sea cual sea el caso, se recomienda implementar el sistema de rotación de los artículos recomendados.

Un trabajo de continuación puede ser, una vez implementadas las medidas propuestas para aumentar las ventas, realizar otra versión del programa que permita imprimir los puntajes promedio. Esto es útil porque se pueden realizar pruebas estadísticas para determinar si, en base al puntaje promedio y la cantidad de reseñas, el artículo es de *buena* o *mala* calidad y merece ser discontinuado del catálogo o no. Naturalmente, esto se debe tratar con más detenimiento y precisión al momento de la implementación.

Ese mismo sistema para discontinuar productos puede usarse para la estrategia de rotación de las recomendaciones: promoviendo aquellos artículos que no tengan suficientes reseñas para determinar si son *buenos* o *malos*. Respecto al problema de las proporciones de las recomendaciones de artículos conocidos y los menos vendidos, se recomienda correr una prueba A/B para determinar cuál es la mejor proporción posible. Eso, nuevamente, se debe tratar con más detenimiento y precisión en un trabajo futuro.

En conclusión, se requieren más datos para determinar la naturaleza del problema de las ventas. A pesar de eso, se puede empezar a implementar las medidas de revisión de la publicidad y la infraestructura de la página cuanto antes. Asimismo, también se puede implementar el estudio de las fluctuaciones de los precios en otras páginas así como las promociones usadas y en qué temporada se lanzan.