

Curso de Data Analytics – CODERHOUSE

“Análisis de Rentabilidad y Eficiencia en Northwind”

Lic. en Economía Juan M. Marquina

21 de septiembre de 2023

Índice

1	Introducción	3
2	Objetivo	3
3	Alcance	3
4	Herramientas tecnológicas implementadas	3
5	Base de datos	4
6	Diagrama entidad-relación:.....	5
7	Listado de tablas:	6
8	Listado de columnas de tablas:	7
9	Transformación de los datos	8
10	Medidas.....	8
11	Segmentadores	8
12	Visualización de los datos.....	9
12.1	Página “Portada”	9
12.2	Página “Overview”	10
12.3	Página “Product and Category”	10
12.4	Página “Employees”	11
12.5	Página “Employees”	11
12.6	Página “Shipping”	12
13	Conclusión	12
14	Futuras líneas	12

1 Introducción

En este trabajo, se ha optado por utilizar la base de datos Northwind, que es un ejemplo desarrollado por Microsoft. Esta base de datos contiene información sobre las ventas de una empresa ficticia llamada "Northwind Traders", la cual se dedica a la importación y exportación de alimentos especiales de todo el mundo.

2 Objetivo

El objetivo principal del análisis de la base de datos Northwind es comprender su estructura, relaciones y datos para mejorar la gestión y toma de decisiones relacionadas con el proceso de venta y distribución de productos. En este análisis también se buscará responder preguntas clave sobre el comportamiento de los clientes. Se determinará la ubicación geográfica donde se concentra la mayoría de los clientes, qué productos compran con mayor frecuencia y en qué cantidades. Esto proporcionará una visión clara de los mercados más activos y permitirá identificar oportunidades de crecimiento y personalizar estrategias de marketing para satisfacer mejor las necesidades de los clientes en diferentes regiones. En este análisis también se evaluará el rendimiento de los empleados para identificar quiénes están logrando un mayor volumen de ventas. Por último, se examinará el desempeño de las empresas transportistas, evaluando su puntualidad en la entrega de productos y los posibles retrasos.

3 Alcance

El tablero de control está diseñado para ser utilizado por líderes responsables de las decisiones estratégicas en la empresa como directores de ventas, gerentes de marketing y ejecutivos de alto nivel. Estos profesionales emplean este análisis para evaluar el desempeño del negocio, descubrir posibles áreas de crecimiento y tomar decisiones tanto tácticas como estratégicas.

4 Herramientas tecnológicas implementadas

En el presente trabajo se utilizaron las siguientes herramientas:

- Editor de Power Query para la limpieza de los datos.
- Canva para la creación del diagrama entidad-relación.
- Power Point para el diseño del tablero de control.
- Power Bi Desktop para la creación del tablero de control.

5 Base de datos

Una vez que la base de datos Northwind fue cargada en Power BI, se llevaron a cabo las siguientes modificaciones en los mismos:

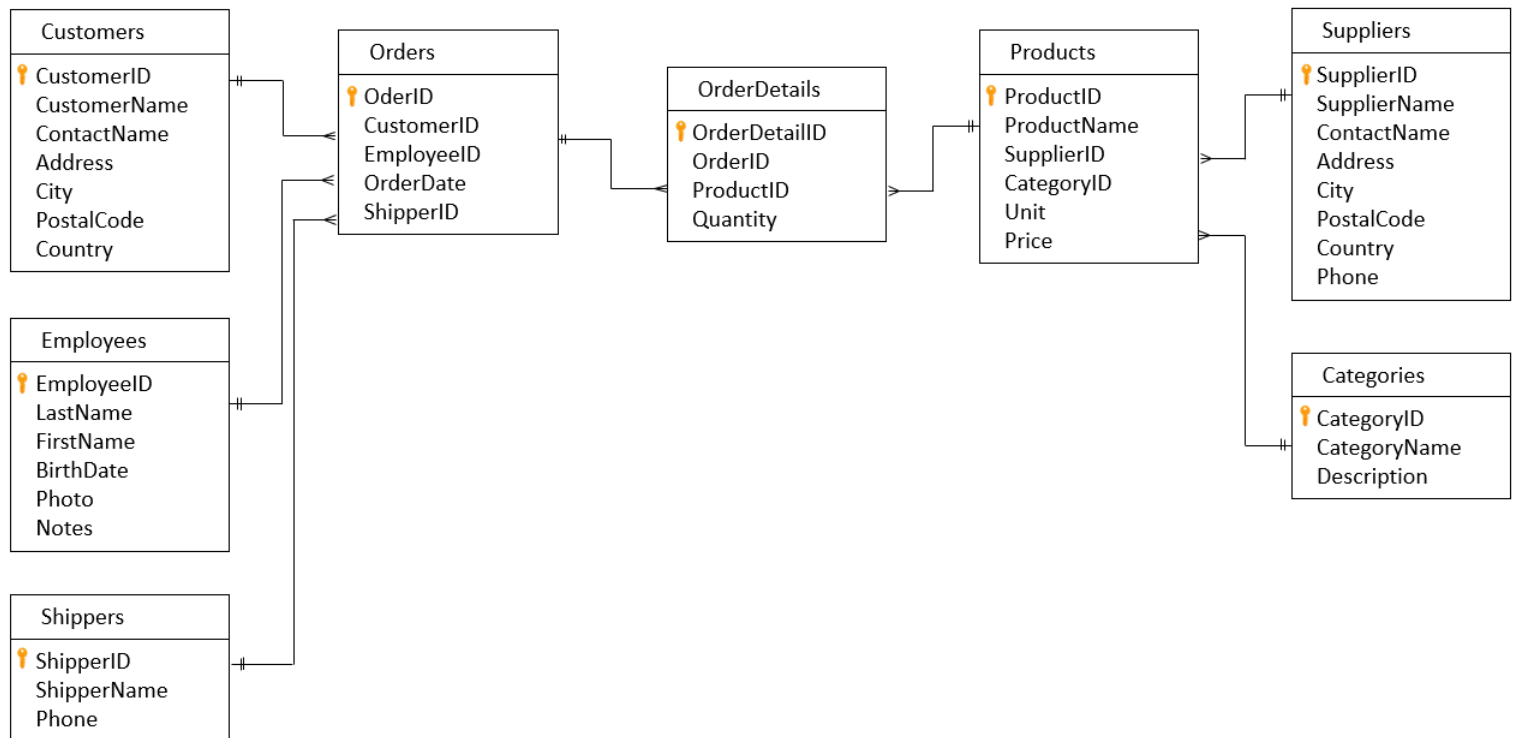
En un principio se examinaron las tablas en busca de valores nulos o errores. Solo se identificaron valores nulos relevantes en la tabla "Orders".

En la tabla "Orders", en la columna "ShippedDate", se encontraron celdas con valores nulos debido a que representaban datos para los cuales los productos aún no habían sido entregados. Por esta razón se procedió a eliminar estas filas, ya que su presencia posteriormente afectaba el cálculo de los días promedio necesarios para la entrega de los productos.

Posteriormente se eliminaron ciertas columnas en todas las tablas que habían sido generadas automáticamente y cuyas celdas contenían la etiqueta "value" o "table", según el caso. Estas columnas carecían de relevancia y no contribuían de manera significativa al análisis.

Las relaciones entre las tablas fueron revisadas y se encontraron en correcto estado, manteniendo una relación de uno a muchos.

6 Diagrama entidad-relación:



7 Listado de tablas:

- Order Details: Este conjunto de datos incluye el número de pedido, un identificador único del producto, el precio y la cantidad de productos solicitados, junto con el descuento aplicado a la orden.

FK: OrderID

FK: ProductID

- Orders: Incluye el número de identificación del pedido, el número de identificación del cliente, el número de identificación del empleado, la fecha del pedido, la fecha requerida y la fecha de envío.

PK: OrderID

FK: CustomerID

FK: EmployeeID

FK: ShipVia

- Shippers: Contiene número de transportista, con el nombre de la empresa y número de teléfono

PK: ShipperID

- Customers: Contiene información sobre los clientes de la empresa, como nombres, direcciones, números de teléfono y detalles de contacto.

PK: CustomerID

- Employees: Almacena datos sobre los empleados de la empresa, incluyendo nombres, direcciones, información de contacto, fechas de contratación y puestos de trabajo.

PK: EmployeeID

- Products: Guarda detalles sobre los productos que Northwind vende, incluyendo nombres, precios, unidades en stock y otras características.

PK: ProductID

FK: SupplierID

FK: CategoryID

- Suppliers: Almacena datos relacionados con los proveedores de productos, como nombres, direcciones y detalles de contacto.

PK: SupplierID

- Categories: Contiene información sobre las categorías de productos ofrecidos por Northwind, incluyendo nombres y descripciones.

PK: CategoryID

8 Listado de columnas de tablas:

Orders	
OrderID	int
CustomerID	varchar
EmployeeID	int
OrderDate	datetime
RequiredDate	datetime
ShippedDate	datetime
ShipVia	int
Freight	decimal
ShipName	varchar
ShipAddress	varchar
ShipCity	varchar
ShipRegion	varchar
ShipPostalCode	varchar
ShipCountry	varchar

Order Details	
OrderID	int
ProductID	int
UnitPrice	decimal
Quantity	smallint
Discount	double

Shippers	
ShipperID	int
CompanyName	varchar
Phone	varchar

Customers	
CustomerID	varchar
CompanyName	varchar
ContactName	varchar
ContactTitle	varchar
Address	varchar
City	varchar
Region	varchar
PostalCode	varchar
Country	varchar
Phone	varchar
Fax	varchar

Employees	
EmployeeID	int
LastName	varchar
FirstName	varchar
Title	varchar
TitleOfCourtesy	varchar
BirthDate	datetime
HireDate	datetime
Address	varchar
City	varchar
Region	varchar
PostalCode	varchar
Country	varchar
HomePhone	varchar
Extension	varchar
Photo	longblob
Notes	mediumtext
ReportsTo	int
PhotoPath	varchar

Products	
ProductID	int
ProductName	varchar
SupplierID	int
CategoryID	int
QuantityPerUnit	varchar
UnitPrice	decimal
UnitsInStock	smallint
UnitsOnOrder	smallint
ReorderLevel	smallint
Discontinued	tinyint

Categories	
CategoryID	int
CategoryName	varchar
Description	mediumtext
Picture	longblob

Suppliers	
SupplierID	int
CompanyName	varchar
ContactName	varchar
ContactTitle	varchar
Address	varchar
City	varchar
Region	varchar
PostalCode	varchar
Country	varchar
Phone	varchar
Fax	varchar
HomePage	mediumtext

9 Transformación de los datos

Las siguientes transformaciones se realizaron en la tabla Order Details.

1. Se realizó una modificación en el tipo de dato de la columna "Discount" para expresarla como porcentaje.
2. A continuación, se generó una columna adicional llamada "Descuento en dólares", la cual se deriva de la multiplicación de la columna "Ingreso Bruto" por el contenido de la columna "Discount".

En la tabla de Orders, se añadió la columna "Días para el Envío", que representa la diferencia entre la fecha de entrega y la fecha de pedido. Además, se modificó el formato para que sea un número entero. Además, se creó la columna "Entrega a Tiempo" utilizando la función IF. Esta columna proporciona una indicación de si el pedido fue entregado dentro del plazo, basándose en la comparación entre la fecha requerida del pedido y la fecha de entrega.

10 Medidas

1. Se introdujo la medida Orders = DISTINCTCOUNT('Order Details'[OrderID]) con el propósito de identificar y eliminar los duplicados en los ID de los pedidos.
2. Se creó de una nueva medida denominada "Ingreso Bruto" que excluye los descuentos en su cálculo. Esta medida se obtiene al multiplicar el precio por la cantidad de cada producto.
3. Se elaboró la medida "Ingreso Neto", en esta ocasión teniendo en cuenta los descuentos ya aplicados. Para su cálculo, se utilizó la siguiente formula: $\text{Net Revenue} = \text{SUMX}('Order Details', ('Order Details'[UnitPrice] * (1 - 'Order Details'[Discount \%])) * 'Order Details'[Quantity])$
4. Luego, se estableció una nueva medida denominada "Net Revenue by Order", la cual realiza la suma total de los ingresos netos y luego los divide por la cantidad de órdenes presentes en los registros.
5. Otra medida fue el calculo de la cantidad de empleados: $\text{Number of Employees} = \text{COUNT}(\text{Employees}[\text{EmployeeID}])$
6. La última función fue la más compleja, sirve para calcular que pedidos fueron entregados a tiempo y cuales no. Se utilizo la siguiente formula con la función IF: $\text{On-Time Delivery} = \text{IF}(\text{MAX}(\text{Orders}[\text{RequiredDate}]) - \text{MAX}(\text{Orders}[\text{ShippedDate}]) \geq 0, \text{"Within Time Limit"}, \text{"Outside of Time Limit"})$

11 Segmentadores

Existen cuatro segmentadores disponibles en la visualización. El primero ofrece opciones de filtrado por categoría y producto. El segundo se enfoca en filtrar por país y ciudad. El tercero está relacionado con las fechas y presenta un estilo de filtrado "entre". Por último, el cuarto es un botón con la función de borrar los filtros que previamente fueron seleccionados Los dos

primeros se presentan en formato de menú desplegable para facilitar la elección, mientras que el segmentador de fechas se configuró con el estilo “entre” para definir un rango específico.

12 Visualización de los datos

12.1 Página “Portada”

La página dispone de cinco botones que permiten acceder a la página deseada.



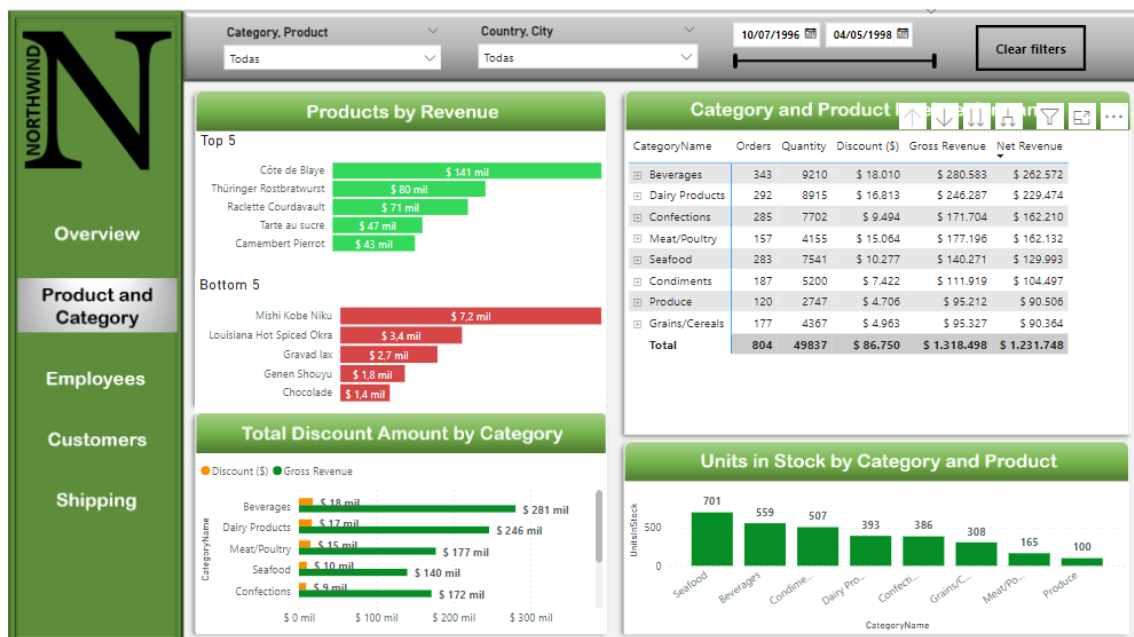
12.2 Página “Overview”

Esta página está diseñada para ofrecer una rápida visualización de los datos más relevantes que se abordarán en las páginas siguientes, tales como el ingreso neto, la cantidad de productos y los días promedio requeridos para su transporte.



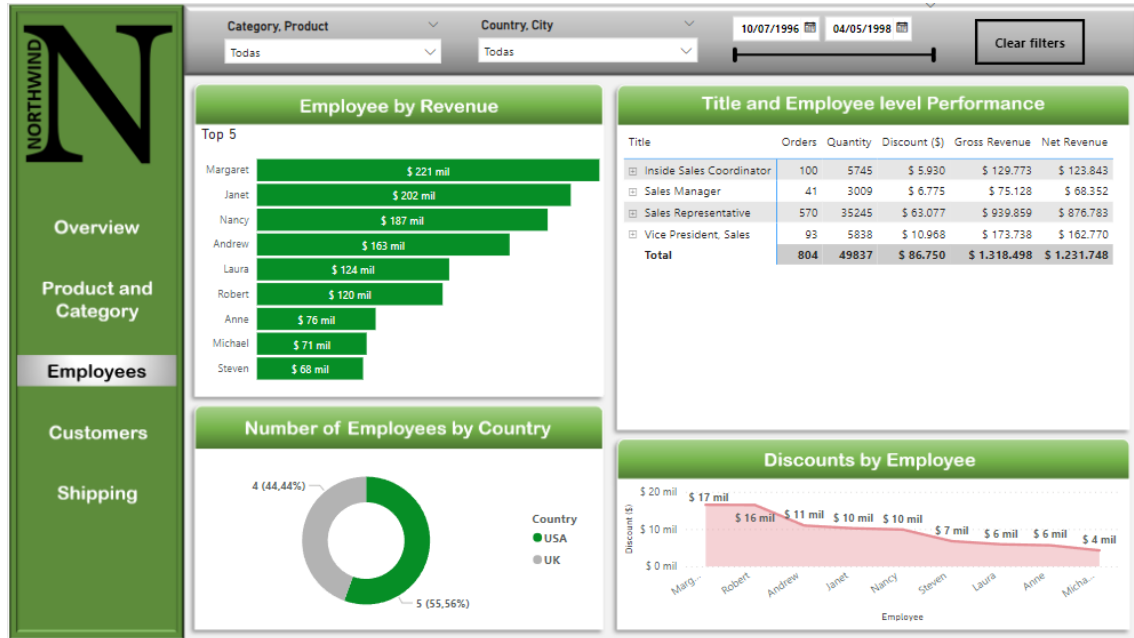
12.3 Página “Product and Category”

En la sección ubicada en la parte superior izquierda, se pueden identificar los cinco productos más populares y los cinco menos demandados. Al lado de esta sección, se encuentra una tabla que detalla los ingresos y las cantidades vendidas, tanto por categoría como por producto. Debajo, existe un gráfico que representa los montos de descuento para cada producto, y el último se enfoca en la gestión de inventario.



12.4 Página "Employees"

La página "Employees" sigue un diseño similar al anterior, pero esta vez el enfoque del análisis se centra en los empleados. Aquí se puede determinar cuáles de ellos generan mayores ingresos, obtener información sobre su ubicación de trabajo y conocer la cantidad de descuentos que aplican.



12.5 Página "Customers"

Esta página presenta un gráfico en la parte superior que muestra a los cinco clientes que han generado los mayores ingresos. En la parte inferior, se encuentran estadísticas sobre la cantidad de clientes agrupados por país.



12.6 Página "Shipping"

En la página "Shipping", se realiza un análisis de las empresas de transporte y sus costos de envío. Se investiga cuáles de estas empresas tienen un tiempo de envío promedio más largo. Además, incluye un gráfico que muestra el porcentaje de pedidos entregados puntualmente con respecto al total de órdenes.



13 Conclusión

Se observa que Estados Unidos y Alemania destacan como los países que generan la mayor cantidad de ingresos a través de las ventas, representando conjuntamente el 30% de los ingresos totales. Además, al analizar el gráfico "Yearly Sales Trends," se aprecia un crecimiento sostenido en los ingresos a lo largo del tiempo.

En cuanto a los productos y categorías, "Beverages" y "Dairy Products" sobresalen como las categorías que más ingresos generan, contribuyendo en conjunto con el 40% de los ingresos totales.

En relación a los empleados, aquellos que residen en Estados Unidos superan en generación de ingresos a sus contrapartes del Reino Unido.

Por último, es notable que Francia cuente con un gran número de clientes, pero sus ingresos no igualan a los de Estados Unidos o Alemania, debido a la menor cantidad de pedidos realizados en este país. Además, en el análisis de las compañías de transporte, se observan datos bastante similares, y es importante destacar que el 95% de los productos se entregan puntualmente.

14 Futuras líneas

Para enriquecer aún más el análisis del desempeño de la empresa Northwind, sería altamente beneficioso obtener información adicional sobre aspectos como los costos asociados a los productos, los salarios de los empleados y otros gastos relacionados. De esta manera se podría llevar a cabo análisis económicos más exhaustivos y detallados.