

ENFOQUE BASADO EN DATOS PARA LA GESTIÓN E INCREMENTO DE VENTAS DE PREMIUM DRINKS

Nombre de los autores:

Bianca de Robertis

Agustin Ledesma

Gisell Mateus

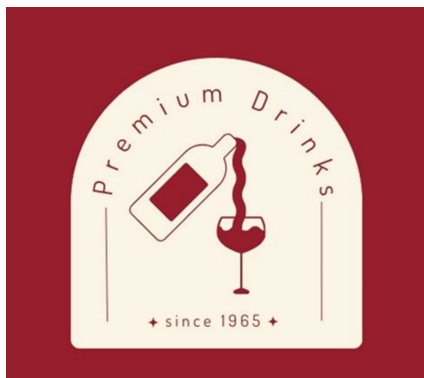
Sandra Ávila

Email: vesta935@gmail.com

Cohorte: da-pt02

Fecha de entrega: octubre de 2024

Institución: Premium Drinks



Premium Drinks se enfoca en ofrecer un servicio personalizado, tanto en sus puntos de venta físicos como a través de su plataforma en línea. Su modelo de negocio incluye la venta al por mayor, dirigida a bares, restaurantes y tiendas especializadas, así como la atención directa a clientes particulares, ofreciendo promociones exclusivas y productos de edición limitada.

Introducción:

La empresa de bebidas alcohólicas "PREMIUM DRINKS" enfrenta problemas en la gestión eficiente de inventarios, lo que genera altos costos o desabastecimientos, afectando la rentabilidad y la satisfacción del cliente.

Las soluciones existentes son costosas o complejas para esta empresa, lo que crea una necesidad de enfoques accesibles basados en análisis de datos para optimizar los niveles de stock y predecir la demanda.

Por lo cual, el objetivo principal de este trabajo es desarrollar un análisis integral de los datos de ventas, compras e inventarios para proporcionar a Premium Drinks insights que le permitan optimizar sus operaciones de inventarios, donde se identifiquen patrones de la demanda, permitiendo ajustar el nivel de stock en el

tiempo adecuado y por ende incrementar sus ingresos y ajustar sus estrategias de marketing.

Para lograr el objetivo en mención se responderán las siguientes preguntas:

- ¿Cuáles son los productos con mayor y menor rotación de inventario?
- ¿Cuál es la correlación entre los precios y el volumen de ventas?
- ¿Cómo varía la demanda según las temporadas?
- ¿Qué productos están generando mayores costos por sobrestock o desabastecimiento?
- ¿Cuál es el costo operativo por unidad almacenada de cada producto?
- ¿Qué productos generan mayor o menor satisfacción en los clientes debido a su disponibilidad?
- ¿Cuáles son los principales factores que influyen en las proyecciones de demanda (precio, promociones, estacionalidad)?
- ¿Qué margen de beneficio se obtiene por producto y cómo impacta en la estrategia de inventario?

Desarrollo del proyecto:

Para llevar a cabo este proyecto trabajamos el marco de trabajo ágil denominado SCRUM, el cual es utilizado para la gestión de proyectos complejos, enfocado en la entrega incremental de valor a través de iteraciones cortas y organizadas en "sprints". Cada sprint tiene una duración fija y persigue un conjunto específico de objetivos, permitiendo la adaptación continua a los cambios y la mejora continua en cada ciclo.

En el proyecto Premium Drinks, Scrum fue implementado para estructurar el trabajo en dos sprints clave:

Sprint 1: Implementación del Origen de los Datos:

Objetivos del sprint:

- Crear el repositorio público y compartirlo con el equipo
- Identificar los archivos a usar y armar la estructura de carpetas a usar en el repositorio
- Crear la base de datos
- Identificar las entidades/tablas que van a usar en la BD
- Crear las tablas y objetos SQL necesarios
- Crear el flujo que importe los datos a la base de datos
- Validar que todos tengan acceso a dicha base de datos y que puedan extraer data sin problemas
- Crear el método de automatización de ingesta de datos nuevos

Actividades realizadas:

ALMACENAMIENTO SEGURO DE DATOS:

- Se crear el repositorio público y se comparte con el equipo:

El primer paso del proyecto, fue la creación de un repositorio público en GitHub, que centraliza todas las actividades de desarrollo y colaboración entre los integrantes del equipo. Este repositorio actúa como un espacio de trabajo compartido donde se almacenan todos los recursos relacionados con el proyecto, como scripts, notebooks, documentación, imágenes y otros elementos esenciales para el avance adecuado.

El repositorio fue creado bajo el nombre PremiumDrinks, accesible en el siguiente enlace: <https://github.com/PremiumDrinks/PremiumDrinks>. Desde este punto, todos los miembros del equipo fueron invitados como colaboradores para asegurar el acceso adecuado.

Se crearon dos logos: El de la consultora "The Borcelle" que será la encargada de ayudar a "DrinksPremium" en la toma de decisiones, y el logo de "DrinksPremium" que es una distribuidora de bebidas alcohólicas.

- Archivos y estructura de carpetas:

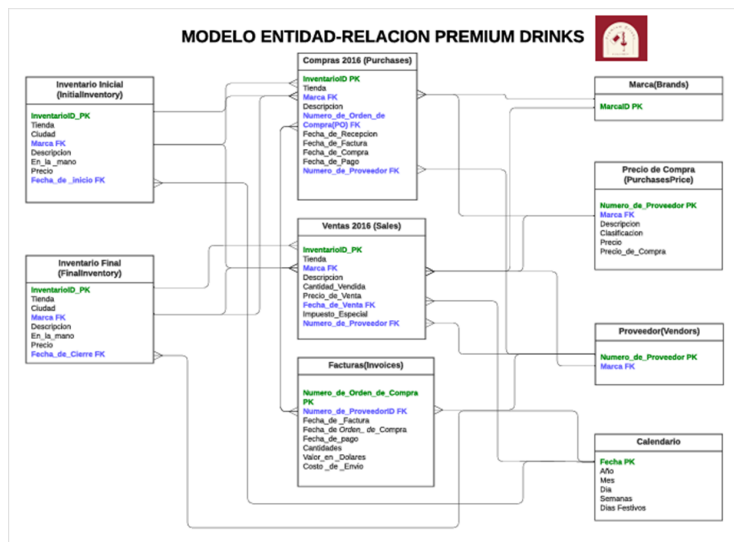
Se crean carpetas de "Dashboard y Presentación", "Data sets", "Queries SQL", "Phyton" con el fin de organizar de manera estructurada el proyecto.

Falta almacenamiento de los datos en una Base de Datos SQL Server y el manejo de claves seguras

ESTRUCTURA INTELIGENTE:

El modelo que estamos utilizando para Premium Drinks se asemeja más a un modelo de copo de nieve (snowflake schema) en lugar de un modelo estrella (star schema), donde Las tablas de hechos están en el centro (Ventas, Compras, Facturas). Las tablas dimensionales están alrededor, ofreciendo descripciones adicionales (Inventario, Marca, Proveedor, Calendario).

La normalización de tablas como Marca, Proveedor y Calendario fue necesaria para reducir la duplicación de datos y mejorar la organización y análisis detallado por dimensiones como tiempo, proveedor y marca. Este enfoque es ideal para análisis complejos y consultas que requieran múltiples niveles de detalle, manteniendo el modelo más ordenado y eficiente.



HALLAZGOS SOBRE EL MODELO ENTIDAD- RELACIÓN:

Hay varios hallazgos importantes para la gestión de **Premium Drinks**:

Relación integral entre inventarios, ventas y compras: Nos da una visión completa del flujo de productos, y nos permite ver con claridad problemas de stock, ya sea exceso o faltantes.

conexión detallada de proveedores: nos facilita negociar mejores condiciones y controlar el costo de compras.

optimización basada en fechas: Al analizar por temporadas o días festivos, podemos anticipar la demanda y ajustar inventarios

clasificación de productos mediante la tabla de Marca: Esto nos permite comparar el rendimiento de diferentes marcas y ajustar nuestras decisiones de compra o venta según su desempeño.

Facturas vinculadas a múltiples dimensiones: Las facturas están conectadas a compras y proveedores, lo que permite hacer un análisis más preciso de los costos y gastos.

Granularidad en la gestión de inventario: Comparar inventario inicial y final nos ayuda a mantener un control riguroso de los niveles de stock y evitar problemas de abastecimiento.

CARGA Y LIMPIEZA DE DATOS:

- Manejo de archivos CSV:

Trabajamos con grandes volúmenes de datos provenientes de archivos CSV, donde se almacenan las tablas de compras, ventas, inventarios entre otras, que representan transacciones clave de la empresa. Utilizando la librería pandas, hemos podido cargar estos archivos de manera eficiente.

- Limpieza de datos con Python:

Realizamos un proceso de limpieza para eliminar duplicados, corregir valores nulos e inconsistencias, asegurando la calidad de los datos, crucial para análisis posteriores, Utilizando las librerías de Pandas, Seaborn, matplotlib y numpy

Cargamos los archivos y creamos los dataframes, Verificamos la cantidad de registros de cada dataset y Verificamos los valores nulos. Con la variable head miramos la información general de las tablas, así las exploramos/

Luego de limpiar los datasets, los guardamos como archivo CSV y les dimos nombre en español para una mejor comprensión.

- Análisis exploratorio en Python:

Hicimos análisis exploratorio para identificar patrones iniciales. Generamos estadísticas descriptivas con pandas que nos ayudaron a comprender mejor las distribuciones de datos, las tendencias de ventas y los puntos de inventario clave. Utilizamos Matplotlib para crear gráficos simples y efectivos, y Seaborn, que nos permitieron hacer visualizaciones más atractivas y complejas, como gráficos de dispersión, boxplots y heatmaps.

AUTOMATIZACIÓN DE LA INGESTA DE DATOS: Desarrollar

- Configuración del flujo de ingesta de datos
- Programación de Tareas (Task Scheduler)
- Automatización del flujo de carga-Monitoreo y carga automática de archivos nuevos
- Mover los archivos procesados

Herramientas utilizadas: SQL Server, Python, Github, Trello

Resultados alcanzados:

- Calidad de los datos mejorada mediante la limpieza y transformación de las fuentes de datos iniciales, eliminando registros duplicados, tratando valores nulos y unificando formatos. Esto permite tener un conjunto de datos consistente y listo para análisis.
- Creación de repositorio en Github donde se almacena todos los archivos relacionados al proyecto y facilita el trabajo colaborativo del equipo
- Elaboración del modelo de entidad-Relación para Premium DRinks
- Creación de la base de datos en SQL Server y guarda de los diferentes datasets (ampliar)
- Elaboración modelo automatizado de ingesta de datos nuevos, para mantener actualizada la información.

Retos enfrentados y soluciones implementadas.

Lecciones aprendidas.

Sprint 2: Desarrollo del informe en Power BI, incluyendo limpieza y transformación de datos, medidas calculadas y filtros

Objetivos del sprint:

- Crear el archivo .pbix
- Importar la data a power BI
- Iniciar el proceso de limpieza y transformación de la data
- Crear las medidas/columnas/relaciones necesarias para el análisis
- Diseñar los reportes y gráficos a usar
- Etapa de Pruebas de calidad del reporte.

Actividades realizadas:

- Informe Interactivo Creado en Power BI:

Se diseña un dashboard completo que incluye visualizaciones dinámicas y segmentaciones que permiten a los usuarios analizar compras, ventas e inventario. Se ha dividido el dashboard en estas tres paginas.

- Medidas Personalizadas Implementadas:

Se crean cálculos específicos, como la rotación de inventario, productos más y menos vendidos, total de compras, relación de proveedores vs compras entre otras medidas que responden a las preguntas de investigación planteadas para facilitar la toma de decisiones informadas.

- Filtros y Segmentaciones:

Se añaden filtros intuitivos que permiten a los usuarios segmentar los datos por **xxxxx**, mejorando la capacidad de explorar los datos desde diferentes perspectivas.

- Identificación de Problemas Clave:

A través del análisis de datos, se identificaron áreas críticas, como productos con baja rotación o ubicaciones con inventario excesivo, proporcionando insights clave para optimizar la gestión de inventarios.

Carga y transformación de datos en Power BI: Explica cómo se importaron los datos y las transformaciones adicionales realizadas.

Medidas y cálculos: Detalla las medidas creadas, como métricas de inventario, predicciones de demanda, márgenes, etc.

Diseño del informe: Describe las visualizaciones utilizadas y su propósito. Incluye gráficos de tendencias, tablas, tarjetas de indicadores, etc.

Filtros y segmentaciones aplicadas: Explica cómo se organizaron los filtros para permitir un análisis interactivo.

Resultados alcanzados:

Visualización del estado del inventario.

Predicción de la demanda.

Retos enfrentados y soluciones implementadas.

Lecciones aprendidas.

Revisión del Proyecto y Retrospectiva

Comparación entre el objetivo inicial y los resultados obtenidos.

Identificación de áreas de mejora para futuros sprints.

Feedback del equipo: Incluye comentarios sobre lo que funcionó bien y lo que podría mejorarse.

Conclusiones y Recomendaciones

Resumen de los hallazgos clave.

Recomendaciones para la gestión de inventario en PremiumDrinks.

Siguientes pasos: Propuestas para mejorar el proyecto en el futuro.

Anexos

Código utilizado para la limpieza de datos y cálculos.

Documentación técnica adicional (por ejemplo, diccionario de datos).

Capturas de pantalla del informe en Power BI.

