

Análisis de Datos en E-Commerce: Olist

Introducción

Contexto de Olist

Olist es una empresa brasileña de e-commerce que opera como marketplace, permitiendo a pequeños y medianos negocios vender sus productos en grandes plataformas como Amazon, Mercado Libre y Americanas. Su modelo de negocio se basa en facilitar infraestructura, logística y soporte a los vendedores.

Propósito del Análisis

El objetivo de este estudio es **extraer insights clave sobre el comportamiento de los clientes, tendencias de ventas, impacto de las entregas y patrones de satisfacción**, utilizando consultas SQL y visualizaciones en Power BI.

Fuente de los Datos

Los datos han sido obtenidos desde Kaggle en el dataset "**Brazilian E-Commerce Public Dataset by Olist**". Contiene información detallada sobre casi 100,000 pedidos realizados en la plataforma entre **2016 y 2018**, incluyendo datos sobre clientes, productos, métodos de pago, tiempos de entrega y reseñas.

Metodología

Para realizar este análisis utilizamos:

- **SQL** para exploración de datos y análisis de patrones.
 - **Power BI** para la visualización y creación de dashboards interactivos.
 - **GitHub** para la documentación y presentación del proyecto.
-

2 Exploración Inicial del Dataset

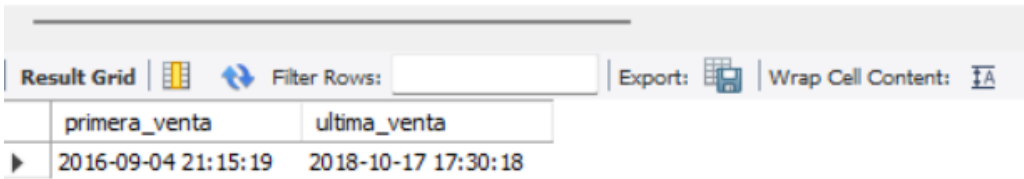
 **Objetivo:** Entender la estructura de los datos y definir el punto de partida.

♦ Rango de Fechas de las Ventas

```
SELECT MIN(order_purchase_timestamp) AS fecha_inicio,  
MAX(order_purchase_timestamp) AS fecha_fin  
FROM orders;
```

 Determinar el período de análisis.

```
1 • SELECT MIN(order_purchase_timestamp) AS primera_venta,  
2         MAX(order_purchase_timestamp) AS ultima_venta  
3 FROM orders;  
4
```




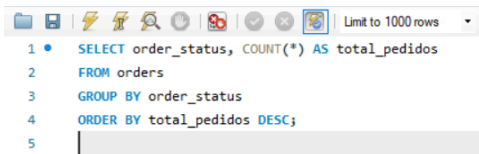
	primera_venta	ultima_venta
▶	2016-09-04 21:15:19	2018-10-17 17:30:18

Podemos ver que la primera venta fue realizada en septiembre de 2016 y la última en octubre de 2018.

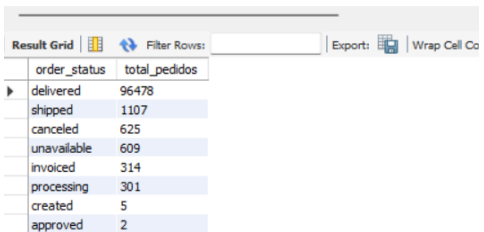
♦ Número Total de Pedidos por Estado

```
SELECT order_status, COUNT(*) AS total_pedidos  
FROM orders  
GROUP BY order_status  
ORDER BY total_pedidos DESC;
```

 Identificar la distribución de pedidos (entregados, cancelados, en espera, etc.).



```
1 • SELECT order_status, COUNT(*) AS total_pedidos  
2 FROM orders  
3 GROUP BY order_status  
4 ORDER BY total_pedidos DESC;  
5
```



	order_status	total_pedidos
▶	delivered	96478
	shipped	1107
	canceled	625
	unavailable	609
	invoiced	314
	processing	301
	created	5
	approved	2

Conclusión principal:

La gran mayoría de los pedidos han sido entregados (96,478 pedidos), lo que representa casi la totalidad de las órdenes procesadas.

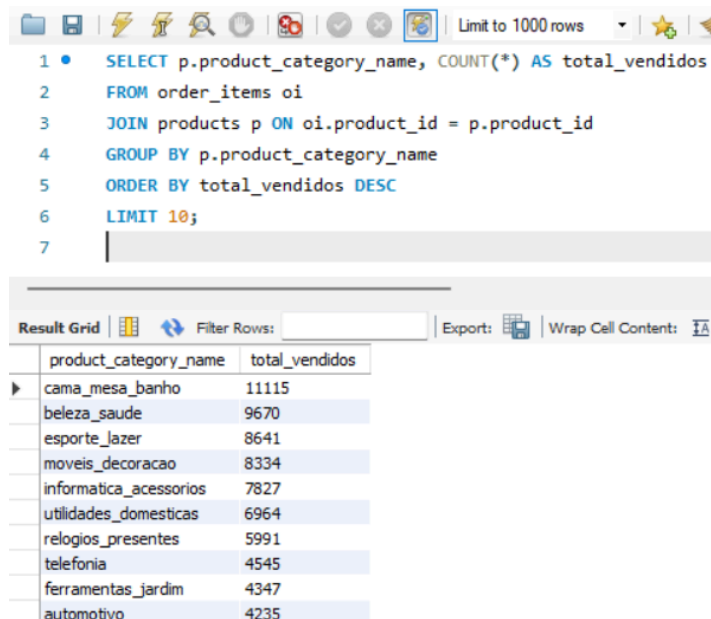
Puntos clave:

- Solo 625 pedidos fueron cancelados, lo que indica una baja tasa de cancelación.
- 1,107 pedidos están en estado "shipped", es decir, en tránsito.
- Hay un pequeño número de pedidos aún en procesamiento (301) o en estado "created" (5), lo que indica que la mayoría del flujo de pedidos ya está completado.
- 609 pedidos están "unavailable", lo que podría representar problemas de stock o cancelaciones internas.

♦ Categorías de Productos Más Vendidas

```
SELECT product_category_name, COUNT(order_id) AS total_pedidos
FROM order_items
GROUP BY product_category_name
ORDER BY total_pedidos DESC
LIMIT 10;
```

📌 Identificar las 10 categorías de productos más populares.



```
1 • SELECT p.product_category_name, COUNT(*) AS total_vendidos
2   FROM order_items oi
3  JOIN products p ON oi.product_id = p.product_id
4  GROUP BY p.product_category_name
5  ORDER BY total_vendidos DESC
6  LIMIT 10;
7
```

product_category_name	total_vendidos
cama_mesa_banho	11115
beleza_saude	9670
esporte_lazer	8641
moveis_decoracao	8334
informatica_acessorios	7827
utilidades_domesticas	6964
relogios_presentes	5991
telefonia	4545
ferramentas_jardim	4347
automotivo	4235

Conclusión principal:

La categoría "cama_mesa_banho" (ropa de cama, mesa y baño) lidera las ventas con 11,115 pedidos, representando aproximadamente un 12.9% del total de los 10 productos más vendidos. Le sigue "beleza_saude" (belleza y salud) con 9,670 ventas, lo que equivale al 11.2%.

Distribución de las categorías más vendidas:

- "cama_mesa_banho" → 11,115 pedidos (12.9%)
- "beleza_saude" → 9,670 pedidos (11.2%)
- "esporte_lazer" → 8,641 pedidos (10.0%)
- "moveis_decoracao" → 8,341 pedidos (9.7%)
- "informatica_acessorios" → 7,827 pedidos (9.1%)
- "utilidades_domesticas" → 5,969 pedidos (6.9%)
- "relogios_presentes" → 5,694 pedidos (6.6%)
- "telefonica" → 4,545 pedidos (5.3%)
- "ferramentas_jardim" → 4,347 pedidos (5.0%)
- "automotivo" → 4,235 pedidos (4.9%)

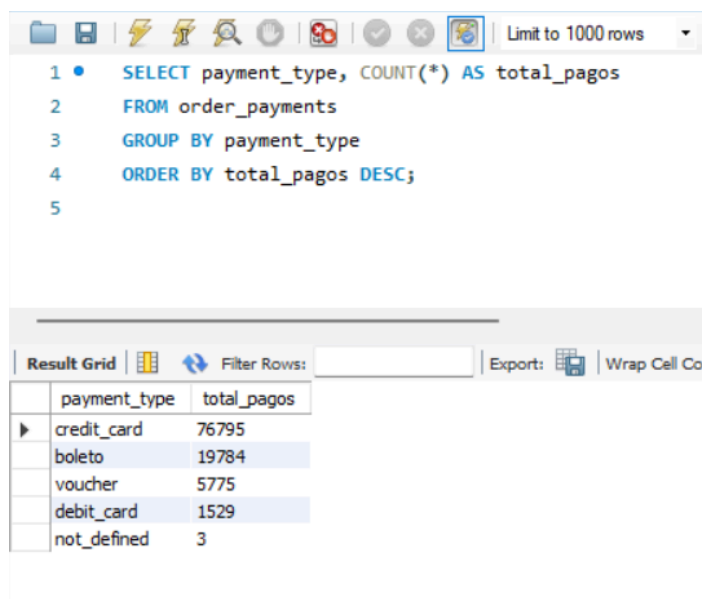
Datos clave:

- Las tres primeras categorías ("cama_mesa_banho", "beleza_saude" y "esporte_lazer") concentran el 34.1% del total de pedidos en el Top 10.
- Las categorías de tecnología ("informatica_acessorios" y "telefonica") suman un 14.4%, mostrando una demanda considerable.
- Productos del hogar y decoración ("moveis_decoracao", "utilidades_domesticas" y "ferramentas_jardim") representan el 21.6%, lo que indica una fuerte inclinación hacia la mejora del hogar.

♦ Métodos de Pago Más Usados

```
SELECT payment_type, COUNT(order_id) AS total_pedidos
FROM order_payments
GROUP BY payment_type
ORDER BY total_pedidos DESC;
```

📌 Analizar qué método de pago es el más utilizado.



The screenshot shows a SQL query editor with a toolbar at the top. The query is as follows:

```
1 • SELECT payment_type, COUNT(*) AS total_pagos
2 FROM order_payments
3 GROUP BY payment_type
4 ORDER BY total_pagos DESC;
5
```

Below the query, there is a 'Result Grid' section. It includes a 'Filter Rows' input field and an 'Export' button. The results are displayed in a table with two columns: 'payment_type' and 'total_pagos'.

payment_type	total_pagos
credit_card	76795
boleto	19784
voucher	5775
debit_card	1529
not_defined	3

Conclusión principal:

El pago con tarjeta de crédito es el método más utilizado, representando un 73.4% de todas las transacciones registradas. Le sigue el método boleto bancario, con un 18.9%, mientras que el uso de voucher y tarjeta de débito es significativamente menor.

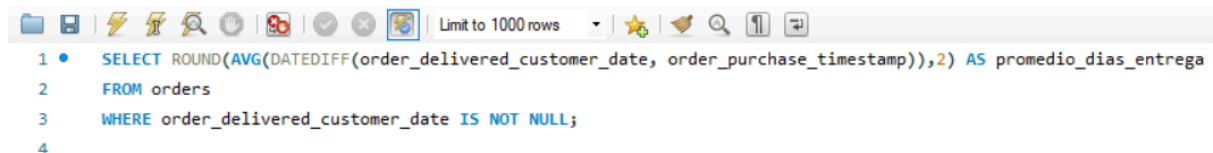
Distribución de métodos de pago:

- Tarjeta de crédito: 76,795 pagos (73.4%)
- Boleto bancario: 19,784 pagos (18.9%)
- Voucher: 5,775 pagos (5.5%)
- Tarjeta de débito: 1,529 pagos (1.5%)
- No definido: 3 pagos (0.003%)

♦ Tiempos de Entrega Promedio

```
SELECT ROUND(AVG(DATEDIFF(order_delivered_customer_date,  
order_purchase_timestamp)), 2) AS tiempo_entrega_promedio  
FROM orders  
WHERE order_delivered_customer_date IS NOT NULL;
```

📌 Evaluar cuánto tarda en promedio la entrega de un pedido.



```
1 • SELECT ROUND(AVG(DATEDIFF(order_delivered_customer_date, order_purchase_timestamp)),2) AS promedio_dias_entrega  
2 FROM orders  
3 WHERE order_delivered_customer_date IS NOT NULL;  
4
```



promedio_dias_entrega
12.50


Conclusión principal:

El tiempo promedio de entrega de los pedidos en la plataforma es de 12.5 días desde la fecha de compra hasta la entrega al cliente.

Datos clave:

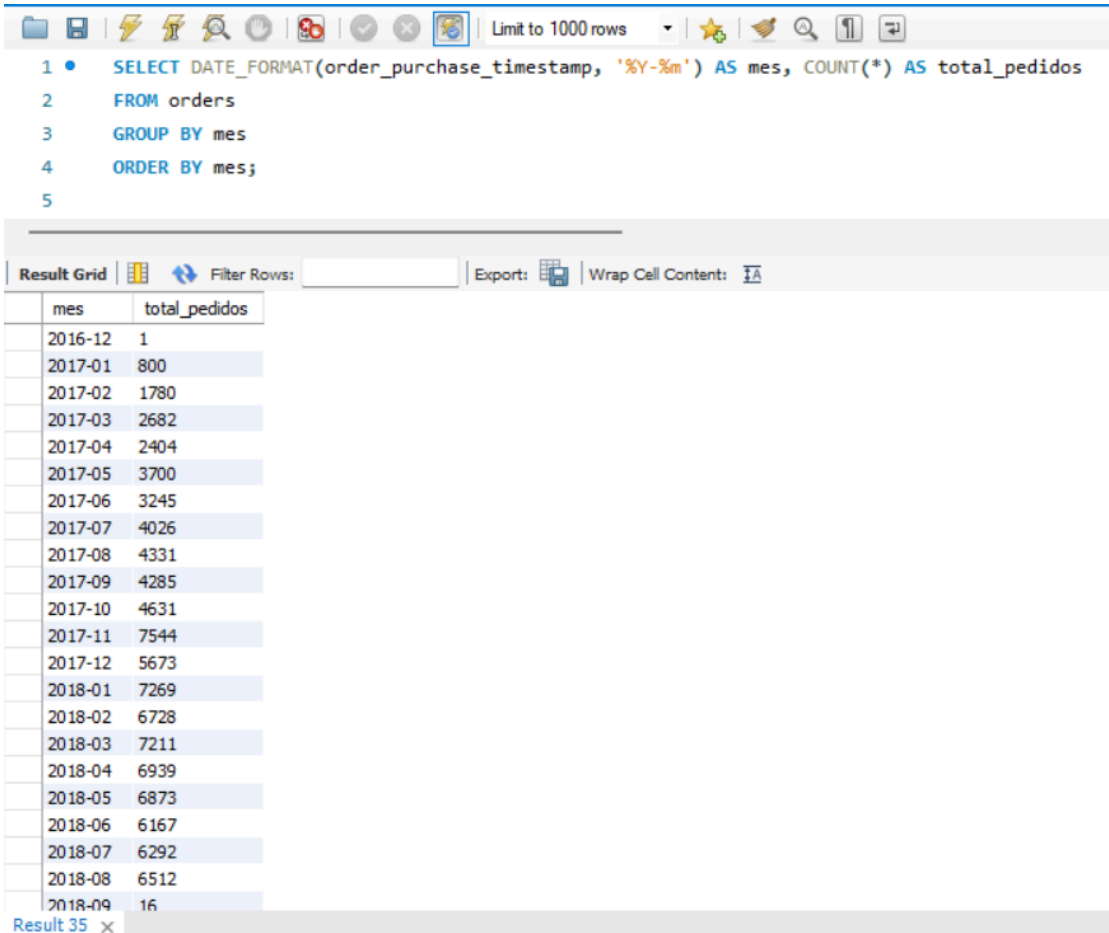
- La media de 12.5 días puede ser un factor relevante en la experiencia del cliente.
- Este valor incluye todas las categorías de productos y ubicaciones de entrega, por lo que podría variar según la región y el tipo de producto.
- La velocidad de entrega puede influir directamente en la satisfacción del cliente y en las reseñas recibidas.

3 Análisis de Ventas y Estacionalidad

 **Objetivo:** Identificar patrones de ventas a lo largo del tiempo.

♦ Análisis de Tendencias de Ventas

```
SELECT DATE_FORMAT(order_purchase_timestamp, '%Y-%m') AS mes, COUNT(*) AS  
total_pedidos  
FROM orders  
GROUP BY mes  
ORDER BY mes;
```



mes	total_pedidos
2016-12	1
2017-01	800
2017-02	1780
2017-03	2682
2017-04	2404
2017-05	3700
2017-06	3245
2017-07	4026
2017-08	4331
2017-09	4285
2017-10	4631
2017-11	7544
2017-12	5673
2018-01	7269
2018-02	6728
2018-03	7211
2018-04	6939
2018-05	6873
2018-06	6167
2018-07	6292
2018-08	6512
2018-09	16

Conclusión principal:

Se observa un crecimiento progresivo en la cantidad de pedidos desde inicios de 2017 hasta mediados de 2018, con picos en noviembre de 2017 (7,544 pedidos) y enero de 2018 (7,269 pedidos).

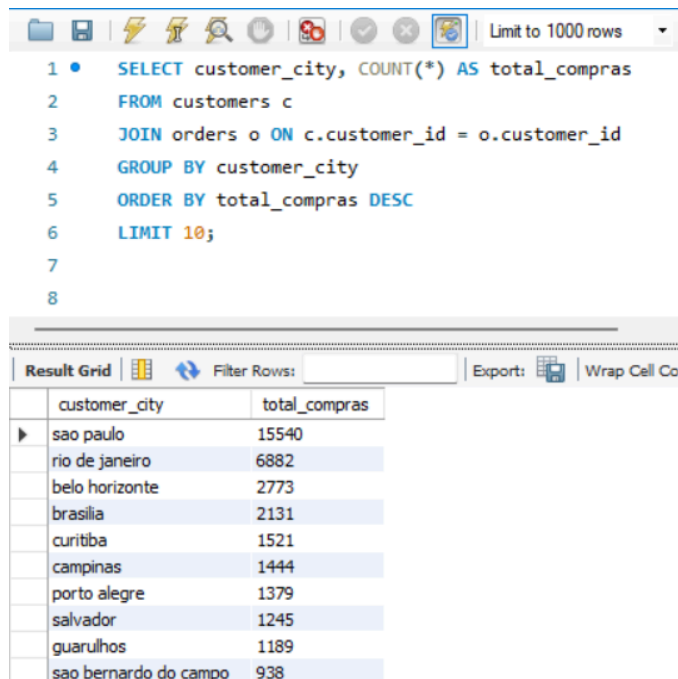
Datos clave:

- A lo largo de los meses, se nota una tendencia creciente en las ventas, lo que indica una expansión del e-commerce de Olist.
- Noviembre y diciembre de 2017 presentan aumentos significativos, lo que podría estar relacionado con eventos de compras como Black Friday y Navidad.
- En 2018, los pedidos siguen altos, con más de 6,000 pedidos mensuales de enero a agosto, lo que demuestra una consolidación en el mercado.

♦ Análisis Geográfico: Ciudades con Más Compras

```
SELECT customer_city, COUNT(order_id) AS total_pedidos
FROM customers
JOIN orders ON customers.customer_id = orders.customer_id
GROUP BY customer_city
ORDER BY total_pedidos DESC;
```

📌 Determinar las ciudades con más pedidos.



```
1 • SELECT customer_city, COUNT(*) AS total_compras
2   FROM customers c
3   JOIN orders o ON c.customer_id = o.customer_id
4   GROUP BY customer_city
5   ORDER BY total_compras DESC
6   LIMIT 10;
7
8
```

customer_city	total_compras
sao paulo	15540
rio de janeiro	6882
belo horizonte	2773
brasilia	2131
curitiba	1521
campinas	1444
porto alegre	1379
salvador	1245
guarulhos	1189
sao bernardo do campo	938

Conclusión principal:

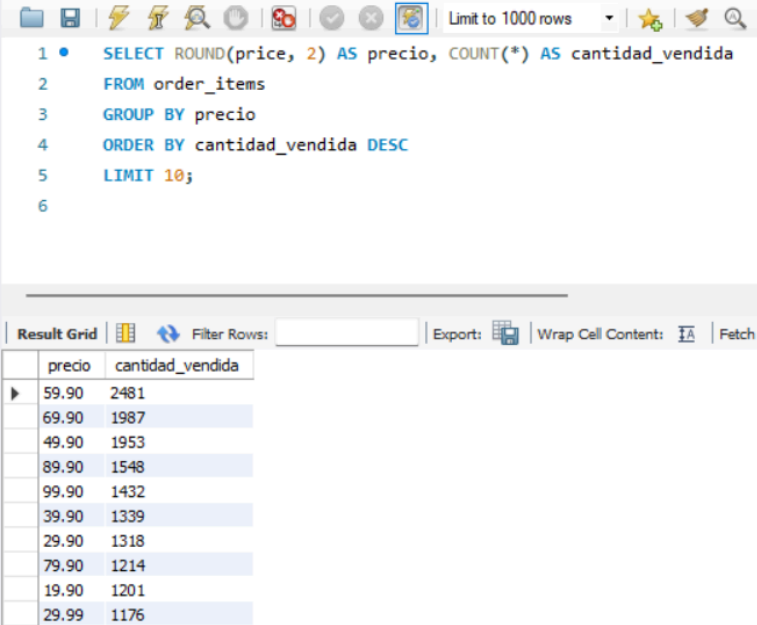
La ciudad con mayor número de compras es São Paulo, con 15,540 pedidos, lo que representa una clara concentración del mercado en esta área.

- Río de Janeiro (6,882 compras) se posiciona en el segundo lugar, aunque con menos de la mitad de los pedidos de São Paulo.
- Belo Horizonte (2,773 compras) y Brasília (2,131 compras) ocupan las siguientes posiciones, mostrando una menor participación en las compras.
- Las tres principales ciudades (São Paulo, Río de Janeiro y Belo Horizonte) concentran más del 50% de las compras en este ranking.

♦ Relación Entre Precio y Ventas

```
SELECT FLOOR(price/10)*10 AS rango_precio, COUNT(order_id) AS total_pedidos
FROM order_items
GROUP BY rango_precio
ORDER BY rango_precio;
```

📌 Identificar los rangos de precios más populares.



The screenshot shows a SQL query editor with the following query:

```
1 • SELECT ROUND(price, 2) AS precio, COUNT(*) AS cantidad_vendida
2 FROM order_items
3 GROUP BY precio
4 ORDER BY cantidad_vendida DESC
5 LIMIT 10;
6
```

Below the query, the results are displayed in a table with columns 'precio' and 'cantidad_vendida'.

precio	cantidad_vendida
59.90	2481
69.90	1987
49.90	1953
89.90	1548
99.90	1432
39.90	1339
29.90	1318
79.90	1214
19.90	1201
29.99	1176


Conclusión principal:

El precio más común en las ventas es 59.90, con 2,481 unidades vendidas, seguido de cerca por precios como 69.90 y 79.90, con más de 1,900 unidades cada uno. Esto sugiere que los productos en estos rangos de precio tienen una mayor demanda.

♦ Black Friday y Navidad: Impacto en Ventas

```
SELECT MONTH(order_purchase_timestamp) AS mes, COUNT(order_id) AS total_pedidos
FROM orders
WHERE MONTH(order_purchase_timestamp) IN (11, 12)
GROUP BY mes
ORDER BY total_pedidos DESC;
```

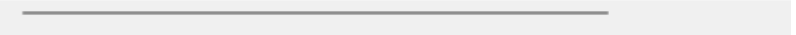
📌 Analizar si las ventas aumentan en noviembre y diciembre debido a Black Friday y Navidad.



```

1 • SELECT DATE_FORMAT(order_purchase_timestamp, '%Y-%m') AS mes,
2     COUNT(*) AS total_pedidos
3 FROM orders
4 WHERE MONTH(order_purchase_timestamp) IN (11, 12)
5 GROUP BY mes
6 ORDER BY mes;
7

```



mes	total_pedidos
2016-12	1
2017-11	7544
2017-12	5673

Conclusión principal:

El mes de noviembre de 2017 presenta un pico de ventas significativo, con 7,544 pedidos, lo que sugiere un impacto positivo del Black Friday. En diciembre, las ventas bajan ligeramente a 5,673 pedidos, pero siguen siendo altas, probablemente debido a las compras navideñas.


♦ Otras Temporadas Fuertes del Año

```

SELECT MONTH(order_purchase_timestamp) AS mes, COUNT(order_id) AS total_pedidos
FROM orders
GROUP BY mes
ORDER BY total_pedidos DESC;

```


📌 Identificar si existen otros picos de ventas en diferentes épocas del año.



```

1 • SELECT DATE_FORMAT(order_purchase_timestamp, '%m') AS mes,
2     COUNT(*) AS total_pedidos
3 FROM orders
4 GROUP BY mes
5 ORDER BY total_pedidos DESC;
6

```



mes	total_pedidos
08	10843
05	10573
07	10318
03	9893
06	9412
04	9343
02	8508
01	8069
11	7544
12	5674
10	4959
09	4305


Conclusión principal:

Los meses con mayor número de pedidos son agosto (10,843 pedidos), julio (10,573 pedidos) y mayo (10,318 pedidos). Esto sugiere que los picos de venta no solo ocurren en noviembre (Black Friday) y diciembre (Navidad), sino también en otros momentos del año.

Datos clave:


- Agosto, julio y mayo son los meses con más pedidos, lo que podría estar relacionado con promociones de mitad de año o periodos de vacaciones en Brasil.
- Noviembre (7,544 pedidos) sigue destacando, aunque no es el mes con más ventas, lo que refuerza la idea de que hay otros picos de demanda importantes en el año.
- Septiembre y octubre tienen los registros más bajos (4,995 y 4,305 pedidos, respectivamente), lo que indica una posible desaceleración en estas fechas.

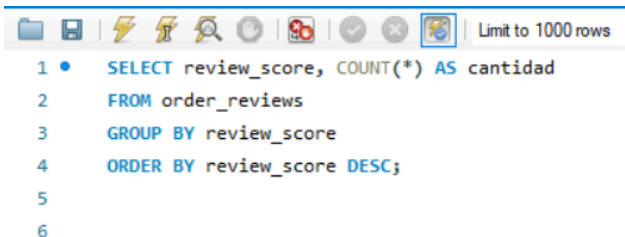
4 Impacto de las Entregas en la Satisfacción del Cliente

 **Objetivo:** Evaluar si los tiempos de entrega afectan la calificación del pedido.

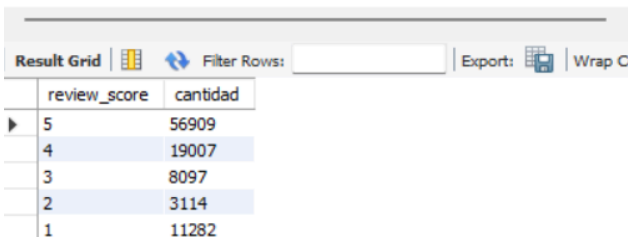
♦ Impacto de las Reseñas en las Ventas

```
SELECT review_score, COUNT(order_id) AS total_reviews
FROM order_reviews
GROUP BY review_score
ORDER BY review_score DESC;
```

 Evaluar cómo se distribuyen las reseñas y su impacto en la satisfacción del cliente.



```
1 • SELECT review_score, COUNT(*) AS cantidad
2   FROM order_reviews
3  GROUP BY review_score
4  ORDER BY review_score DESC;
5
6
```



review_score	cantidad
5	56909
4	19007
3	8097
2	3114
1	11282

Conclusión principal:

La mayoría de las reseñas son positivas, con calificaciones de 5 estrellas representando el 57.3% del total, lo que indica un alto nivel de satisfacción general entre los clientes.

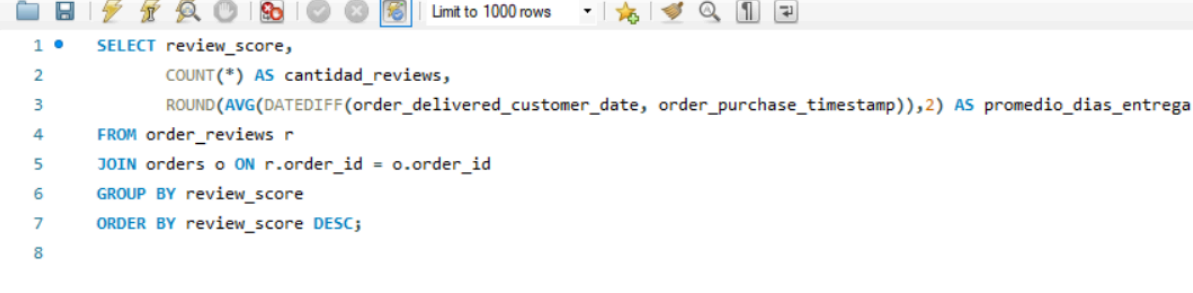
Datos clave:

- Calificación de 5 estrellas: 56,909 reseñas (57.3% del total)
- Calificación de 4 estrellas: 19,007 reseñas (19.1%)
- Calificación de 3 estrellas: 8,097 reseñas (8.1%)
- Calificación de 2 estrellas: 3,114 reseñas (3.1%)
- Calificación de 1 estrella: 11,282 reseñas (11.3%)

♦ Relación entre Tiempo de Entrega y Reseñas


```
SELECT DATEDIFF(order_delivered_customer_date, order_purchase_timestamp) AS
dias_entrega, ROUND(AVG(review_score), 2) AS promedio_review
FROM orders
JOIN order_reviews ON orders.order_id = order_reviews.order_id
GROUP BY dias_entrega
ORDER BY dias_entrega;
```

📌 Comparación de satisfacción en entregas rápidas vs. lentas.



```
1 • SELECT review_score,
2       COUNT(*) AS cantidad_reviews,
3       ROUND(AVG(DATEDIFF(order_delivered_customer_date, order_purchase_timestamp)),2) AS promedio_dias_entrega
4 FROM order_reviews r
5 JOIN orders o ON r.order_id = o.order_id
6 GROUP BY review_score
7 ORDER BY review_score DESC;
8
```



	review_score	cantidad_reviews	promedio_dias_entrega
▶	5	56909	10.62
	4	19007	12.25
	3	8097	14.20
	2	3114	16.60
	1	11282	21.29

 **Objetivo:** Evaluar si los tiempos de entrega afectan la calificación del pedido.

Conclusión principal:

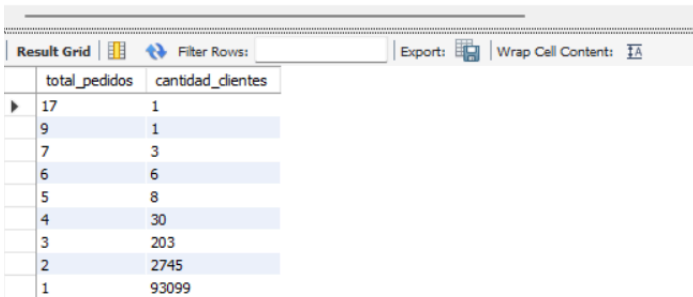
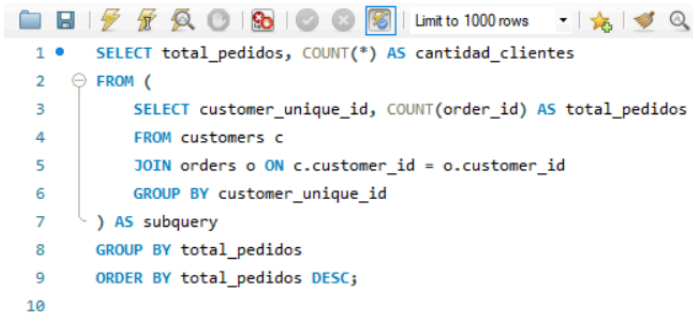
Existe una relación clara entre los tiempos de entrega y la satisfacción del cliente. A menor tiempo de entrega, mayor es la calificación del pedido.

5 Análisis de Clientes y Fidelización

 **Objetivo:** Identificar si Olist tiene clientes recurrentes o depende de nuevos compradores.

♦ Clientes recurrentes vs. Clientes de una Sola Compra

```
SELECT total_pedidos, COUNT(*) AS cantidad_clientes
FROM (
SELECT customer_unique_id, COUNT(order_id) AS total_pedidos
  FROM customers c
  JOIN orders o ON c.customer_id = o.customer_id
  GROUP BY customer_unique_id
) AS subquery
GROUP BY total_pedidos
ORDER BY total_pedidos DESC;
```



	total_pedidos	cantidad_clientes
▶	17	1
	9	1
	7	3
	6	6
	5	8
	4	30
	3	203
	2	2745
	1	93099

Conclusión principal:




La mayoría de los clientes en Olist son compradores únicos, con una mínima proporción de clientes recurrentes.

Datos clave:

- 93,099 clientes realizaron solo 1 compra (más del 93% del total).
- 2,745 clientes hicieron 2 compras.
- 203 clientes realizaron 3 compras.
- Solo 30 clientes han realizado 4 compras.
- Apenas 6 clientes han hecho 5 compras.
- Un solo cliente ha comprado 17 veces.

♦ Análisis de Fidelización de Clientes

```
SELECT
CASE
    WHEN total_pedidos = 1 THEN 'Clientes de una sola compra'
    ELSE 'Clientes recurrentes'
END AS tipo_cliente,
COUNT(DISTINCT customer_unique_id) AS cantidad_clientes,
SUM(price) AS ingresos_totales,
ROUND(SUM(price) * 100 / (SELECT SUM(price) FROM order_items), 2) AS
porcentaje_ingresos
FROM (
    SELECT
        o.customer_unique_id,
        COUNT(o.order_id) AS total_pedidos,
        SUM(oi.price) AS price
    FROM orders o
    JOIN order_items oi ON o.order_id = oi.order_id
    GROUP BY o.customer_unique_id
) AS subquery
GROUP BY tipo_cliente;
```

Result Grid  Filter Rows: <input type="text"/> Export:  Wrap Cell Content: 				
	tipo_cliente	cantidad_clientes	ingresos_totales	porcentaje_ingresos
▶	Clientes de una sola compra	83551	10990284.22	80.86
	Clientes recurrentes	11869	2601359.48	19.14

Conclusión principal:

Los clientes de una sola compra generan el 80.86% de los ingresos, mientras que los clientes recurrentes contribuyen solo con el 19.14%.

Datos clave:

- 83,551 clientes realizaron solo una compra, generando aproximadamente 10.99 millones en ingresos.
- 11,869 clientes son recurrentes, aportando 2.6 millones en ingresos.
- Aunque los clientes recurrentes son solo el 12.4% del total, representan casi el 19.14% de los ingresos.

◆ Comparación entre Tasa de Recompra y Ticket Promedio por Categoría

Consulta de recompra por categoría

```
SELECT
  p.product_category_name,
  COUNT(DISTINCT CASE WHEN total_pedidos > 1 THEN o.customer_unique_id END)
AS clientes_recurrentes,
  COUNT(DISTINCT o.customer_unique_id) AS total_clientes,
  ROUND(
    COUNT(DISTINCT CASE WHEN total_pedidos > 1 THEN o.customer_unique_id END)
    * 100.0 /
    COUNT(DISTINCT o.customer_unique_id), 2
  ) AS porcentaje_recurrentes
FROM orders o
JOIN order_items oi ON o.order_id = oi.order_id
JOIN products p ON oi.product_id = p.product_id
JOIN (
  SELECT customer_unique_id, COUNT(order_id) AS total_pedidos
  FROM orders
  GROUP BY customer_unique_id
) AS subquery ON o.customer_unique_id = subquery.customer_unique_id
GROUP BY p.product_category_name
ORDER BY porcentaje_recurrentes DESC;
```

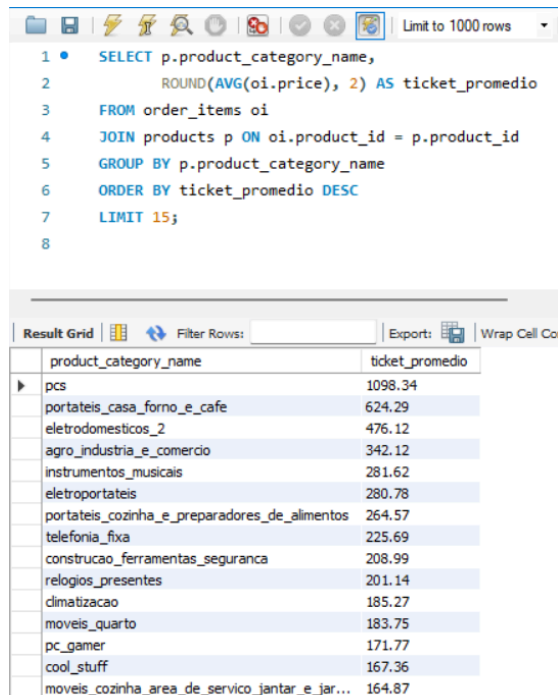
	product_category_name	clientes_recurrentes	total_clientes	porcentaje_recurrentes
▶	la_cuisine	2	13	15.38
	artesanato	3	21	14.29
	fashion_roupa_feminina	5	39	12.82
	fashion_roupa_infanto_juvenil	1	8	12.50
	pc_gamer	1	8	12.50
	eletrodomesticos	77	703	10.95
	moveis_colchao_e_estofado	4	38	10.53
	musica	4	38	10.53
	moveis_quarto	9	91	9.89
	fashion_bolsas_e_acessorios	164	1798	9.12
	fashion_roupa_masculina	10	110	9.09
	fashion_calcados	19	238	7.98
	climatizacao	20	251	7.97
	bebidas	23	292	7.88
	casa_conforto	31	396	7.83

Interpretaciones:

- Productos de consumo frecuente o reemplazo como moda, accesorios y electrodomésticos parecen tener mayor tasa de recompra.
- Baja retención en productos de compra única o duraderos como muebles y artículos de decoración.

Consulta de ticket promedio por categoría

```
SELECT p.product_category_name,  
       ROUND(AVG(oi.price), 2) AS ticket_promedio  
FROM order_items oi  
JOIN products p ON oi.product_id = p.product_id  
GROUP BY p.product_category_name  
ORDER BY ticket_promedio DESC  
LIMIT 15;
```



SQL Query:

```
1 • SELECT p.product_category_name,  
2       ROUND(AVG(oi.price), 2) AS ticket_promedio  
3 FROM order_items oi  
4 JOIN products p ON oi.product_id = p.product_id  
5 GROUP BY p.product_category_name  
6 ORDER BY ticket_promedio DESC  
7 LIMIT 15;  
8
```


Result Grid:

product_category_name	ticket_promedio
pcs	1098.34
portateis_casa_forno_e_cafe	624.29
eletrodomesticos_2	476.12
agro_industria_e_comercio	342.12
instrumentos_musicais	281.62
eletroportateis	280.78
portateis_cozinha_e_preparadores_de_alimentos	264.57
telefonias_fixas	225.69
construcao_ferramentas_seguranca	208.99
relogios_presentes	201.14
climatizacao	185.27
moveis_quarto	183.75
pc_gamer	171.77
cool_stuff	167.36
moveis_cozinha_area_de_servico_jantar_e_jar...	164.87

Interpretaciones:

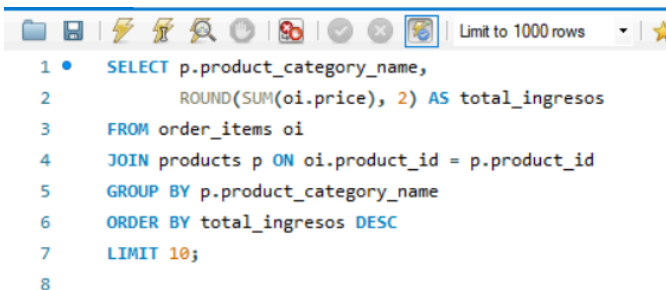
- Equipos electrónicos y electrodomésticos tienen un ticket promedio significativamente alto, lo que indica que son productos costosos.
- Categorías relacionadas con tecnología y música tienen precios elevados, lo que puede reflejar una menor frecuencia de compra pero un alto valor por transacción.
- Algunas categorías con menor ticket promedio, como muebles y utensilios para el hogar, pueden tener un margen de rentabilidad más bajo.

6 Análisis de Factores Clave en las Ventas

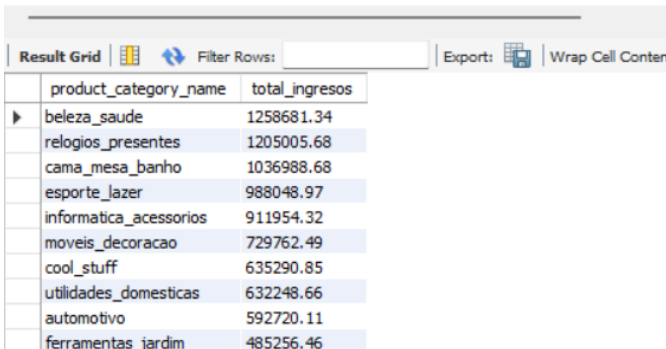
 **Objetivo:** Analizar tendencias según distintos factores.

♦ Análisis de Productos Más y Menos Rentables

```
SELECT p.product_category_name,  
       ROUND(SUM(oi.price), 2) AS total_ingresos  
FROM order_items oi  
JOIN products p ON oi.product_id = p.product_id  
GROUP BY p.product_category_name  
ORDER BY total_ingresos DESC  
LIMIT 10;
```



```
1 • SELECT p.product_category_name,  
2       ROUND(SUM(oi.price), 2) AS total_ingresos  
3 FROM order_items oi  
4 JOIN products p ON oi.product_id = p.product_id  
5 GROUP BY p.product_category_name  
6 ORDER BY total_ingresos DESC  
7 LIMIT 10;  
8
```



product_category_name	total_ingresos
beleza_saude	1258681.34
relogios_presentes	1205005.68
cama_mesa_banho	1036988.68
esporte_lazer	988048.97
informatica_acessorios	911954.32
moveis_decoracao	729762.49
cool_stuff	635290.85
utilidades_domesticas	632248.66
automotivo	592720.11
ferramentas_jardim	485256.46

Categorías con mayores ingresos totales:

- "beleza_saude": 1,258,681.34 BRL
- "relogios_presentes": 1,205,005.68 BRL
- "cama_mesa_banho": 1,080,368.68 BRL
- "esporte_lazer": 988,048.97 BRL
- "informatica_acessorios": 911,954.32 BRL

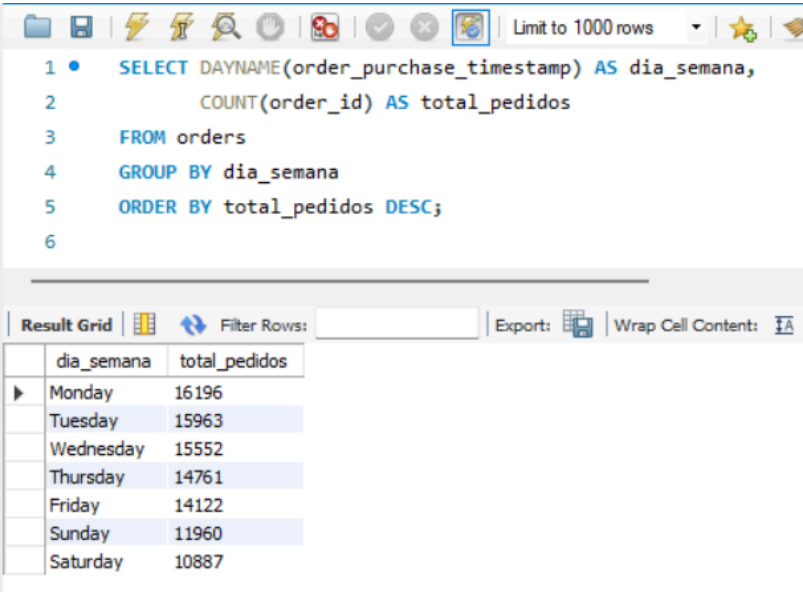
Interpretaciones:

- Productos de belleza y salud ("beleza_saude") lideran en ingresos, lo que indica una alta demanda en este segmento.
- La categoría de "relogios_presentes" (relojes y regalos) tiene un volumen de ventas significativo, probablemente impulsado por compras en fechas especiales.

- Las categorías de hogar y mobiliario ("cama_mesa_banho", "moveis_decoracao") están entre las más rentables, reflejando una fuerte demanda en este sector.
- Los productos de informática y accesorios ("informatica_acessorios") generan ingresos elevados, lo que sugiere una combinación de alto ticket promedio y volumen de ventas.

♦ Días de la Semana con Más Ventas

```
SELECT DAYNAME(order_purchase_timestamp) AS dia_semana,
       COUNT(order_id) AS total_pedidos
FROM orders
GROUP BY dia_semana
ORDER BY total_pedidos DESC;
```



The screenshot shows a SQL query execution interface. The query is:


```
1 SELECT DAYNAME(order_purchase_timestamp) AS dia_semana,
2       COUNT(order_id) AS total_pedidos
3 FROM orders
4 GROUP BY dia_semana
5 ORDER BY total_pedidos DESC;
6
```

 The results are displayed in a table with two columns: 'dia_semana' and 'total_pedidos'. The rows are ordered by 'total_pedidos' in descending order.

dia_semana	total_pedidos
Monday	16196
Tuesday	15963
Wednesday	15552
Thursday	14761
Friday	14122
Sunday	11960
Saturday	10887

Días con mayor cantidad de pedidos:

- Lunes: 16,196 pedidos (día con mayor actividad).
- Martes: 15,963 pedidos.
- Miércoles: 15,552 pedidos.

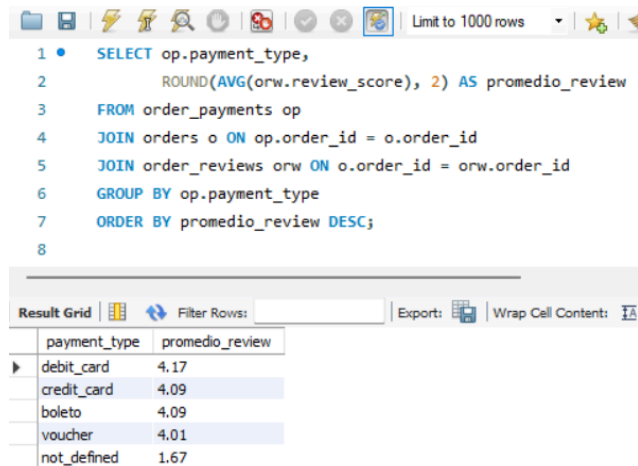
Días con menor cantidad de pedidos:

- Sábado: 10,887 pedidos (día con menor actividad).
- Domingo: 11,960 pedidos.

♦ Métodos de Pago y su Relación con la Satisfacción

```
SELECT payment_type, ROUND(AVG(review_score), 2) AS promedio_review
FROM order_payments
JOIN order_reviews ON order_payments.order_id = order_reviews.order_id
GROUP BY payment_type
ORDER BY promedio_review DESC;
```

📌 ¿Ciertos métodos de pago están relacionados con mejores reseñas?



```

1 • SELECT op.payment_type,
2     ROUND(AVG(orw.review_score), 2) AS promedio_review
3 FROM order_payments op
4 JOIN orders o ON op.order_id = o.order_id
5 JOIN order_reviews orw ON o.order_id = orw.order_id
6 GROUP BY op.payment_type
7 ORDER BY promedio_review DESC;
8

```

payment_type	promedio_review
debit_card	4.17
credit_card	4.09
boleto	4.09
voucher	4.01
not_defined	1.67

Métodos de pago con mejor calificación promedio:

- Tarjeta de débito: 4.17 (mayor satisfacción).
- Tarjeta de crédito: 4.09.
- Boleto bancario: 4.09.
- Voucher: 4.01.

Método de pago con peor calificación:

- "Not defined": 1.67 (posibles problemas con pedidos no completados o datos mal registrados).

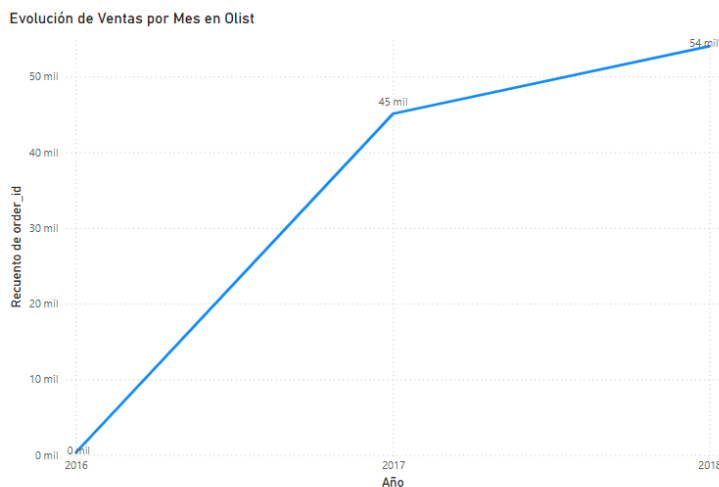
7 Visualización y Comparación SQL vs. Power BI

 **Objetivo:** Mostrar cómo Power BI complementa el análisis en SQL.

♦ Principales Visualizaciones en Power BI

A continuación, se presentan las principales visualizaciones realizadas en Power BI, con el objetivo de analizar patrones de ventas, tendencias de clientes y la relación entre diversas métricas.

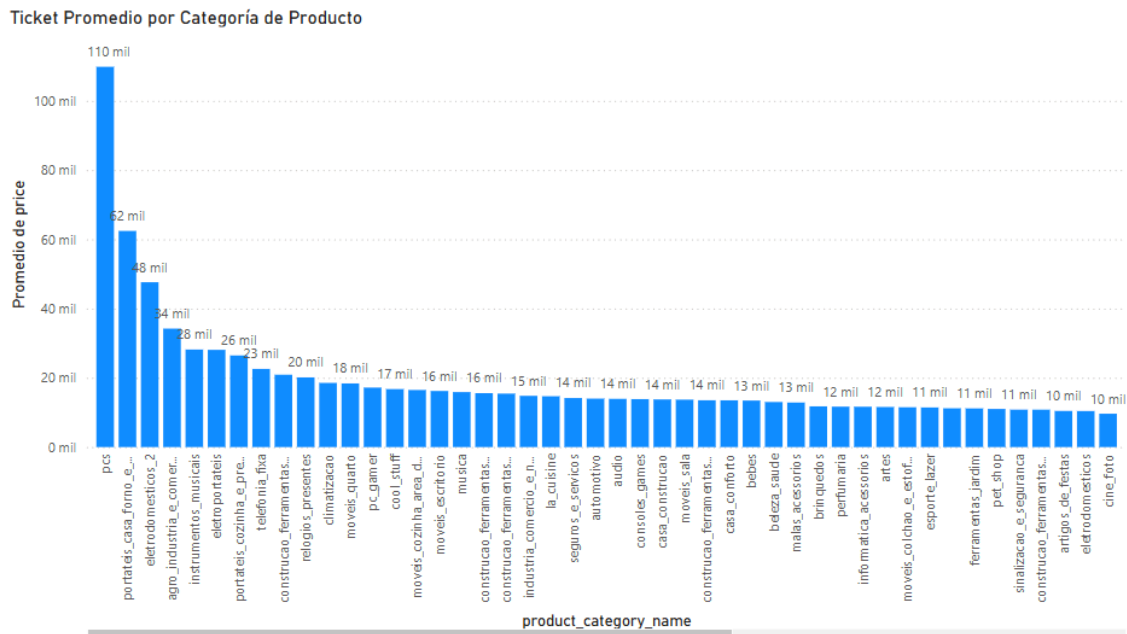
Primera Visualización: Evolución de Ventas por Mes



Evolución de Ventas por Mes en Olist

- Muestra el crecimiento de las ventas a lo largo del tiempo.
- Se evidencia un aumento progresivo en la cantidad de pedidos.

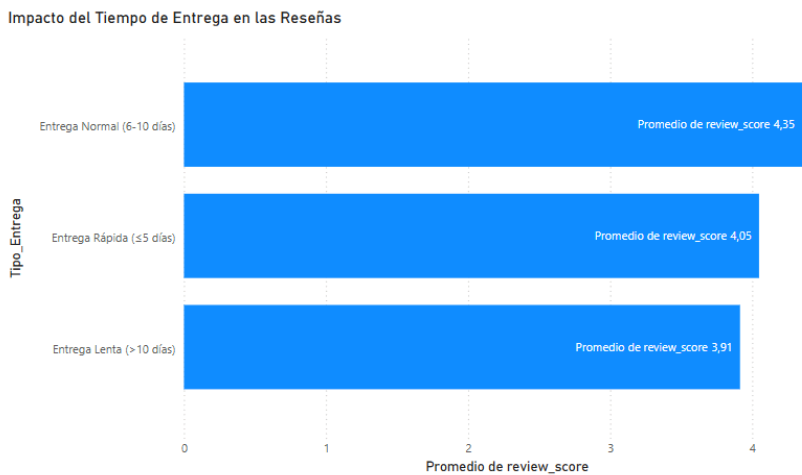
Segunda Visualización: Ticket Promedio por Categoría de Producto



Ticket Promedio por Categoría de Producto

- Se observa que algunas categorías tienen un ticket promedio significativamente más alto que otras.
- Destaca la categoría "pcs" con el mayor ticket promedio.

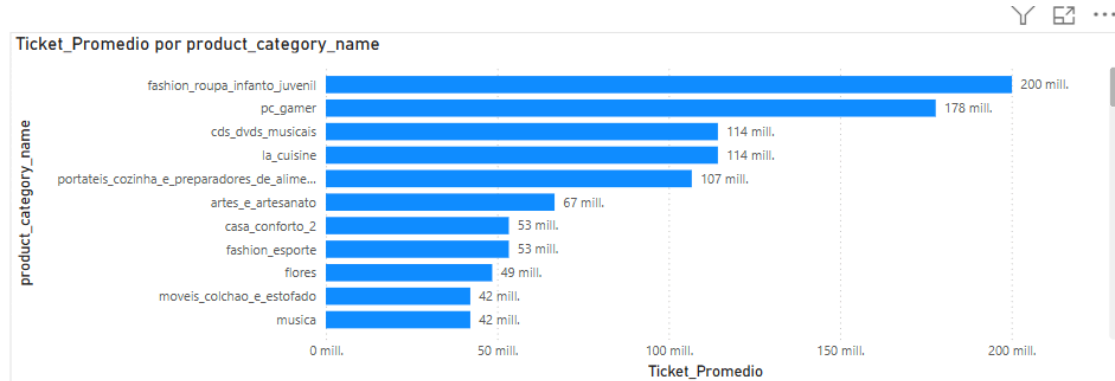
Tercera Visualización: Impacto de las Entregas en las Reseñas



Impacto del Tiempo de Entrega en la Satisfacción del Cliente

- Se verifica que las entregas más rápidas tienden a recibir mejores calificaciones.
- Los pedidos con entregas en más de 10 días tienen una calificación promedio más baja.

Cuarta visualización: Análisis de Ticket Promedio por Categoría de Producto

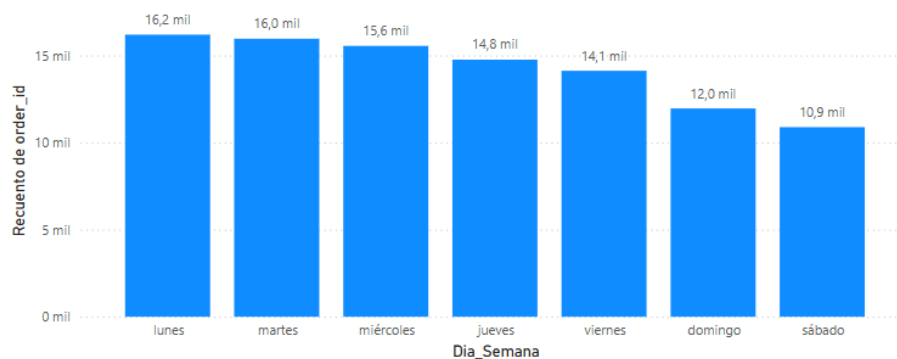


Categorías con Ticket Promedio Más Alto

- "fashion_roupa_infanto_juvenil" lidera con un ticket promedio de 200 mil.
- "pc_gamer" ocupa el segundo lugar con 178 mil.
- "cds_dvds_musicais" y "la_cuisine" comparten el tercer puesto con 114 mil cada uno.

Gráfico de Ventas por Día de la Semana

Total de Pedidos por Día de la Semana

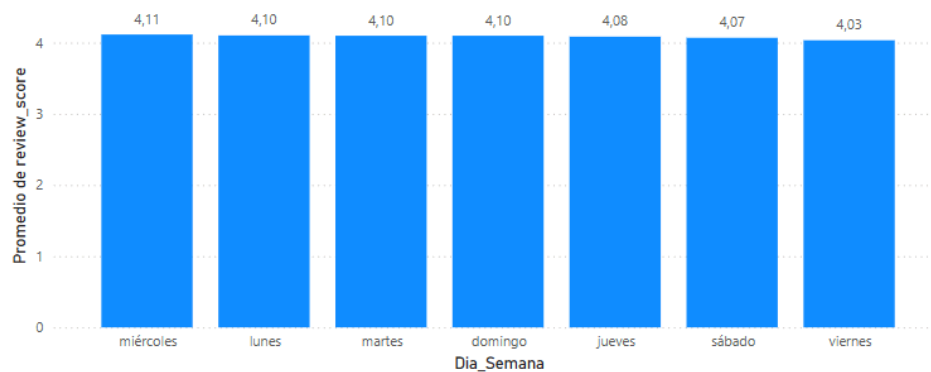


Total de Pedidos por Día de la Semana

- Los días lunes y martes registran el mayor volumen de pedidos.
- Sábados y domingos presentan una menor actividad comercial.

Promedio de Reseñas por Día de la Semana

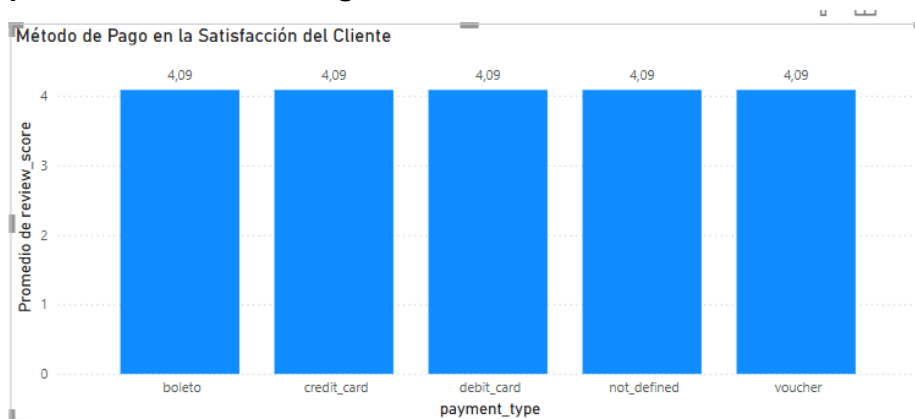
Total de Pedidos por Día de la Semana



Promedio de Reseñas por Día de la Semana

- Las calificaciones de los clientes se mantienen relativamente estables sin grandes variaciones.
- No se observan diferencias significativas entre los distintos días.

Impacto del Método de Pago en la Satisfacción del Cliente



Método de Pago en la Satisfacción del Cliente

- No hay variaciones notables en la satisfacción del cliente dependiendo del método de pago.
- Todos los métodos muestran un promedio de reseñas similar.

♦ Comparación de Resultados entre SQL y Power BI


- SQL permitió realizar análisis detallados sobre tendencias de ventas, satisfacción del cliente y comportamiento de compra mediante consultas específicas.
- Power BI complementa el análisis al proporcionar visualizaciones interactivas que facilitan la interpretación de los datos.
- Diferencias clave:

SQL: Ideal para exploración detallada de datos, cálculos exactos y extracción de insights específicos.

Power BI: Excelente para visualizar tendencias, identificar patrones y comunicar hallazgos de manera intuitiva.

Con esta comparación, se concluye que ambas herramientas se complementan en el análisis de datos de e-commerce, permitiendo obtener conclusiones más precisas y visualmente accesibles.

8 Conclusiones Finales y Recomendaciones

 **Objetivo:** Resumir hallazgos clave y proponer mejoras basadas en los datos.

Resumen de Hallazgos

- ◆ **Tendencias de Ventas y Estacionalidad**
 - Se identificó un crecimiento sostenido en las ventas de Olist entre 2016 y 2018, con una tendencia al alza en la cantidad de pedidos.
 - Se observó un incremento de ventas en noviembre, lo que confirma la influencia del Black Friday en el comercio electrónico. Sin embargo, diciembre no mostró un aumento tan marcado, lo que sugiere que la campaña navideña no es tan relevante para esta plataforma.
 - Los meses con mayor volumen de ventas fueron julio y agosto, lo que sugiere que existen otros picos de demanda más allá del Black Friday.
- ◆ **Impacto de los Tiempos de Entrega en la Satisfacción del Cliente**
 - Existe una correlación clara entre tiempos de entrega y calificación de los clientes.
 - Los pedidos entregados en menos de 5 días tienen una calificación promedio de 4.05, mientras que los que tardan más de 10 días reciben una calificación media de 3.91, lo que representa una diferencia significativa.
 - Las entregas dentro de un rango de 6 a 10 días presentan la mejor calificación promedio (4.35), lo que sugiere que no siempre la entrega más rápida es la mejor percibida por los clientes.
- ◆ **Clientes Recurrentes vs. Nuevos**
 - El 80.86% de los ingresos proviene de clientes que solo han realizado una compra, mientras que solo el 19.14% proviene de clientes recurrentes.
 - La tasa de recompra es baja, lo que indica una alta dependencia de nuevos clientes en lugar de una base consolidada de compradores recurrentes.
 - La recompra varía según la categoría de producto; algunas como electrodomésticos y moda tienen mayor retención de clientes en comparación con otras categorías.

♦ Categorías de Productos Más Rentables

- Las categorías con mayores ingresos son "beleza_saude" (salud y belleza), "cama_mesa_banho" (ropa de cama y baño) y "relogios_presentes" (relojes y regalos).
 - Sin embargo, no todas las categorías con alto volumen de ventas generan los mayores ingresos.
 - El ticket promedio varía considerablemente, con productos electrónicos y de informática teniendo valores más altos que moda u hogar.
-

Recomendaciones Estratégicas

♦ Optimización del rendimiento en los días con menos ventas

Implementar estrategias promocionales específicas para los fines de semana, como descuentos flash o campañas de email marketing, para incentivar las compras en estos días con menor actividad.

♦ Segmentación de campañas por categoría de producto

Para categorías con ticket promedio bajo pero alto volumen de ventas, explorar estrategias de "venta cruzada" recomendando productos complementarios en el checkout. En categorías con ticket alto pero bajo volumen de ventas, como tecnología o informática, se podría implementar financiación a plazos o promociones con envío gratuito para incentivar la compra.

♦ Optimización de envíos en regiones con alta demanda

Identificar las ciudades con más pedidos (como São Paulo y Río de Janeiro) y establecer centros de distribución más eficientes en estas zonas para reducir tiempos de entrega y costos logísticos.

♦ Estrategias para mejorar la retención de clientes

Dado que el 80% de los clientes son de una sola compra, una estrategia clave es crear programas de recompra mediante descuentos o beneficios exclusivos para quienes vuelvan a comprar dentro de un período determinado. Se podría implementar un sistema de puntos, donde cada compra genera puntos canjeables por descuentos en futuras compras.

♦ Optimización de los Tiempos de Entrega

- Reducir la cantidad de pedidos con tiempos de entrega superiores a 10 días, ya que impactan negativamente en la satisfacción del cliente.
- Priorizar entregas eficientes dentro del rango de 6 a 10 días, que es el punto óptimo según las calificaciones de los clientes.
- Explorar alianzas con proveedores logísticos para mejorar los tiempos de entrega en regiones con mayores demoras.

♦ Estrategias de Precios Según Categoría

- Optimizar la estrategia de precios para categorías con ticket promedio alto, asegurando que se mantenga un balance entre volumen de ventas y rentabilidad.
 - Para categorías con bajo ticket promedio pero alto volumen de ventas, considerar estrategias de ventas cruzadas y bundles de productos.
 - Ajustar campañas de marketing en función de las temporadas más rentables, enfocándose no solo en Black Friday, sino también en julio y agosto, que mostraron altos volúmenes de ventas.
-