



ProfitLens: Visualizing E-Commerce Trends

A data-driven analysis of product performance, regional sales patterns, and discount strategies to optimize business profitability.

Business Challenge

Understanding Profitability Drivers

E-commerce businesses face complex decisions about product mix, regional expansion, and pricing strategies. Without clear insights, companies risk investing in underperforming segments while missing high-potential opportunities.

This analysis addresses a critical question: Which products, regions, and discount strategies drive the highest profitability?



Project Objectives



Data Preparation

Clean and organize raw e-commerce data, handling missing values and inconsistencies for accurate analysis.



Pattern Discovery

Identify correlations between sales, profit, and discount strategies across product categories.



Regional Analysis

Determine which geographic regions and categories generate the highest profitability.



Visual Insights

Create clear, actionable visualizations to communicate findings to stakeholders.



Team Responsibilities



Vikrama

Business proposal development, dataset naming and organization, data cleaning pipelines, and comprehensive code review.



Iryna

Dataset selection and sourcing, exploratory data analysis, KPI development, and establishing business objectives aligned with project goals.



Paul

Repository management, project tracking and coordination, data exploration and cleaning, visualization design, and presentation delivery.

Data Sources & Structure

Kaggle E-Commerce Dataset

Our analysis leverages a comprehensive e-commerce dataset from Kaggle, containing transaction records with key variables:

- Product categories and pricing
- Order quantities and amounts
- Geographic data (state, city, postal code)
- Fulfillment methods and order status
- Promotional strategies and discounts
- Temporal patterns (dates, seasons)



Data Cleaning Process



Remove Redundancies

Eliminated unnecessary index columns and promotional IDs with minimal analytical value.



Handle Missing Values

Filled courier status gaps with "Unknown" and removed records with missing critical sales amounts.



Convert Data Types

Transformed dates to datetime format and standardized postal codes for time-series analysis.



Standardize Text

Converted categorical fields to lowercase and removed extra spaces for consistency.



Deduplicate Records

Retained only the most recent status for duplicate order IDs to ensure data integrity.

Analytical Framework

Product Analysis

Which categories generate the highest revenue? Compare performance across product lines like kurtas, tops, and dresses.

Discount Impact

Analyze how promotional strategies and discounts influence sales volume and profit margins.



Regional Insights

Identify which states and cities drive the most orders and assess regional profitability patterns.

Fulfillment Performance

Compare Amazon-fulfilled versus merchant-fulfilled orders to understand operational efficiency.

Key Research Questions



Product Performance

Which categories sell the most units and generate the highest revenue per transaction?



Geographic Distribution

Are certain regions more profitable than others, and what drives these differences?



Temporal Patterns

How do sales trends vary by month, season, or holiday periods?



Strategy Optimization

What data-driven decisions can improve overall business profitability and growth?

Project Risks & Constraints

Known Limitations

Profit Margin Calculations

Dataset lacks cost and bill-of-materials data needed for precise margin analysis.

Time Constraints

Project timeline requires focused scope and efficient resource allocation.

Data Quality

Missing values and inconsistencies require careful cleaning and validation.



Expected Deliverables

01

Clean Dataset

Fully processed data ready for analysis with documented cleaning procedures.

02

Exploratory Analysis

Comprehensive insights on sales patterns, regional performance, and discount effectiveness.

03

Visual Dashboard

Interactive charts showing category performance, regional trends, and temporal patterns.

04

Predictive Model

Statistical model to forecast future profit and sales based on historical patterns.

05

Business Recommendations

Actionable insights for optimizing product mix, regional focus, and pricing strategies.