

Project Report: Amazon Review Sentiment Analysis

Title: Amazon Review Sentiment Analysis

Subtitle: A data driven approach to understanding customer feedback

Author: Ankush P Gowda

Date: 11-02-2025

Table of Content:

1. Introduction
2. Objectives
3. Features
4. System Design
 - a. Architecture
 - b. Data Structure
5. Implementation
6. Results & Graphical Analysis

1. Introduction

This project focuses on analysing customer sentiments from Amazon product reviews. The goal is to classify sentiments as Positive, Negative, or Neutral and extract key insights through various visualizations.

2. Objectives

- Extract and process product reviews from Amazon.
- Perform sentiment analysis using NLP techniques.
- Visualize key insights with meaningful graphs.
- Provide a summarized review using OpenAI's GPT model.

3. Features

- Sentiment classification (Positive, Neutral, Negative)
- Review summary generation
- Yearly Sales Trend
- Yearly Sentiment Distribution
- Monthly Review Trends
- Word Cloud of Positive Reviews
- Overall Sentiment Distribution
- Sentiment Trends Over Time
- Interactive dashboard using Streamlit

4. System Design

a. Architecture

The system follows a pipeline-based architecture:

- **Data Collection:** Extracts review using web scraping tool.
- **Pre-processing:** Cleans and structures data.
- **Sentiment Analysis:** Uses a vader sentiment to sentiment classification.
- **Visualisation & Reporting:** Generates graphs and summary.

b. Data Structure

- reviewTitle (string).
- reviewDescription (string).
- Sentiment (Positive, Neutral, Negative).
- ratingScore (numeric).
- Date (datetime).

5. Implementation

- **Libraries Used:** apify_client, pandas, openpyxl, nltk, vaderSentiment, langdetect, streamlit, matplotlib, seaborn, wordcloud, openai.

6. Results & Graphical Analysis

To demonstrate the analysis, we take an example product: **boAt Rockerz 450**.

- **Product Image:**



- Sentiment analysed chart and score of this product:

Amazon Review Sentimental Analysis

Enter your product link:

https://www.amazon.in/Rockerz-450-Wireless-Bluetooth-Headphone/dp/B07PR1CL3S/?_encoding=UTF8&pd_rd_w=5Fu9G&content-id=amzn1.sym.50f

Submit

	ratingScore	reviewTitle	reviewDescription	date	Sentiment
1	5	Detailed Review: compared with Sony headphones in the same price range	I am using Sony Headphones for over 2 years now, which were sufficiently good for m	2020-01-23 00:00:00	Positive
2	2	Good one as we can upgrade to 450 from 400 but these are not strong enough broke	Hi guys, as expected both 450 and 400 are alike but slight high at 230hz and thump fo	2020-01-24 00:00:00	Neutral
3	2	Sound quality not upto the mark, but can be considered for a deal price of 999	Initially I have received a faulty product. While used it for the 1st time, I felt the sounc	2020-10-27 00:00:00	Neutral
4	2	Neither Good nor bad but buy it at your own risk	I am writing this review 10 months after buying these headphones. The reason of writ	2020-11-02 00:00:00	Negative
5	3	Audio quality just average, broke quickly after warranty period.	Packaging was good and things seemed to be fine initially apart from a little discomf	2021-01-11 00:00:00	Neutral
6	2	Poor Quality Product, a complete waste of Money	This is my review after using Boat Rockerz 450. This review is of the usage with a year	2022-09-13 00:00:00	Negative
7	2	Honest review on the boat rockerz 450	Please read the whole review.The product I recieved was fine, working and packaging	2022-09-27 00:00:00	Neutral
8	4	A Sonic Marvel for Wireless Audio Enthusiasts - boAt Rockerz 450 Bluetooth	The boAt Rockerz 450 Bluetooth headphones have left a lasting impression on me as	2023-01-13 00:00:00	Positive
9	1	Bad build quality, cheap plastic ,design flaw	* After 2 months of usage the right side broke into 2 pieces the build quality is bad the	2023-08-07 00:00:00	Negative
10	2	BOAT sound =Happy in Budget but ear cushion heat soon	Overall Good headphoneExplains :-1. Good Sound Quality2. Good cousin for ear (Lor	2023-10-04 00:00:00	Neutral

Overall Sentiment: Negative

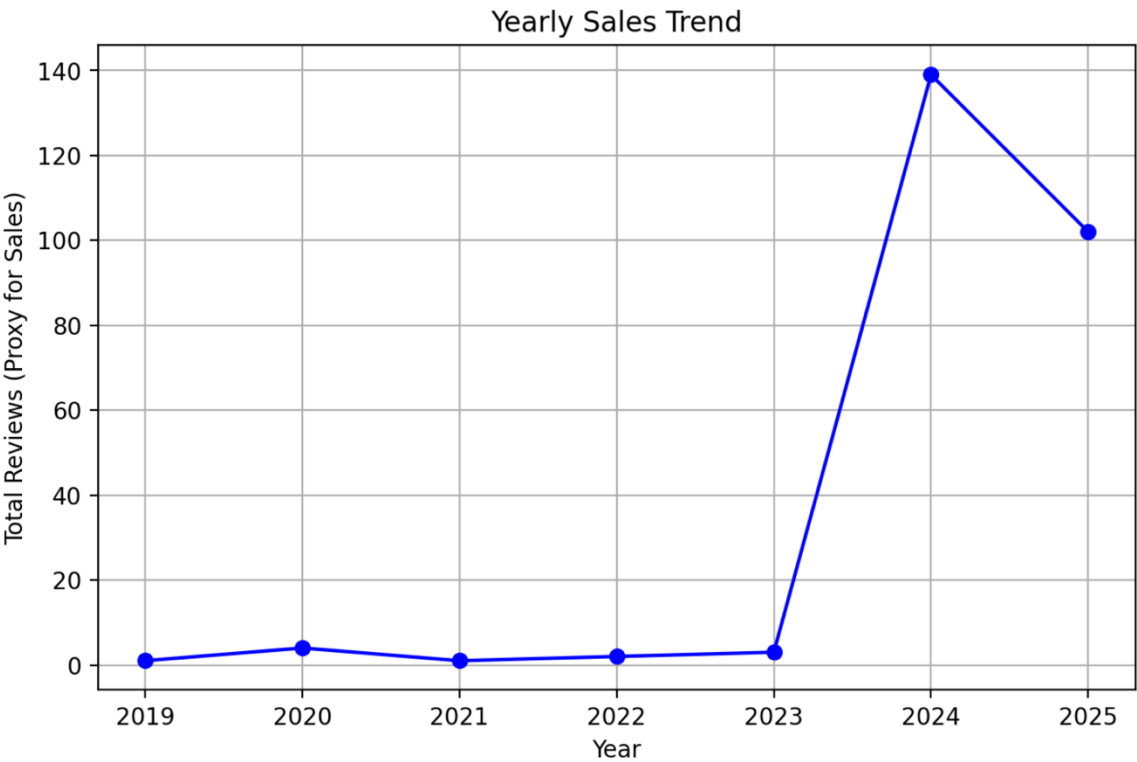
Score: -0.22

- AI- Generated overview of all reviews:

AI-Generated Review Summary:

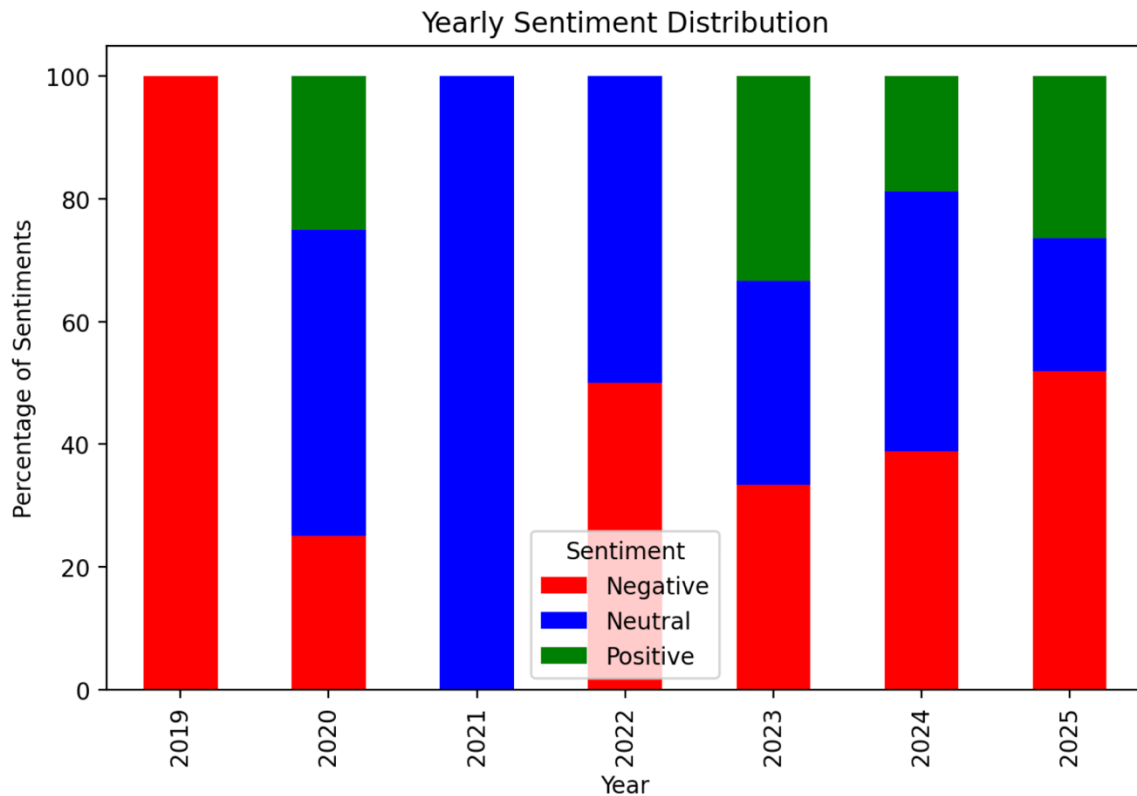
The boAt Rockerz 450 headphones receive overwhelmingly negative reviews for poor sound quality, uncomfortable design, and low build quality. Users report issues with sound clarity, discomfort, flimsy plastic construction, and a weak mic. Many recommend avoiding this model due to its subpar performance and suggest looking at alternative options for better value and durability. Despite a few positive comments on aspects like bass and Bluetooth connectivity, the consensus is to steer clear of the b

- Yearly Sales Trend (based on total reviews posted each year):



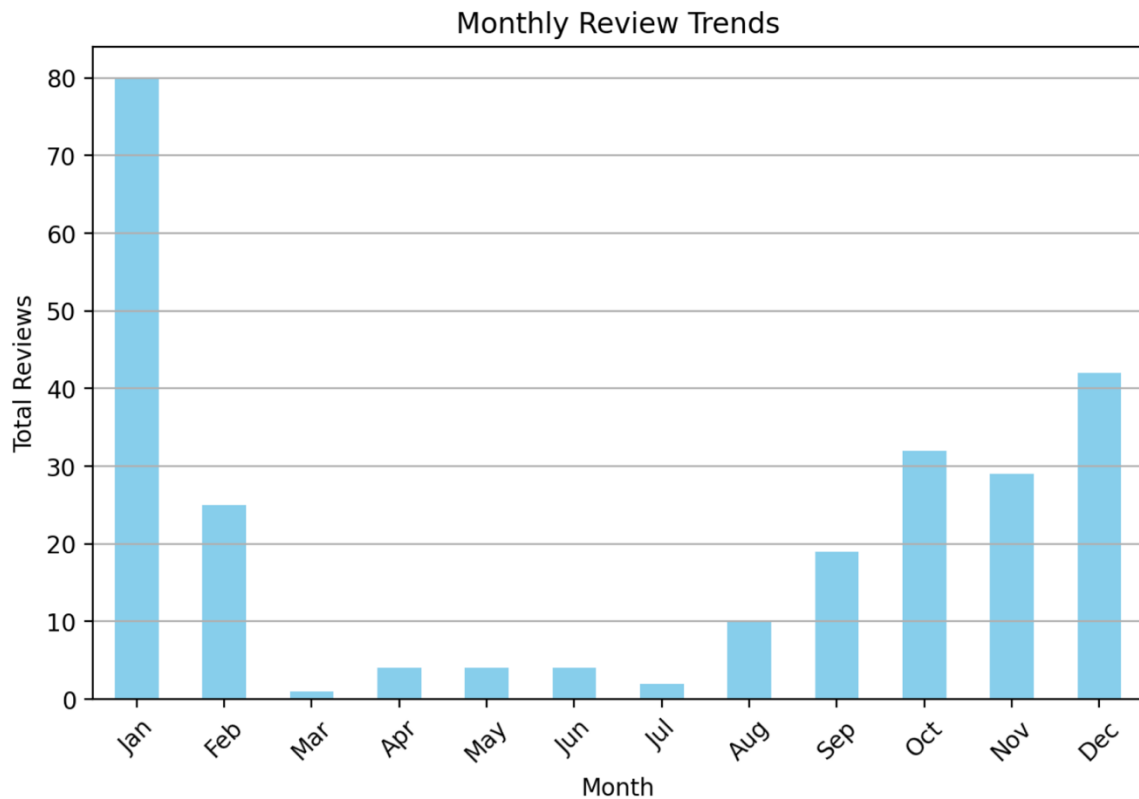
- The product had minimal customer engagement from 2019 to 2022, with very few reviews indicating low sales or market awareness.
- A significant spike in reviews in 2024 suggests a major increase in sales, possibly due to marketing efforts, improved product features, or increased demand.
- Since 2025 has just begun, its data may not be fully indicative of a trend yet.

- **Yearly Sentiment Distribution:**



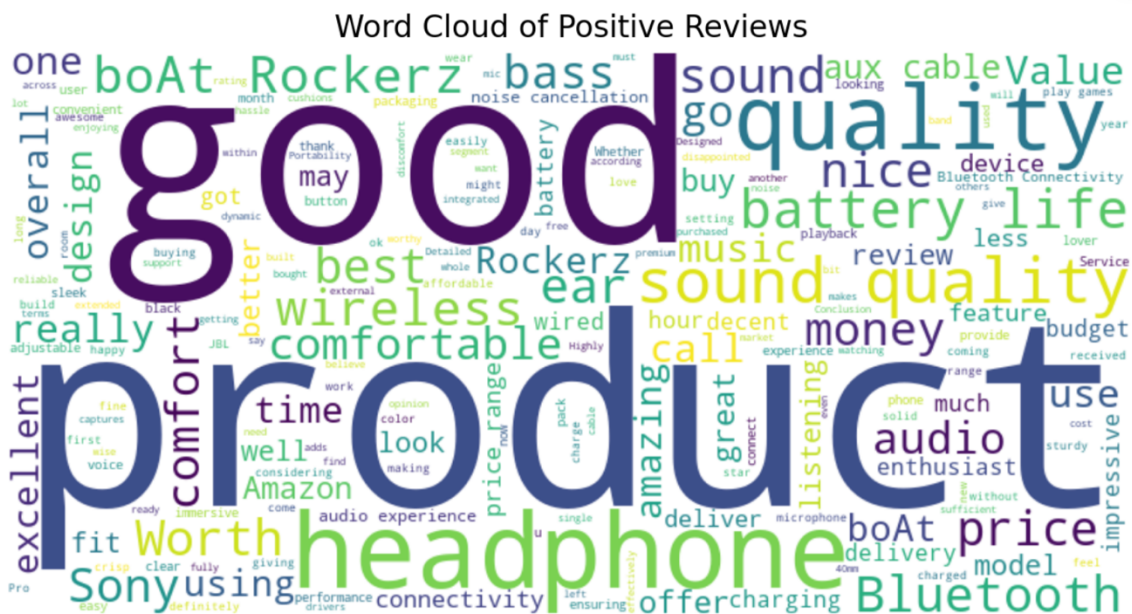
- Negative sentiment has significantly decreased from 2019 to 2024, with a corresponding increase in neutral and positive sentiments.
- Positive sentiment shows a notable upward trend starting from 2022, indicating a shift towards a more favourable outlook.

- **Monthly Review Trends:**

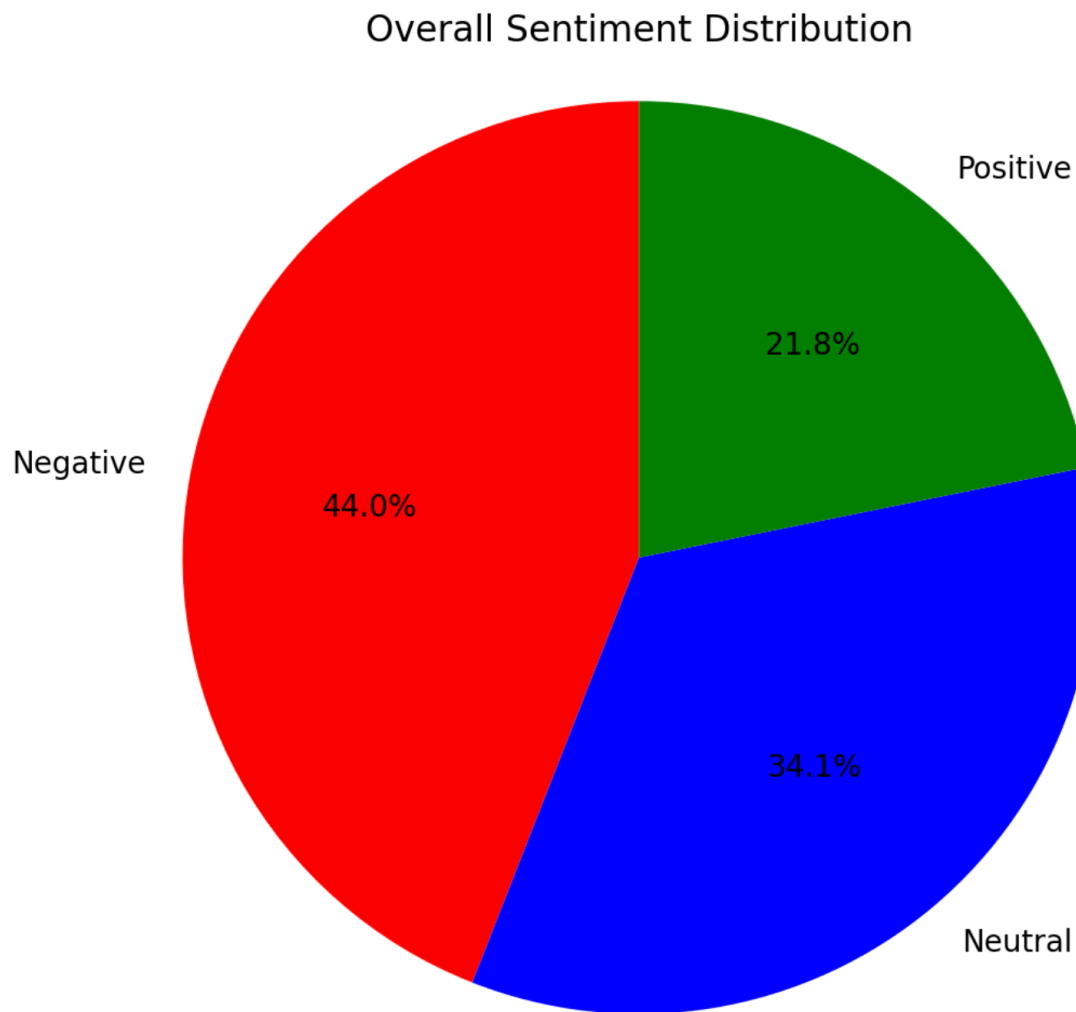


- The graph shows a spike in reviews during January and February, followed by a sharp decline from March through August.
- There is a steady increase in reviews from September to December, indicating rising activity towards the year's end.

- **Word Cloud of Positive Reviews:**

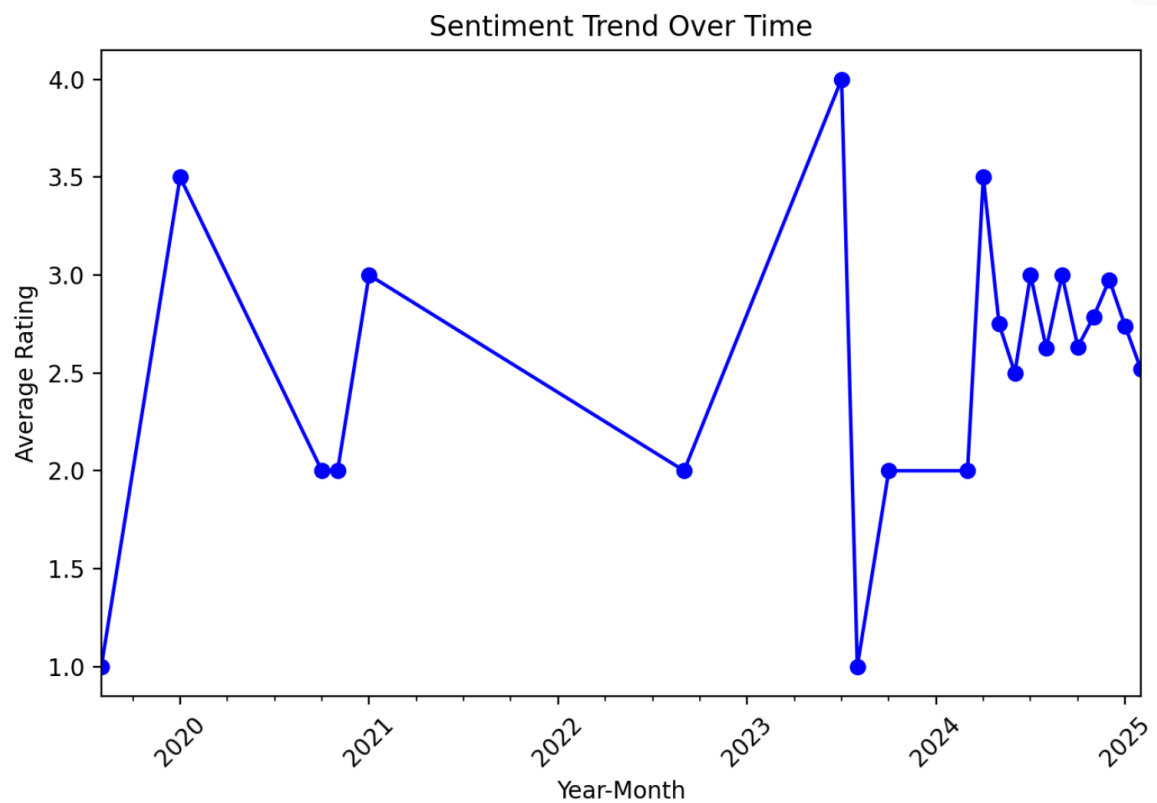


- **Overall Sentiment Distribution:**



- The pie chart represents the overall sentiment distribution, divided into three categories: **Positive (21.8%)**, **Neutral (34.1%)**, and **Negative (44.0%)**.
- The **Negative sentiment (red)** holds the largest portion, followed by **Neutral (blue)** and then **Positive (green)**, indicating an overall leaning towards negative sentiment in the dataset.

- **Rating Over Time:**



- The sentiment exhibits **high variability**, with multiple sharp spikes and drops, indicating fluctuating opinions over time.
- There is a significant dip around **2023**, followed by a rapid recovery in **2024**, and then more **stabilized fluctuations** moving into **2025**.