E-commerce Data Insights: Order Trends, Payments & Customer Reviews

Author: Haythem Abouda

Date: 16/03/2025

Summary

This case study explores key trends in an e-commerce business, analyzing **order volumes, payment behaviors, cancellations, delivery delays, and customer reviews** to uncover actionable business insights. Using **SQL for data extraction, Python for advanced analysis, and Tableau for visualization**, we dive into the dataset to identify patterns that impact revenue and customer satisfaction.

📌 **Tableau Dashboard:**
You can explore the **interactive visualizations** [here](#).
Click through different views to analyze key insights!

📌 **SQL & Python Scripts:**
All queries, data cleaning, and analysis scripts are available in my **GitHub repository** [here](#).
This includes raw SQL queries and Python scripts for data cleaning and preprocessing.

Introduction

Business Problem

An e-commerce company wants to optimize its **sales, payments, and customer experience** by understanding:

- How order volumes & revenue fluctuate over time.
- Which payment methods generate the most revenue vs. cancellations.
- Which states experience the most delivery delays.
- What customers complain about the most in 1-star reviews.

Goals

- Analyze **trends in orders & revenue** to inform sales strategy.
- Identify **high-risk payment methods** linked to cancellations.
- Detecting **delivery inefficiencies** to improve logistics.
- Perform **NLP analysis on reviews** to extract key pain points.

Tools & Skills Used

- **SQL:** Data extraction, cleaning, aggregations, joins.
- **Python (Pandas, Matplotlib, NLP):** Advanced analysis & text mining.
- **Tableau:** Interactive data visualization & dashboarding.
- **Business Impact Analysis:** Translating data into actionable insights.

Exploratory Data Analysis (EDA)

Before diving into visualizations, we conducted a structured **Exploratory Data Analysis (EDA)** using SQL and Python to uncover key business insights. This phase helped us understand patterns, trends, and potential problem areas in the dataset.

Key Findings from SQL Analysis

Using SQL, we performed a series of queries to analyze different aspects of the dataset. Below are the main takeaways:

1 Order & Revenue Trends

- Monthly orders and revenue **steadily increased** from Jan 2017 to mid-2018, peaking around November 2017.
- Revenue follows the same trend as total orders, indicating **consistent pricing** with no major shifts in average order value (AOV).
- The **AOV remains stable** across the dataset, with slight fluctuations.

2 Payment Methods & Revenue Contribution

- **Credit cards dominate** transactions, contributing the highest revenue.
- **Boleto (bank slip)** is the second most common but has a slightly lower AOV.
- **Voucher payments** have the **lowest AOV**, suggesting they are used for discounts or promotional orders.

3 Order Cancellations by Payment Type

- **Credit cards account for the highest number of cancellations** (likely due to chargebacks or customer disputes).
- **Voucher payments have the highest cancellation rate (2%)**, indicating possible issues with coupon-based transactions.

4 Delivery Delays by State

- Certain states experience **longer delivery delays** (up to 20 days), impacting customer satisfaction.
- The worst-affected states are **Acre (AC), Roraima (RR), and Amazonas (AM)**, possibly due to **logistical challenges** in remote areas.

5 Customer Review Patterns

- The dataset is **heavily skewed towards 5-star ratings** (common in e-commerce).
- **1-star reviews** focus mainly on **product issues (49%)** and **delivery delays (11%)**.

Python Analysis: Customer Reviews

To further analyze negative reviews, we used **Python (NLTK)** to extract the **most common words** in 1-star reviews.

- The most frequent complaints include **"produto" (product), "veio" (came), "entrega" (delivery), and "prazo" (deadline)**.
- This confirms that **delivery delays and product quality issues** are the top pain points for customers.

Most Frequent Words in 1-Star Reviews
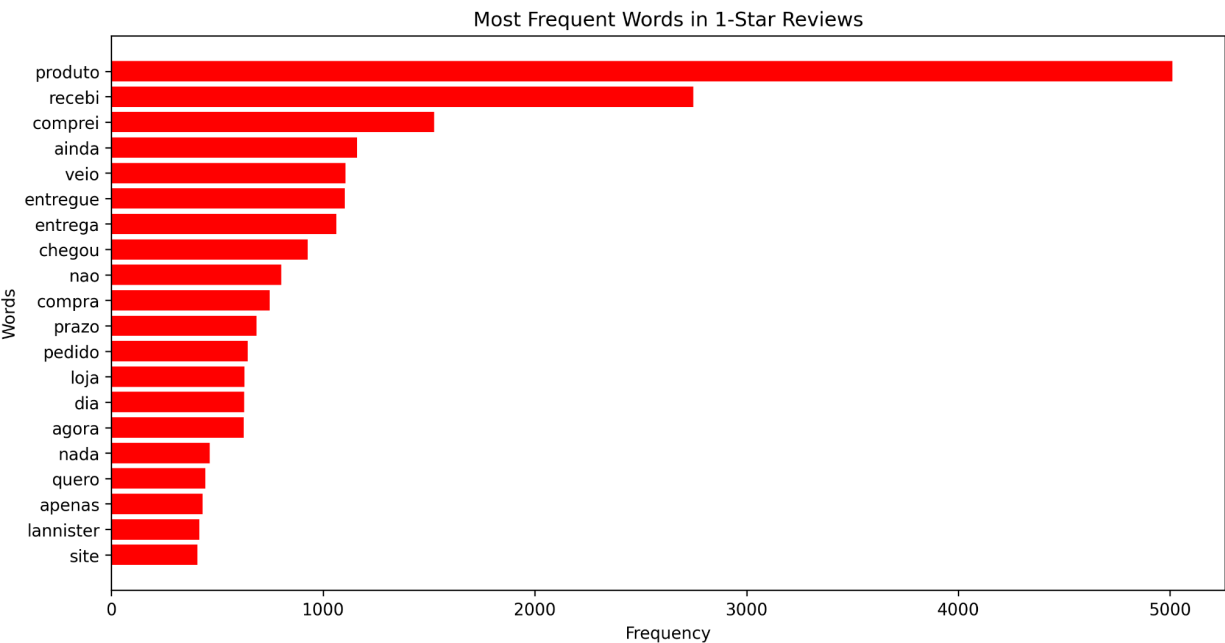
Preparing for Visualization

From the SQL & Python findings, we identified **six key areas** for visualization:

✅ **Order & Revenue Trends** (Total Orders, Total Revenue, AOV over time)
✅ **Payment Method Analysis** (Payment type distribution, AOV per method)
✅ **Cancellation Trends** (Cancellations & rates by payment type)
✅ **Delivery Delays** (Average delay by state)
✅ **Review Score Distribution** (Total count of ratings 1-5)
✅ **Review Issues Breakdown** (Major themes in 1-star reviews)

With these insights, we built a **Tableau dashboard** to present the findings in a clear, data-driven way.

Data Visualization

Revenue, Order Count, and AOV Trends

Revenue, AOV & Orders Over Time
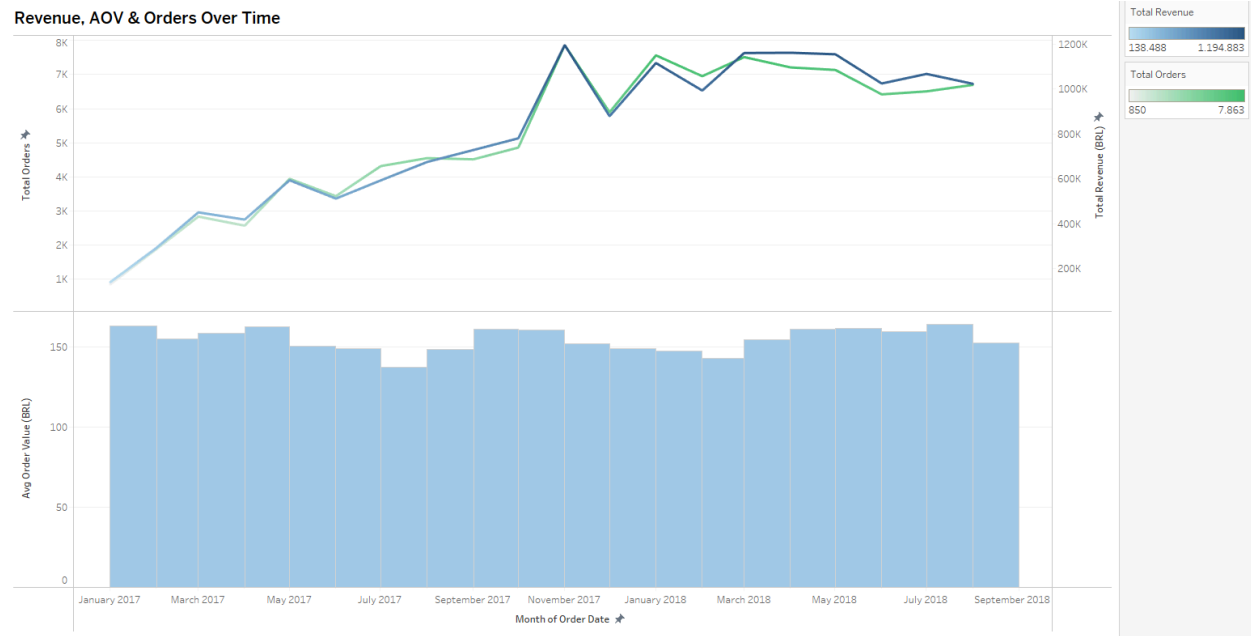
📌 Chart Summary:
This visualization shows the trend of monthly revenue, order volume, and average order value (AOV) from January 2017 to August 2018.
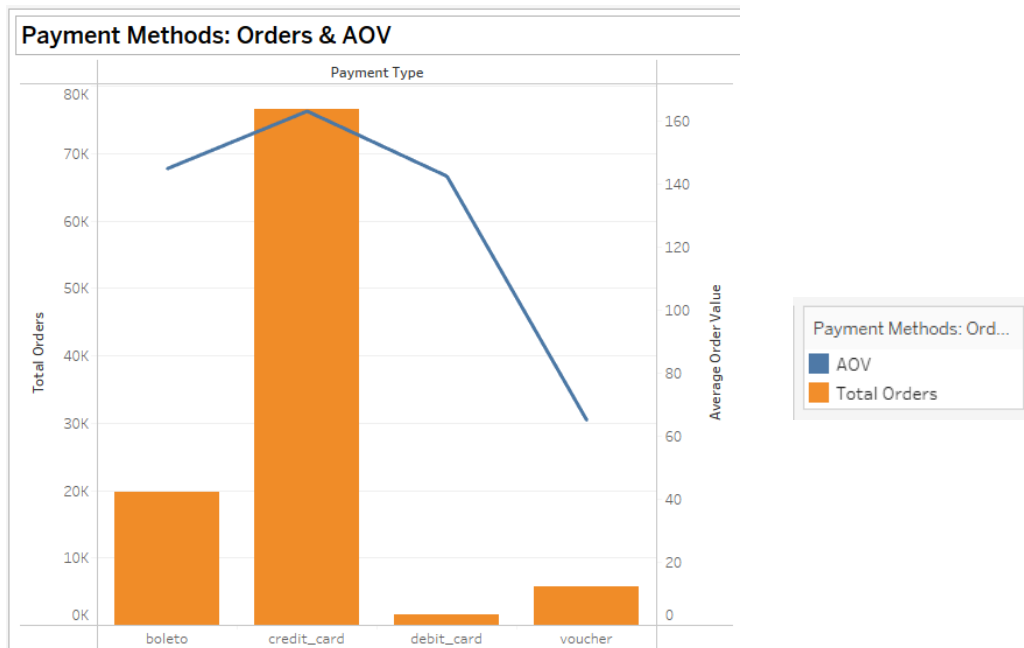
📊 Key Insights:

- Orders & revenue steadily increased, peaking in November 2017 (likely due to seasonal events like Black Friday).
- After the peak, the growth stabilized but remained strong.
- AOV remained consistent, showing pricing stability despite order volume fluctuations.

🧐 Business Implications:

- The peak in November suggests potential seasonal marketing opportunities.
- Stable AOV means customers spend about the same per order, so increasing order volume is key for revenue growth.

Payment Methods Analysis

Payment Methods: Orders & AOV

📌 Chart Summary:

This visualization breaks down **total orders, total revenue, and average order value (AOV)** by **payment method**, showing how different payment types contribute to overall sales.
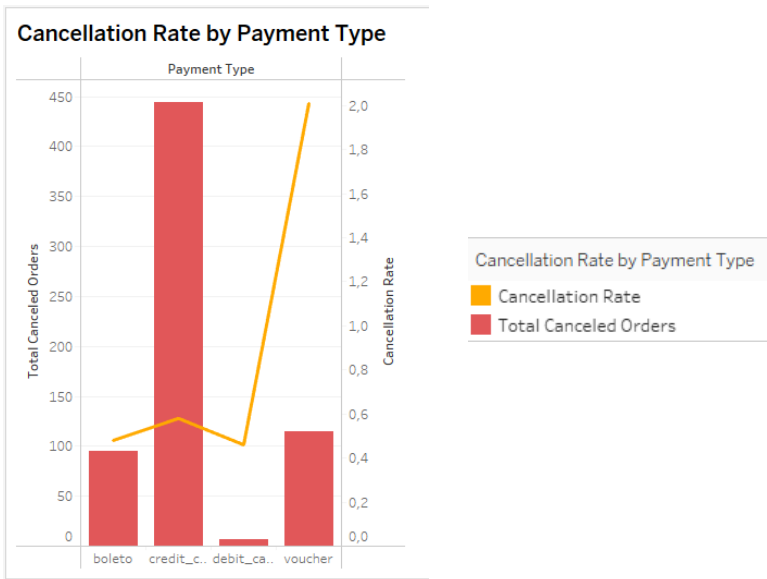
📊 Key Insights:

- Credit cards are the dominant payment method, accounting for the majority of orders and revenue.
- Boleto payments rank second in total orders but generate lower revenue per transaction compared to credit cards.
- Voucher payments have the lowest AOV (~65 BRL), indicating they are used for low-value purchases.
- Debit cards have a relatively small share of transactions but a high AOV (~142 BRL), suggesting they are preferred for larger purchases.

🤔 Business Implications:

- Credit cards should remain the primary focus for optimizing checkout experiences and fraud prevention.
- Boleto payments could be promoted further, particularly for customers with high cancellation risks.
- Vouchers should be strategically managed to ensure they are driving additional purchases rather than just discounting existing customers.

Cancellations

📌 Chart Summary:

This visualization examines **order cancellations** by payment method, highlighting the **total number of canceled orders** and the **cancellation rate (%)** for each payment type.
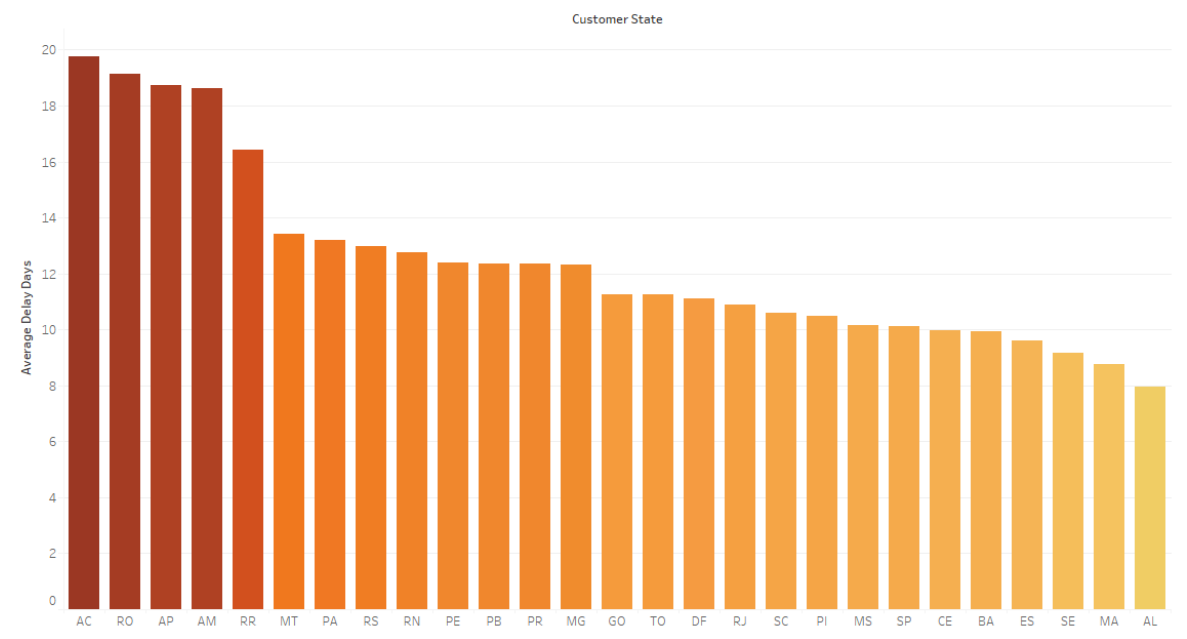
📊 Key Insights:

- Credit card payments have the highest cancellations in absolute numbers, but the cancellation rate is moderate.
- Voucher payments show the highest cancellation rate (~2%), suggesting higher risk of order drop-offs when using promotional payment methods.
- Boleto and debit card payments have the lowest cancellation rates, likely due to their non-reversible nature once paid.

🤔 Business Implications:

- Monitor and optimize voucher-based orders, they drive cancellations disproportionately.
- Improve fraud detection & customer verification for credit card transactions.
- Encourage boleto payments for customers with higher cancellation tendencies, as they are more secure.

Delivery Delays

Delivery Delays by State

📌 Chart Summary:

This visualization examines **average delivery delays across different Brazilian states**, measuring the number of days between the estimated and actual delivery date.
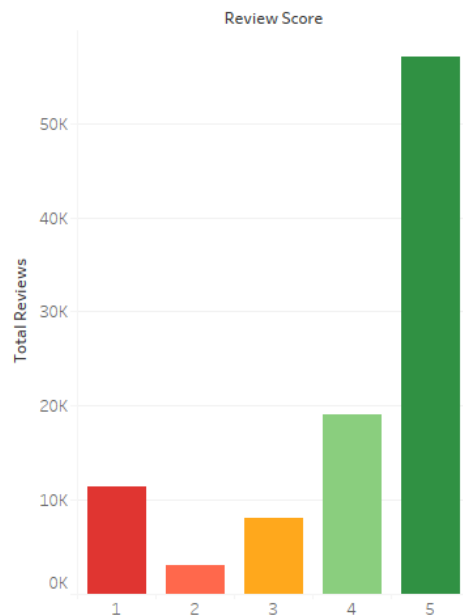
📊 Key Insights:

- Acre (AC), Rondônia (RO), and Amapá (AP) experience the longest delays, averaging nearly 20 days. These are remote regions with limited logistics infrastructure.
- Amazonas (AM) and Roraima (RR) also have high delays, likely due to geographic challenges and reliance on long-distance transportation.
- São Paulo (SP), Bahia (BA), and Espírito Santo (ES) have some of the shortest delays, likely due to better logistics networks and high shipping volumes.
- There is a clear trend of longer delays in Northern and Central states compared to more developed Southeastern regions.

🤔 Business Implications:

- Improve logistics partnerships in high-delay states, especially in the North.
- Offer realistic delivery estimates for customers in remote areas to set better expectations.
- Consider regional fulfillment centers to reduce delays in key states with recurring shipping challenges.

Review Score Distribution

📌 Chart Summary:

This visualization presents the **distribution of review scores** given by customers, ranging from **1 (worst) to 5 (best)**.
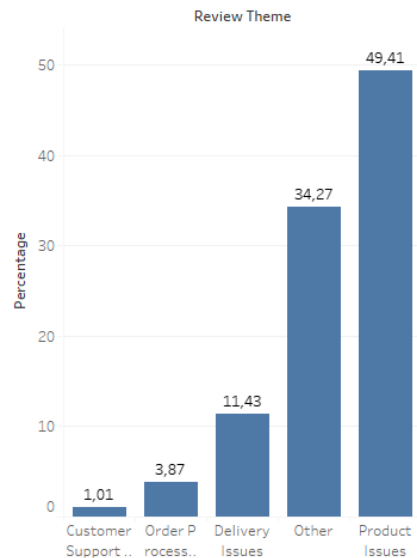
📊 Key Insights:

- Most reviews are highly positive, with over 55K reviews (more than half) receiving a 5-star rating.
- 1-star reviews account for a significant portion, with over 10K negative reviews, indicating dissatisfaction among a subset of customers.
- Mid-range reviews (2, 3, and 4 stars) are less common, suggesting that customers tend to either be highly satisfied or highly dissatisfied rather than neutral.
- The stark contrast between 5-star and 1-star reviews suggests polarized customer experiences, either excellent or frustrating.

🤔 Business Implications:

- Investigate common issues in 1-star reviews to understand and address major pain points.
- Leverage insights from 5-star reviews to reinforce strengths and maintain high customer satisfaction.
- Encourage neutral reviewers (3-4 stars) to share more detailed feedback, potentially uncovering areas for improvement.

Review Issues Breakdown

Top Complaints in 1-Star Reviews

Review Theme

📌 Chart Summary:

This visualization categorizes **the most common themes in 1-star reviews**, breaking down customer complaints into **Product Issues, Delivery Issues, Order Process Issues, Customer Support Issues, and Other**.

📊 Key Insights:

- Product Issues dominate customer complaints, accounting for nearly 50% of all 1-star reviews, indicating frequent dissatisfaction with the items received.
- "Other" issues make up 34%, suggesting a wide range of miscellaneous complaints that may require further investigation.
- Delivery Issues represent 11%, highlighting shipping delays, lost packages, or unmet expectations in delivery times.
- Order Process Issues (3.87%) and Customer Support Issues (1.01%) are less frequent but still relevant concerns.

🤔 Business Implications:

- Product quality control should be a top priority, understanding common defects or misleading product descriptions can help reduce complaints.
- Investigating "Other" issues could reveal hidden pain points not currently addressed.
- Delivery processes can be optimized to improve customer satisfaction and prevent shipping-related frustrations.
- Enhancing customer support responsiveness may help mitigate negative experiences before they escalate into 1-star reviews.

Conclusion & Recommendations

🔍 Key Findings

- 📈 **Revenue & Order Growth**: Orders and revenue **steadily increased**, peaking in **November 2017**, likely due to seasonal factors like Black Friday.
- 💳 **Payment & Cancellations**: While **credit card** is the most used payment method, it also has the **highest cancellation rate**.
- 🚚 **Delivery Delays**: Certain states experience **significant shipping delays**, with some averaging **over 19 days**.
- 📦 **Review Issues**: **49% of 1-star reviews** cite **product quality problems**, while **delivery issues (11%)** are the second-largest complaint.
- ⭐ **Review Distribution**: Majority of reviews are **5-star**, but a noticeable **1-star segment highlights customer pain points**.

🚀 Business Recommendations

✅ Leverage Seasonal Trends

- Increase marketing spend before peak months like November to maximize revenue.
- Consider limited-time promotions or bundles to capitalize on increased traffic.

✅ Reduce Credit Card Cancellations

- Investigate why credit card cancellations are so high, possible reasons could be fraud detection, failed payments, or unclear refund policies.
- Offer alternative payment incentives for lower-risk options like boleto.

✅ Improve Logistics & Delivery Time

- Identify courier inefficiencies in high-delay states and consider new partnerships to reduce shipping times.
- Implement real-time tracking notifications to improve customer satisfaction.

✅ Address Product & Service Quality

- Since product quality is the top complaint, ensure stricter supplier quality control.
- Improve customer support response time, especially for order issues & delivery delays.

✅ Turn Reviews into Insights

- Use customer review analytics to spot recurring issues early.
- Implement a follow-up system for 1-star reviews to offer solutions and improve brand trust.