Propuesta de proyecto final

Análisis y Optimización de la Gestión de Inventarios en   
Alcoholic Beverage***.***

Integrantes del equipo:

1. *Vidal, Ezequiel* - Coordinador EDA, ETL y Generación Data Python

2. *García, Felipe* - Diseñador de Visualizaciones y Documentación Final Power BI

3. *Maldonado, Camilo* - Responsable de Modelado Predictivo

4. *Peña, Jaime* - Analista de Optimización y Gestión de Inventarios

1. Caso de negocio

| **Industria y contexto del caso de negocio a desarrollar:** | La industria en la que nos enfocaremos es la venta minorista de bebidas alcohólicas, enfocada en la gestión de inventario y rentabilidad en una tienda ubicada en HARDERSFIELD. El negocio enfrenta competencia intensa, cambios en preferencias de los consumidores y regulaciones estrictas. El proyecto analizará datos de inventario, precios y proveedores para optimizar operaciones y mejorar rentabilidad. Se utilizarán técnicas de análisis de datos para identificar oportunidades, adaptarse a tendencias del mercado y respaldar decisiones estratégicas que aumenten la eficiencia operativa y la competitividad en el sector. |
| --- | --- |
| **Descripción del Problema:** | El problema es la gestión ineficiente de inventario y la búsqueda de optimización de la rentabilidad. La falta de datos sobre ventas y compras dificulta el análisis preciso de los niveles de stock, márgenes de beneficio y desempeño de proveedores. Esto impide tomar decisiones informadas para equilibrar la oferta con la demanda, reducir costos operativos y mejorar la competitividad. El objetivo es utilizar análisis EDA, ETL, entre otros, para identificar oportunidades de mejora en la gestión del inventario y estrategias comerciales. |
| **Objetivo principal del proyecto:** | El objetivo general del proyecto es analizar y optimizar la ***gestión del inventario y la rentabilidad***. Mediante un análisis detallado, se busca identificar patrones y tendencias en los niveles de stock, márgenes de beneficio y desempeño de proveedores, con el fin de mejorar la eficiencia operativa, maximizar las ganancias y respaldar decisiones estratégicas que aumenten la competitividad y el éxito sostenible del negocio. |
| **Resultados esperados & Impacto:** | Se espera mejorar la eficiencia ***operativa y aumentar la rentabilidad*** de la tiendas. Los objetivos concretos son:   * Optimizar los niveles de inventario para reducir costos de almacenamiento en un 15%, evitando excesos y faltantes. * Incrementar el margen de beneficio promedio en un 10% al identificar y promover productos más rentables. * Reducir el costo de adquisición mediante la negociación con proveedores, logrando un ahorro del 5%. * Mejorar la rotación de productos de baja demanda mediante estrategias de promoción específicas.   El impacto será un negocio más competitivo y rentable, con decisiones informadas basadas en datos, aumentando la satisfacción del cliente y favoreciendo el crecimiento sostenible. |

1. Fuentes de datos

| **Fuentes de Datos:** | Las fuentes de datos son archivos en formato .csv que contienen información detallada sobre inventario inicial y final, productos, precios de venta y compra, proveedores y niveles de stock de las tiendas. Para incorporar nuevos registros:   * Estandarización: Asegurar que los nuevos datos siguen la misma estructura y formato que los archivos existentes. * Validación: Verificar la calidad de los datos, comprobando que no haya errores o valores faltantes. * Integración: Añadir los nuevos registros a los archivos .csv correspondientes, manteniendo la coherencia de la información. * Actualización Regular: Implementar un proceso periódico para incorporar y actualizar datos, permitiendo análisis continuos y precisos.   Esto garantiza que el conjunto de datos esté completo y sea útil para el análisis y la toma de decisiones. |
| --- | --- |
| **Descripción de los Datos:** | 1. Inventario Inicial y Final:    * Inventario Inicial (*startDate*: 2016-01-01): Incluye campos como *InventoryId*, *Store*, *City*, *Brand*, *Description*, *Size*, *onHand* (unidades en stock), *Price* y *startDate*.    * Inventario Final (*endDate*: 2016-12-31): Contiene los mismos campos que el inventario inicial, pero refleja las unidades en stock al final del año (*endDate*). 2. Información de Productos:    * Campos como *Brand* (identificador del producto), *Description* (descripción detallada), *Price* (precio de venta), *Size* (tamaño del producto), *Volume*, *Classification* (categoría), *PurchasePrice* (precio de compra al proveedor), *VendorNumber* (identificador del proveedor) y *VendorName* (nombre del proveedor).   Estos datos permiten:   * Analizar cambios en los niveles de inventario durante el año. * Calcular márgenes de beneficio comparando precios de compra y venta. * Evaluar el desempeño y la relación con diferentes proveedores. * Identificar oportunidades para optimizar la gestión del inventario y mejorar la rentabilidad del negocio.   Los datos están estructurados para facilitar su análisis, con identificadores comunes (***Brand*, *InventoryId***) que permiten la integración entre los diferentes conjuntos y ofrecen una visión completa del estado y movimiento de los productos en la tienda. |

1. Plan o estrategia de análisis

| **Requerimientos del caso:** | Integración de Datos: Fusionar los conjuntos de datos utilizando identificadores comunes.   1. Análisis Exploratorio: Examinar los datos para identificar patrones en inventario, precios y proveedores. 2. Cálculo de Márgenes: Calcular márgenes de beneficio por producto comparando precios de compra y venta. 3. Segmentación de Productos: Clasificar productos según rentabilidad y niveles de stock. 4. Detección de Oportunidades: Identificar productos con exceso o falta de stock y proveedores menos eficientes. 5. Propuestas de Mejora: Recomendar acciones para optimizar el inventario y aumentar la rentabilidad, pese a las limitaciones de datos. |
| --- | --- |
| **Tareas y tecnologías** | Para llevar a cabo el proyecto de análisis de datos y optimización de la gestión de inventario, se utilizarán diversas tecnologías y herramientas que faciliten el procesamiento, análisis y visualización de datos, como las siguientes:   1. Recopilación y Gestión de Datos    * Tecnologías:      + Python: Lenguaje de programación principal para el procesamiento de datos.      + Pandas: Biblioteca para manipulación y análisis de datos estructurados.      + NumPy: Biblioteca para cálculos numéricos y manejo de matrices.      + CSV: Formato de archivo utilizado para almacenar y transferir los conjuntos de datos.    * Tareas:      + Importar los archivos .csv con los datos de inventario, productos y proveedores.      + Estandarizar y combinar los datos de diferentes fuentes en un único conjunto coherente.      + Manejar valores faltantes, duplicados o inconsistencias en los datos. 2. Análisis Exploratorio de Datos (EDA)    * Tecnologías:      + Jupyter Notebooks: Entorno interactivo para escribir y ejecutar código, documentar hallazgos y visualizar resultados.      + Matplotlib: Biblioteca para crear gráficos y visualizaciones básicas.      + Seaborn: Biblioteca para visualizaciones estadísticas más avanzadas.    * Tareas:      + Generar estadísticas descriptivas de los datos (medias, medianas, desviaciones estándar).      + Crear visualizaciones para entender distribuciones, tendencias y patrones en los datos.      + Identificar correlaciones entre variables como precios de compra, precios de venta y niveles de inventario. 3. Procesamiento y Preparación de Datos    * Tecnologías:      + Python (Funciones y Scripts personalizados): Para automatizar procesos de limpieza y transformación de datos.    * Tareas:      + Transformar formatos de fecha y hora para facilitar el análisis temporal.      + Crear nuevas variables o características, como márgenes de beneficio o cambios en el inventario.      + Normalizar y escalar datos si es necesario para análisis avanzados. 4. Visualización y Presentación de Resultados    * Tecnologías:      + Power BI: Herramientas de inteligencia de negocios para crear dashboards interactivos.    * Tareas:      + Diseñar gráficos y dashboards que muestren insights clave de manera clara.      + Visualizar tendencias temporales, patrones en el inventario y rentabilidad por producto.      + Preparar informes visuales para presentar a los stakeholders del negocio. 5. Automatización y Actualización de Datos    * Tecnologías:      + Python (Automatización con scripts): Para programar tareas de actualización periódica de datos.      + **Task Scheduler** (en Windows): Para programar la ejecución automática de scripts.      + Git y GitHub: Para control de versiones y colaboración en el código.    * Tareas:      + Configurar procesos automáticos para incorporar nuevos registros en los datos.      + Implementar control de versiones en el código para mantener un historial de cambios.      + Asegurar la reproducibilidad del análisis y facilitar futuras actualizaciones. 6. Seguridad y Cumplimiento    * Tecnologías:      + Buenas prácticas de manejo de datos: Asegurar la confidencialidad y protección de los datos sensibles.    * Tareas:      + Garantizar que los datos se manejen de acuerdo con las políticas de seguridad y privacidad. |