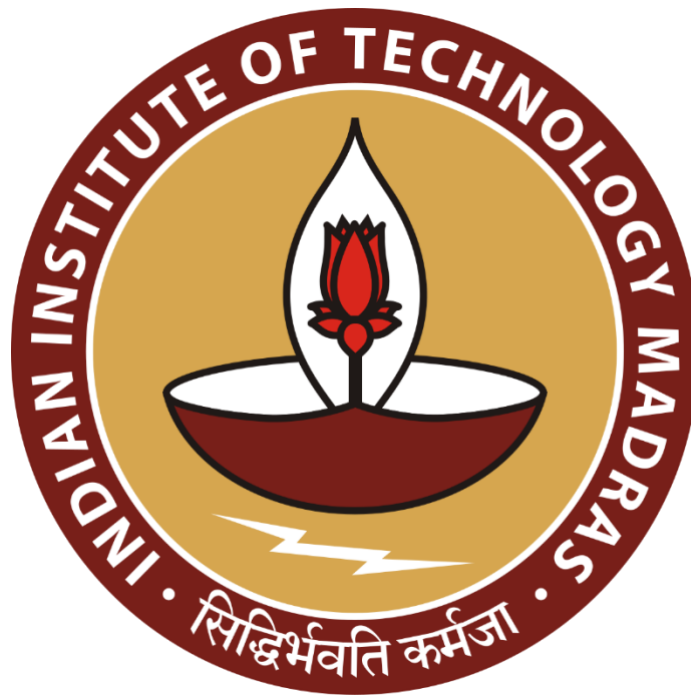


Case Study of Amazon Sales
A Proposal report for the BDM capstone Project

Submitted by

Name: Saudeep Chattopadhyay
Roll number: 23f1001295



IITM Online BS Degree Program,
Indian Institute of Technology, Madras, Chennai
Tamil Nadu, India, 600036

Contents

1	Executive Summary and Title	3
2	Organization Background	3
3	Problem Statement (Listed as objectives)	4
3.1	Problem statement 1- To determine the impact discounts and prices have on the product ratings and reviews: Changes in ratings of a product will be analyzed to see whether the discounts actually drive higher satisfaction or whether they in fact create expectations that go unmet on occasion.	4
3.2	Problem statement 2- To examine the correspondence of review sentiment with product performance based on category, pricing and discounting: By categorizing reviews and then analyzing the sentiments of those categories in relation to other parameters, one can understand how such factors affect the rating of customer satisfaction within a particular price point or product category which usually flourishes or buckles under pressure.	4
4	Background of the Problem	4
5	Problem Solving Approach	4
6	Expected Timeline	5
6.1	Work Breakdown Structure:	5
6.2	Gantt chart	6
7	Expected Outcome	6

Declaration Statement

I am working on a Project titled “Case Study of Amazon Sales A Proposal report for the BDM capstone Project”. I extend my appreciation to [Amazon Sales \(Kaggle\)](#) for providing the necessary resources that enabled me to conduct my project.

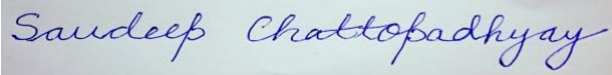
I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered from secondary source(Kaggle) and carefully analyzed to assure its reliability.

Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the principles of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I understand that all recommendations made in this project report are within the context of the academic project taken up towards course fulfillment in the BS Degree Program offered by IIT Madras. The institution does not endorse any of the claims or comments.

Signature of Candidate: 

Name: Saudeep Chattopadhyay

Date: 10/11/2024

1 Executive Summary and Title

This project aims to find critical business insights based on Amazon's product ratings and reviews dataset. This dataset is B2B and B2C - mostly e-commerce, but I'll focus only on the B2C aspect of it. It is a multinational e-commerce, as well as a tech corporation, with leadership in the e-commerce world. The data applied in this particular project is very extensive; it exceeds 1000 different products presented on the very popular e-commerce website that is Amazon. The dataset applies crucial elements such as product ratings, present discounts, detailed descriptions, and insightful reviews created by customers. The purpose of this project is thus to research and discover the factors controlling the performance of the products to understand which of the factors cause satisfaction among the customers who eventually lead to sales.

The two paramount objectives this paper addresses over a full scope are the first to reach such a fundamental understanding of the relationship between various product features and the sentiments expressed by consumers.

Such knowledge will be useful for Amazon as well as other similar businesses that are actively competing in the online platform. It will enable them to know with whom their customers may find satisfaction and happiness in meeting their respective buying needs. Besides, it will help them monitor adjustments in their pricing policies so they could get closer to consumers' expectations. Moreover, it will allow these organizations to strategically make their products more attractive through effective marketing strategies. For these purposes, a range of advanced methodologies such as data processing, exploratory data analysis, sentiment analysis, and predictive modeling would be used. These methods together will yield actionable insights that could drastically inform and guide significant business decisions.

2 Organization Background

Amazon is really one of the best in e-commerce that has its vast operations of buying and storing inventory. Here, they are performing many crucial tasks such as sending products, finding strategies to set prices, handling customer service, and return processing with efficiency. Most significantly, Amazon is considered to be a leading global technology company that acts at a global level. It has more influence in those growing sectors like cloud computing and artificial intelligence. They really are strong and impressive there. The central strength that Amazon boasts has its roots in the focused effort to remain customer-centric. Through this approach, the platform has provided an incredible and vast variety of products that cover a stupendous variety, ranging from electronics to everyday groceries. Because of this customer focus, it allows customers to find nearly anything they may be interested in searching for and conveniently shop for these items from virtually anywhere they wish. Such capability is due to Amazon's pioneering efforts in delivering extraordinary value in areas of vastly competitive pricing, highly insightful user reviews, and an excellent set of highly comprehensive logistics capabilities. All these strategies have been very effective in enabling the firm gain and maintain a significant level of market share, putting the company far ahead of any other competitor.

The company has always strived to use technology and data analytics as very necessary tools toward meaningful improvements at every level, from their delivery times and competitive pricing strategies to their overall user experience. In mastering a greater understanding of the product offerings by Amazon and continued work to enhance them in line with customer needs, this project becomes very relevant and instrumental in reaching the targeted business goals that the company strives for.

3 Problem Statement (Listed as objectives)

This project's objective is to explore which factors strongly influence customer ratings and reviews of the products on Amazon. In particular, it addresses the following objectives:

- 3.1 **Problem statement 1-** To determine the impact discounts and prices have on the product ratings and reviews: Changes in ratings of a product will be analyzed to see whether the discounts actually drive higher satisfaction or whether they in fact create expectations that go unmet on occasion.
- 3.2 **Problem statement 2-** To examine the correspondence of review sentiment with product performance based on category, pricing and discounting: By categorizing reviews and then analyzing the sentiments of those categories in relation to other parameters, one can understand how such factors affect the rating of customer satisfaction within a particular price point or product category which usually flourishes or buckles under pressure.

4 Background of the Problem

Customer satisfaction is a key factor to succeed in e-commerce segment, with ratings and reviews sometimes cause changes in product visibility and influence a customer buying decision. Even though there are ample customer feedback available, online firms frequently fail to recognize precisely what factors play the largest role in promoting the process of making positive or negative reviews.

This dataset available for the project contains useful information: product prices, discounts, customer ratings, and some very detailed reviews on products. The opportunity arises here to explore these relationships for determining the factors that ensure higher customer satisfaction and feedback. Just for example discounts might entice customers but could potentially impact their perception of quality. Similarly, detailed reviews can shed light on consumer expectations and preferences.

The challenge is to find associations that indicate how pricing strategies and product attributes create or diminish customer satisfaction. This analysis will provide data-enabled insights into how product listings, marketing, and pricing may be optimized to gain higher customer satisfaction and higher sales.

Dataset Link: <https://www.kaggle.com/datasets/karkavelrajaj/amazon-sales-dataset/data>

5 Problem Solving Approach

Following is the problem - solving approach which comprises a seven-stage process, from data preprocessing to insights formulation:

- 1) **Data Preprocessing:** The primary step will include data cleaning to fill in missing values, remove duplicate entries, or rectify any inconsistent values. Product categories, price ranges, and customer reviews will be standardized to synthesize and standardize the attributes appropriately.
- 2) **Exploratory Data Analysis (EDA):** The EDA will help visualize the distribution of ratings, discounts, and review counts across different product categories. Statistical summaries will allow the analysis of the average product price, average product rating, or discount trends. Histograms, box plots, and scatter plots will visualize the trends and dependencies among the variables, detecting positive or negative relationships and highlighting specific patterns or outliers.
- 3) **Correlation Analysis:** The correlation analysis will help us to establish the extent to which the discount percentage, rating, or price of the product is related to the satisfaction of consumers. The result will illustrate which key product variables increase or decrease product ratings.
- 4) **Sentiment Analysis:** Leveraging Natural Language Processing (NLP) techniques, we will decode customer reviews and categorize them as positive, negative, or neutral. The analysis helps identify topics driving satisfaction and uncover the dissatisfaction leading drivers.
- 5) **Regression and Classification Modeling:** Regression models will predict product ratings based on their pricing and discount percentages. Classification models will categorize the sentiment of customer ratings based on the content of their reviews. The Key driver analysis provides insights into how different features contribute to consumer satisfaction ratings.
- 6) **Visualization and Reporting:** The analysis will be communicated with the stakeholders through interactive visualizations. Graphs, charts, and heatmaps will be created to highlight the three patterns and illustrate the key insights have been highlighted. Pivot tables, KPI dashboards, and summary reports will facilitate communication in comprehensible and easy-to-understand format.
- 7) **Recommendations:** Based on these insights, we will provide actionable recommendations on optimizing product listings, revising pricing strategies, or improving product characteristics to improve ratings or reviews.
Tools used would be Python program and MS Excel.

6 Expected Timeline

6.1 Work Breakdown Structure:

The following is the work breakdown structure:

Week 1: Collect and clean and preprocess data.

Week 2: Exploratory Data Analysis and Correlation Analysis.

Week 3: Sentiment analysis using NLP techniques.

Week 4: Regression and Classification model development also Report(Final Submission) writing and submission.

6.2 Gantt chart

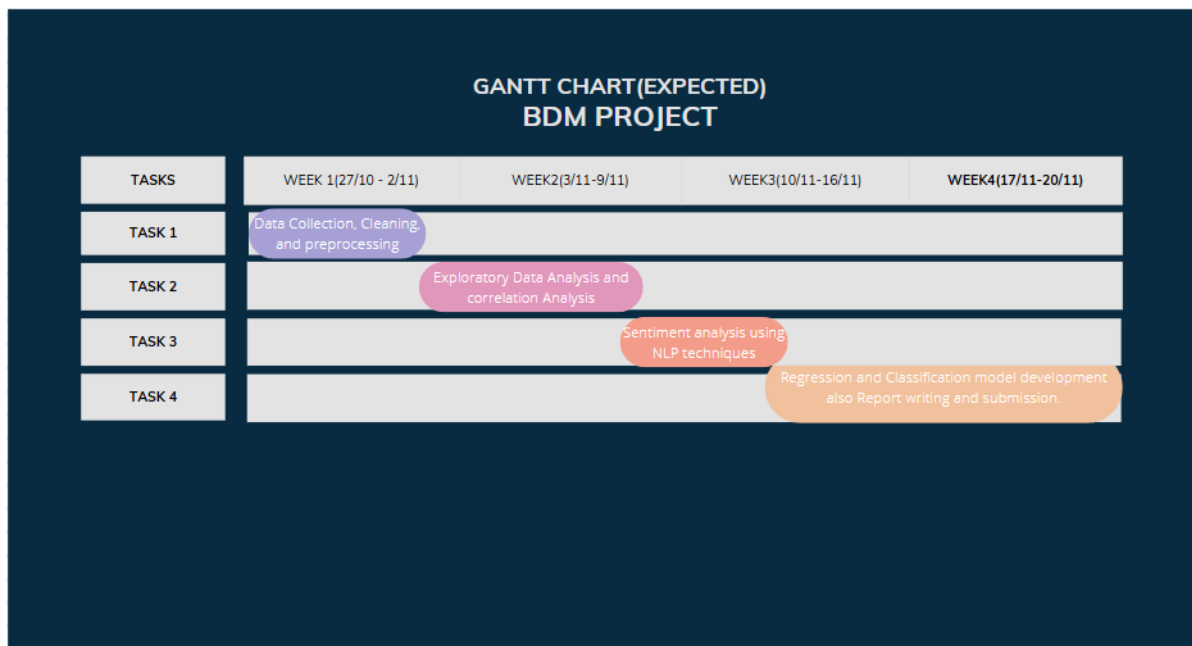


Figure 1 Expected timeline for completion of project.

7 Expected Outcome

Expected outcome of the project would be a comprehensive report with all the details focusing on aspects that determine the ratings and reviews of products on Amazon. With deeper insights from the analysis of how various pricing strategies and specific features of a product depict overall satisfaction among customers, the end result will provide e-commerce platforms with invaluable information that can be gained for their advantage. In particular, this analysis will help in understanding aspects to optimize discount strategies, quality, and effectiveness of product descriptions, and even derive better ways of handling customer reviews as they arise.

Sentiment analysis will not only reflect on prominent and critical issues that customers are facing at this time so that Amazon can take proactive measures and work on these areas more vigorously. This analysis will also reveal insights that could result in sound conclusions because it supports strategic planning about the management of product listings, the alteration of pricing structures, and focused marketing campaigns that lead to superior customer satisfaction and subsequently higher sales numbers.