

Utilizando la Ciencia de Datos en una organización

Proyecto individual evaluado por pares

José Alejandro López Quel

EDX - TecdeMonterreyX

Introducción a la Ciencia de Datos y el Big Data

Enero 2022

Índice

1	Introducción	3
2	Identificación de indicadores	4
3	Tipo de análisis y justificación	4
4	Propuesta de decisiones	5
5	Proceso de evaluación	5
6	Conclusiones	6

1. Introducción

La Librería Iztaccihuatl ubicada en la ciudad de Monterrey, Nuevo León, México, es una empresa con más de 57 años en el mercado literario, la cual posee un compromiso con la comunidad creando experiencias únicas, impulsando la lectura, la cultura y la educación a nivel nacional. Esta empresa busca ser líder en la distribución y comercialización de material bibliográfico a nivel nacional y convertirse en una referencia cultural y de lectura mediante la difusión, apoyándose de editoriales, autores y asociaciones; siempre pensando en la responsabilidad hacia la comunidad.

El objetivo de este trabajo es analizar las fuentes de información externas e internas de la organización a través de herramientas computacionales para genera un modelo basado en ciencia de datos como apoyo para la toma de decisiones.

Para este análisis se cuenta con la siguiente información:

- Datos generales de cada libro existente en la librería. Junto con esta información se tiene también el promedio de calificación de cada libro de acuerdo a las votaciones y compras del cliente.
- Datos de los 20 libros más vendidos de acuerdo a una calificación general.
- Datos de los libros más votados por los clientes dentro del sitio web de la librería.
- Datos de las recomendaciones que cada cliente o usuario realiza en el sitio web sobre libros para leer.

Con el fin de lograr el objetivo mencionado se utiliza la herramienta computacional llamada Power BI para gestionar los datos proporcionados en archivos de Excel, lo cual permitió obtener una visión general de la información. De esta manera se descartan columnas que no tienen mayor relevancia y se agrupan datos para generar nuevas columnas las cuales se detallan más adelante. Gracias a este análisis se logra extraer más información sobre los libros que se comercializan, tal como:

- Libros que generan un mayor importe de ventas.
- Libros que más se vendieron.
- Valoraciones de cada libro y su promedio.
- Detalles de venta dada la clasificación por categoría.

2. Identificación de indicadores

Posteriormente a analizar la información se determina que los indicadores de desempeño más importantes que ayudarán al área administrativa de la empresa para la toma de decisiones son los siguientes:

- **Indicador de tendencias por categorías de libros.** Con base en los datos proporcionados se puede determinar la cantidad de ventas por cada categoría. Este calculo se realizaría de la siguiente manera:

$$\text{Tendencia por categoría} = \frac{\text{Ventas del mes de la categoría deseada}}{\text{Ventas totales de la misma categoría}}$$

- **Indicador de utilidad por libro.** Este indicador mide la relación entre la ganancia y lo invertido por cada libro. La relación numérica estaría dada por:

$$\text{Utilidad por libro} = \frac{\text{Importe total de ventas}}{\text{Importe invertido en la compra del libro}}$$

- **Indicador de compras por cliente.** Este indicador mide la actividad del cliente, en otras palabras, sirve para cuantificar la cantidad de compras realizadas por el cliente desde la fecha de su primer compra. Por lo que su expresión esta dada por:

$$\text{Compras por cliente} = \frac{\text{Compras totales del cliente}}{\text{Meses transcurridos desde su primer compra}}$$

- **Indicador de libros más vendidos.** Este lo conforman las unidades vendidas de cada libro.

3. Tipo de análisis y justificación

Los datos con los que se disponen de la librería son de carácter histórico por lo que con ellos se puede realizar un análisis descriptivo para determinar el estado actual del negocio. Con ello se podría determinar cuales son las categorías mas vendidas, los clientes más frecuentes y la ganancia por cada libro. Con estos datos se obtienen una visión de lo que ha pasado y lo que esta pasando.

Adicionalmente, se propone un análisis predictivo ya que con las calificaciones de los clientes y el comportamiento de compra de los mismos se puede generar mejores sugerencias para ellos, así como programas de fidelización para obtener mejores resultados de venta. Además, se puede anticipar la demanda de ciertas categorías de libros.

4. Propuesta de decisiones

A continuación se describe las decisiones que se podrían tomar con base en los datos analizados, incluyendo la estrategia de implementación en donde corresponde.

- Con base en el indicador de tendencias por categoría de libros, se puede implementar campañas perfiladas para generar mayor promoción a dichas categorías, lo que permitiría aumentar el número de ventas y ahorro de gastos de la compra de categorías menos solicitadas, orientando de mejor manera los recursos.
- El indicador de libros más vendidos, permitiría a la administración de la biblioteca gestionar de manera mas eficiente el inventario de libros de acuerdo a la demanda de los clientes.
- Mediante el uso del indicador de utilidad por libro se obtendría el beneficio actual por la compra y venta de los mismos, aportando a la toma de decisiones la opción de adquirir los libros con mayor margen de ganancia.

Además de estos datos obtenidos mediante el análisis descriptivo, se puede tomar decisiones basadas en las predicciones que se podrían obtener del modelo propuesto en el análisis predictivo. Algunas de las decisiones que se podrían tomar son:

- Preparar y disponer del inventario necesario para ventas, basado en el modelo predictivo.
- Ofrecer experiencias personalizadas a los clientes.

5. Proceso de evaluación

Para obtener mejores resultados se propone que durante el proceso del desarrollo del modelo se evalúe el mismo para comprender la calidad de los datos obtenidos y garantizar que la toma de decisiones se abordará de manera adecuada, de acuerdo con los parámetros del negocio. Esta evaluación implica el calculo de métricas de diagnostico para que sean comparadas con los datos reales y otros resultados, generando gráficas y tablas para que sean visualizados de una forma ordenada y sencilla de entender. Esto permitirá interpretar y analizar la calidad del modelo generado y su eficiencia para cumplir con los objetivos trazados por el negocio.

6. Conclusiones

El desarrollo de este caso de estudio permite profundizar sobre el estado actual del negocio analizado. A través de la ciencia de datos y sus conceptos se puede evaluar y generar mayor información para apoyar a la toma de decisiones, por lo a continuación se presentan las conclusiones.

- Se propone un modelo basado en ciencia de datos de acuerdo a los datos proporcionados por el negocio.
- Se identifica 3 indicadores que aportarán valor a la toma de decisiones del negocio.
- Se realiza dos tipos de análisis, que permiten extraer el máximo valor de los datos compartidos.
- De acuerdo con los indicadores se proporciona las decisiones clave que se pueden llegar a tomar y su estrategia de implementación.
- Se detalla un proceso de evaluación para garantizar la calidad del modelo generado.