International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

Analyzing Amazon Sales Dataset with Tableau: A Visualization Approach

Malgaram Punith Teja[1], Srishti Singh[2], Baddula Sreeya Yadav[3],Mrs.Sumayya Samreen[4]
[123]UG Scholar, Department of AI&DS,Methodist College Of Engineering and Technology, Hyderabad,India
[4]Assistant Professor of CSE(AI&ML), Methodist College Of Engineering and Technology, Hyderabad,India
1Malgarampunith8055@gmail.com,2shstisingh@gmail.com,3Sreeyabaddula380@gmail.com, 4sumayyasamreen01@gmail.com

Abstract- This research paper examines what makes an e-commerce business succeed. They use a data set on fifty thousand Amazon transactions to see the effect of all sorts of different variables on sales. What factors we considered was the sale price, discount percentage, ratings and amount of reviews. We examined the mechanics of pricing ,examined the level of consumer confidence in a product , we studied preferences we considered payment mode It shows how offs, and customer ratings influence earned money of the businessThis results from suggesting discounts might increase sales over a specific term but influence long term revenues via performing well on customer ratings and consumer trust. We find that both markets of North America and Middle-East have the revenue. The findings of this research may be accessed by merchants and marketers who wish to set their prices to helps consumers in their purchase decisions and at the same time, increase their revenues. Based on these results, they are able decide about their business.

Keywords – E-commerce analytics, Amazon sales dataset, data visualization, Tableau, customer behavior, revenue analysis, discount modeling, customer ratings, social proof, regional segmentation, sales performance, predictive analytics.

I. INTRODUCTION

Availability of insurmountable grounds of variations of markets experienced from the change-over from traditional brick-and-mortar stores to Everywhere Digital Market Place. As the purchasing public of the network makes the sales cycle intricate forcasting models increasingly sophisticated for inventory at the potential to reduce financial hazards. The research article called "E-commerce Retail Sales Trend Analysis and Prediction Based on the ARIMA Model(Fang, 2025)" makes an inkling into the operations of retail sales trend analysis. Basic consciousness of ground realities behind Amazon Complex is essential for the designed as well as the third fourth merchant. Amazon Still Continues to worldwide dominate the existing marketplace.

Central to this paper is the "Discount-trust Paradox." High discount levels are necessary to grab attention when there is extreme competition, but But, without high ratings, goods might be taken as inferior. Using a database of 50,000 records across different category-products, this
paper ascertains the inverse relationship of discount and consumer trust across categories electronics, beauty products and so on. With specific emphasis on how regional preferences and modes of payments (wallets, UPI) introduce moderator variables like income levels or earnings capacity.

II. LITERATURE SURVEY

Fang (2018) analyzed retail sales patterns in e commerce and found that sales are highly fluctuating online. The author suggests that forecast models are more critical in managing inventories given fluctuating online retail environment, with recommended focus on order date analysis for sales pattern and seasonality detection.

Pandey (2025) investigated the growth of Amazon and found that for effective e-commerce marketing, a search for consumer preferences in local regions is a necessity.

This justifies the reasoning to look for customer regions in the dataset. Wang (2021) found that customer ratings and reviews provide additional information to buyers and lack of customer feedback, that enhances online shopping trust and lowers buyer risk. This research highlights the significance of social proof in purchase behavior.

As illustrated by Alsmadi et al. (2023), the analysis of big data allows companies to understand more about customer behaviors and new sales trends. This is part of the rationale for analysing the fairly large 50,000 row e-commerce dataset which has been used in this case study.

Kumar et al. (2025) shows that a machine learning approach can be used to forecast e-commerce sales effectively based on past prices and discounts. This work serves as an

important reference for sales forecasting and prediction in e-commerce applications.
UNCTAD (2025) Cross-Border E-commerce: Regional Challenges and LogisticsWe looked at trade and logistics data.

Regional performance varies a lot depending on infrastructure and how people like to pay.

This helps you understand how payment methods work in regions.

IEEE (2023) Dynamic Pricing in E-commerce: A Survey This is a survey of automated pricing algorithms.

Dynamic pricing or offering discounts works well for things like electronics that people are very sensitive to price about.

This explains how discounts and product categories affect sales.

MDPI (2024) The Influence of Online Ratings on Purchase Intentions We did consumer surveys and statistical testing.

High star ratings are like a gatekeeper. If you don't have them discounts often don't work.

This supports what we found about the "Rating Floor", in
our results.

Lund University (2024) Online Consumer Behavior in the Beauty Industry Case study and consumer behavior analysis. "Self-care" and beauty items have the highest conversion and repurchase rates on Amazon. Explains why Beauty was the top revenue generator in your file. of global e-commerce. Validates why the Middle East was your top-performing region.

III. DATA VISUALIZATION

Table I: Dataset Attributes and Variable Descriptions:

Attribute Name	Data Type	Description	Unit/ Example
Product Id	String/I	A unique identifier for every individual item in the inventory	e.g., 1931
Product Name	String	The descriptive title of the item being sold on the platform.	e.g., Leather Bag
Price	Numerical	The original manufacturer's suggested retail price before any markdowns.	USD ($)
Product Category	Categorical	The primary group the product belongs to (used for sector analysis).	Beauty, Fashion, Electronics
Discount Percent	Percentage	The promotional markdown applied to the product price.	0% to 50%
Discounted Price	Numerical	The final price paid by the customer after the discount is subtracted.	USD ($)
Quantity Sold	Integer	The total number of units purchased in	e.g., 2 units

		a single transaction.	
Rating	Numerical	The average star rating (out of 5) given by customers for the product.	e.g., 4.2
Review Count	Integer	The total volume of customer feedback/review s for the product.	e.g., 1,250 reviews
Total Revenue	Numerical	The total financial value generated (Discounted Price × Quantity).	USD ($)
Customer Region	Categorical	The geographical location of the buyer (used for market analysis).	Middle East, N. America
Payment Method	Categorical	The medium through which the transaction was processed.	UPI, Wallet, Credit Card
Order Date	Date/Time	The specific timestamp when the transaction was completed.	DD-MM-YYYY

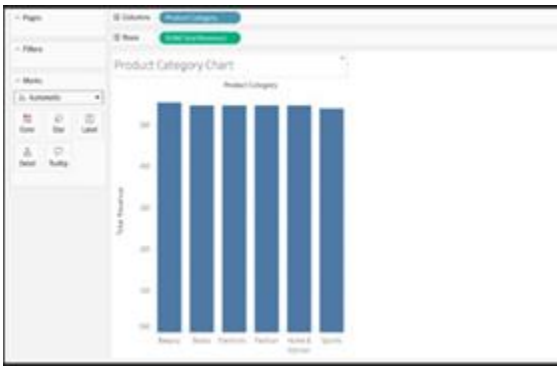

Fig 1: Total Revenue by Product Category

Goal:- My main aim here is to find out which product categories on Amazon make the money. I want to look at the sales data and see how different categories compare financially.

Observation:-The bar chart shows that categories like Beauty, Books, Electronics, Fashion, Home & Kitchen and Sports all make a lot of money over $5 million each. They are all pretty close. Beauty is the top one. This tells me that people really like buying beauty products and that category is doing well in terms of sales.

The product categories are Beauty, Books, Electronics, Fashion, Home & Kitchen and Sports. The Beauty category is making the money. Beauty products are in demand. The sales of Beauty products are very high.

The categories generate similar levels of revenue. Each category exceeds $5 million. The Beauty category slightly outperforms the others. The strong customer demand and higher sales contribution are, in care and beauty-related products.

Technical Implementation:-This visualisation was developed in Tableau by dropping the Product Category on Columns shelf and, Total Revenue on Rows shelf. The bar chart format, and sorting was selected for better comparison between categories.

Figure 2: Regional Revenue Distribution

Goal:-The objective of this visualization is to see regional sales performance, and see how various product categorizations influence revenue for worldwide markets.

Observation:-The result indicated by the stacked bar charts comparing total revenues across all regions reveals that Middle East and North America yield the overall highest revenues, followed by Asia and Europe.

Each region has at least 2product categories with relatively balanced sales, So suggesting stable demand across categories like Electronics, Beauty and Fashion.

Technical Implentation:-This chart was built using Tableau by dragging Customer Region on to the Columns shelf and SUM(Total Revenue) on to the Rows shelf.

Dragged Product Category to the Color mark to have a stacked category wise revenue breakup in each of the regions.

Figure 3: Top Performing Products by Revenue

Goal:-This visualization aims to illustrate the most profitable articles of the dataset and the profit contributed by the top payment articles.

Observation:-It shows the Product ID 1931 most revenue out of all product IDs, subsequent by other products with a little less sale revenue.

The gap of revenue does not seem to be big between them, it indicates there are some other contributions to sales and support a healthy level of Competition may exist.

Technical implementation:-The visualization was built in Tableau by dragging Product ID to Columns and SUM(Total Revenue) to Rows.

Top 10 filter was implemented to result only products with the highest revenue and labels were turned on to see the specific value of revenue.

Figure 4: Total Revenue Trend Over Time

goal:-this visualization allows us to look at the overall monthly revenue throughout the year and assess the variation from season to season.

observation:-the revenue for each month has varied in the line graph. the revenue has been high at the beginning of the year in the month of January and fallen down in the month of February and then has been grown gradually into these months,it has grown exceeding to high in the month of august.

then revenue has started fall in the months of October and November and again grown in the month of December. these variations in the revenue indicate the change in customer behaviour as the seasonal effects.

Technical Implementation:-This visualization was created in Tableau. MONTH(Order Date) was dragged to the Columns shelf, and put as a continuous date.

Sum(Total Revenue) was dragged on the Rows shelf. It was shown as a Line Chart, and show the values of each month.

IV. CONCLUSION

This project used Amazon sales data to see how product discounts affect customer ratings and how different regions and payment methods impact shopping. We found out that people buy a lot of Beauty, Electronics and Books and that the Middle East and North America are markets for online shopping.

One thing we learned from this project is that customers need to trust a company to keep buying from them over

time. When products are on sale people buy more. Products with good reviews and high ratings sell steadily over a long period.

We also saw that more and more people are using payments like UPI and Wallet to buy things online. We used Tableau to look at sales trends see when people buy things and understand how customers shop. The results of this project can help companies make decisions, about prices, marketing and taking care of customers. Maybe someday we can use machine learning to predict what people will buy and suggest products to customers.

Acknowledgment

M. Punith Teja would like to thank miss Sana Mateen, for her direction and support with Data Visualization Course and along this project work. M. Punith Teja is also thankful to the colleagues at AI&DS Department for their help and cooperation. Thanks to Kaggle for dataset and Tableau public for visualization tools to carry out this project.

REFERENCES

1. Alsmadi, M., et al. (2023). Big Data Analytics and Innovation in E-commerce. Journal of Business Analytics. https://pmc.ncbi.nlm.nih.gov/articles/PMC10214350/pdf/10660_ 2023_Article_9726.pdf
2. Fang, X. (2025). E-commerce Retail Sales Trend Analysis and Prediction Based on the ARIMA Model. Mathematics. https://www.mdpi.com/2227-7390/13/11/1838/pdf
3. IEEE Survey. (2023). Dynamic Pricing in E-commerce: A Survey. IEEE Xplore. https://ieeexplore.ieee.org/document/8710321
4. Kumar, A., et al. (2025). Sales Forecasting for Future Trends: A Machine Learning Approach. IJFMR. https://www.ijfmr.com/papers/2025/2/39954.pdf
5. Lund University. (2024). Online Consumer Behavior in the Beauty Industry. LUP Student Papers. https://lup.lub.lu.se/luur/download?func=downloadFile&recordO Id=8910000&fileOId=8910005
6. MDPI Sustainability. (2024). The Influence of Online Ratings on Purchase Intentions. Sustainability Journal. https://www.mdpi.com/2071-1050/13/16/9338/pdf
7. Pandey, D. (2025). Amazon's E-Commerce Growth Analysis: A Study on International Marketing. ResearchGate. https://www.researchgate.net/publication/385242812_Examining _the_Influence_of_Digital_Transformation_on_ Consumer_Beha vior_A_Study_of_Amazon's_E-commerce_Tactics_in_the_Retail_Sector
8. UNCTAD. (2025). Cross-Border E-commerce: Regional Challenges and Logistics. United Nations Conference on Trade and Development. https://unctad.org/system/files/official-document/der2024_en.pdf
9. Végh, Z., et al. (2025). Market Expansion Strategies: Amazon's Entry into Emerging Markets. ResearchGate. https://www.researchgate.net/publication/392019283_Amazons_ E-Commerce_Growth_Analysis_A_Study_on_International_Marke ting
10. Wang, Y. (2021). Trust and Risk in E-commerce: The Role of Social Proof. ResearchGate https://www.researchgate.net/publication/317634691_Trust_and_ Risk_in_E-Commerce_A_Re-examination_and_Theoretical_Integration

AUTHOR PROFILES

Malgaram Punith Teja
An undergraduate student pursuing his degree in the field of Artificial Intelligence and Data Science in Methodist college of Engineering and Technology

Srishti Singh
An undergraduate student pursuing her degree in the field of Artificial Intelligence and Data Science in Methodist college of Engineering and Technology

Baddula sreeya yadav
An undergraduate student pursuing her degree in the field of Artificial Intelligence and Data Science in Methodist college of Engineering and Technology

Mrs. Sumayya Samreen
Mrs. Sumayya Samreen is an Assistant Professor at Methodist College of Engineering and Technology.She guided the research work and supported the authors throughout the project development process.