

DATA

ANALYTICS

### Alumno:

## Montero Guzman Alvaro Jair

### Profesor:

## Luciano Julián Gómez Olivera

Comisión: **40000**

**Tabla de contenido**

1. [Introducción 3](#_bookmark0)
2. [Descripción de la temática de los datos 4](#_bookmark1)
3. [Alcance 5](#_bookmark2)
4. [Hipótesis 6](#_bookmark3)

# Introducción



En el dinámico panorama empresarial actual, la capacidad de tomar decisiones informadas y estratégicas es esencial para el éxito de cualquier organización. La facturación, como uno de los indicadores clave de rendimiento, desempeña un papel fundamental en la evaluación de la salud financiera y la dirección futura de una empresa. Con el advenimiento de herramientas tecnológicas avanzadas, como Power BI, las empresas tienen la oportunidad de transformar datos crudos en información valiosa y visualmente impactante.

Este documento se centra en el análisis de facturación empresarial utilizando Power BI como herramienta principal. Exploraremos cómo Power BI, una plataforma de análisis de datos de Microsoft, puede permitir a las organizaciones desglosar, comprender y aprovechar los datos de facturación de manera efectiva. Desde la creación de visualizaciones perspicaces hasta la identificación de tendencias y patrones ocultos, el análisis de facturación en Power BI ofrece a las empresas la capacidad de tomar decisiones más fundamentadas y estratégicas.

A lo largo de este documento, examinaremos las etapas clave para diseñar y desarrollar un tablero de análisis de facturación efectivo en Power BI. Además, exploraremos cómo la interactividad, la visualización de datos y las capacidades de generación de informes en tiempo real de Power BI pueden empoderar a los equipos de gestión y a los stakeholders para comprender mejor el rendimiento financiero de la empresa y formular estrategias basadas en evidencia.

En última instancia, este documento busca resaltar la importancia del análisis de facturación en el contexto empresarial actual y cómo Power BI se erige como una herramienta valiosa para transformar datos de facturación en conocimientos accionables.

# Descripción de la temática de los datos

El Dataset aborda la temática de la facturación de la empresa, proporcionando una visión integral y detallada de los aspectos financieros clave de una organización. Este conjunto de datos recopila información sobre las transacciones de ventas y los ingresos generados por la empresa en un período de tiempo específico, lo que permite analizar y comprender la dinámica económica de la organización.

El Dataset contiene una variedad de variables relevantes que capturan diversos aspectos de la facturación, como la fecha de cada transacción, los productos o servicios vendidos, los precios unitarios, las cantidades vendidas, los descuentos aplicados (si los hubiera) y los clientes involucrados en cada transacción. Además, el conjunto de datos podría incluir información contextual adicional, como la categoría de productos, la ubicación de la venta o el canal de distribución utilizado.

Al analizar este Dataset, es posible identificar tendencias y patrones en los ingresos a lo largo del tiempo, comprender qué productos son los más populares entre los clientes, evaluar la eficacia de las estrategias de fijación de precios y descuentos, y detectar posibles oportunidades de crecimiento o áreas que requieran mejoras en términos de ventas y facturación.

El objetivo principal de este Dataset es permitir a los analistas financieros, gerentes y stakeholders de la empresa examinar el rendimiento financiero desde una perspectiva centrada en la facturación. Con la aplicación de técnicas de análisis de datos y herramientas como Power BI, se puede transformar este conjunto de datos en información visualmente atractiva y fácilmente comprensible, lo que ayuda a tomar decisiones estratégicas informadas y a impulsar el éxito empresarial.

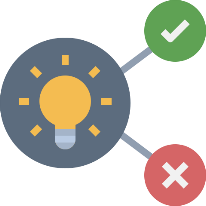
# Alcance

El alcance del proyecto de Power BI centrado en el análisis de la facturación de una empresa puede abarcar diversas etapas y objetivos clave para obtener una comprensión completa y valiosa de los aspectos financieros. A continuación, se describe un alcance típico para un proyecto de este tipo:

* 1. **Definición de objetivos**: establecer objetivos claros para el proyecto, como identificar patrones de ventas, analizar la rentabilidad de los productos y comprender las tendencias de facturación a lo largo del tiempo.
  2. **Recopilación de datos**: obtener y preparar los datos de facturación de la empresa, que incluirán información sobre transacciones, productos, precios, clientes y fechas.
  3. **Limpieza y transformación de datos**: realizar limpieza de datos para manejar valores nulos, datos duplicados o incoherentes. Transformar los datos según sea necesario para crear una estructura adecuada para el análisis, como la creación de tablas de hechos y dimensiones.
  4. **Creación de un modelo de datos**: diseñar un modelo de datos en Power BI que refleje las relaciones entre las tablas y permita un análisis coherente y eficaz.
  5. **Desarrollo de visualizaciones**: crear visualizaciones interactivas, como gráficos de barras, líneas, tortas y mapas, para mostrar la facturación por período, productos, categorías, etc. Desarrollar tableros interactivos que permitan a los usuarios explorar los datos y obtener información relevante.
  6. **Análisis y descubrimiento**: identificar tendencias en los ingresos a lo largo del tiempo y realizar análisis comparativos entre productos, categorías o segmentos de clientes. Evaluar la influencia de los descuentos, promociones u otros factores en la facturación.
  7. **Implementación de métricas clave**: calcular y mostrar métricas financieras clave, como ingresos totales, promedio de ventas por cliente, margen de beneficio, etc.
  8. **Generación de informes y Dashboards**: diseñar informes ejecutivos y Dashboards que presenten los hallazgos y permitan a los usuarios acceder rápidamente a los datos relevantes.
  9. **Capacitación y documentación**: proporcionar capacitación a los usuarios finales sobre cómo interactuar con el tablero y realizar análisis personalizados. Documentar el proceso de preparación de datos, modelado y visualización para futuras referencias.
  10. **Entrega y seguimiento**: presentar el proyecto a los stakeholders y usuarios finales, recopilando comentarios y realizando ajustes si es necesario.
  11. **Mantenimiento continuo**: actualizar regularmente el tablero con nuevos datos para mantener la relevancia y precisión de las visualizaciones.

Este alcance proporciona una estructura general para abordar un proyecto de Power BI relacionado con la facturación de la empresa. Sin embargo, el alcance específico puede variar según las necesidades y metas de la organización.

# Hipótesis



En el análisis de la facturación de la empresa, pueden plantearse diversas hipótesis para explorar y validar. Estas hipótesis son suposiciones preliminares que se pueden someter a pruebas utilizando los datos disponibles. A continuación, se presentan algunos ejemplos de hipótesis que podrían ser relevantes en este contexto:

**Temporada**: la facturación de la empresa varía según las estaciones del año. Por ejemplo, se podría hipotetizar que la empresa experimenta un aumento significativo en las ventas durante las temporadas festivas.

**Producto estrella**: existe un producto o conjunto de productos que contribuyen de manera desproporcionada a la facturación total de la empresa. Se podría investigar si algunos productos generan la mayoría de los ingresos.

**Día de la semana**: la facturación puede ser diferente según el día de la semana. Podría plantearse la hipótesis de que los fines de semana generan un mayor volumen de ventas en comparación con los días laborables.

**Segmentación de clientes**: los diferentes segmentos de clientes pueden tener un impacto significativo en la facturación. Se podría investigar si ciertos grupos de clientes contribuyen más que otros a los ingresos de la empresa.

**Eficacia de descuentos**: se podría hipotetizar que la implementación de descuentos en productos específicos puede aumentar la facturación total, al atraer a más clientes o estimular compras adicionales.

**Tendencias a lo largo del tiempo**: podría plantearse la hipótesis de que la facturación ha estado creciendo o disminuyendo de manera constante a lo largo de un período de tiempo específico.

**Relación entre productos**: se podría investigar si existe una correlación entre la venta de ciertos productos. Por ejemplo, se podría hipotetizar que la venta de un producto específico impulsa la venta de otro producto relacionado.

**Canales de venta**: se podría hipotetizar que ciertos canales de venta (por ejemplo, tiendas físicas vs. comercio electrónico) tienen un impacto diferente en la facturación total.

**Respuesta a eventos externos**: podría plantearse la hipótesis de que eventos externos, como cambios en la economía o tendencias del mercado, tienen un impacto en la facturación de la empresa.

**Influencia geográfica**: se podría investigar si la ubicación geográfica tiene algún efecto en los patrones de facturación, como si las ventas varían según la región o el país.

Estas son solo algunas ideas de hipótesis que podrían ser relevantes para un análisis de facturación empresarial. Es importante recordar que las hipótesis deben ser respaldadas por datos y sometidas a pruebas estadísticas para determinar su validez.

**Comparación entre ciudades grandes y pequeñas**:

Las ciudades con mayor población tienen una mayor facturación en comparación con las ciudades con menor población. Se podría investigar si la densidad poblacional correlaciona positivamente con la facturación.

**Tendencias geográficas a lo largo del tiempo**:

La facturación en ciertas ciudades ha crecido o disminuido de manera constante a lo largo de un período de tiempo específico. Por ejemplo, se podría investigar si hay un crecimiento constante en la facturación en ciudades emergentes como Denver.

**Influencia de eventos locales**:

La facturación se ve afectada por eventos locales específicos en ciertas ciudades. Por ejemplo, se podría analizar si las ventas aumentan en Philadelphia durante eventos deportivos importantes o en Minneapolis durante festivales locales.

**Segmentación geográfica de clientes**:

Los diferentes segmentos de clientes en diversas ciudades tienen un impacto significativo en la facturación. Se podría investigar si ciertos grupos de clientes en ciudades específicas contribuyen más a los ingresos de la empresa.

**Relación entre productos y ciudades**:

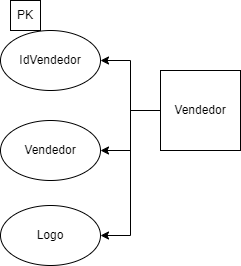
La venta de ciertos productos es más fuerte en algunas ciudades en comparación con otras. Se podría investigar si ciertos productos son más populares en ciudades específicas, por ejemplo, productos tecnológicos en San Francisco o productos de moda en New York.

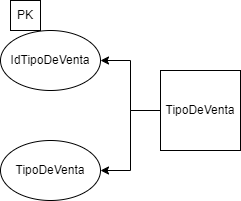
# Herramientas tecnológicas utilizadas

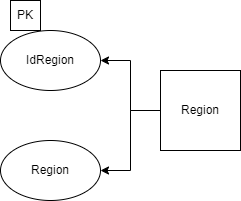
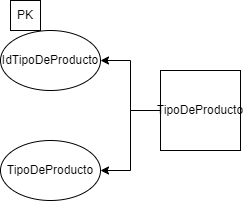
En el contexto de este proyecto en particular se seleccionaron y emplearon las siguientes aplicaciones y herramientas tecnológicas con el propósito de llevar a cabo las tareas y análisis requeridos:

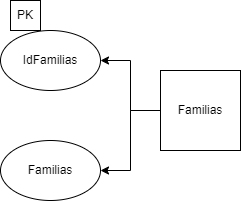
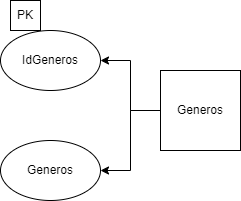
* Microsoft Excel para la lectura de los archivos CSV que conforman el Datashet
* Draw IO para la generación de los diagramas E-R
* Word para la creación del documento
* GitHub para guardar la información en el repositorio

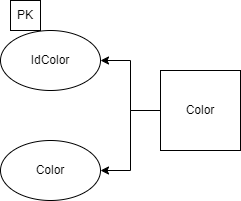
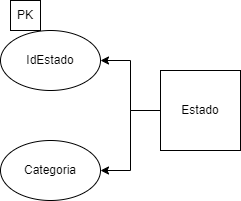
# Diagrama entidad-relación

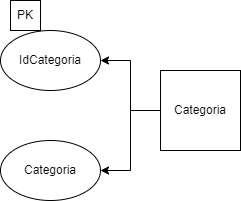
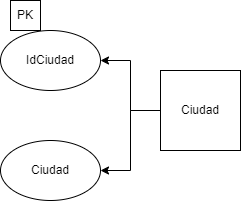


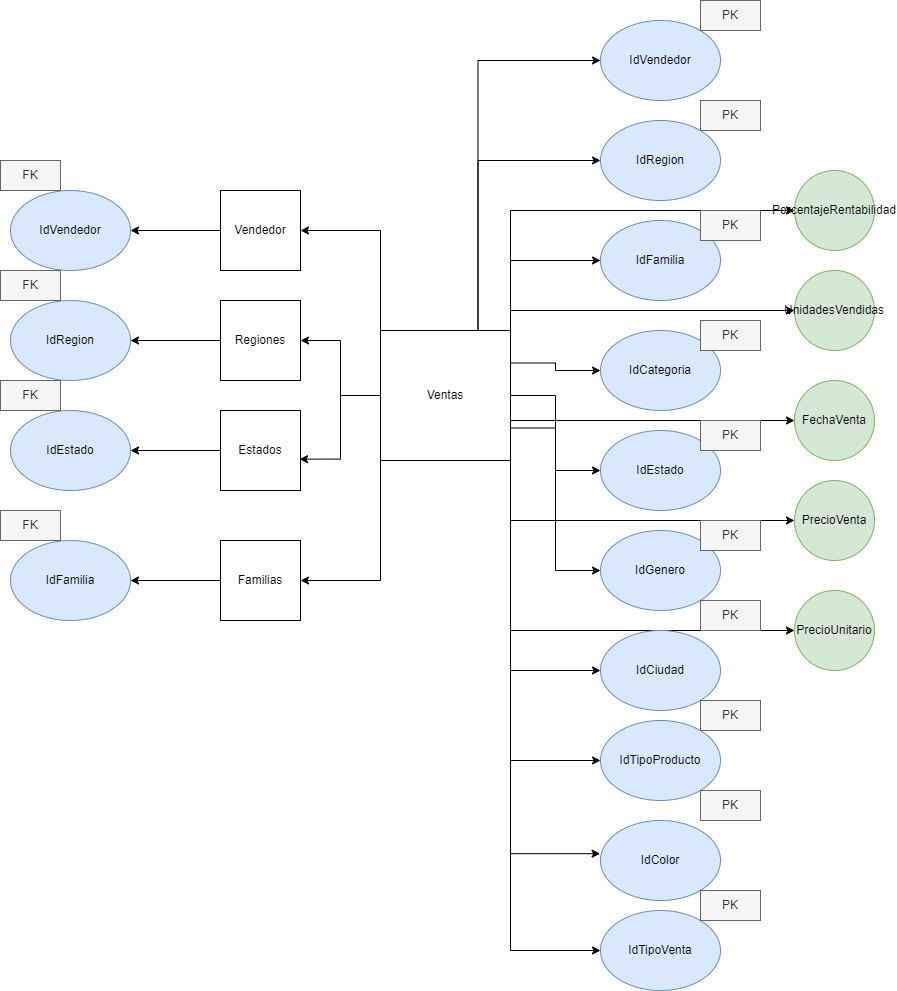












# Listado de campos por tablas

# En el siguiente apartado, procedemos a proporcionar un exhaustivo desglose de la estructura de cada tabla, incluyendo la enumeración de sus columnas, la especificación detallada de los tipos de datos asignados a cada una de ellas y, además, la claraidentificaciond e los tipos de clave que se han implementado en dichas tablas,Este análisis minucioso de la disposición y características de los datos permitirá una comprensión mas profunda y completa de la base de datos en cuestión, brindando una visión integral de su diseño y funcionamiento

# 

# Interfaz de usuario gráfica, Texto, Aplicación Descripción generada automáticamente

Tabla

Descripción generada automáticamente

Tabla

Descripción generada automáticamente

Tabla

Descripción generada automáticamente

Tabla

Descripción generada automáticamente

Tabla, Excel

Descripción generada automáticamente

Tabla

Descripción generada automáticamente

Tabla, Excel

Descripción generada automáticamente

Tabla, Excel

Descripción generada automáticamente

Tabla, Excel

Descripción generada automáticamente

# Generación de fondos para los tableros

En el desarrollo de los informes, se utiliza Power BI como la herramienta principal para crear visualizaciones y analizar los datos de manera eficiente. Esta elección se fundamenta en la capacidad de Power BI para integrar y procesar grandes volúmenes de datos, permitiendo la creación de informes interactivos que ofrecen un análisis profundo y detallado.

Después de realizar el proceso de análisis y diseño en Power BI, se presentan a continuación dos informes que han sido cuidadosamente elaborados. Estos informes se han creado con un enfoque en la precisión y la claridad, asegurando que se alineen con los objetivos específicos del proyecto, facilitando así la toma de decisiones basada en datos.





# Importacion de tablas

**Tabla CATEGORIAS**

En la tabla **CATEGORIAS**, se realizó una serie de modificaciones importantes. Se ajustó el nombre de la columna **CATEGORIA**, agregando el acento correspondiente y renombrándola como **CATEGORÍA**. Además, se llevaron a cabo cambios en los valores de esta columna, reemplazando **CORRER** por **RUNNING**, **ENTRENAR** por **TRAINING**, y **FUTBOL** por **SOCCER**, asegurando una correcta categorización de los productos.

**Tabla COLORES**

Para la tabla **COLORES**, se realizó un cambio en la columna **COLOR**, donde el valor **ROSA** fue actualizado a **FUCSIA**. Este ajuste se hizo para reflejar una mejor diferenciación cromática en la base de datos.

**Tabla FAMILIAS**

En la tabla **FAMILIAS**, se procedió a reemplazar el término **ROPA** por **INDUMENTARIA** en la columna **FAMILIA**. Este cambio tuvo como objetivo lograr una descripción más precisa y profesional de las categorías de productos.

**Tabla GENEROS**

En la tabla **GENEROS**, se realizaron varios cambios en la columna **GENERO**. Los valores **BEBE (NENE)** y **BEBE (NENA)** fueron modificados a **NIÑO** y **NIÑA**, respectivamente. Además, se corrigió el nombre de la columna, añadiendo el acento necesario y renombrándola como **GÉNERO**.

**Tabla REGIONES**

Dentro de la tabla **REGIONES**, se creó una columna condicional denominada **ZONA**. Dependiendo de los valores en la columna **REGION**, se realizaron las siguientes conversiones: **NORTHEAST** se transformó en **NORESTE**, **SOUTH** en **SUR**, **WEST** en **OESTE**, **MIDWEST** en **MEDIO OESTE**, y **SOUTHEAST** en **SURESTE**. En caso de que el valor no coincidiera con ninguna de estas opciones, se completó con la palabra **DESCONOCIDA**. Posteriormente, se cambió el tipo de dato de la columna **ZONA** a **TEXTO** y se eliminó la columna **REGION** para evitar duplicidades.

**Tabla TIPO DE PRODUCTO**

En la tabla **TIPO DE PRODUCTO**, se renombró la columna **TIPOPRODUCTO** a **TIPO PRODUCTO** para mejorar la claridad. Además, se corrigió la palabra **FUTBOL** en la columna correspondiente, agregando el acento para que se lea **FÚTBOL**.

**Tabla TIPO DE VENTA**

Para la tabla **TIPO DE VENTA**, se realizó un ajuste en el nombre de la columna **TIPOVENTA**, cambiándolo por **TIPO VENTA**. Asimismo, el valor **IN-STORE** fue traducido a **TIENDA** en esta columna, reflejando mejor el contexto local.

**Tabla VENTAS**

En la tabla **VENTAS**, se generó una nueva columna llamada **TOTAL**, que fue calculada multiplicando los valores de **PRECIO UNITARIO** por **UNIDADES VENDIDAS**. Luego, se ajustó el tipo de dato de la columna **TOTAL** a **NUMERO ENTERO** y se eliminaron las columnas originales de **PRECIO UNITARIO** y **UNIDADES VENDIDAS**.

Además, se realizó una verificación en la columna **PORCENTAJERENTABILIDAD** para asegurar que los valores se interpretaran correctamente. Si se detectaron inconsistencias, se creó una nueva columna personalizada con la fórmula necesaria y se renombró como **PORCENTAJE RENTABILIDAD**. Posteriormente, se generó una columna llamada **RENTABILIDAD**, donde se multiplicaron los valores de **TOTAL** por **PORCENTAJE RENTABILIDAD**. Finalmente, se cambió el tipo de dato de **RENTABILIDAD** a **NUMERO DECIMAL** y se eliminó la columna **PORCENTAJE RENTABILIDAD** para simplificar la estructura de datos.

Una vez completadas todas estas modificaciones, se aplicaron los cambios y se guardó el archivo de Power BI con el nombre **PROYECTO FINAL** en la carpeta **DATA ANALYTICS - PROYECTO FINAL**.

El diagrama queda de la siguiente manera:

# Aplicación de los fondos creados en Pixlr en PoweBi

Después de diseñar meticulosamente los gráficos en Excel, el siguiente paso crucial es su implementación en el archivo de Power BI. Esta transición es fundamental para crear un informe visualmente impactante y funcional. Los gráficos, creados con un enfoque en la claridad y precisión, se convierten en componentes visuales clave para mejorar la presentación de los datos en nuestro proyecto de Power BI.

La integración de estos gráficos no solo optimiza la representación visual de la información, sino que también añade un nivel de personalización que se alinea perfectamente con los objetivos y la narrativa de cada página en el archivo de Power BI. De este modo, los gráficos actúan como un elemento unificador en el diseño del informe, haciendo que la información sea más accesible y comprensible para el público destinatario.

# Alcane, hipótesis, glosario y herramientas tecnologicas

La segunda página del informe es un componente fundamental del contenido, ya que ofrece una visión detallada del análisis realizado en el proyecto. Esta sección no solo contextualiza la investigación, sino que también establece una base clara para entender los objetivos y resultados previstos. Aquí se presenta la hipótesis principal que se está evaluando, lo que proporciona una visión clara del enfoque y propósito del proyecto.

Además de la explicación del análisis y la hipótesis, esta página incluye un glosario esencial, donde se definen los términos técnicos o específicos del proyecto, facilitando la comprensión del lector y evitando malentendidos. También se destaca un listado exhaustivo de todas las herramientas tecnológicas empleadas en el proyecto, lo que aporta transparencia sobre la infraestructura utilizada y puede servir como guía para quienes deseen replicar o explorar más a fondo el análisis.

Para asegurar una navegación sencilla y eficiente a lo largo del informe, se ha añadido una imagen específica en esta página, acompañada de un hipervínculo. Al hacer clic en esta imagen, los usuarios pueden regresar rápidamente a la primera página del informe, llamada PORTADA, lo que mejora la navegación y permite un acceso rápido a los contenidos clave. Este elemento de diseño es crucial para la comodidad y accesibilidad del informe, garantizando que los lectores puedan moverse de manera fluida entre las distintas secciones.

La botonera quedo de la siguiente manera:

Imagen que contiene Forma

Descripción generada automáticamente

Texto

Descripción generada automáticamente

# Columnas calculadad / medidad /visualizaciones

En Power BI, las MEDIDAS son elementos cruciales para realizar cálculos específicos o agregar datos según ciertas condiciones. Estas medidas se construyen mediante el lenguaje de fórmulas de Power BI, conocido como DAX (Data Analysis Expressions).

Las medidas permiten llevar a cabo análisis avanzados y extraer información específica que no está directamente disponible en los conjuntos de datos originales. A continuación, se presentan algunas situaciones comunes en las que las medidas resultan ser particularmente útiles:

* **Cálculos personalizados**: Las medidas permiten la creación de cálculos personalizados mediante fórmulas DAX, adaptándose a necesidades específicas.
* **Indicadores clave de rendimiento (KPI)**: Son esenciales para definir y calcular KPIs, que proporcionan un resumen del rendimiento organizacional en áreas específicas.
* **Análisis de tendencias y variaciones**: Las medidas ayudan a calcular tasas de crecimiento, variaciones porcentuales, y otras métricas para analizar tendencias a lo largo del tiempo.
* **Filtrado dinámico**: Permiten aplicar filtros dinámicos en los informes, ajustando los cálculos según las selecciones realizadas por el usuario.
* **Segmentación y agrupación**: Facilitan la segmentación o agrupación de datos según ciertos criterios, ofreciendo un análisis más detallado y personalizado.
* **Cálculos condicionales**: Se pueden crear medidas para realizar cálculos basados en condiciones específicas, proporcionando información relevante según diferentes escenarios.
* **Comparación con objetivos o benchmarks**: Las medidas permiten comparar el rendimiento real con objetivos predefinidos o benchmarks, evaluando el desempeño frente a metas establecidas.

En resumen, las medidas en Power BI son herramientas poderosas que permiten realizar análisis avanzados y personalizados, ofreciendo información clave para la toma de decisiones empresariales. Estas medidas pueden integrarse en tablas, gráficos y otros elementos visuales en informes y paneles.

Para este proyecto, se han definido las siguientes medidas:

* **Total Facturado**: Calcula el total facturado por la empresa sumando los valores de la columna TOTAL en la tabla VENTAS.

Se genero la tabla medidas utilizando el formato de coma para la separación del numero

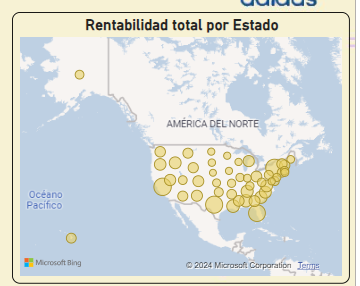
Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

* Generación de mapa, el cual representa la rentailidad por estado:



* Generación de grafico de columnas, el cual represena la rentabilidad por tipo de venta:

Gráfico, Gráfico de barras

Descripción generada automáticamente

* Se genero un tipo de grafico scrollerpara complementar las venas y la rentabilidad por estado:



El motivo de este grafico es complementar al mapa de rentabilidad por estado y generar en el usuario una mayor visibiidad de los datos.