OLIST E-COMMERCE DATA ANALYSIS USING SQL & EXCEL

View the full project and dashboard on GitHub: https://github.com/aswin-panengal/olist-sales-analysis-sql-excel

Business Objective

The goal of this project is to analyze customer behavior, seller performance, and product trends using the Brazilian Olist E-Commerce dataset. The insights aim to support better decision-making in areas like sales strategy, logistics optimization, and customer satisfaction.

Tools Used

- SQL Server for data analysis
- Excel for data visualization

Dataset Overview

- **olist_orders**: Contains order info like status and timestamps
- **olist_customers**: Customer demographics
- **olist_order_items**: Product-level order data
- **olist_products**: Product details
- **olist_product_category**: Category translations
- **olist_sellers**: Seller information
- **olist_payments**: Payment details
- **olist_reviews**: Customer reviews

Key Business Questions & SQL Insights

Metric	Value
Total Orders	99441
Total Customers	99441
Total Products	32951
Total Sellers	3095

Top Selling Product Categories

Results Messages

	category_name	total_orders
1	bed_bath_table	11115
2	health_beauty	9670
3	sports_leisure	8641
4	furniture_decor	8334
5	computers_accessories	7827
6	housewares	6964
7	watches_gifts	5991
8	telephony	4545
9	garden_tools	4347
10	auto	4235
11	toys	4117
12	cool_stuff	3796
13	perfumery	3419
14	baby	3065
15	electronics	2767
16	stationery	2517

Top Categories:

- Bed & Bath
- Health & Beauty
- Sports & Leisure
- Computers & Accessories
- Furniture

Average Payment Value per Order

	avg_payment_per_order
1	154.1

The average payment per order is around **154.1**, indicating general customer spending.

Average Delivery Time

	avg_delivery_days
1	12

Average delivery time is **12 days** — this can be used to assess logistics performance.

Top Sellers by Rating

	seller_id	avg_rating	total_reviews
1	48efc9d94a9834137efd9ea76b065a38	5	34
2	a08692680c77d30a0b4280da5df01c5a	5	17
3	0b36063d5818f81ccb94b54adfaebbf5	5	15
4	c8c1bea22194a4eefa2dc9a9fa89f536	5	13
5	2addf05f476d0637864454e93ba673d5	5	12
6	404e1ba01358af4cd63f679b2c4d1fa1	5	12
7	b2eecf5ea250510da76590ca79d60e5d	5	11
8	0570350b23eda5444f8d1d9544932058	5	11
9	ab27bbbad5239bc31a34709275a70db4	5	11
10	cc419e0650a3c5ba77189a1882b7556a	4	1811
11	da8622b14eb17ae2831f4ac5b9dab84a	4	1568
12	955fee9216a65b617aa5c0531780ce60	4	1489
13	7a67c85e85bb2ce8582c35f2203ad736	4	1166
14	4869f7a5dfa277a7dca6462dcf3b52b2	4	1148

Sellers with high ratings and consistent reviews help build customer trust. Only sellers with >10 reviews are considered to ensure reliability.

Customer Distribution by State

	customer_state	total_customers
1	SP	41746
2	RJ	12852
3	MG	11635
4	RS	5466
5	PR	5045
6	SC	3637
7	BA	3380
8	DF	2140
9	ES	2033
10	GO	2020
11	PE	1652
12	CE	1336
13	PA	975
14	MT	907

Helps identify where most customers are from. Useful for region-wise campaigns or logistics hubs.

Revenue by Product Category

	product_category	total_revenue
1	health_beauty	1441248.07094423
2	watches_gifts	1305541.60778558
3	bed_bath_table	1241681.72044591
4	sports_leisure	1156656.47915162
5	computers_accessories	1059272.3975535
6	furniture_decor	902511.792542586
7	housewares	778397.771007238
8	cool_stuff	719329.951636389
9	auto	685384.320795551
10	garden_tools	584219.212224483
11	toys	561372.550566981
12	baby	480118.000149572
13	perfumery	453338.70862484
14	telephony	394883.319746591

Shows which product categories bring in the most revenue. Used later in Excel for visualization.

Monthly Order Trends

	order_month	total_orders
1	2016-09	4
2	2016-10	324
3	2016-12	1
4	2017-01	800
5	2017-02	1780
6	2017-03	2682
7	2017-04	2404
8	2017-05	3700
9	2017-06	3245
10	2017-07	4026
11	2017-08	4331
12	2017-09	4285
13	2017-10	4631
14	2017-11	7544

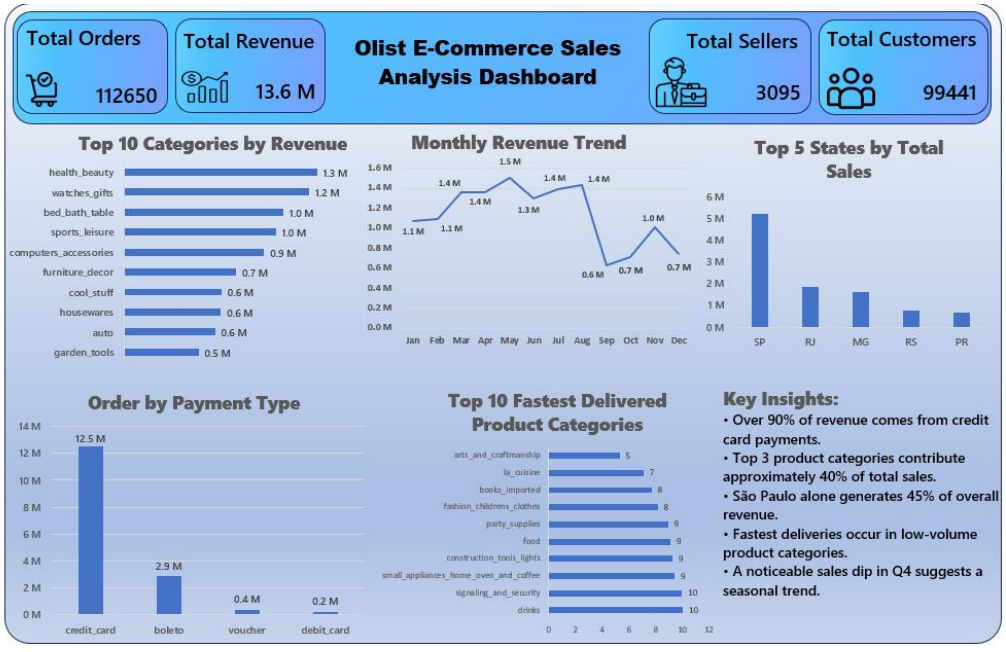
Reveals seasonal spikes or drops in orders. Can help optimize marketing and stock management.

Summary of SQL Insights

The SQL analysis uncovered valuable metrics such as top-selling categories, delivery delays, high-revenue products, and trusted sellers. These insights support strategic planning across marketing, sales, and logistics.

Excel Dashboard Visuals

1. **KPI Cards**
    - Total Orders
    - Total Revenue
    - Total Sellers
    - Total Customers
2. **Monthly Revenue Trend (Line Chart)**
    - Revenue movement over the year to detect seasonal patterns.
3. **Top 10 Categories by Revenue (Bar Chart)**
    - Highlights which product categories drive the highest sales.
4. **Order by Payment Type (Column Chart)**
    - Breakdown of payment methods (Credit Card dominant).
5. **Top 10 Fastest Delivered Product Categories (Bar Chart)**
    - Reflects efficiency in logistics.
6. **Top 5 States by Total Sales (Column Chart)**
    - Geographic contribution to total revenue.
7. **Key Insights Box**
    - Consolidated business-level conclusions.



Key Business Insights

- **Over 90%** of revenue is from **credit card payments**.
- **Top 3 categories** account for **~40%** of total sales.
- **São Paulo** alone contributes **45%+** of total revenue.
- Fastest deliveries are seen in **low-volume categories**.
- Noticeable **Q4 dip** may indicate seasonal variation or delivery challenges.

Challenges Faced

- Managing multiple joins with duplicate keys in SQL.
- Building relationships in Excel's Power Pivot due to non-unique fields.
- Matching product categories with translated names.

Outcome

This project demonstrates the ability to:

- Translate raw data into actionable insights.
- Use SQL to extract and prepare data.
- Build clean dashboards in Excel.
- Tell a compelling story for stakeholders.

Conclusion

This project demonstrates my ability to apply SQL and Excel to uncover key business insights from real-world e-commerce data. From data extraction to dashboard creation, I translated raw data into meaningful visual summaries and trends.

The outcome highlights core strengths in analysis, attention to detail, and visual storytelling making this a solid foundation in my data analytics journey and a showcase of my ability to deliver clean, insight-driven work.