Análisis Comparativo de Ventas y Facturación por Tienda y Ciudad

Vamos a examinar dos conjuntos de datos que son críticos para comprender el desempeño de nuestras tiendas a lo largo del tiempo, entre el 29 de enero de 2011 y el 24 de abril de 2016. La diapositiva se divide en dos paneles comparativos que reflejan la suma de ventas y la facturación total en las ciudades de Boston, Nueva York y Filadelfia.

Este análisis comparativo sirve como una herramienta valiosa para evaluar el impacto de nuestras estrategias de venta a lo largo del tiempo y cómo las diferentes regiones contribuyen al éxito general de la compañía. Con estos datos, podemos profundizar en la planificación estratégica y la toma de decisiones para impulsar el crecimiento y la eficiencia operativa en todas nuestras ubicaciones.

CANTIDAD DE VENTAS Y FACTURACIÓN

Análisis de Datos por Tienda, Eventos Especiales y Categorías de Producto

La presentación a continuación se centra en el desempeño financiero de nuestras tiendas, examinando la facturación total y las ventas por categoría de producto durante un periodo específico, del 29 de enero de 2011 al 24 de abril de 2016. Se destaca la contribución de eventos especiales a nuestras ventas y cómo varían las ventas durante los fines de semana y días laborales.

Facturación Total por Ciudad: Nueva York lidera con una facturación impresionante de más de $103 millones. Boston y Filadelfia siguen con aproximadamente $66 y $60 millones, respectivamente.

Facturación por Tienda: Tribeca resalta con una facturación que supera los $39 millones, seguida de Greenwich y Roxbury, ambos con ventas sustanciales superiores a $20 millones.

Número de Ventas por Tienda: La tienda de Tribeca no solo es líder en facturación, sino también en número de ventas, con más de 11 millones de transacciones. Greenwich Village y Roxbury muestran también un rendimiento fuerte en términos de volumen con 7 y 7.2 millones de ventas, respectivamente.

Facturación en Días de Eventos Especiales: Se evidencia el impacto de las festividades en nuestras ventas, siendo la Pascua la más rentable en las categorías de Accesorios y Hogar y Jardín. El Día de la Independencia es significativo para el Supermercado. El total combinado en estos días especiales asciende a más de $3.5 millones. (¡¡¡¡¡Comprobar al info!!!!!)

Media de Facturación: Los fines de semana presentan una media de facturación excepcional de casi $146 mil, mientras que los días laborales mantienen una media sólida de más de $111 mil. Esto puede deberse a que en los fines de semana las personas tienen más disponibilidad para ir a hacer las compras, por razones de trabajo.

Ventas por Categoría: En la distribución de ventas por categoría, el Supermercado domina claramente con un 68.63%, seguido por Hogar y Jardín con un 22.03%, y Accesorios con un 9.32%.

Este análisis detallado revela patrones importantes de consumo y preferencias de compra que se pueden aprovechar para optimizar estrategias de ventas y marketing. Por ejemplo, la alta facturación en eventos específicos podría impulsarnos a diseñar campañas promocionales dirigidas que coincidan con estas fechas. Además, la significativa contribución del Supermercado a nuestras ventas totales sugiere un enfoque en la expansión y promoción de esta categoría.

PRODUCTOS

Los tres productos más vendidos que tenemos son de Supermarket\_3\_090, en segundo lugar, el Supermarket\_3\_586 y, en tercer lugar, el Supermarket\_3\_252. Esto sería pudiendo incluir todo el rango de fechas, pero podemos ajustarlo por ejemplo para ver los productos más vendidos en el último año… Además de eso, podemos mirar por una tienda en concreto… Y también por ciudad… Al ir variando y filtrando entre diferentes casos de fechas, tiendas y ciudades, podemos apreciar que los productos más vendidos se suelen mantener en el top más vendido, lo que nos puede sugerir lo demandados que son esos productos. También podemos pensar que son de primera necesidad.

En la dashboard de abajo, podemos ver cada producto con rango de precio, y lo que se ha vendido con ese precio.

CLUSTERING

A continuación, vamos a ver los grupos que hemos identificado y segmentado. Estos constan de tres categorías: productos de supermercado, de casa y jardinería y accesorios. Además de estos grupos, cada uno se divide en varios tipos de productos: Ítems Top Sales, Ítems Medium Sales y Ítems Premium. Para la categoría de Supermercado, hemos añadido un tipo adicional el cuál serían los ítems complementarios (bienes que satisfacen conjuntamente un mismo tipo de necesidad y hacen que su demanda fluctúe a la par, puesto que dependen uno de otro, cuando el precio de un producto complementario sube notablemente, es probable que baje la demanda de ambos productos.).

Las gráficas presentadas muestran el análisis de estos tres grupos de categoría. Cada gráfica consta de dos partes principales: una barra de distribución porcentual y un scatter plot (diagrama de dispersión) que muestra la dispersión de los tipos de productos basada en las ventas y el precio. El eje de las X representa el precio, y el eje Y, la cantidad de ventas.

* El grupo de productos de supermercado representan la mayoría de nuestro inventario, siendo la categoría más amplia. La mayoría de las ventas provienen de productos complementarios, seguidos por los ventas medianas y premium.
* El grupo de productos de casa y jardinería representa el segundo lugar de nuestro inventario. Muestra una predominancia en productos premium, seguido por las ventas medias.
* El grupo de accesorios, constituyen una porción pequeña de nuestro inventario, pero tienen un buen desempeño, especialmente los de ventas medianas, lo que sugiere una oportunidad para potenciar esta categoría.

PAGINA 2, CLUSTERS

Los ítems de Top Sales se caracterizan por tener muchas ventas de estas, pero su precio es bajo.

Los ítems Premium son los ítems que tienen un precio elevado, pero tienen menos ventas.

Los ítems de Medium Sales son ítems que se han mantenido a nivel intermedio entre el precio y ventas.

Los ítems Complementarios tienen un precio más elevado que el promedio, pero se han generado menos ventas, igualándose a los ítems Premium.

Visualización: Tipología de los Clusters

Esta visualización detalla la tipología de los ítems en términos de precio medio, cuántos productos componen el tipo y el total de ingresos. Sería el precio medio y total de ingresos de todos los productos de un tipo, no de una categoría en concreto.

Los Ítems Complementarios, con un precio medio de $3.17, son los líderes en ventas, lo cual resalta su importancia en la estrategia de precios y promociones.

Los Ítems de Medium Sales, se quedan en segundo lugar de ventas totales, con un precio medio de 4,17$.

Los Ítems Premium, aunque menos en cantidad con un precio medio de $8.73, contribuyen significativamente a las ventas totales, lo que refleja un mercado dispuesto a invertir en productos de alto valor.

Los Ítems Top Sales, tienen el menor volumen, pero con un precio medio similar a los complementarios, lo que indica una fuerte lealtad de los consumidores hacia ciertos productos específicos. Tener en cuenta que tenemos solamente 72 ítems en este grupo.

Visualización: Conjunto de Productos

Los Ítems Complementarios tienen una facturación máxima diaria de más de $3794,51 y una media de ventas medias diarias que alcanza las 27,29 unidades.

Por otro lado, los ítems Premium han alcanzado una facturación máxima impresionante de $11504,46 con menos unidades vendidas, lo que enfatiza su alto valor y demanda.

Los ítems Top Sales tienen ventas medias diarias que superan las del resto, con un 118,21 por producto, pero con la menor facturación que el resto.

Los ítems de Medium Sales tienen una facturación máxima de $2789,78 , y una media de ventas de 40,62 por día, vendiéndose más que los Ítems Premium y Complementarios.

MLOPS

Ahora, daremos una propuesta para aplicar los modelos de ventas a un caso de uso de abastecimiento de las tiendas. Podríamos migrar los datos de ventas que se generen en un espacio como el de Google Cloud Storage para que almacenen y se traten de una manera más segura y controlada, disminuyendo la pérdida o modificación de estos datos.

Los datos nuevos se almacenan en Google Cloud Storage para su posterior elaboración.

Dataflow se usa para extraer, validar y transformar los datos crudos (ETL). Este proceso limpia los datos y los prepara para el análisis, asegurando que sean consistentes y estén en el formato adecuado.

Usando Google AI Platforms, se desarrolla y entrena el modelo de predicción de ventas.

Los modelos entrenados y los insights generados se integran con Power BI a través de Google App Engine, que actúa como una API para servir las predicciones y facilitar la visualización de los datos. Crea archivos en formato csv con las predicciones, para que sean fácil de visualizar y utilizar en PowerBI.

PowerBI Desktop se utiliza para crear dashboards detallados que permiten visualizar las predicciones de las ventas. Los dashboards desarrollados en PowerBI Desktop se publican a través de PowerBI Service. Esto permite compartir las visualizaciones con los usuarios de la empresa en una plataforma accesible y segura. Para darnos de alta en la suscripción de PowerBI Service, necesitamos el correo electrónico de la empresa para crear la cuenta.

Los usuarios finales de la empresa acceden a los dashboards de Power BI para tomar decisiones informadas. Esto incluye evaluar las predicciones de ventas para cada tienda y optimizar los niveles de inventario. Pueden hacerlo a través de la aplicación móvil o portátil.

Google Cloud Scheduler es un servicio completamente administrado que permite programar la ejecución de tareas recurrentes. Con Google Pub/Sub, además con la ayuda de Cloud Scheduler, podemos programar que nos entre una notificación, en el momento en donde las predicciones estén listas, y enviadas al PowerBI.También podemos crear triggers, que nos alerten cuando el modelo no va por el buen camino en cuanto a métricas y predicciones.

Hemos hecho un presupuesto con los gastos de suscripciones que supondría el uso de los servicios de Google Cloud y PowerBI. Estimamos que un total de 426 euros anuales, pero ese importe puede variar dependiendo de los limites de uso de cada servicio de Google Cloud ya que este ofrece una suscripción especial de pagar el coste especifico por el uso que se le da.

Objetivos:

Evitar el exceso de inventario: Las predicciones ayudan a la empresa a mantener solo el stock necesario, evitando el costo de almacenar productos no vendidos.

Minimizar el riesgo de agotar stock: El modelo pronostica la demanda futura para que la empresa pueda prepararse y reaccionar a tiempo, asegurando que siempre haya suficiente producto disponible para satisfacer las ventas.

Estos objetivos son parecidos a los mencionados anteriormente en la parte de predicciones de venta, pero además de eso, es un proceso de fácil acceso a los resultados visuales preparados para los usuarios de la empresa. Eso facilitaría la toma de decisiones de los usuarios a la hora de gestionar el inventario ya que tendrían a disposición los datos necesarios y actualizados.

Una API, o Interfaz de Programación de Aplicaciones, es un conjunto de reglas y definiciones que permite a diferentes programas de software comunicarse entre sí. Actúa como un intermediario que permite que una aplicación haga uso de los servicios o datos de otra.

CONCLUSION

Respondiendo al reciente pedido recibido por email para disminuir el stock en los almacenes regionales, aplicar las conclusiones obtenidas a partir del modelo de datos puede optimizar considerablemente el ciclo completo, abarcando la adquisición de artículos de venta al detalle hasta la entrega de estos. Esto nos llevaría a una gestión más eficaz de los activos, que es esencial para nuestra empresa dado que es imprescindible recortar gastos operativos con el fin de fortalecer nuestra posición en el mercado.

POSIBLES PREGUNTAS A JOSEPH:

En el caso de que queramos reducir costes de suscripción, podemos excluir los servicios de PowerBI, y utilizar otro servicio que ofrece Google, que es el llamado Data Studio. Sería de la misma manera, generando un archivo csv con los resultados de las predicciones, y los metemos en Data Studio. Tiene el mismo fin que PowerBI (creación de elementos visuales, tablas, gráficos interactivos…). Una vez esté listo, el informe, se puede compartir con otros usuarios de la empresa. Data Studio permite controlar el acceso de lectura y edición. Los informes pueden ser incrustados en sitios web internos o enviados directamente a los stakeholders mediante programación de entregas periódicas por correo electrónico.

**. Prepara tu modelo**

Antes de pensar en la infraestructura, asegúrate de que tu modelo esté bien entrenado, validado y listo para ser empaquetado. Esto incluye haberlo probado localmente para asegurarte de que funciona como se espera.

**2. Containeriza tu modelo con Docker**

El primer paso para desplegar tu modelo es empaquetarlo en un contenedor Docker. Esto te permite correr tu modelo de manera consistente en cualquier entorno.

**3. Automatiza el proceso con Apache Airflow**

Airflow puede ser utilizado para automatizar el despliegue de tu modelo, incluyendo la construcción de imágenes Docker, el despliegue en un servidor o un entorno de nube, y la ejecución de pruebas de integración.

**. Ejecuta y monitoriza el DAG**

Asegúrate de que Airflow está ejecutando el scheduler y el webserver. Puedes usar la interfaz de usuario web de Airflow para disparar y monitorizar el DAG que maneja el despliegue de tu modelo.

**4. Consideraciones adicionales**

* **Escalabilidad**: Considera utilizar servicios de orquestación como Kubernetes si esperas necesitar escalabilidad.
* **Seguridad**: Asegúrate de manejar correctamente las credenciales y datos sensibles.
* **Monitoreo y Logging**: Configura adecuadamente el monitoreo y logging para tu modelo en producción para poder reaccionar a posibles fallos o degradaciones del desempeño.