

Northwind

Contexto De Proyecto

Este proyecto de análisis de datos tiene como objetivo generar insights estratégicos para Northwind mediante la evaluación de su desempeño en los últimos dos años. En colaboración con los stakeholders clave, abordaremos 13 preguntas críticas que permitirán identificar oportunidades de mejora y optimizar procesos.

Los hallazgos se presentarán a través de un dashboard interactivo, complementado con un resumen ejecutivo que destacará los insights más relevantes. Este documento servirá como base para diseñar estrategias que incrementen la rentabilidad y eficiencia operativa de Northwind en el próximo año, asegurando decisiones basadas en datos para impulsar su crecimiento sostenible.

A tener en cuenta

“las preguntas serán desarrolladas en el dashboard, aquí en el documento estará la parte de los hallazgos más importantes y las recomendaciones (para la optimización de recursos de la empresa y el mejoramiento de ella misma) y preguntas de clarificación (para la continuación del proyecto de la mano con los stakeholders)”

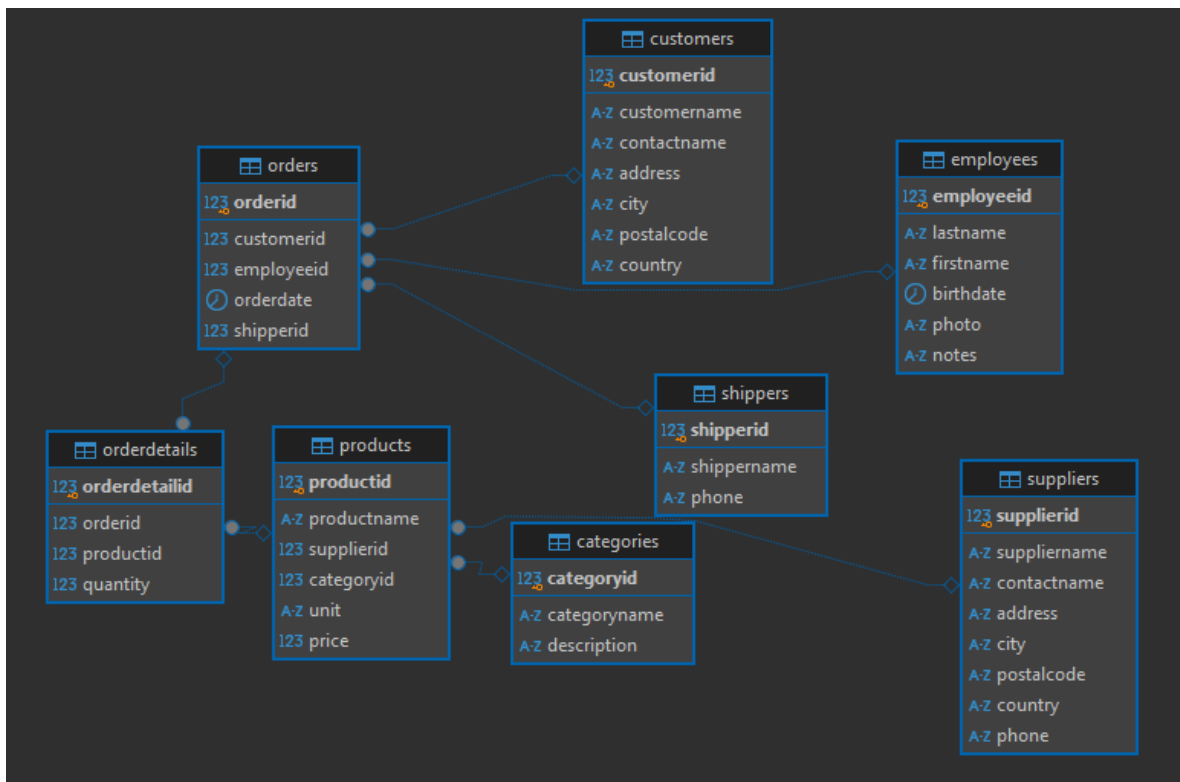
Dashboard link :

Objetivos del Análisis---dashboard

- 1. Mostrar todos los productos de la compañía segmentados por categoría*
- 2. Identificar los 5 productos más rentables según margen de ganancia.*
- 3. Evaluar el desempeño por empleado, considerando: volumen de ventas, unidades vendidas y contribución a las ganancias.*
- 4. Clasificar a los clientes más valiosos según su gasto total en la empresa.*
- 5. Analizar las ganancias por categoría de producto para priorizar líneas de negocio.*
- 6. Diistribucion de las ganancias por categoria de producto*
- 7. Examinar el rendimiento por proveedor: diversidad de productos ofertados y su impacto en las ganancias.*
- 8. Estudiar la estacionalidad de ventas, midiendo ganancias y transacciones por mes.*
- 9. Determinar la capacidad de gestión de clientes por empleado (clientes únicos atendidos).*
- 10. Calcular la contribución porcentual de cada producto a las ganancias de su categoría.*
- 11. Calcular el valor promedio por orden (ticket medio) y el número promedio de productos por pedido*
- 12. Segmentar clientes mediante clustering, utilizando variables como: gasto total, frecuencia de compra, variedad de productos adquiridos, días transcurridos desde su última compra (recency) y días únicos de compra.*

13. Desarrollar un modelo predictivo para estimar el gasto total de un cliente en función del volumen de productos solicitados.

Base De Datos Northwind



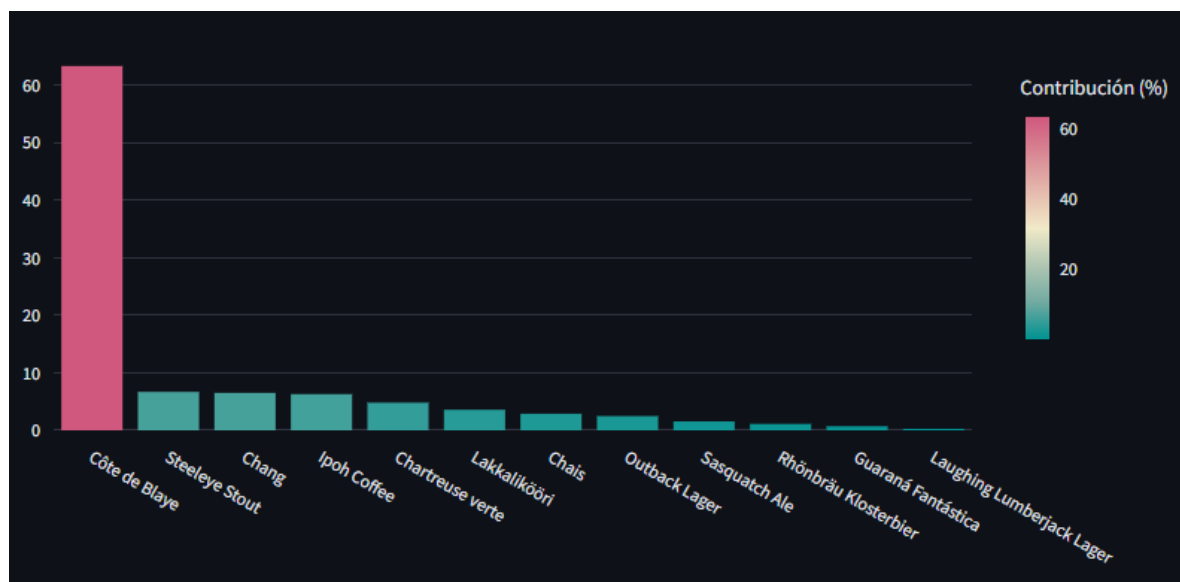
Contenido del proyecto

| | |
|--|----|
| Contexto De Proyecto..... | 1 |
| A tener en cuenta | 1 |
| Dashboard link :..... | 1 |
| Objetivos del Análisis---dashboard..... | 2 |
| Base De Datos Northwind | 3 |
| Resumen ejecutivo..... | 5 |
| Cote de blaye “beaverages”--- Mejores productos de northwind | 5 |
| Los 5 mejores productos en ganancias..... | 6 |
| Cuales son los productos lideres en su categoria por su mayor aporte en ganancia dentro de su categoria?..... | 9 |
| Cuales son los productos que menos aportaron dentro de su categoria en ganancia?..... | 10 |
| Grain cereals---Meat/poultry---produce | 11 |
| Categorias overviev | 12 |
| ingresos por categoria y mes..... | 13 |
| Products | 14 |
| Distribucion de ganancia por categoria | 15 |
| Empleados --- Clientes..... | 16 |
| Segmentacion de clientes | 21 |
| Suppliers | 24 |
| MES – MES Northwind..... | 26 |
| Modelo predictivo del total de gasto que tendria un cliente en la empresa según el numero de productos comprados..... | 27 |
| Recomendaciones | 29 |
| Preguntas de Clarificación, Supuestos y Advertencias..... | 35 |
| Cientifico de datos :..... | 37 |

Resumen ejecutivo

“hallazgos mas importantes para la ayuda de toma de decisiones en cuanto la optimizacion y rendimiento de la empresa”

Cote de blaye “beaverages”--- Mejores productos de northwind

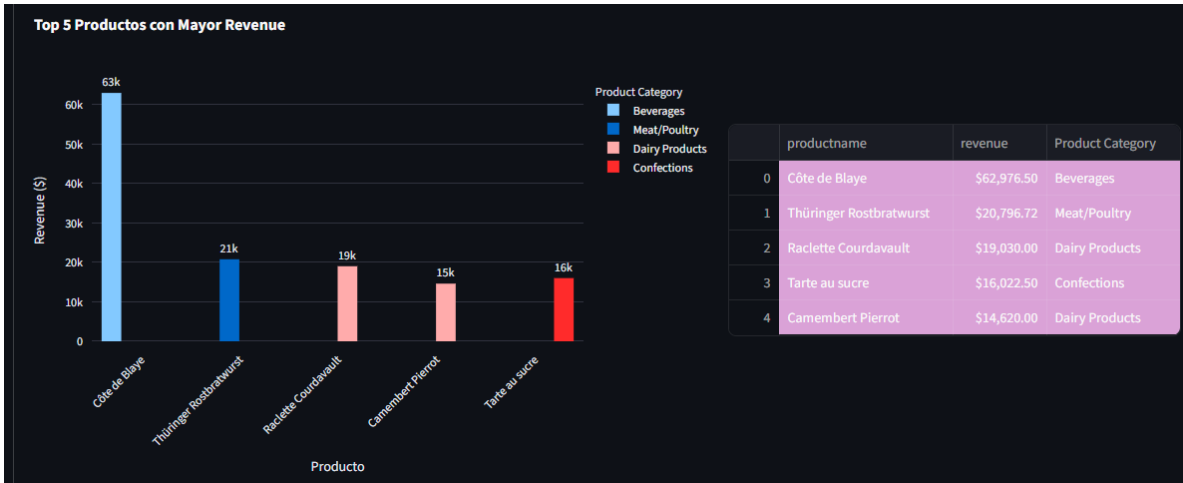


El producto más rentable de la base de datos Northwind fue Côte de Blaye categoria beverages con una ganancia total de \$62,976, representando aproximadamente el 63% de las ganancias dentro de su categoría.

Sin embargo, también se identificó una alta dispersion en ganancia dentro de esta misma categoría. Productos como Guaraná Fantástica y Laughing Lumberjack Lager aportaron menos del 1% de las ganancias, aunque la categoría beaverages es la que mas ganancias genera tambien es la que tiene los productos que menos aportan dentro de una categoria.

"La diferencia relativa entre el producto top (Côte de Blaye) y el promedio de los demás productos es de aproximadamente 1,758%. Esto significa que el producto líder genera 18 veces más ganancias que el promedio del resto de productos en su categoría.

Los 5 mejores productos en ganancias



el promedio de ganancias de los cuatro productos más rentables después de Côte de Blaye fue de \$17,610 lo que significa que en Côte de Blaye hay un 270.59% más de ganancias que el promedio de los otros 4 productos top. En otras palabras, es 3.7 veces más rentable que el promedio de los demás (\$63K vs. \$17K).. Esto sugiere una alta dependencia de un único producto, lo que podría representar un riesgo comercial si no se diversifica la oferta o se refuerzan las ventas de otros productos con potencial.

| | productname | revenue | Product Category |
|---|-------------------------|-------------|------------------|
| 0 | Côte de Blaye | \$62,976.50 | Beverages |
| 1 | Thüringer Rostbratwurst | \$20,796.72 | Meat/Poultry |
| 2 | Raclette Courdavault | \$19,030.00 | Dairy Products |
| 3 | Tarte au sucre | \$16,022.50 | Confections |
| 4 | Camembert Pierrot | \$14,620.00 | Dairy Products |

Kpis ⇄

Total de Ganancias Recaudadas

\$386,424.23

Revenue Total top 5

\$133,445.72

En el top 5 mejores productos se recaudo un revenue total de 133,445 dolares de los cuales de la categoria dairy products entra 2 productos, las categorias que estan afuera del top 5 de productos son las que menor dinero recaudo en total (seafood, produce, grain/cereals) estas 3 categorias en total recaudaron 75.381,45 dolares lo cual en conjunto superan solo en un 19 porciento a el mejor producto "cote de blaye"

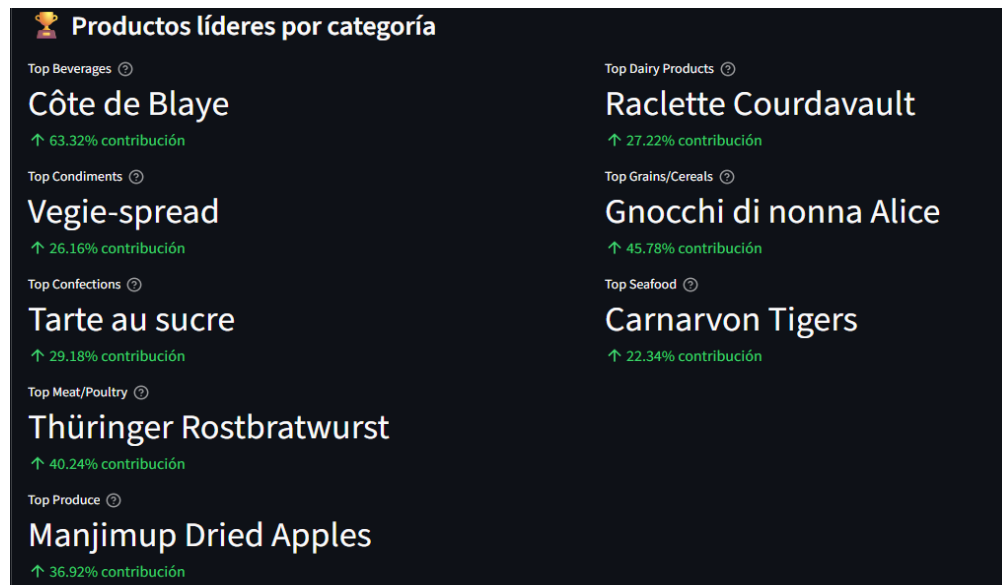
El total de ingresos recaudados fue de \$386,424 dólares, de los cuales:

- Top 5 productos (6.5% del portafolio): Generaron \$133,445, lo que representa 34% de las ganancias totales.*
- 72 productos restantes (93.5% del portafolio): Contribuyeron con \$252,979 (66%) del total.*

Conclusión:

Una mínima parte de los productos (solo el 6.5%) concentra un tercio de las ganancias (34%), mientras que la gran mayoría (93.5%) genera dos tercios (66%) de los ingresos. Esto sugiere que, aunque unos pocos artículos tienen alta rentabilidad, el grueso de las ventas proviene de una amplia variedad de productos.

Cuales son los productos lideres en su categoria por su mayor aporte en ganancia dentro de su categoria?



Cada categoría muestra una fuerte dependencia hacia un único producto que concentra la mayor parte de sus ganancias. Este nivel de concentración implica un riesgo comercial elevado: si ese producto enfrenta problemas de distribución, calidad o una caída en la demanda, podría afectar negativamente a toda la categoría de producto.

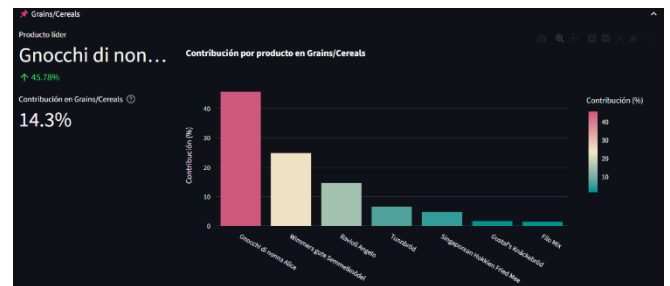
Sin embargo, algunas categorías como Seafood y Dairy Products presentan una distribución de ingresos más equilibrada. Por ejemplo, en Dairy Products, el producto líder (Raclette Courdavault) representa el 27.22% de las ganancias, aunque el promedio de contribución de los demás productos de esa categoría es de aproximadamente 8.11 lo cual, si bien indica una menor concentración comparada con otras categorías, sigue siendo un margen considerable.

Cuales son los productos que menos aportaron dentro de su categoria en ganancia?

| categoryname | productname | ↑ percent_contrib |
|----------------|---------------------------|-------------------|
| Beverages | Laughing Lumberjack Lager | 0.07% |
| Seafood | Røgede sild | 0.48% |
| Beverages | Guaraná Fantástica | 0.71% |
| Condiments | Genen Shouyu | 1.10% |
| Beverages | Rhönbräu Klosterbier | 1.12% |
| Dairy Products | Geitost | 1.13% |
| Dairy Products | Queso Manchego La Pastora | 1.47% |
| Grains/Cereals | Filo Mix | 1.50% |
| Beverages | Sasquatch Ale | 1.55% |
| Confections | Chocolade | 1.63% |

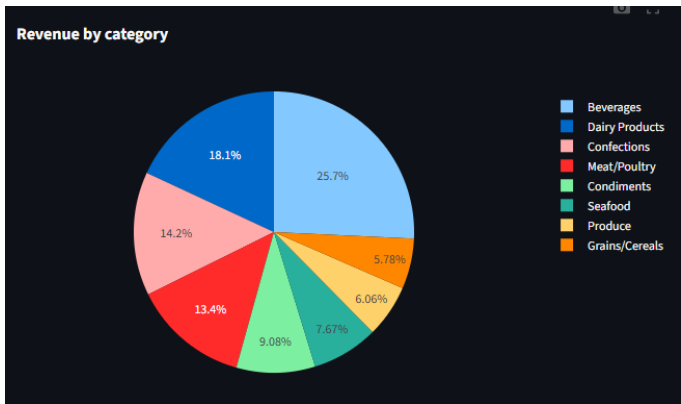
Como se puede ver hay 4 productos de beverages que en total seria el 33 % de los productos individuales de beverages que no aportan mas del 1 %, la segunda categoria con mas productos que menos aportan fue dairy products con dos productos o un total del del 20 % de los productos individuales que no aportan mas del 1.5 % , las categorias meat/poultry y produce no tuvieron productos que estuvieran en este top de los 10 productos que menos aportan a su categorias

Grain cereals---Meat/poultry---produce



En la imagen de arriba se puede ver las categorías que en promedio tienen los productos que tienen el porcentaje mas elevado en participacion al total de su categoria en cuanto ganancias, estas categorías son las que mas productos lideres tienen, el margen promedio de participacion en ganancia de cada categoría en cuanto sus productos va en promedio desde 14.3 hasta 20 % siendo las categorías mas consistentes en cuanto su stock

Categorias overview

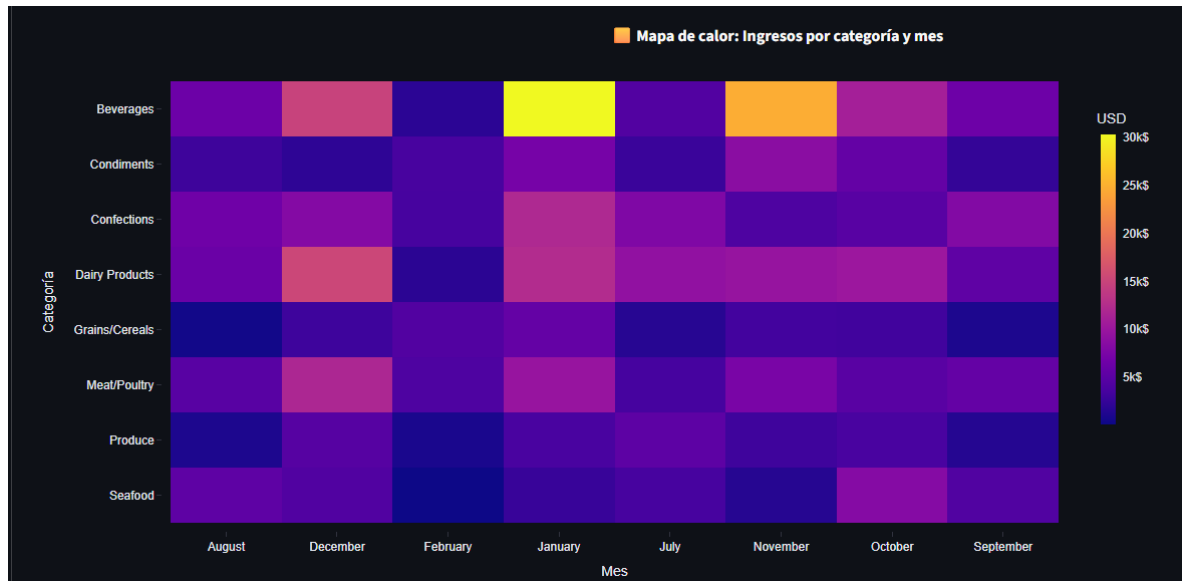


| | categoryname | revenue |
|---|----------------|-------------|
| 0 | Beverages | \$99,464.50 |
| 1 | Dairy Products | \$69,921.00 |
| 2 | Confections | \$54,909.16 |
| 3 | Meat/Poultry | \$51,676.52 |
| 4 | Condiments | \$35,071.60 |
| 5 | Seafood | \$29,652.30 |
| 6 | Produce | \$23,401.40 |
| 7 | Grains/Cereals | \$22,327.75 |

si miramos bien las categorias (produce y grain cereals) son las que menos revenue fue obtenido lo que nos muestra que las categorias que mas dispersion tiene en cuanto a un producto lider y los demas productos son las categorias lider en ganancia total de la empresa

/

ingresos por categoría y mes



Aquí podemos ver como se comporto las categorías en ganancias por mes
dairy/products fue la mas consistente en ganancia en el transcurso de los meses
tambien vemos que beverages tuvo un gran performance en enero como tambien en
noviembre, el mes de diciembre fue bueno en general pero el mes de febrero y
agosto—septiembre fueron meses muy regulares en cualquier categoria

Products

Los 20 productos con menor ganancia

| | | |
|---------------------------------|---------|----------------|
| Valkoinen suklaa | 1137.50 | Confections |
| Rhönbräu Klosterbier | 1116.00 | Beverages |
| Singaporean Hokkien Fried Mee | 1078.00 | Grains/Cereals |
| Queso Manchego La Pastora | 1026.00 | Dairy Products |
| Gravad lax | 1014.00 | Seafood |
| NuNuCa Nuß-Nougat-Creme | 994.00 | Confections |
| Zaanse koeken | 959.50 | Confections |
| Grandma's Boysenberry Spread | 900.00 | Condiments |
| Chocolade | 892.50 | Confections |
| Aniseed Syrup | 800.00 | Condiments |
| Geitost | 790.00 | Dairy Products |
| Uncle Bob's Organic Dried Pears | 750.00 | Produce |
| Guaraná Fantástica | 711.00 | Beverages |
| Röd Kaviar | 675.00 | Seafood |
| Konbu | 552.00 | Seafood |
| Genen Shouyu | 387.50 | Condiments |
| Gustaf's Knäckebröd | 378.00 | Grains/Cereals |
| Filo Mix | 336.00 | Grains/Cereals |
| Røgede sild | 142.50 | Seafood |
| Laughing Lumberjack Lager | 70.00 | Beverages |

los 20 productos con mayor ganancia

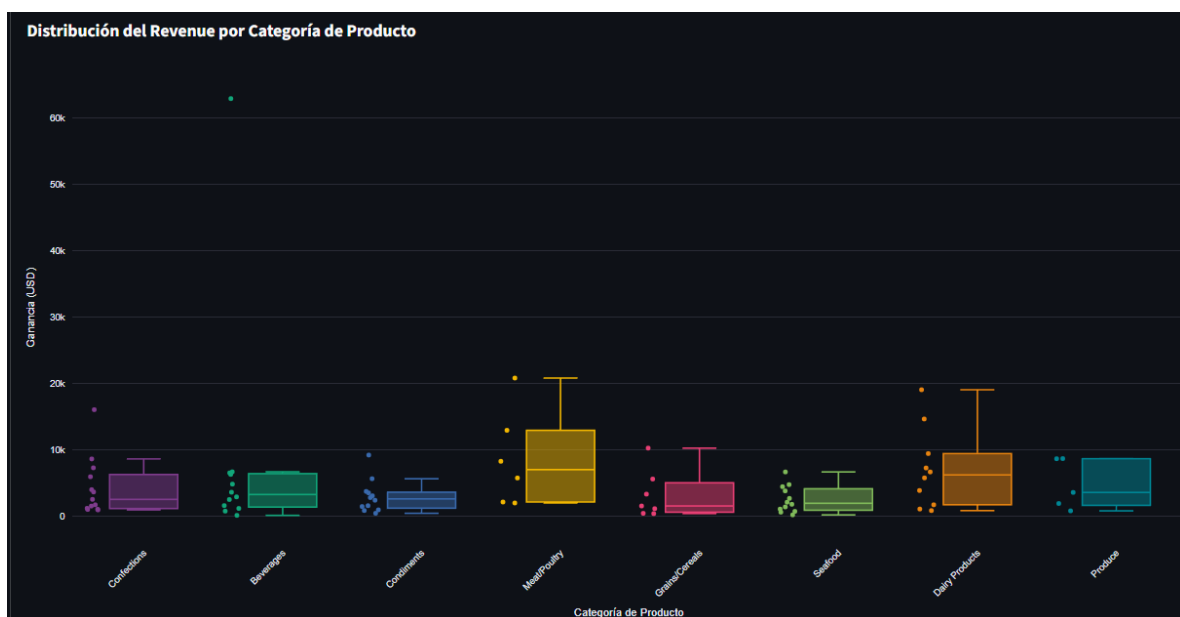
| | | |
|-------------------------|----------|----------------|
| Côte de Blaye | 62976.50 | Beverages |
| Thüringer Rostbratwurst | 20796.72 | Meat/Poultry |
| Raclette Courdavault | 19030.00 | Dairy Products |
| Tarte au sucre | 16022.50 | Confections |
| Camembert Pierrot | 14620.00 | Dairy Products |
| Alice Mutton | 12909.00 | Meat/Poultry |
| Gnocchi di nonna Alice | 10222.00 | Grains/Cereals |
| Mozzarella di Giovanni | 9396.00 | Dairy Products |
| Vegie-spread | 9175.10 | Condiments |
| Manjimup Dried Apples | 8639.00 | Produce |
| Rössle Sauerkraut | 8618.40 | Produce |
| Sir Rodney's Marmalade | 8586.00 | Confections |
| Perth Pasties | 8232.80 | Meat/Poultry |
| Gumbär Gummibärchen | 7245.36 | Confections |
| Fløtemysost | 7224.00 | Dairy Products |
| Steeleye Stout | 6642.00 | Beverages |
| Carnarvon Tigers | 6625.00 | Seafood |
| Gudbrandsdalsost | 6624.00 | Dairy Products |
| Chang | 6479.00 | Beverages |
| Ipoh Coffee | 6256.00 | Beverages |

Como se puede observar, las categorías Seafood y Grains/Cereals se destacan por tener una alta presencia entre los productos con menor ganancia generada, lo que sugiere una baja rentabilidad en varios de sus productos. Por otro lado, categorías como Confections y Beverages presentan una amplia cantidad de productos en ambas muestras (tanto en las de mayor como menor ganancia), lo cual evidencia una alta dispersión en el desempeño individual de sus productos.

En contraste, la categoría Meat/Poultry no aparece entre los 20 productos con menores ingresos, lo que indica un mejor rendimiento general dentro de su línea. Finalmente, la categoría Dairy Products se destaca por su fuerte presencia en el top 20 de productos más rentables, con un total de 5 productos dentro de ese grupo, lo que refleja una alta consistencia en su desempeño comercial.

El 10% de los productos (los top premium) genera el 65% de las ganancias totales, mientras que el 50% de productos económicos aporta solo el 15%

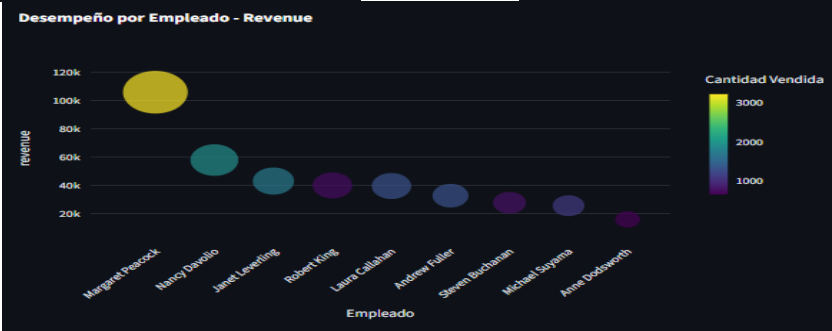
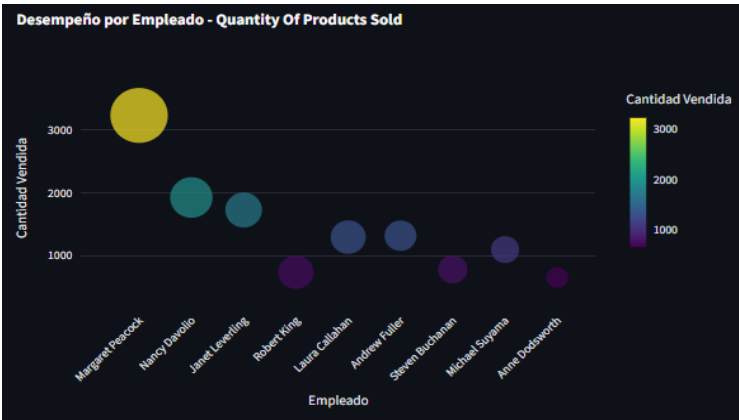
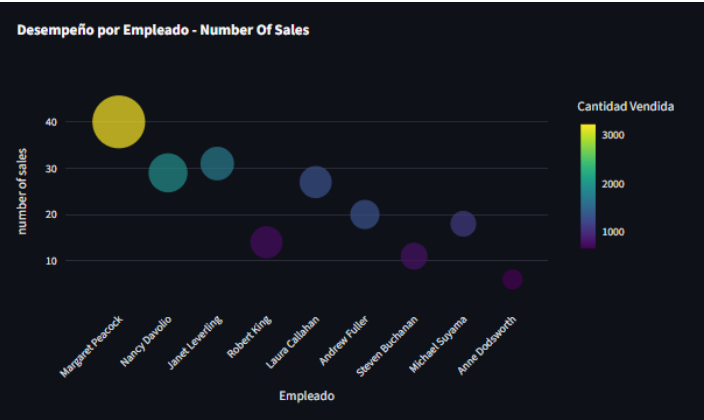
Distribucion de ganancia por categoria



"Como se puede observar, los productos de meat/poultry y dairy products son los que alcanzan los precios más altos. Por otro lado, los más económicos pertenecen a las categorías de seafood y condiments. Finalmente, los productos con mayor dispersión de precios son los de confections, beverages y confections .

Empleados --- Clientes

| | lastname | firstname | number of sales | quantity of products sold | revenue | employee_name |
|---|-----------|-----------|-----------------|---------------------------|--------------|------------------|
| 0 | Peacock | Margaret | 40 | 3232 | \$105,696.50 | Margaret Peacock |
| 1 | Davolio | Nancy | 29 | 1924 | \$57,690.39 | Nancy Davolio |
| 2 | Leverling | Janet | 31 | 1725 | \$42,838.35 | Janet Leverling |
| 3 | King | Robert | 14 | 733 | \$39,772.30 | Robert King |
| 4 | Callahan | Laura | 27 | 1293 | \$39,309.38 | Laura Callahan |
| 5 | Fuller | Andrew | 20 | 1315 | \$32,503.16 | Andrew Fuller |
| 6 | Buchanan | Steven | 11 | 778 | \$27,480.80 | Steven Buchanan |
| 7 | Suyama | Michael | 18 | 1094 | \$25,399.25 | Michael Suyama |
| 8 | Dodsworth | Anne | 6 | 649 | \$15,734.10 | Anne Dodsworth |



1. Top Performers (Empleados Destacados)

- *Margaret Peacock:*

- *Líder absoluta en todas las métricas:*
 - *Ventas: hizo un 83 % mas de ventas que el promedio que fue 21,7 ventas*
 - *Productos vendidos: 3,232 unidades (67% más que Nancy, la segunda).*
 - *Ganancia generada: \$105,696.50 (83% más que Nancy).*
- *Dato clave: Genera el 28% de los ingresos actuales*

- *Nancy Davolio y Janet Leverling:*

- *Segundo y tercer lugar, respectivamente, pero con brechas significativas:*
 - *Nancy vende 29 órdenes (\$57,690),*
 - *Janet vende 31 órdenes, pero con menor ganancia (\$42,838), lo que sugiere ventas de menor valor promedio.*

2. Empleados con bajo rendimiento:

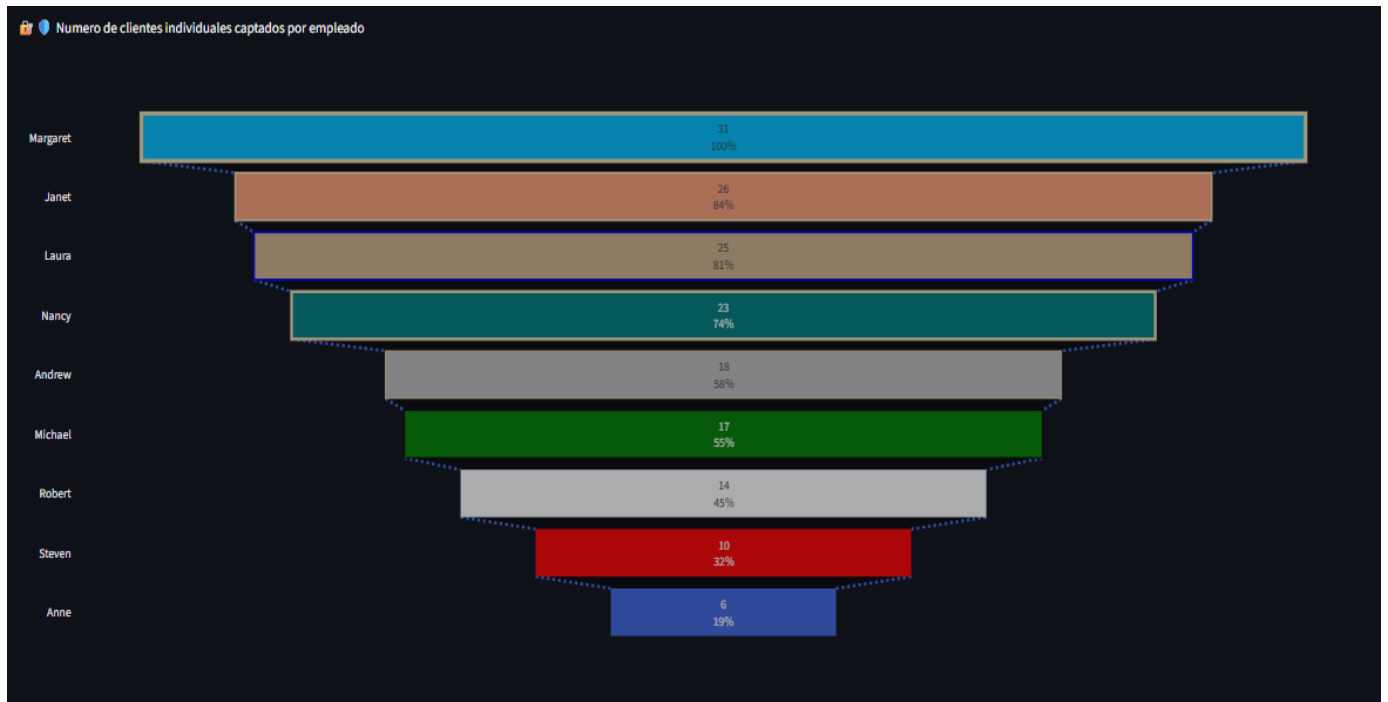
- *Anne Dodsworth: Solo 6 ventas (\$15,724), lo que representa:*
 - *6.7 veces menos revenue que Margaret.*
 - *3 veces menos órdenes que el promedio (excluyendo a Margaret).*
- *Steven Buchanan y Michael Sayama: Aunque tienen más ventas que Anne, su revenue por orden es bajo (\$2,498/orden y \$1,411/orden, respectivamente) en comparacion con anne que se obtuvo por orden de revenue 2622 dolares.*

3. Eficiencia Revenue por orden

| Empleado | Revenue por Orden | Ranking |
|------------------|-------------------|---------|
| Robert King | \$2,840 | 1° |
| Margaret Peacock | \$2,642 | 2° |
| Anne Dodsworth | \$2,622 | 3° |
| Steven Buchanan | \$2,498 | 4° |
| Nancy Davolio | \$1,989 | 5° |
| Andrew Fuller | \$1,625 | 6° |
| Laura Callahan | \$1,455 | 7° |
| Janet Leverling | \$1,381 | 8° |
| Michael Suyama | \$1,411 | 9° |

Aunque ane fue la de menor rendimiento en general en cuanto a revenue por orden esta en el top3 lo que significa que su participacion ha estado mas en las ventas de alto valor , robert king fue el que tuvo Mayor eficiencia con Pocas órdenes (14), leverling fue el que mas ganancia baja tuvo por orden teniendo en cuenta la ganancia

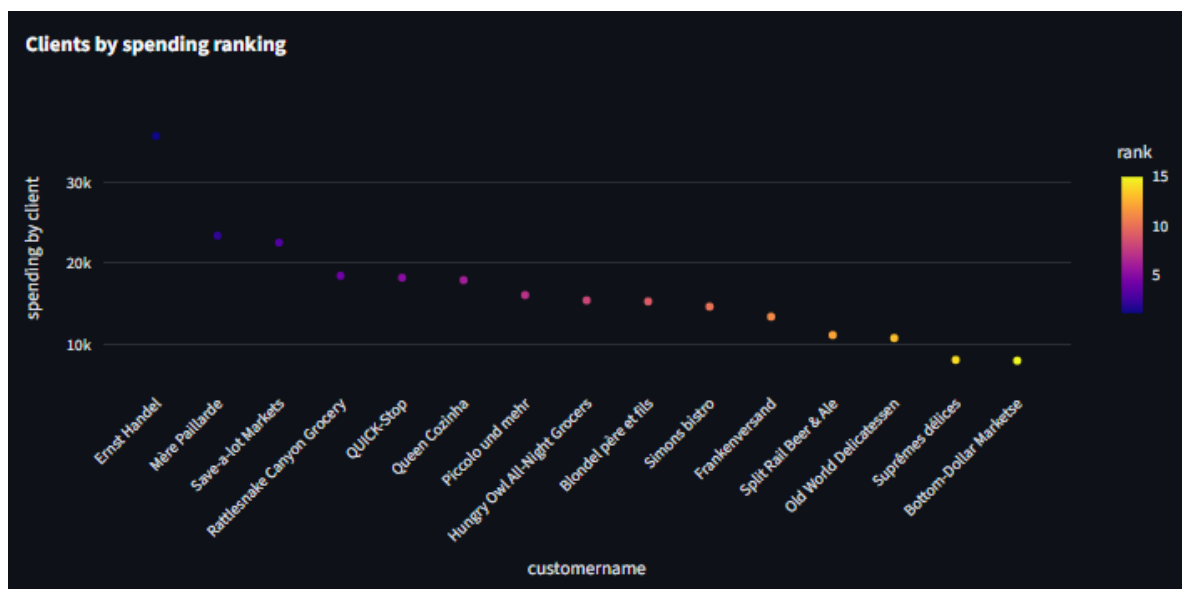
total que dio que fue de 42.838 que lo posiciono en el top 3 osea este empleado estuvo mas enfocado a los productos de bajo valor

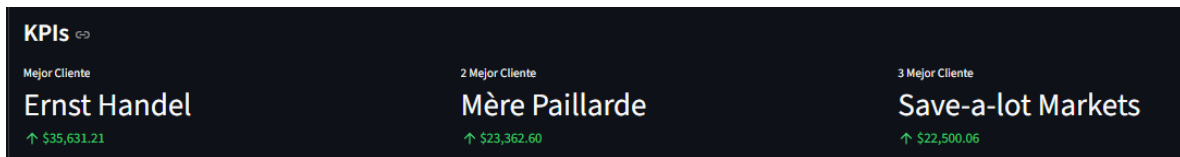


Janet leverling al estar enfocada a las ventas de bajo valor fue de las que mas clientes individuales tuvo en total 26, robert el mas eficiente con pocas ordenes tuvo solamente 14 clientes individuales lo que significa que aunque esta en el top 3 pocos clientes fue el que tuvo los clientes con mas valor , margaret fue la que mas clientes obtuvo con 31 osea tuvo mas clientes que robert , steven , anne juntos en total margaret tuvo el 18 % de clientes totales teniendo en cuenta que el total de clientes que la empresa tiene es de 170

Contraste clave:

- *Mientras Janet (26 clientes) prioriza volumen con bajo valor,*
- *Robert (14 clientes) optimiza valor con bajo volumen.*
- *Margaret logra ambos: máximo volumen (31 clientes) + alto valor por transacción.*

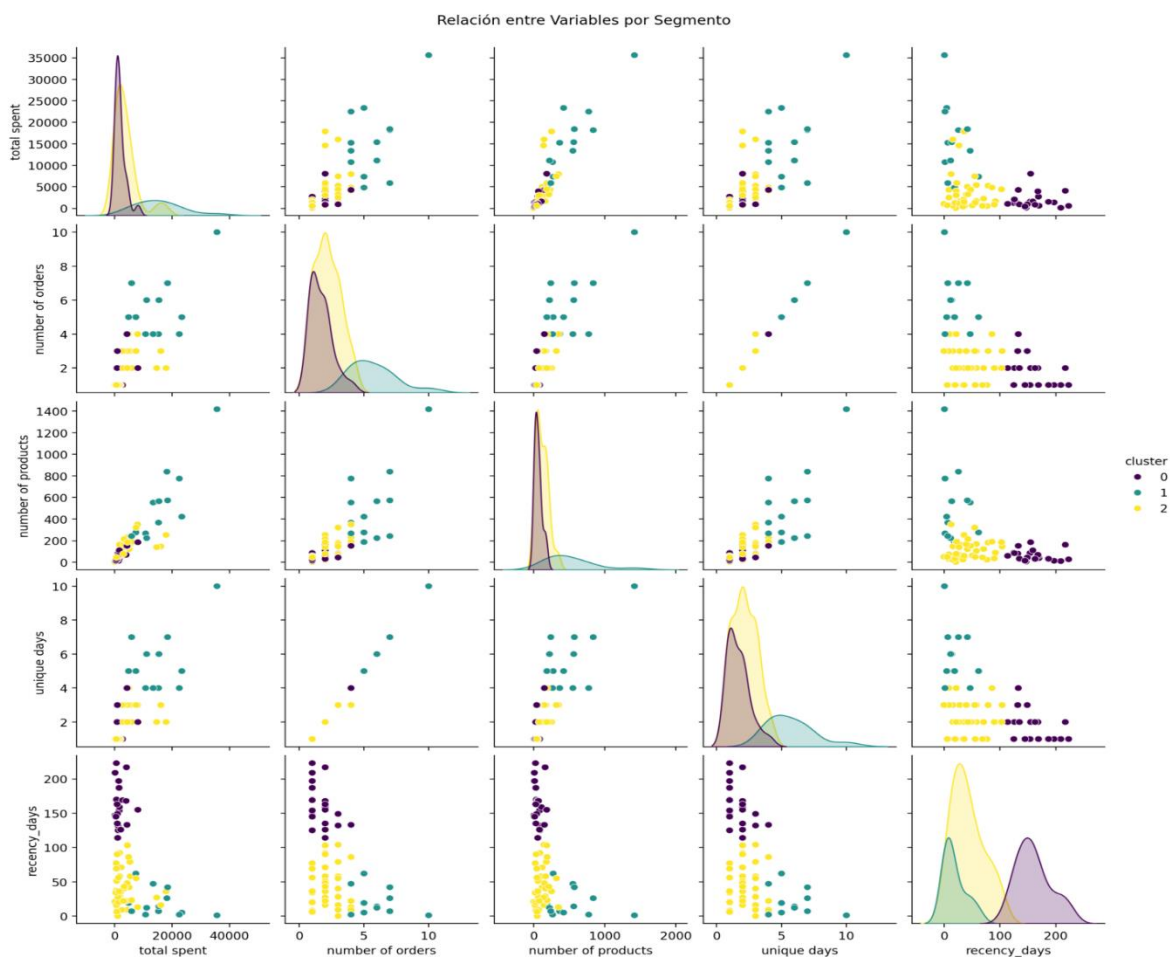




Hay un total de 74 clientes individuales de los cuales los 15 que mas gastan en la empresa generan 255.659 dolares que seria el 66 porciento del total generado el top 3 de mejores clientes ha generado mas dinero 81.493 que 48 clientes individuales en conjunto 77.192 lo que nos muestra que hay en la empresa mucho cliente nuevo (no hay rentencion de clientes “flujo de clientes constante”)

o que la ganancia en dinero que se le ha aportado a la empresa por cliente individual es muy baja


Segmentacion de clientes





Interpretación de Segmentos


Análisis de los clusters basado en el comportamiento de compra:

- Cluster 1 (Clientes Premium)

 Gasto promedio: \$15,541.51


 Pedidos: 5.69 por cliente


 Productos distintos: 516


 Frecuencia de compra: Compra cada 18 días en promedio


Perfil: Nuestros mejores clientes, con alto volumen de compras y máxima fidelidad.

- Cluster 2 (Clientes Regulares)

 Gasto promedio: \$3,718.98


 Pedidos: 2.18 por cliente


 Productos distintos: 119


 Frecuencia de compra: Compra cada 43 días


Perfil: Clientes con potencial de crecimiento, buena frecuencia pero menor volumen.

- Cluster 0 (Clientes Ocasionales)

 Gasto promedio: \$1,783.39

 Pedidos: 1.68 por cliente

 Productos distintos: 64

 Frecuencia de compra: Última compra hace 160 días

Perfil: Clientes inactivos o de bajo compromiso, requieren estrategias de reactivación.

Características de Cada Segmento

| cluster | total spent | number of orders | number of products | unique days | recency_days |
|---------|-------------|------------------|--------------------|-------------|--------------|
| 1 | \$15,541.51 | 5.7 | 516.0 | 5.7 | 18.9 días |
| 2 | \$3,718.98 | 2.2 | 118.8 | 2.2 | 42.8 días |
| 0 | \$1,788.39 | 1.7 | 63.7 | 1.7 | 160.1 días |

Los clientes de la empresa northwind estan segmentados en 3 categorias la primera categoria (premium) son clientes que tienen su frecuencia de compra de cada 18 dias

en promedio con una media de ordenes de 5 y un gasto promedio de 15 k dolares en total ,son 13 clientes los que entran en esa categoria o del total de clientes un 17 %

Esta categoria de clientes tienden a comprar en promedio muchos productos distintos

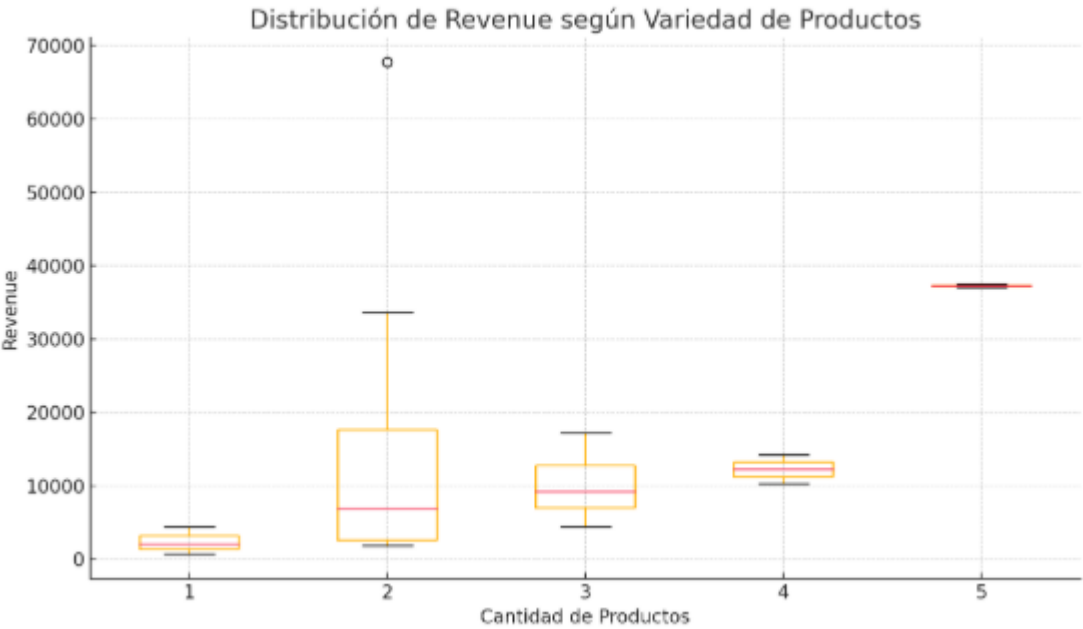
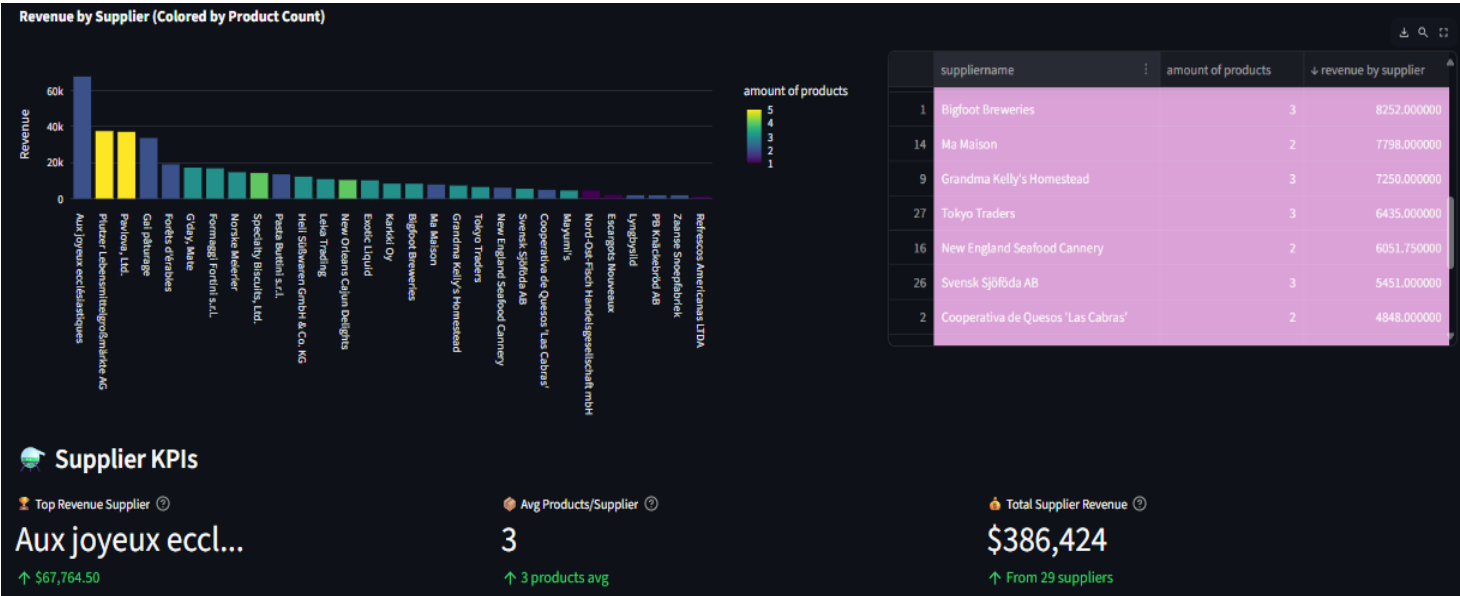
En la segunda categoria estan los clientes regulares que gastan en promedio 3718 \$ con un promedio de 2 ordenes y una frecuencia de compra de 43 dias (frecuencia de compra alta a pesar de ser clientes regulares) , son 39 clientes los que pertenecen a esta categoria o un total de 52 % de los clientes

En la ultima categoria serian los clientes ocasionales que en promedio gastan 1783 \$ con un promedio de ordenes de 1 y una frecuencia de compra de 160 dias no suelen tener muchos productos distintos 64 que los de la segunda categoria que aunque sean clientes regulares compran mas productos distintos 119

Conclusion

Se puede ver en la grafica que los clientes premium tienden a comprar en mas dias distintos que las demas categorias de clientes ,sin segmentacion de clientes el promedio general de productos por orden fue 24 y la ganancia promedio por orden fue de 1971 \$, el cliente que es mas concurrente es ernst hadel (total gastado 35631) con una concurrencia en promedio diaria en compra en la empresa ,tambien save-a-lot-markets (total gastado 22500) con una concurrencia en promedio de cada 2 dias

Suppliers

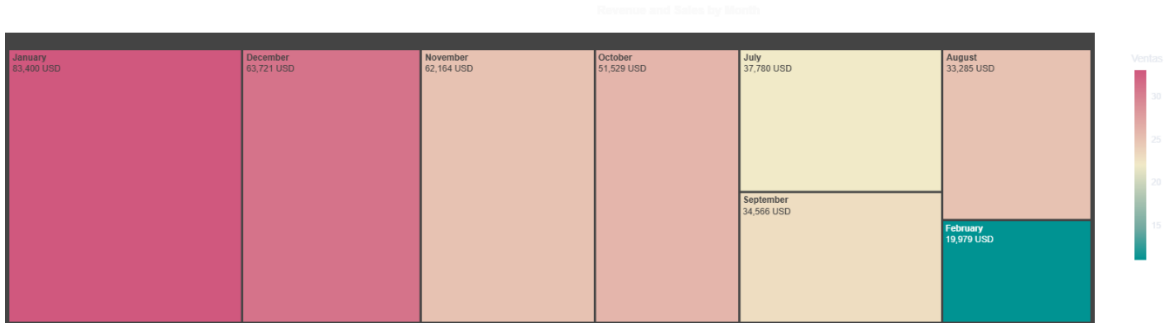


El 13,8 % de los proveedores (4 de 29) concentran más de la mitad del ingreso total, lo que indica una alta dependencia en un grupo muy reducido de socios. Además, estos proveedores ofrecen en promedio solo tres productos cada uno,

Aux joyeux ecclésiastique, con apenas dos productos en su portafolio, aporta \$67 764,50, es decir, más de cinco veces el ingreso medio por proveedor (\$13 324). Este resultado subraya la calidad y demanda específica en cada artículo

Por otra parte, los proveedores que suministran dos productos—que representan el 45 % del total—muestran una gran dispersión en sus ingresos, oscilando desde cifras muy bajas hasta casi \$34 000. En contraste, los proveedores con uno, cuatro o cinco artículos (un grupo que suma siete socios, alrededor del 29 % del total) registran una variabilidad mucho menor, manteniendo un rango de ganancia entre estos grupos mas centrado

MES – MES Northwind



El mes de enero se lleva con el mayor protagonismo, aportando \$83 400, que representan el 21,6 % del ingreso anual total. Este pico coincide además con el mayor número de ventas reportadas (35), por lo que no es solo un mes de grandes montos unitarios, sino también de alto volumen de transacciones.

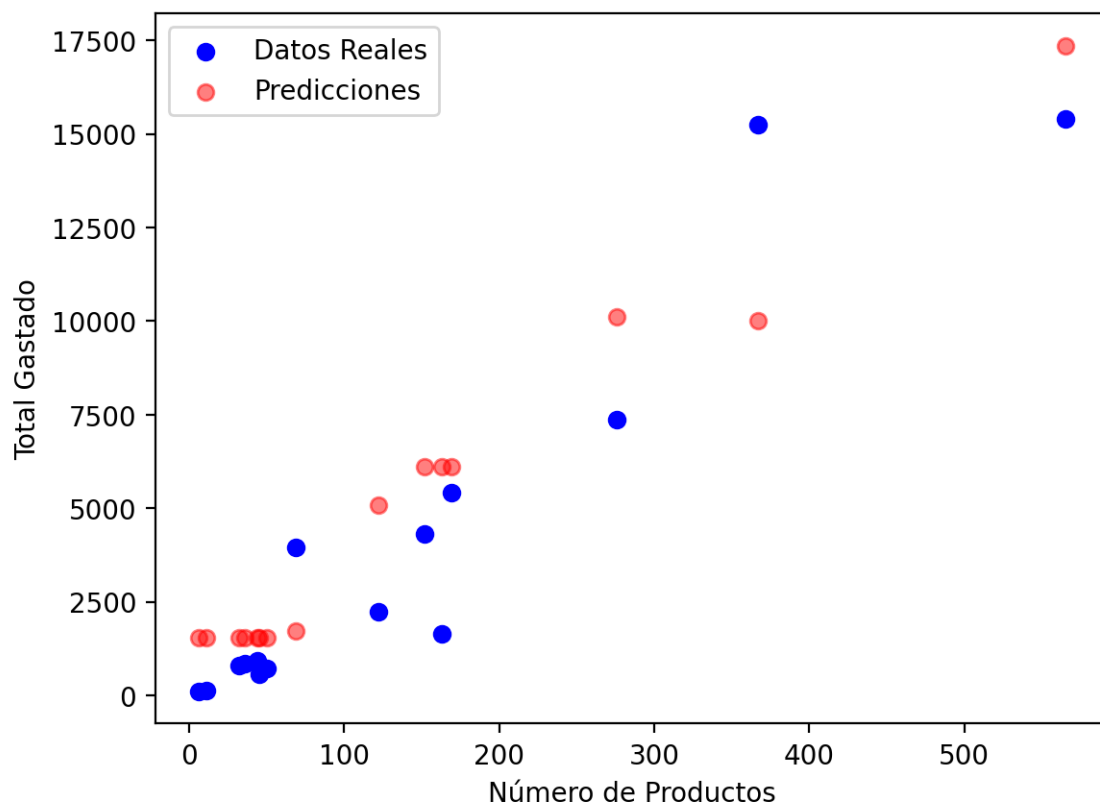
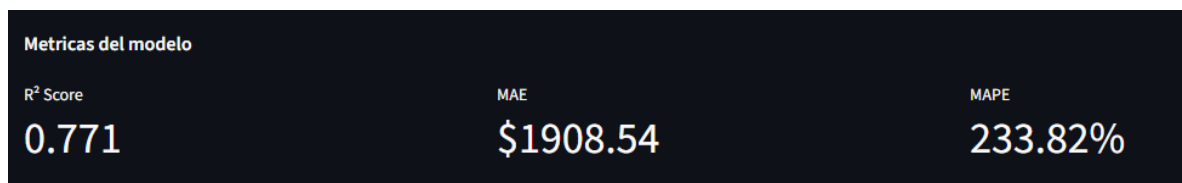
febrero aparece como el punto más bajo de la serie, con solo \$19 979 (5,2 % del total) y el número de ventas más reducido (11).

julio a septiembre se aprecia un valle moderado: julio alcanza \$37 780 (9,8 %) y septiembre \$34 566 (8,9 %). El color de las casillas en tonos amarillos claros confirma que estos meses mantienen un volumen de ventas medio (23–24 transacciones), lo que sugiere un comportamiento estable pero sin picos de demanda

A partir de octubre, la tendencia vuelve a escalar: \$51 529 en octubre (13,3 %), \$62 164 en noviembre (16,1 %) y \$63 721 en diciembre (16,5 %). Aunque diciembre no supera a enero en su alto nivel de ventas (31 transacciones)

En conjunto, el 76 % del ingreso anual se concentra en solo cuatro meses: enero, octubre, noviembre y diciembre. Esta concentración estacional marca dos “ventanas” de oportunidad —inicio y fin de año—, frente a un resto del periodo donde los ingresos se dispersan y oscilan entre \$20 000 y \$38 000.

Modelo predictivo del total de gasto que tendria un cliente en la empresa según el numero de productos comprados



Después de probar varios modelos tanto lineales como no lineales, la conclusión es que todavía no tenemos un modelo realmente confiable para predecir el gasto total de los clientes según el número de productos. El principal problema es que faltan datos históricos - necesitaríamos al menos dos años más de información para que el modelo aprenda bien.

Por ahora, lo que tenemos es un modelo con un R^2 de 0.771, que no está mal para lo que tenemos, y que capta la tendencia general aceptablemente. Donde mejor funciona es cuando los clientes tienen pocos productos (menos de 80), porque es donde tenemos más datos y ahí las predicciones se acercan bastante a la realidad.

El problema serio viene cuando los clientes tienen muchos productos (más de 80). Ahí el modelo empieza a fallar porque simplemente no tenemos suficientes ejemplos de ese tipo de clientes en nuestros registros. En términos prácticos, el modelo se equivoca en promedio unos 1,900 dólares en sus predicciones, lo cual es un margen de error considerable.

En resumen: el modelo actual sirve como primera aproximación, especialmente para clientes con pocos productos, pero para que sea realmente útil y confiable necesitamos recopilar más datos, especialmente de clientes con grandes carteras de productos. Con la información actual, no lo recomendaría para tomar decisiones críticas de negocio.

Recomendaciones

1. Optimización de inventario stock

Objetivo:

Eliminar productos con bajo rendimiento (<2% de ganancia por categoría) para reducir costos de almacenamiento y reenfocar el stock en artículos con mayor potencial, basándose en datos de preferencias del cliente.

Acciones propuestas:

- *Análisis de rentabilidad: Identificar y retirar gradualmente SKUs que generen menos del 2% de ganancia en su categoría .*
- *Estudio de mercado: Realizar encuestas a clientes para entender:*
 - *Criterios de compra (precio, calidad, marca, empaque, etc.).*
 - *Productos preferidos y razones por las que eligen (o no) ciertos artículos.*
 - *Expectativas no cubiertas (sabores, presentaciones o beneficios deseados).*

Ajuste de stock basado en datos:

- *Reemplazar productos eliminados con alternativas alineadas a las preferencias del cliente.*
- *Introducir variantes de productos líderes (ej. nuevos tamaños, sabores o formatos demandados).*

Beneficios esperados:

- *Mayor diversificación de ingresos, disminuyendo la dependencia de unos pocos productos líderes.*
- *Mejor experiencia del cliente, alinear el catálogo con sus necesidades reales.*

Métrica de éxito:

- *Incrementar la participación de los productos no líderes en al menos un 10% en 6 meses.*

2. Optimización de la Frecuencia de Compra por Segmento

Objetivo:

Reducir el intervalo promedio entre compras en un 40% en 6 meses para maximizar ingresos recurrentes, con estrategias segmentadas:

Acciones Clave:

A. Para Clientes Regulares (Cluster 2):

- *Programa "Compra Recurrente":*
 - *Descuento del 15% en el tercer pedido dentro de un trimestre.*
 - *Envío automatizado de recordatorios a los 30 días de la última compra usando historial de Orders al correo.*

B. Para Clientes Ocasionales (Cluster 0):

- *Campaña "Te Extrañamos":*

- *Cupón de reactivación (25% descuento) tras 90 días de inactividad.*
- *Kit de muestras gratis de productos de la empresa “requisito compras mayores a 50 dolares”*

C. Para Clientes Premium (Cluster 1):

Programa vip

- *Acceso a tarjeta premium que permita a estos clientes premium tener descuentos por el semestre especiales “para tener esta tarjeta el cliente tiene que haber gastado en la empresa mas de 3900 dolares en el semestre”*

| Segmento | Frecuencia Actual | Meta (6 meses) | KPI Clave |
|----------------------|-------------------|----------------|--------------------------|
| Clientes Regulares | 43 días | 26 días | +30% pedidos recurrentes |
| Clientes Ocasionales | 160 días | 96 días | Reactivación del 40% |
| Clientes Premium | 18 días | 15 días | Retención 95% |

3. Programa de retención de clientes "pirámide de valor"

Objetivo:

Retener y escalar clientes de alto valor, mientras se convierte a clientes de bajo rendimiento en medianos contribuyentes.

Acciones Segmentadas:

A. Para el Top 3 (Clientes "premium top 3"):

- *Gestión 1 a 1:*
 - *Asignar un account manager dedicado.*
 - *Análisis trimestral personalizado de sus compras con recomendaciones de ahorro.*
- *Beneficios Exclusivos:*
 - *Descuentos "cerrados" (ej: 12% en grandes volúmenes de compras).*
 - *Tarjeta premium exclusiva del top 3 con beneficios personalizados*

B. Para el Top 4-15 (Clientes "premium"):

- *Incentivo: "de devolución de dinero según monto de compra ".*
- *Envíos personalizados :*

Entrega en 24/48 horas sin costo adicional.

Horario de entrega a elección (mañana, tarde, noche).

C. Para los 48 Clientes de Bajo Rendimiento:

- *Campaña "Ascenso de Valor":*
 - *Oferta: " recibir un % de descuento permanente por la compra de tantos productos (el descuento ajustarlo según monto de compra pero que no pase del 5 %)".*

D. Para Nuevos Clientes:

- *Kit de bienvenida con:*
 - *Cupón para segunda compra (15% dto. si gasta >500 USD).*

Impacto Global:

- *Reducir dependencia del top 3 del 21% al 15% del total.*

4. *Gestion de estacionalidad*

Objetivo:

gestionar la caída en febrero en ganancia

A. "Reactivación Post-Pico"

- *Diagnóstico:*
 - *Mes más bajo (\$19,979; solo 11 transacciones).*
 - *Causa probable: "Resaca financiera" post-navideña.*
- *Acciones:*
 - *Paquetes de "Desintoxicación Saludable":*
 - *Combinar Produce (vegetales) + Dairy (yogures) + Beverages (jugos naturales).*
 - *Descuento del 25% para pedidos recurrentes de enero (usar CustomerID de tabla Orders).*

objetivo : enfocar este mes en utilizar bundling con productos que estén cerca de su periodo de expiración. Con este enfoque, los clientes verán este mes como una oportunidad para adquirir paquetes de productos a buen precio, lo que hará que febrero sea un mes muy dinámico en ventas. Además, esta estrategia nos permitirá liberar stock que quedó sin vender después del pico de fin de año, evitando así pérdidas por productos que caduquen. La idea es

crear bundles atractivos que combinen artículos perecederos con otros productos, ofreciendo descuentos especiales para incentivar las compras.

Esto posicionará a febrero como un mes de oportunidades para los clientes, manteniendo el flujo de ventas incluso en temporada baja. La clave será comunicar estas ofertas como promociones por tiempo limitado, creando urgencia y atrayendo a más compradores.

Preguntas de Clarificación, Supuestos y Advertencias

--Preguntas para el stakeholder previo a la continuacion del avance del proyecto

- 1. Por que en la base de datos no hay registrado movimiento en los meses de desde marzo hasta junio ¿hay un leak de informacion ? o la empresa en esos meses no esta funcionando ?*

Supuestos y advertencias :

1. *Actualmente, la columna Unit en la tabla Products utiliza un formato de "bundle" (ej: "10 boxes x 20 bags" o *"24 - 12 oz bottles"*). Esto genera ambigüedad en el cálculo del precio por unidad real y complica el análisis de costos, inventario y rentabilidad.*

Recomendación:

Modificar el formato de Unit para que refleje unidades individuales estandarizadas (ej: "200 bags" en lugar de "10 boxes x 20 bags"). Esto permitirá:

1. *Cálculos precisos: Facilitará determinar el precio por unidad (ej: Price / Total de unidades) de una manera mas precisa.*
 2. *Gestión de inventario: Evitará confusiones al contar stock (¿se registra por cajas o por unidades?).*
-
2. *No hay variables de costos en la base de datos lo cual es muy grave teniendo en cuenta que si se quiere saber la rentabilidad de un producto o empleado como de un supplier no se podria llegar a discernir. Se recomienda empezar a*

recopilar estos datos historicos y asi en un periodo de 6 meses se tendria una mejor percepcion de la rentabilidad de la empresa y desde ahí empezar a implementar estrategias de optimizacion

Cientifico de datos :

Nicolas Felipe Mogollon Granda