

E-commerce Order Analysis: Key Insights and Trends

Analyzing Order Processing, Delivery, and Performance
Metrics

Contents

1. Introduction
2. Dataset Overview
3. Data Preparation and Cleaning
4. Key Analysis Questions
5. Insights From Analysis
6. Recommendations
7. Conclusion

1. Introduction

The goal of this project is to analyze the efficiency and performance of an e-commerce supply chain, with a focus on understanding order fulfillment processes and delivery times. I wanted to identify key factors such as delivery speed, order processing, and the volume of orders, providing actionable insights to optimize this supply chain.

The dataset used for this analysis consists of e-commerce order records, including details such as order placement dates, approval times, delivery dates, order statuses, and customer locations. This data provides a comprehensive view of the order lifecycle, from purchase to delivery, and covers multiple aspects of the fulfillment process.

By examining trends in processing times, delivery performance, and order volume across different regions and time periods, this analysis strives to uncover bottlenecks and areas for improvement in the supply chain, enabling more efficient operations and better service.

2. Dataset Overview

The dataset we will be conducting our analysis from is the Ecommerce Order & Supply Chain Dataset from Kaggle. Containing approximately 89,000 entries spanning from 2016 to 2018, it includes various attributes of the order lifecycle that will provide us insights into the supply chain performance.

Key features of the dataset include order details, order status, and locations of delivery.

The dataset covers a variety of factors that influence the efficiency of the supply chain, making it suitable for analyzing order fulfillment performance, identifying trends, and recommending improvements.

3. Data Preparation and Cleaning

The e-commerce order and supply chain dataset was prepared for analysis by first importing and merging two CSV files—customer and order data—using Pandas.

Data cleaning involved checking for missing values, particularly in delivery and approval timestamps. Null values were addressed by replacing missing delivery timestamps with "pending" and approval timestamps with "unknown" for orders still in processing. The dataset was then checked for duplicates and converted to ensure correct data types, with all date fields formatted as datetime. Finally, the cleaned dataset was exported as a CSV file for further analysis in an SQL editor.

4. Key Analysis

With the dataset cleaned, we examine several critical aspects of order fulfillment and delivery performance using our new dataset using MySQL.

We start by evaluating the average processing times, specifically focusing on the average time from order purchase to approval and from order placement to delivery. We also analyze the distribution of orders based on their status, which provides insight into how many orders are completed, pending, or canceled, as well as the percentage of late deliveries compared to on-time deliveries. Additionally, we assess the monthly order volume to identify trends over time, and finally, we explore the number of orders placed by state to understand regional variations in order activity.

Together, these analyses offer a comprehensive overview of our supply chain efficiency and help pinpoint areas for improvement.

5. Insights From Analysis

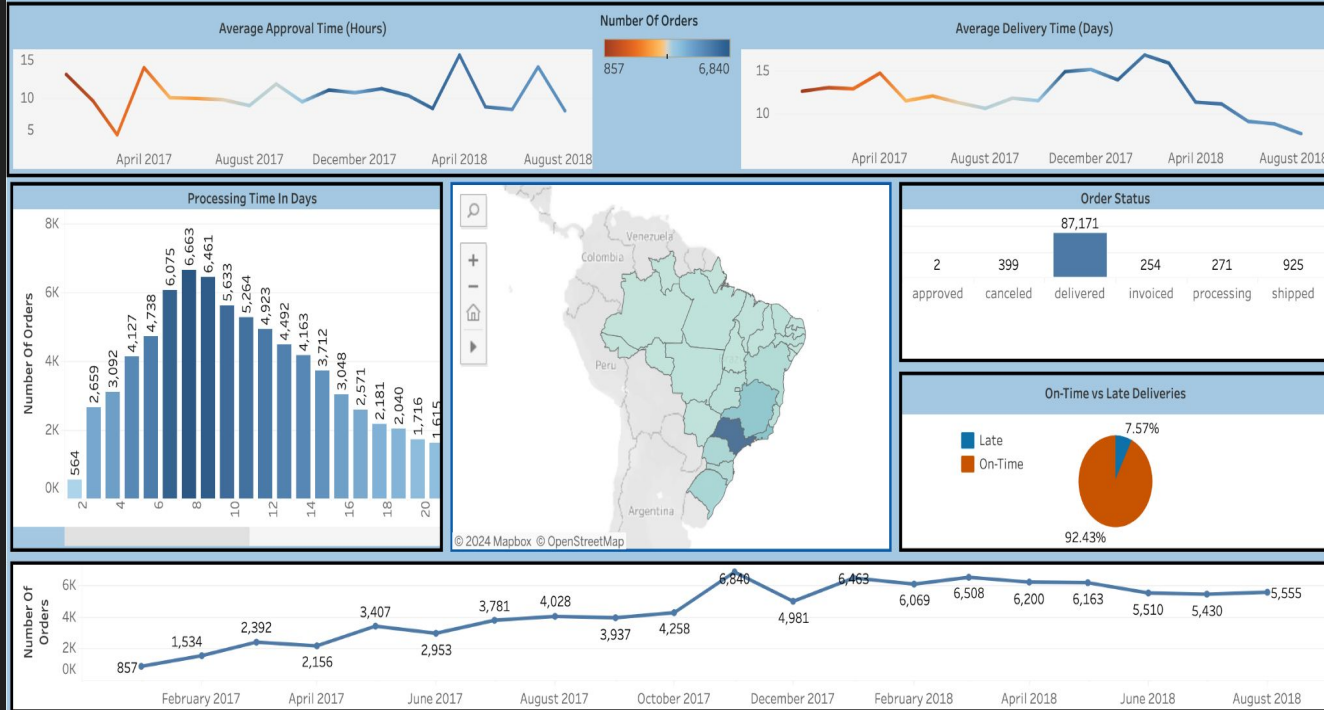
1. The average approval time for orders is approximately 10 hours, reflecting an efficient order management process but indicating room for improvement in handling outliers.
2. The average delivery time is about 12 days from order approval, suggesting potential delays in the shipping and logistics processes that could be optimized.
3. A concerning 93% of orders took over 72 hours for fulfillment, highlighting possible bottlenecks in the supply chain that need to be addressed.
4. A high 97% delivery success rate indicates the reliability of the logistics network and instills customer confidence in the service.
5. Only 7% of orders arrived later than expected, demonstrating an effective delivery system, but further analysis of the delayed cases could enhance punctuality.
6. There was a dramatic increase in order volume in November 2017, which stabilized in subsequent months, likely correlating with seasonal trends or successful promotions.
7. The majority of customers ordered from Sao Paulo, suggesting that targeted marketing strategies could capitalize on this regional demand.

6. Recommendations

1. Optimize Approval Processes: Streamline communication and automate routine approvals to reduce the average 10-hour order approval time.
2. Enhance Delivery Efficiency: Analyze logistics to cut the average delivery time of 12 days through better route planning and carrier partnerships.
3. Address Fulfillment Bottlenecks: Identify and resolve bottlenecks in the supply chain to reduce the 93% of orders taking over 72 hours for fulfillment.
4. Maintain High Delivery Success Rates: Monitor and enhance the 97% delivery success rate by investigating failed deliveries.
5. Improve Punctuality on Expected Deliveries: Develop strategies to improve the 7% of orders arriving late, enhancing communication on potential delays.
6. Leverage Seasonal Trends: Capitalize on the November 2017 order volume increase with targeted marketing campaigns during peak seasons.
7. Focus on Sao Paulo Market: Invest in localized promotions and partnerships in Sao Paulo, where most orders originate, to boost sales.

7. Conclusion

Insights into eCommerce Order Fulfillment and Delivery



This e-commerce supply chain analysis can provide us insights to suggest opportunities for improving logistics and optimizing operations. I created a dashboard using Tableau to go even further in my exploration of this data which can be found here https://public.tableau.com/app/profile/tyler.valdez/viz/EcommerceDataset_17297308296490/Dashboard1