# Pizza Sales Data Analysis – Documento (ESP)

Agradecimientos a: Data Tutorials



# Indice de contenidos

- 1 Introducción ...... 3
  2 Descripción de la temática de los datos ....... 4
  3 Alcance ...... 5
  4 Importación de archivo csv a SQL (y consultas)
  5 Hipótesis Posibles
  6 Lista de columnas y tipo de datos
- 7 Fondos de tablero e iconos usados
- 8 Importación de tablas a Power BI (Dataset)
- 9 Alcance y Herramientas Tecnológicas Implementadas
- 10 Columnas calculadas/Medidas/Visualizaciones
- 11 Conclusiones

#### 1. Introducción

El éxito de cualquier organización en el ambiente empresarial dinámico de la actualidad depende de la capacidad de tomar decisiones estratégicas informadas. Como uno de los principales indicadores de desempeño, la facturación es esencial para evaluar el futuro y la salud financiera de una empresa.

Este artículo se centra en el análisis de facturación empresarial utilizando Power BI como herramienta principal. Examinaremos cómo las organizaciones pueden usar Power BI, una plataforma de análisis de datos de Microsoft, para organizar, comprender y aprovechar los datos de ventas. Este análisis no exhaustivo facilita la toma de decisiones de las empresas.

A continuación, la página introductoria al reporte de ventas:



En la página introductoria del reporte hacemos una breve introducción al análisis, enumeramos las herramientas tecnologicas utilizadas, hacemos una descripción del dataset y los pasos seguidos en la investigación. Además describimos de forma sucinta el alcance del análisis y las hipótesis planteadas.

#### 2. Descripcion de la tematica del set de datos

Al abordar el tema de las Ventas de una empresa, el dataset proporciona una visión completa de los aspectos financieros de la empresa. Este conjunto de datos recopila información sobre ingresos y ventas en un período de tiempo determinado, lo que permite analizar y comprender la dinámica económica de la empresa. El conjunto de datos contiene una combinación de variables pertinentes que capturan aspectos de las ventas, como la fecha de cada transacción, los bienes vendidos, los precios unitarios y las cantidades vendidas, entre otros aspectos importantes.

Al analizar este conjunto de datos, se pueden encontrar tendencias y patrones en las ventas a lo largo del tiempo, conocer qué productos son los más populares entre los clientes, y descubrir posibles oportunidades de crecimiento o áreas que requieren mejoras en términos de ventas.

El objetivo principal de este conjunto de datos es permitir a los analistas financieros, gerentes y stakeholders de la empresa examinar el desempeño financiero de la empresa desde una perspectiva centrada en las ventas. Utilizar técnicas de análisis de datos y herramientas como Power BI ayuda a tomar decisiones estratégicas informadas y aumenta el éxito comercial.

#### 3. Alcance

Un proyecto de SQL/Power BI centrado en el análisis de las ventas de una empresa puede abarcar una variedad de etapas y objetivos clave para obtener una comprensión completa y valiosa de los aspectos financieros de la compañía. Estos son algunos ejemplos del alcance típico de este tipo de proyecto:

- 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.

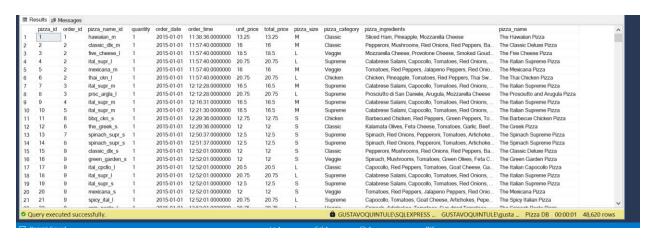
- 6. *Análisis y descubrimiento*: identificar tendencias en los ingresos a lo largo del tiempo y realizar análisis comparativos entre productos, categorías o segmentos de clientes. Evaluar la influencia de los descuentos, promociones u otros factores en la facturación.
- 7. *Implementación de métricas 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 ofrece una estructura general para abordar un proyecto de Power BI de facturación de la empresa. Sin embargo, el alcance exacto puede variar en función de las necesidades y objetivos de la organización.

# 4. Importacion de archivo csv a SQL, consultas y resultados

A continuacion, incluimos la importacion del archivo csv y las consultas en SQL para visualizar, analizar y organizar los datos.

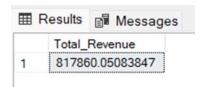
Importamos el archivo csv a SQL y generamos una base de datos para utilizar en el proyecto. Una vez importado el archivo escribimos: **SELECT \* FROM pizza\_sales**,para ver el archivo completo con todas las columnas consideradas.



Luego, escribimos las siguientes consultas y obtuvimos los siguientes resultados:

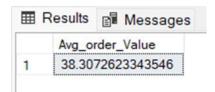
A) Ingresos Totales (Total Revenue):

SELECT SUM(total\_price) AS Total\_Revenue from pizza\_sales;



B) Monto (\$) promedio por cada orden (Average Order Value):

# SELECT SUM(total\_price) / COUNT(DISTINCT order\_id) as Avg\_Order\_Value from pizza\_sales;



C) Cantidad Total de Pizzas Vendidas (Total Pizzas Sold):

# SELECT SUM(quantity) AS Total\_Pizza\_Sold from pizza\_sales;



D) Cantidad Total de Pedidos (Total Orders):

# SELECT COUNT(DISTINCT order\_id) AS Total\_orders FROM pizza\_sales;

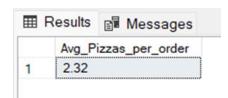


E) Promedio de cantidad de pizzas por pedido (Average Pizzas Per Order):

#### **SELECT**

CAST(SUM(quantity) AS decimal(10,2)) /

CAST(COUNT(DISTINCT order\_id) AS decimal(10,2)) AS average\_quantity\_per\_order FROM pizza\_sales;



F) Tendencia diaria de pedidos totales (Daily Trend for Total Orders)

SELECT DATENAME(DW, order\_date) AS order\_day,

**COUNT(DISTINCT order\_id) AS Total\_orders** 

FROM pizza\_sales

**GROUP BY DATENAME(DW, order\_date)** 



G) Tendencia Mensual de Pedidos Totales (Monthly Trend for Total Orders)

SELECT DATENAME(MONTH, order\_date) AS Month\_Name,

COUNT(DISTINCT order\_id) AS Total\_Orders

FROM pizza\_sales

**GROUP BY DATENAME(MONTH, order\_date)** 

# ORDER BY Total\_Orders DESC



H) Porcentaje de Ventas por Categoria (Percentage of sales by pizza category)

SELECT pizza\_category,

sum(total\_price) AS Total\_Sales,

sum(total\_price) \* 100 / (SELECT sum(total\_price) from pizza\_sales WHERE MONTH(order\_date) = 1) AS PctgeTotal\_Sales

FROM pizza\_sales

WHERE MONTH(order\_date) = 1

**GROUP BY pizza\_category** 



I) Porcentaje de Ventas por Tamaño (Percentage of Total Sales by Pizza Size)

SELECT pizza\_size,

sum(total\_price) AS Total\_Sales,

CAST(sum(total\_price) \* 100 /

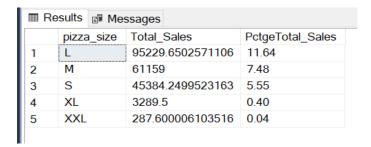
(SELECT sum(total\_price) from pizza\_sales) AS DECIMAL(10,2)) AS PctgeTotal\_Sales

FROM pizza\_sales

WHERE DATEPART(quarter, order\_date) = 1

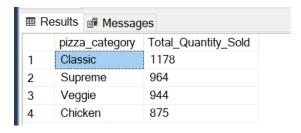
**GROUP BY pizza\_size** 

ORDER BY PctgeTotal\_Sales DESC



J) Cantidad Total de Pizzas Vendidas por Categoria (Total Quantity Sold by Pizza Category)

SELECT pizza\_category, SUM(quantity) as Total\_Quantity\_Sold FROM pizza\_sales WHERE MONTH(order\_date) = 2 GROUP BY pizza\_category ORDER BY Total\_Quantity\_Sold DESC



K) Top 5 Mas Vendidas Por Ingresos (Top 5 pizzas by Revenue)

SELECT TOP 5 pizza\_name, SUM(total\_price) AS Total\_Revenue

FROM pizza\_sales

**GROUP BY pizza\_name** 

**ORDER BY Total Revenue DESC** 

Results				
	pizza_name	Total_Revenue		
1	The Thai Chicken Pizza	43434.25		
2	The Barbecue Chicken Pizza	42768		
3	The California Chicken Pizza	41409.5		
4	The Classic Deluxe Pizza	38180.5		
5	The Spicy Italian Pizza	34831.25		

L) Peores 5 Pizzas Vendidas por Ingreso (Bottom 5 pizzas sold by Revenue):

SELECT TOP 5 pizza\_name, SUM(total\_price) AS Total\_Revenue FROM pizza\_sales

**GROUP BY pizza\_name** 

ORDER BY Total\_Revenue ASC



M) Top 5 Ventas por Cantidad (Top 5 Sales by Quantity):

# SELECT TOP 5 pizza\_name, SUM(quantity) AS Total\_Quantity FROM pizza\_sales GROUP BY pizza\_name

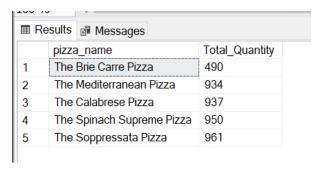
# **ORDER BY Total\_Quantity DESC**



N) Peores 5 ventas por cantidad (Bottom 5 by Quantity):

SELECT TOP 5 pizza\_name, SUM(quantity) AS Total\_Quantity
FROM pizza\_sales
GROUP BY pizza\_name

# ORDER BY Total\_Quantity ASC



O) Top 5 Ventas por Pedidos Totales (Top 5 sales by total orders)

SELECT TOP 5 pizza\_name, COUNT(DISTINCT order\_id) AS Total\_Orders FROM pizza\_sales

**GROUP BY pizza\_name** 

#### ORDER BY Total\_Orders DESC



P) Peores 5 por Pedidos Totales (Bottom 5 by total orders):

SELECT TOP 5 pizza\_name, COUNT(DISTINCT order\_id) AS Total\_Orders FROM pizza\_sales

# **GROUP BY pizza\_name**

# ORDER BY Total\_Orders

■ Results				
	pizza_name	Total_Orders		
1	The Brie Carre Pizza	480		
2	The Mediterranean Pizza	912		
3	The Spinach Supreme Pizza	918		
4	The Calabrese Pizza	918		
5	The Chicken Pesto Pizza	938		

#### 5. Posibles hipótesis:

En el análisis de las ventas de la empresa, hay muchas hipótesis que se pueden explorar y validar. Estas hipótesis son suposiciones preliminares que se pueden probar con los datos disponibles. Intentaremos demostrar y explicar algunas hipótesis pertinentes.

Temporada: las ventas de la empresa varían según la estación del año. Por ejemplo, se podría suponer que la empresa aumenta significativamente sus ventas durante las vacaciones.

*Producto Estrella*: Un producto que contribuye de manera desproporcionada a las ventas totales de una empresa se conoce como producto estrella. Se podría realizar una investigación sobre si ciertos productos generan la mayor parte de los ingresos.

Canales de venta: se podría suponer que ciertos canales de venta (como el comercio electrónico en lugar de las tiendas físicas) tienen un impacto diferente en la facturación total.

Segmentación de clientes: estas divisiones pueden afectar significativamente las ventas de la empresa. Se podría investigar si un grupo de clientes contribuye más o menos a los ingresos de la empresa que un grupo de clientes en particular.

Respuesta a eventos externos: es posible que la facturación de una empresa se vea afectada por eventos externos, como cambios en la economía o las tendencias del mercado.

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

Estos son solo algunos ejemplos de hipótesis que podrían estar relacionadas con un análisis de ventas de una empresa. En nuestro análisis intentaremos demostrar que productos son los más vendidos y cuales los menos vendidos, según categoría y tamaño, pero también por día, mes y estación del año. Con esta información los usuarios del reporte podrán tomar decisiones basadas en un análisis objetivo, completo y entendible.

## 6. Listado de campos por tabla:

En el siguiente apartado proporcionamos un desglose de la estructura de cada columna de la tabla, incluyendo la especificación de los tipos de datos asignados a cada una de ellas. Este análisis de los datos permitirá una comprensión más profunda y completa de la base de datos en cuestión, brindando una visión integral de su diseño y funcionamiento.

Field Names	Data Type	
Pizza_ID	Int	
Order_ID	Int	
Pizza_name_ID	Varchar(50)	
Quantity	Tinyint	
Order_Date	Date	
Order_Time	Time(7)	
Unit_Price	Float	
Total_Price	Float	
Pizza_Size	Varchar(50)	
Pizza_Category	Varchar(50)	
Pizza_Ingredients	Varchar(200)	
Pizza_Name	Varchar(50)	

# 7. Listado de imágenes e iconos utilizados:

A continuación, un pequeño listado de las imágenes e iconos utilizados en el dashboard de Power BI.



Para los fondos utilizamos la paleta de colores provista por Power BI. A continuacion proveemos los códigos hexadecimales que representan los colores utilizados.

Front Page: #118DFF: "blue"; #252423 "black"; imagen de pizza en PNG.

Introduction: #0D6ABF "light blue"; fuente color blanco.

KPI's: #252423 "black"; #0D6ABF "blue"; fuente color blanco; graficos con #26ADB6 "green".

KPI's 2: #252423 "black"; #0D6ABF "blue"; fuente color blanco; graficos con #26ADB6 "green".

5 Best/Worst Sellers: #252423 "black"; #0D6ABF "blue"; fuente color blanco; graficos con #26ADB6 "green, colores en formato condicional.

Conclusions: #0D6ABF "blue" en el cuadro de texto con las conclusiones. Fuente blanca. #E6E6E6 en el fondo de las tarjetas con los valores calculados y blanco en el fondo de las tarjetas. Ademas, una imagen PNG de una pizza. Tipo de letra: Courier New.

# 8. Importacion del dataset a Power BI y transformacion de datos (Power Query)

Importamos nuestra base de datos (llamada PizzaDB) desde SQL a Power BI y hacemos transformaciones sobre los datos para ordenarlos, organizarlos y poder utilizarlos.

Primero, modificamos la columna "pizza\_size" en la cual se representan los tamaños de las pizzas con una sola letra ("L" en el caso de "Large", "M" para medium, etc.). Lo que hacemos, en primer lugar, es abrir Power Query para poder transformar los datos. A continuacion modificamos la columna mencionada y reemplazamos los valores (letras iniciales) en cada registro por valores que describen el tamaño escrito por completo, esto es, "L" se reemplaza por "Large", "M" es reemplazada por "Medium", y asi sucesivamente con los tamaños restantes. Nuestro objetivo es usar la palabra completa para describir las ventas de cada tamaño de pizza.

En Power Query agregamos una nueva columna, basándonos en "order\_date", para que se muestren los nombres de los dias. Llamamos a esta columna "Day Name". De esta última columna solo necesitamos las tres primeras letras, así que una vez que salimos de Power Query agregamos una medida, llamada "Order Day" que describimos en la seccion 10.

Otra columna que generamos fue "Month Number" para ordenar los meses de manera ascendente o descendente.

# 9. Alcance, Glosario y Herramientas Tecnológicas Implementadas:

Un proyecto de SQL/Power BI centrado en el análisis de las ventas de una empresa puede abarcar una variedad de etapas y objetivos clave para obtener una comprensión completa y valiosa de los aspectos financieros de la compañía. Entre estos aspectos se hallan: la definición de objetivos; la recopilación de datos; la limpieza y transformación de datos; creación de modelos; etc.

En la primera página del proyecto "Front Page" incluimos una portada que indica el nombre del proyecto y datos personales del analista. En la segunda proporcionamos una detallada recorrida del proyecto. En la tercera página del proyecto proporcionamos una detallada descripción del análisis del proyecto. Aquí contextualizamos la investigación en curso, y una base sólida para comprender los objetivos y los resultados esperados. Además de describir el análisis y la hipótesis, incorporamos un glosario para definir los términos técnicos y específicos del proyecto, facilitando la comprensión y evitando confusiones.

Agregamos un listado con las herramientas tecnológicas utilizadas en el proyecto, que sirva como una guía transparente y referencia para quienes lean o deseen profundizar en el análisis. Estas herramientas son:

- Microsoft Excel para la lectura de los archivos CVS que conforman el set de datos.
- SQL Server Management Studio para levantar los archivos que conforman el Dataset (archivos CVS).
- Power BI para la generación de los tableros.
- Microsoft Word y Adobe PDF para elaborar la documentación que describe y explica el











proyecto.

#### 10. Columnas calculadas/Medidas/Visualizaciones:

A continuacion, demostraremos la generacion de columnas calculadas, medidas y visualizaciones implementadas en el reporte.

Primero generamos las medidas usando funciones DAX.

La primera medida es "Ingresos Totales":

#### Total Revenue = SUM(pizza\_sales[total\_price]),

cuyo valor es \$817.860.-

La segunda es "Promedio de valor por pedido". Antes de obtener el promedio, necesitamos generar una medida con el total de pedidos. Esta se llama "Pedidos Totales":

# Total Orders = DISTINCTCOUNT(pizza\_sales[order\_id]),

cuyo numero es 21.350.-

A continuacion obtenemos "Promedio de valor por pedido":

### Average Order Value = [Total Revenue] / [Total Orders],

por un valor de \$38.31 por cada pedido.-

Luego, obtenemos el "Total de pizzas vendidas":

#### Total Pizzas Sold = SUM(pizza sales[quantity]),

con un total de 49.574 pizzas vendidas.-

Obtenemos ahora el "Promedio de cantidad de pizzas por pedido":

# Avg Pizzas Per Order = [Total Pizzas Sold] / [Total Orders],

arrojando un promedio de 2,32 pizzas por pedido.-

Estas 5 medidas generadas se visualizan en tarjetas en la pagina de las conclusiones:



Otra medida generada fue la llamada "Order Day", con la finalidad de obtener las tres primeras 3 letras de cada día de la semana, y así lograr que los gráficos se vean claros y prolijos. Para esta medida escribimos el código:

#### Order Day = UPPER(LEFT(pizza sales[Day Name]), 3)),

Obteniendo una nueva columna llamada "Order Day" en la que se registran los días de la semana, pero solo considerando las primeras 3 letras de cada día (MON para Monday, FRI para Friday, etc.)

Otra medida que generamos fue la de las estaciones del año, con el siguiente codigo:

Esto nos generó una columna llamada Seasons, para que podamos analizar las ventas de la empresa en las distintas estaciones del año.

Una última medida es la llamada "Order Month" en la cual registramos los meses considerados, pero utilizando solo las tres primeras letras en mayúscula, para incluirlas en los gráficos de análisis de manera ordenada. Escribimos el siguiente codigo:

# Order Month = UPPER(LEFT(pizza\_sasles[Month Name], 3)),

Y obtenemos una columna con los meses representados por las 3 primeras letras: JAN para January (Enero), FEB para February (Febrero), y así con el resto de los meses.

## Visualizaciones:

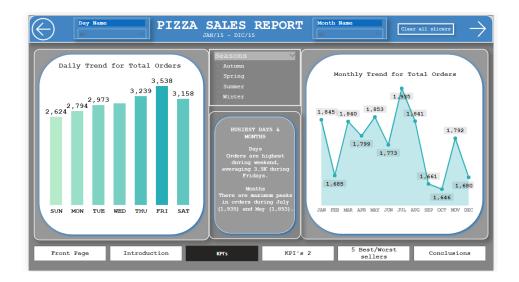
A continuación, vemos la página del reporte que llamamos "KPI's", en la cual analizamos las ventas durante los días de la semana, los meses, y las estaciones del año. Para eso

usamos un gráfico de barras en el cual se incluyen en el eje X la columna "Order Day" y la medida generada llamada "Total Orders" en el eje Y. (grafico izquierda)

En la parte derecha de la página insertamos un gráfico de líneas para analizar la tendencia mensual de las ventas. En este grafico incluimos a "Order Month" y "Total Orders" (medida) en el eje Y. Además de los gráficos incluimos filtros (slicers) para poder segmentar por día de la semana, mes y estación del año. Se incluyó un botón para desactivar los filtros cuando sea necesario (Clear all slicers).

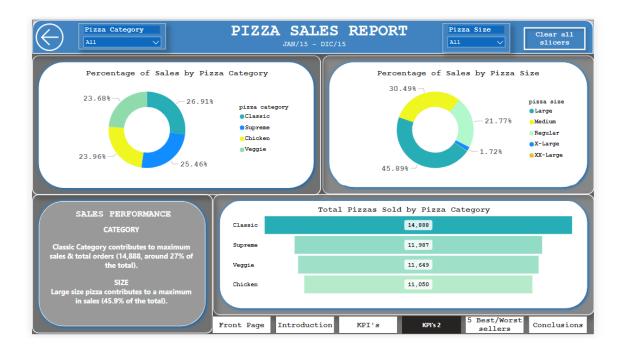
Se incorporó un texto en el cual se describen el/los día/s, mes/es, y estación con mayores ventas.

En todas las páginas del reporte se incluye botones navegadores que permiten acceder a las distintas partes del informe.



En la cuarta página, a la que llamamos "*KPI*'s 2", incluimos dos gráficos de anillos en los cuales se definen los porcentajes de ventas, clasificados por tamaño y por categoría.

Mas abajo incluimos un gráfico de tipo embudo (funnel) para mostrar las ventas totales por categoría. Y en el lado inferior izquierdo un texto en el cual describimos los resultados mostrados en los gráficos.



A continuación, describimos las visualizaciones en la página llama "5 Best/Worst Sellers", en la cual definimos los mejores productos de acuerdo al número de ventas.



Se puede observar que la pizza con mejor venta por ingresos es la Thai Chicken, aportando \$43.434 al total de ingresos; la Classic Deluxe fue la más vendida en términos de cantidad total y pedidos totales, con números de 2.453 y 2.329, respectivamente.

En el otro extremo, observamos que la pizza con menos ventas, por ingresos totales, fue la Brie Carre, aportando \$11.588; además, aporto 490 y 480 a la cantidad total y pedidos totales, respectivamente. Esta categoría de pizza fue la peor en términos de desempeño en ventas.

En la parte inferior de la página, añadimos textos que describen los resultados de los gráficos.

Se incluyó un botón de navegación de página en la parte inferior de la página.

#### 11. Conclusiones

La última página del informe describe las conclusiones obtenidas. Incluimos las tarjetas con las primeras medidas creadas para mostrar los KPI de la empresa durante 2015.



Nuestra hipótesis era mostrar cuáles pizzas eran las más vendidas y cuáles las menos, con el fin de ayudar a los tomadores de decisiones a mejorar las ventas o aplicar decisiones de marketing estratégico con datos analizados y procesados.

Los resultados que obtuvimos fueron claros, los más vendidos son la Thai Chicken (por ingresos) y la Classic Deluxe (por tamaño y categoría). La menos vendida es la Brie Carre. Nuestros datos muestran que esos son los productos clave en los que la empresa debe enfocarse.