

# **Inclusive Analysis of Amazon Cell Phone Reviews**

A project submitted by

**DESHMUKH NIRANJAN KHUSHAVANTRAO (20BCT0220 - VELLORE)**

**ADEEBA RASHID (20BAI10107 - BHOPAL)**

**INSHA FATIMA (20BEE0372 - VELLORE)**

**AMARISH KUMAR SRIVASTAVA (20BCE10292 –BHOPAL)**

Prepared For

Smartbridge - Data Analytics Externship



**VIT<sup>®</sup>**  
**Vellore Institute of Technology**  
(Deemed to be University under section 3 of UGC Act, 1956)

June, 2023

# **CONTENTS**

## **1 INTRODUCTION**

1.1 Overview

1.2 Purpose

## **2 LITERATURE SURVEY**

2.1 Existing problem

2.2 Proposed solution

## **3 THEORETICAL ANALYSIS**

3.1 Block diagram

3.2 Hardware / Software Designing

## **4 EXPERIMENTAL INVESTIGATIONS**

## **5 FLOWCHART**

## **6 RESULT**

## **7 ADVANTAGES & DISADVANTAGES**

## **8 APPLICATIONS**

## **9 CONCLUSION**

## **10 FUTURE SCOPE**

## **11 BIBLIOGRAPHY**

## **APPENDIX**

A. Source Code

# INTRODUCTION

## 1.1. Overview

Our ability to purchase goods online has been revolutionized by mobile devices, which put all the information at our fingertips. More and more customers will turn to other customers for product information instead of the seller's information as information access gets easier. Examples of this kind of information include reviews and ratings left by customers, which have already influenced many customers' purchasing decisions. Customers may make educated decisions and feel confident about them thanks to the transparent system created by the review and ratings platforms offered by eCommerce businesses.

Product reviews may be found in abundance on Amazon.com, and their review system is available through all available channels and presents reviews intuitively. The product reviewer gives the item a rating between 1 and 5, along with their own opinion based on their whole experience. To get the final product rating, the mean value from all the ratings is determined. By allowing others to vote on whether a review is useful, both the review and the reviewer gain credibility. For this study, we studied over 40 thousand reviews of unlocked cell phones purchased from Amazon.com to gain an understanding of how reviews, ratings, and prices relate to one another.

## 1.2. Purpose

The process of obtaining and translating user-level data into insights that show how customers interact with certain goods is known as product analytics. An organization can track and analyze the user journeys from user activation through all other phases of use using product analytics to learn what motivates people to use and keep using the product.

Additionally, this aids businesses in comprehending the value that buyers receive from a product.

The most popular features of a product, the average amount of time users spend on a given action, the most effective marketing channels, and the number of users who return to the product on a daily, weekly, or monthly basis are just a few examples of the usage data that can be gathered through product analytics. With this data, businesses may examine how people engage with the products they develop and utilize these insights to enhance user experiences (UX).

Product analytics reveals to businesses their customers' actions, namely what consumers do as opposed to what they claim to accomplish. Building efficient and useful goods requires an understanding of the customer and their demands. Product managers can get a deeper understanding of their customers and create more profitable products by using the incredibly precise information acquired through product analytics.

Customer surveys, conversations, and interviews can yield qualitative data, but this information isn't necessarily reliable and frequently doesn't fully describe the user experience. Product analytics, in contrast, provide unbiased and conclusive customer data that was acquired by software monitoring actual user behavior with the product. As a result, product teams can analyze certified data more thoroughly using product analytics than they could have done using human-error-prone interviews, surveys, and discussion methods.

Product analytics can help businesses better understand how their consumers use their current features, but they can also be useful for testing out new features and evaluating user experience. The data produced by-product analytics can be used by a team to work towards a goal for how frequently a new feature should be used. Product analytics has emerged as one of the most effective strategies for increasing user retention and enhancing a company's position in a cutthroat market in recent years, according to tech businesses that specialize in application development.

# **LITERATURE SURVEY**

## **2.1. Problem statement**

Mobile phones have revolutionized the way we purchase products online, making all the information available at our fingertips. As access to information becomes easier, more and more consumers will seek product information from other consumers apart from the information provided by the seller. Reviews and ratings submitted by consumers are examples of such type of information and they have already become an integral part of customer's buying-decision process. The review and rating platform provided by e-commerce platforms creates a transparent system for customers to take informed decisions and feel confident about it.

## **2.2. Proposed solution**

It is difficult to read all feedback for a particular item, especially for popular items with many comments. So in our project, we will attempt to understand the factor that contributes to telling us if we should take a product or not.

We are using Excel for basic cleaning of data like removing duplicates or blank values. Once we are done with cleaning we will move forward and store the data in MySQL database and perform some basic exercises in order to understand the data.

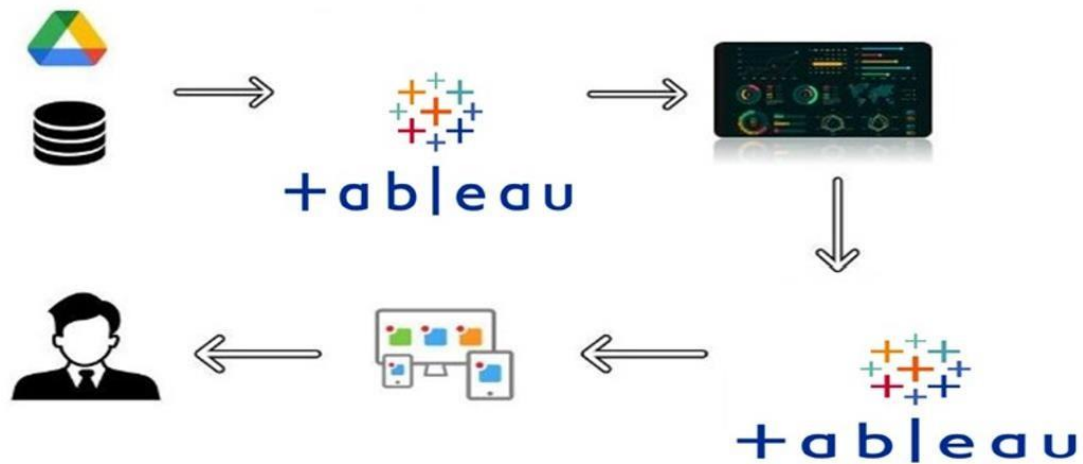
The next step will move to Tableau and create an interactive dashboard and story while answering questions that our stakeholders might ask.

We deploy the dashboard and story on the server and create a website using Bootstrap and Flask deploy our dashboard and story in that.

Bootstrap is a site that provides us with website templates and according to our needs, we can make changes using HTML files. Flask is an API of Python that allows us to build web-applications.

# THEORETICAL ANALYSIS

## 3.1. Block diagram



## 3.2. Software Designing

1. **EXCEL**- Used for cleaning the data like removing duplicate values and filling the blank cell with similar data.
2. **MySQL Workbench**- Database is used to store our data and find insights from our data using SQL Query.
3. **Tableau**- A visualization tool is used to visually represent the insights that we have got and our stakeholders might utilize it for getting information about the business.

4. **Bootstrap**- Bootstrap is a free, open source front-end development framework for the creation of websites and web apps. Designed to enable responsive development of mobile-first websites, Bootstrap provides a collection of syntax for template designs.
5. **Python**- Used for the backend web integration (flask) and the prediction of the reviews.
6. **VS Code**- Used for coding and running the application.

## Uses of Python (Flask) In Model-

Flask, a popular web framework in Python, can be used to build a web application for analyzing and predicting Amazon phone reviews. Here's a general overview of how Flask can be utilized in the process:

**Data Collection:** The first step is to gather the Amazon phone reviews data. This can be done by scraping the reviews from the Amazon website or using publicly available datasets.

**Data Preprocessing:** The collected data needs to be preprocessed before it can be used for analysis and prediction. This may involve tasks such as removing irrelevant information, handling missing values, and performing text cleaning operations like removing stop words and punctuation.

**Analysis and Prediction Model:** Once the data is preprocessed, you can build a machine learning model to analyze and predict sentiments or other characteristics of the phone reviews. This model can be trained using various techniques, such as sentiment analysis, natural language processing, or machine learning algorithms like Naive Bayes, support vector machines (SVM), or deep learning models.

**Flask Web Application:** Flask can be used to create a web application that allows users to interact with the analysis and prediction model. The application can have a user interface where users can input a review, and the model will analyze and predict the sentiment or other relevant information associated with the review.

**Backend Development:** Using Flask, you can define routes and endpoints for different functionalities of the web application. For example, you can create a route to handle the form submission of the review and trigger the analysis and prediction process.

**Integration of the Model:** Flask allows you to integrate the trained model into the web application. You can load the model weights or use a pre-trained model and use it to process the user input, perform the analysis, and generate predictions.

**Displaying Results:** Once the analysis and prediction are done, Flask enables you to display the results back to the user. This can be done by rendering a template that shows the predicted sentiment, sentiment scores, or any other relevant information extracted from the review.

**Deployment:** After developing and testing the web application locally, you can deploy it to a server or a cloud platform like Heroku or AWS Elastic Beanstalk to make it accessible to users over the internet.

By using Flask, you can create a web application that utilizes machine learning techniques to analyze and predict sentiments or other characteristics of Amazon phone reviews, providing valuable insights to users.



# EXPERIMENTAL INVESTIGATION

1. Which Brand has made the highest rating overall or Count of ratings overall?

[https://drive.google.com/file/d/1MlvpbLAtAPxVo2FXheuW7NTm4TsqK66j/view?usp=drive\\_link](https://drive.google.com/file/d/1MlvpbLAtAPxVo2FXheuW7NTm4TsqK66j/view?usp=drive_link)

2. Were the review help another user to buy products and what was the proposition of helpful votes to total votes/reviews?

[https://drive.google.com/file/d/1Mo08bjJnfil1YKnNuN90bzW4KEGr-TZT/view?usp=drive\\_link](https://drive.google.com/file/d/1Mo08bjJnfil1YKnNuN90bzW4KEGr-TZT/view?usp=drive_link)

3. Did the brand have a verified tag to it or not and which brand had not been verified?

[https://drive.google.com/file/d/1MoZpytl-pNxZru9mxyH6EOf3-zdFzelo/view?usp=drive\\_link](https://drive.google.com/file/d/1MoZpytl-pNxZru9mxyH6EOf3-zdFzelo/view?usp=drive_link)

4. How are all brands doing over the years and which brand is no more selling?

[https://drive.google.com/file/d/1MuuUH0x\\_i7c1DR9ZCBAJCW\\_Wqv0-srQkH/view?usp=drive\\_link](https://drive.google.com/file/d/1MuuUH0x_i7c1DR9ZCBAJCW_Wqv0-srQkH/view?usp=drive_link)

5. How is the original price different from the price we are selling in?

[https://drive.google.com/file/d/1MyC8OE-fzb5EK2aQK1lilf\\_c7h1mXp81/view?usp=drive\\_link](https://drive.google.com/file/d/1MyC8OE-fzb5EK2aQK1lilf_c7h1mXp81/view?usp=drive_link)

6. Which brand has the highest sale overall?

[https://drive.google.com/file/d/1N2CINiACACnFHwnKbfBVF3WEmfTvcvjf/view?usp=drive\\_link](https://drive.google.com/file/d/1N2CINiACACnFHwnKbfBVF3WEmfTvcvjf/view?usp=drive_link)

7. Which brand made the most profit of all?

[https://drive.google.com/file/d/1N58fzrMf2jEGiXJ-SCf0iRY\\_mvGZZXz4/view?usp=drive\\_link](https://drive.google.com/file/d/1N58fzrMf2jEGiXJ-SCf0iRY_mvGZZXz4/view?usp=drive_link)

## **FLOWCHART (WORKFLOW)**

1. Data Collection.
  - a. Collect the dataset or Create the dataset
2. Data cleaning.
  - a. Convert the CSV file to Excel
  - b. Do the necessary cleaning and convert back to CSV
3. Tableau.
  - a. Get insights from the data
  - b. Dashboard creation
  - c. Story creation
4. Text Preprocessing.
  - a. Import the Libraries.
  - b. Importing the dataset.
  - c. Remove Punctuations
  - d. Convert each word into a lowercase.
  - e. Stemming.
  - f. Splitting Data into Train and Test.
5. Model Building
  - a. Import the model building Libraries
  - b. Initializing the model
  - c. Adding Input Layer
  - d. Adding Hidden Layer
  - e. Adding Output Layer
  - f. Configure the Learning Process
  - g. Training and testing the model
  - h. Optimize the Model
  - i. Save the Model
6. Application Building
7. Create an HTML file
8. Build a Python Code

# IMPLEMENTATION VIDEO LINKS

## 1. Data Cleaning and Preprocessing

<https://drive.google.com/file/d/1uLiJii0FPhRzusvqOKAGLxQoxCGBAaRE/view?usp=sharing>

## 2. Creating Visualizations in Tableau

1. [https://drive.google.com/file/d/1MlvpbLAtAPxVo2FXheuW7NTm4TsqK66j/view?usp=drive\\_link](https://drive.google.com/file/d/1MlvpbLAtAPxVo2FXheuW7NTm4TsqK66j/view?usp=drive_link)
2. [https://drive.google.com/file/d/1Mo08bjJnfil1YKnNuN90bzW4KEGr-TZT/view?usp=drive\\_link](https://drive.google.com/file/d/1Mo08bjJnfil1YKnNuN90bzW4KEGr-TZT/view?usp=drive_link)
3. [https://drive.google.com/file/d/1MoZpytl-pNxZru9mxyH6EOf3-zdFzelo/view?usp=drive\\_link](https://drive.google.com/file/d/1MoZpytl-pNxZru9mxyH6EOf3-zdFzelo/view?usp=drive_link)
4. [https://drive.google.com/file/d/1MuuUH0x\\_i7c1DR9ZCBAJCWVqv0-srQkH/view?usp=drive\\_link](https://drive.google.com/file/d/1MuuUH0x_i7c1DR9ZCBAJCWVqv0-srQkH/view?usp=drive_link)
5. [https://drive.google.com/file/d/1MyC8OE-fzb5EK2aQK1lilf\\_c7h1mXp81/view?usp=drive\\_link](https://drive.google.com/file/d/1MyC8OE-fzb5EK2aQK1lilf_c7h1mXp81/view?usp=drive_link)
6. [https://drive.google.com/file/d/1N2CINiACACnFHwnKbfBVF3WEmfTvcvjf/view?usp=drive\\_link](https://drive.google.com/file/d/1N2CINiACACnFHwnKbfBVF3WEmfTvcvjf/view?usp=drive_link)
7. [https://drive.google.com/file/d/1N58fzrMf2jEGiXJ-SCf0iRY\\_mvGZZXz4/view?usp=drive\\_link](https://drive.google.com/file/d/1N58fzrMf2jEGiXJ-SCf0iRY_mvGZZXz4/view?usp=drive_link)

### 3. Creating Dashboard in Tableau

Dashboard 1:

[https://drive.google.com/file/d/1qmfUutny81IRTKYWI0AkOhPaLUMmYhXx/view?usp=drive\\_link](https://drive.google.com/file/d/1qmfUutny81IRTKYWI0AkOhPaLUMmYhXx/view?usp=drive_link)

Dashboard 2 & uploading to Tableau Public Server:

[https://drive.google.com/file/d/1nHFOGLOdYwusGxfPY9COoCRUdbxwXoKH/view?usp=drive\\_link](https://drive.google.com/file/d/1nHFOGLOdYwusGxfPY9COoCRUdbxwXoKH/view?usp=drive_link)

### 4. Creating a Story in Tableau

Story Creation:

[https://drive.google.com/file/d/1n308JBBPC3LQyp1n8CJ\\_FFp8DdSTIIAt/view?usp=drive\\_link](https://drive.google.com/file/d/1n308JBBPC3LQyp1n8CJ_FFp8DdSTIIAt/view?usp=drive_link)

Story Formatting & uploading to Tableau Public Server:

[https://drive.google.com/file/d/1ryRxH0M8od6r\\_MlmWTyNq31uhUu8g2yM/view?usp=drive\\_link](https://drive.google.com/file/d/1ryRxH0M8od6r_MlmWTyNq31uhUu8g2yM/view?usp=drive_link)

### 5. Uploading the Dashboards & Story to Tableau Public Server

[https://drive.google.com/file/d/1nHFOGLOdYwusGxfPY9COoCRUdbxwXoKH/view?usp=drive\\_link](https://drive.google.com/file/d/1nHFOGLOdYwusGxfPY9COoCRUdbxwXoKH/view?usp=drive_link)

[https://drive.google.com/file/d/1ryRxH0M8od6r\\_MlmWTyNq31uhUu8g2yM/view?usp=drive\\_link](https://drive.google.com/file/d/1ryRxH0M8od6r_MlmWTyNq31uhUu8g2yM/view?usp=drive_link)

### 6. Downloading a Bootstrap template and editing it for our Webpage

[https://drive.google.com/file/d/1LdAqXh1nMfCb9zPQ8-OVMcK2jN8nQ1Pz/view?usp=drive\\_link](https://drive.google.com/file/d/1LdAqXh1nMfCb9zPQ8-OVMcK2jN8nQ1Pz/view?usp=drive_link)

## 7. Creating a backend file using Python - Flask

[https://drive.google.com/file/d/10oAnt1Pu\\_37w8\\_NSYRgNbO-YjdBfUIM7/view?usp=drive\\_link](https://drive.google.com/file/d/10oAnt1Pu_37w8_NSYRgNbO-YjdBfUIM7/view?usp=drive_link)

## 8. Embedding uploaded Dashboards & Story from Tableau Public Server to our Webpage

[https://drive.google.com/file/d/1f-dV2tNnSIKmnssBz4KULN5jtXWgrEES/view?usp=drive\\_link](https://drive.google.com/file/d/1f-dV2tNnSIKmnssBz4KULN5jtXWgrEES/view?usp=drive_link)

## 9. Review Prediction

[https://drive.google.com/file/d/1sPitOrsqdR5LgLNmnBIG\\_nevs5Uu-oS\\_/view?usp=drive\\_link](https://drive.google.com/file/d/1sPitOrsqdR5LgLNmnBIG_nevs5Uu-oS_/view?usp=drive_link)

[https://drive.google.com/file/d/1gOtFOcHz1LsWPNZ21C0sM\\_Txna\\_LZTZ/view?usp=drive\\_link](https://drive.google.com/file/d/1gOtFOcHz1LsWPNZ21C0sM_Txna_LZTZ/view?usp=drive_link)

# APPLICATIONS

- 1. Sentiment analysis:** Analyze the sentiment expressed in cell phone reviews using natural language processing techniques in Python. Extract sentiment scores or classify reviews as positive, negative, or neutral. Visualize the sentiment distribution and trends in Tableau to understand customer perceptions and identify potential issues or strengths of different cell phone models.
- 2. Customer feedback analysis:** Explore customer feedback in Amazon cell phone reviews to identify common complaints, features that customers appreciate, or areas for improvement. Use Python to analyze the text data, extract keywords or topics, and quantify customer sentiments towards specific aspects of cell phones. Visualize the findings in Tableau to highlight key insights.
- 3. Review rating distribution:** Analyze the distribution of review ratings (e.g., 1 to 5 stars) for different cell phone models. Use Python to calculate summary statistics, such as average rating or rating frequency, and visualize the distribution using histograms or box plots in Tableau. Identify patterns or discrepancies in ratings across different brands or models.
- 4. Comparative analysis:** Compare the performance, features, or customer satisfaction levels of different cell phone brands or models. Use Python to calculate metrics like average ratings, review counts, or sentiment scores for each brand or model. Create visual comparisons using bar charts, heatmaps, or scatter plots in Tableau to identify trends, strengths, or weaknesses.
- 5. Time series analysis:** Analyze the temporal aspects of cell phone reviews. Use Python to extract the review date, group the reviews by time periods (e.g., monthly or quarterly), and calculate metrics like review counts or average ratings over time. Visualize the temporal patterns using line

charts or area charts in Tableau to observe changes in customer sentiment or review volume over the years.

**6. Interactive dashboards:** Combine the above analyses and visualizations into interactive dashboards in Tableau. Create filter controls, drill-down options, or tooltips to enable users to explore the data and gain insights dynamically. Incorporate Python-generated visualizations or summary statistics into the Tableau dashboards for a comprehensive view of the Amazon cell phone review data.

These applications leverage the strengths of Tableau for visual exploration and Python for data analysis to derive meaningful insights from Amazon cell phone reviews.



# CONCLUSION

1. **Feature analysis:** Analyzing customer feedback and reviews can provide insights into specific features or functionalities that customers appreciate or dislike in cell phones. You can identify key features that customers mention frequently in positive reviews and areas for improvement mentioned in negative reviews. This information can guide product development and marketing strategies.
2. **Brand or model comparison:** Comparing different cell phone brands or models based on review ratings, sentiment analysis, or other metrics can lead to conclusions about their relative performance. Identify the brands or models that consistently receive higher ratings or positive sentiment and those that have room for improvement. This information can be valuable for consumers making purchasing decisions and for manufacturers to assess their competitive positioning.
3. **Temporal analysis:** Analyzing the trends and patterns of cell phone reviews over time can provide insights into the evolution of customer sentiment and preferences. You can identify if certain models have gained or lost popularity over time, if customer satisfaction has improved or declined, or if there are any seasonal or cyclical patterns in review volume or sentiment.
4. **User segmentation:** By segmenting customers based on their review patterns or sentiments, you can conclude different customer groups and their preferences. Identify segments of customers who are highly satisfied or dissatisfied, early adopters, or price-sensitive. Tailor marketing strategies, product positioning, or customer support based on these segments.
5. **Impact of product updates or launches:** Analyzing reviews before and after product updates or new product launches can provide insights into the impact of these events on customer sentiment and satisfaction. Determine if updates or new launches have positively or negatively affected customer perceptions, and identify areas for improvement or success in product development.

## FUTURE SCOPE

1. **Real-time analysis:** Develop a system to collect and analyze Amazon cell phone reviews in real time using Python. This would enable businesses to monitor customer sentiment and feedback continuously and respond promptly to emerging trends or issues.

2. **Social media integration:** Extend the analysis beyond Amazon reviews and include data from social media platforms like Twitter or Facebook. By incorporating sentiment analysis of social media conversations related to cell phones, a more comprehensive understanding of customer perception can be achieved.

3. **Text mining and topic modeling:** Apply advanced text mining techniques in Python, such as topic modeling or text clustering, to extract latent topics and identify key themes in cell phone reviews. Visualize these topics in Tableau to gain deeper insights into customer preferences, pain points, or emerging trends.

4. **Advanced sentiment analysis:** Enhance sentiment analysis in Python by using more sophisticated techniques such as aspect-based sentiment analysis or emotion detection. This would allow for a more granular understanding of sentiment towards specific features or aspects of cell phones, helping businesses prioritize product improvements or marketing messages.

5. **Predictive analytics:** Employ predictive modeling techniques in Python to forecast customer sentiment or review ratings based on historical data. By integrating these predictive models with Tableau, businesses can anticipate future trends and make data-driven decisions accordingly.

## APPENDIX

Source Code:

<https://github.com/NIRANJAN-K-DESHMUKH/Amazon-Cell-Phone-Reviews>

Drive Link containing all the Video Recordings:

[https://drive.google.com/drive/folders/1pRp6mEwkamUoE\\_l3-Tv5g\\_ixOPIZVmTK?usp=drive\\_link](https://drive.google.com/drive/folders/1pRp6mEwkamUoE_l3-Tv5g_ixOPIZVmTK?usp=drive_link)