**SUMMER INTERNSHIP PROJECT REPORT**

**Project Title:** Customer Sentiment Analysis of i-Phone Models: Insights for Strategic Business Decisions

**Student Names** **:**

1. Suropriya Chakravarti ( Lady Brabourne College , Calcutta University )

2. Dhruv Ghag ( NMIMS )

3. Arghyadeep Biswas ( Ramkrishna Mission Residential College)

4. Chirantan Mahapatra ( Jadavpur University )

5. Ashmita Mukherjee ( Loreto College )

6. Arghya Bhowal ( Central Institute of Technology, Kokrajhar )

**Mentored by** Prof. Surojit Mukherjee, IIT KGP (VGSoM)

**Period of Internship:** 19th May 2025 - 15th July 2025

**Report submitted to:** IDEAS – Institute of Data Engineering, Analytics and Science Foundation, ISI Kolkata

**Abstract:**

The project aims to extract actionable business insights by analysing customer sentiment towards various iPhone models based on online reviews. The review data was collected through web scraping from major e-commerce platforms like Amazon and Flipkart. We used a Chrome extension designed for real-time web scraping. The collected data was cleaned using Python code and Excel functionalities.

The central aspect of the project involves applying an AI-powered sentiment analysis code to process the review data, yielding sentiment scores and counts of positive and negative for each review. The detailed results, along with the extracted review metadata, were stored in separate files for each product model. Finally, all processed data was consolidated into a single Excel file for further visualization and insights using Power BI.

**Introduction:**

Amidst the rapid growth of e-commerce, gaining a deep understanding of customer sentiment is vital for businesses striving to make data-driven decisions about product development, marketing strategies, and customer satisfaction initiatives. This project focuses on applying sentiment analysis to uncover valuable insights from the customer reviews available in retail platforms such as Amazon and Flipkart.

By transforming large volumes of unstructured text into measurable sentiment indicators, this project highlights how businesses can harness customer feedback to evaluate product performance and brand perception better. Our project demonstrates how sentiment analysis can be integrated into a broader business analytics framework, empowering businesses to turn customer opinions into strategic insights that drive smart decision-making and give them a competitive advantage.

Prerequisite knowledge acquired for this project :

1. Fundamentals of Business Analytics
2. Web Scraping Techniques
3. Data Cleaning and Preprocessing using Microsoft Excel
4. Concepts of Natural Language Processing
5. Sentiment Analysis Algorithms
6. Data visualization with Power BI

**Objectives:**

1. Collect customer reviews of different iPhone models from Amazon and Flipkart using a data scraper extension in Chrome.
2. Clean and pre-process the raw text data to ensure accuracy and consistency.
3. Apply sentiment analysis techniques to classify reviews as positive, negative or neutral.
4. Compare sentiment scores across different iPhone models and price segments to identify trends in perceived value.

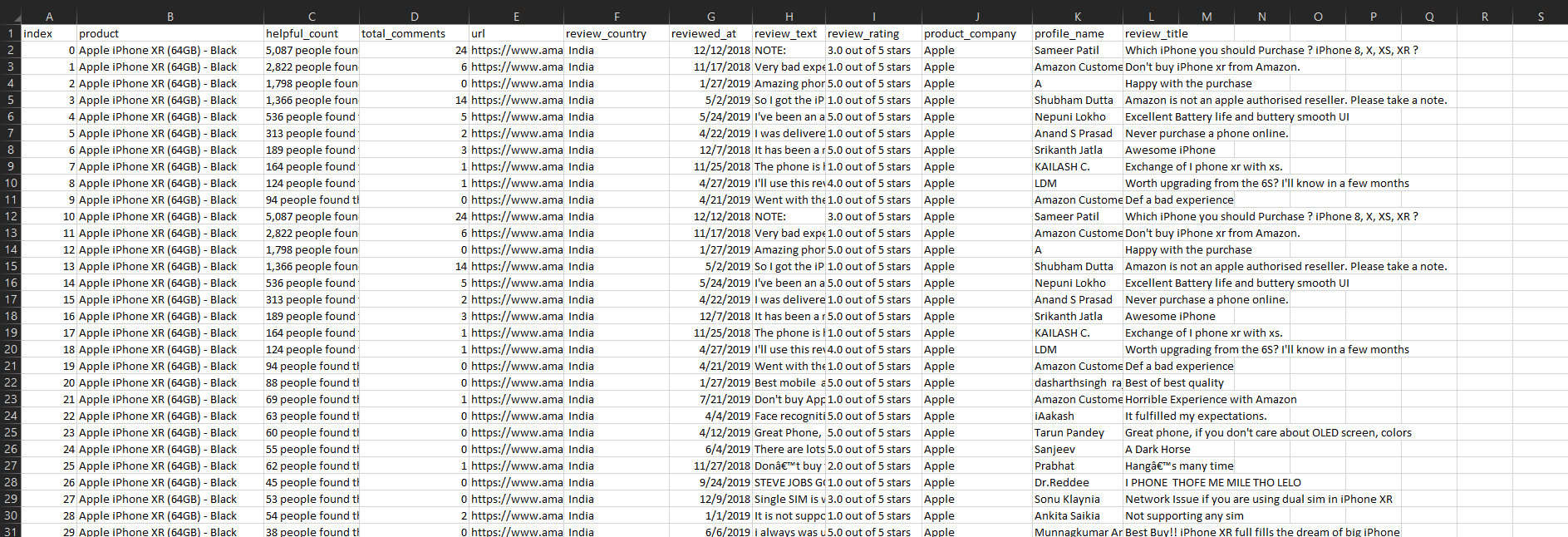
Compile the analysed data into a structured Excel sheet and visualize insights using Power BI dashboards for clear business interpretation.

**Methodology:**

The project followed a systematic methodology encompassing data acquisition, cleaning, sentiment analysis, data consolidation and visualization.  
*\*Please refer to the link given below in the References to showcase the data collected.*

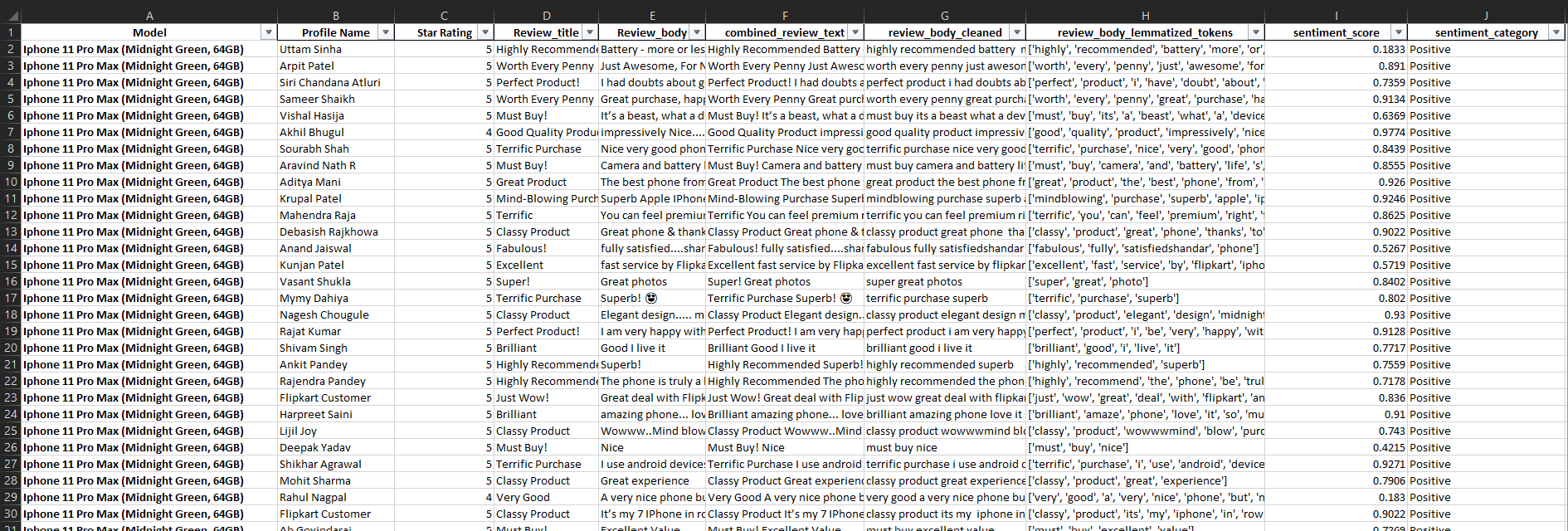
The detailed processes and steps undertaken are as follows:

1. **Data collection:**
   1. Raw customer reviews for various product models were acquired from Amazon and Flipkart
   2. A Chrome extension designed for real-time data scraping was used to extract the review data.
   3. The scraped data was then saved directly into an Excel or CSV file format.



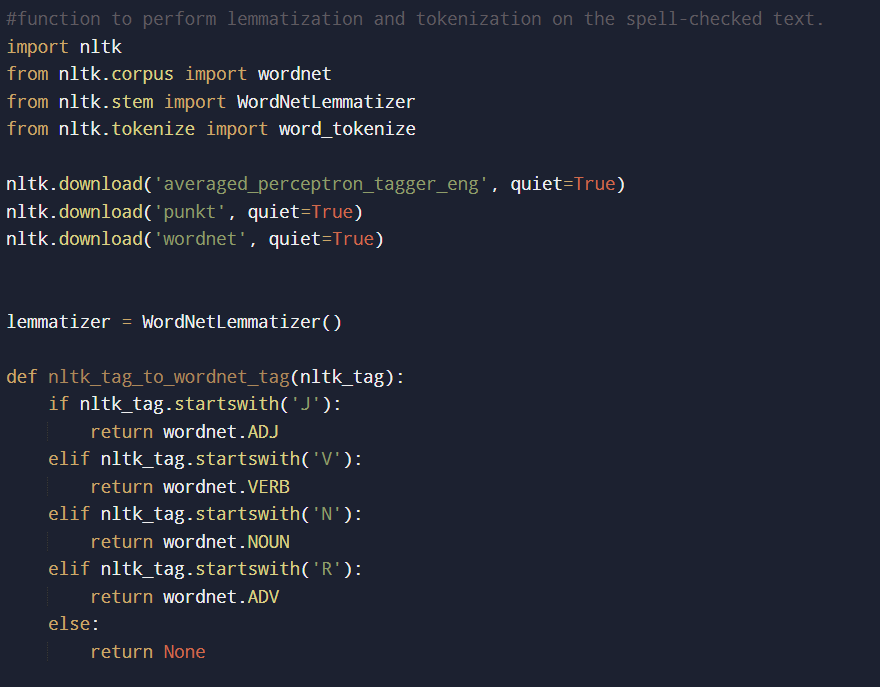
\**Raw data after extraction (uncleaned).*

1. **Data Cleaning and Preprocessing:**
2. The initial raw data containing irrelevant rows or formatting inconsistencies was cleaned using various Excel formulas and functions.
3. This involved trimming unnecessary spaces, removing blank rows and standardizing data formats to ensure data quality and consistency.
4. **Sentiment analysis:**
5. A natural language processing technique was applied to extract from the customer reviews was applied
6. Used lemmatization to standardize words, improving the accuracy of sentiment detection.
7. Leveraged the VADER sentiment analysis model to assign sentiment scores to each review.
8. The reviews were categorized into positive, negative, and neutral based on compound sentiment scores.
9. The unstructured textual data was converted into structured insights for business interpretation and visualization.

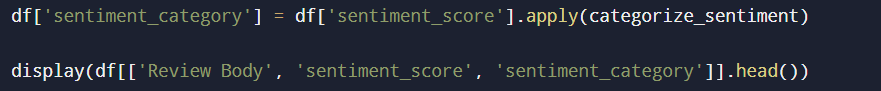
****

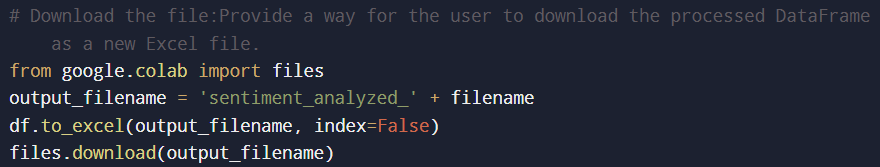
*\*Excel data after cleaning and Sentiment Analysed.*

1. **Data Consolidation:**
2. All the individually generated files containing processed sentiment analysis results for different product models were combined into a single, comprehensive Excel file.
3. Further modifications were performed within this consolidated file to ensure data integrity, such as omitting any remaining blank spaces.
4. **Data Visualization and Power BI:**
5. Power BI is a business intelligence platform developed by Microsoft that allows users to visualize and analyze data from various sources
6. This step involved importing the cleaned and analysed data into Power BI. Various charts, graphs and summary tables were used to represent the sentiment analysis insights clearly.

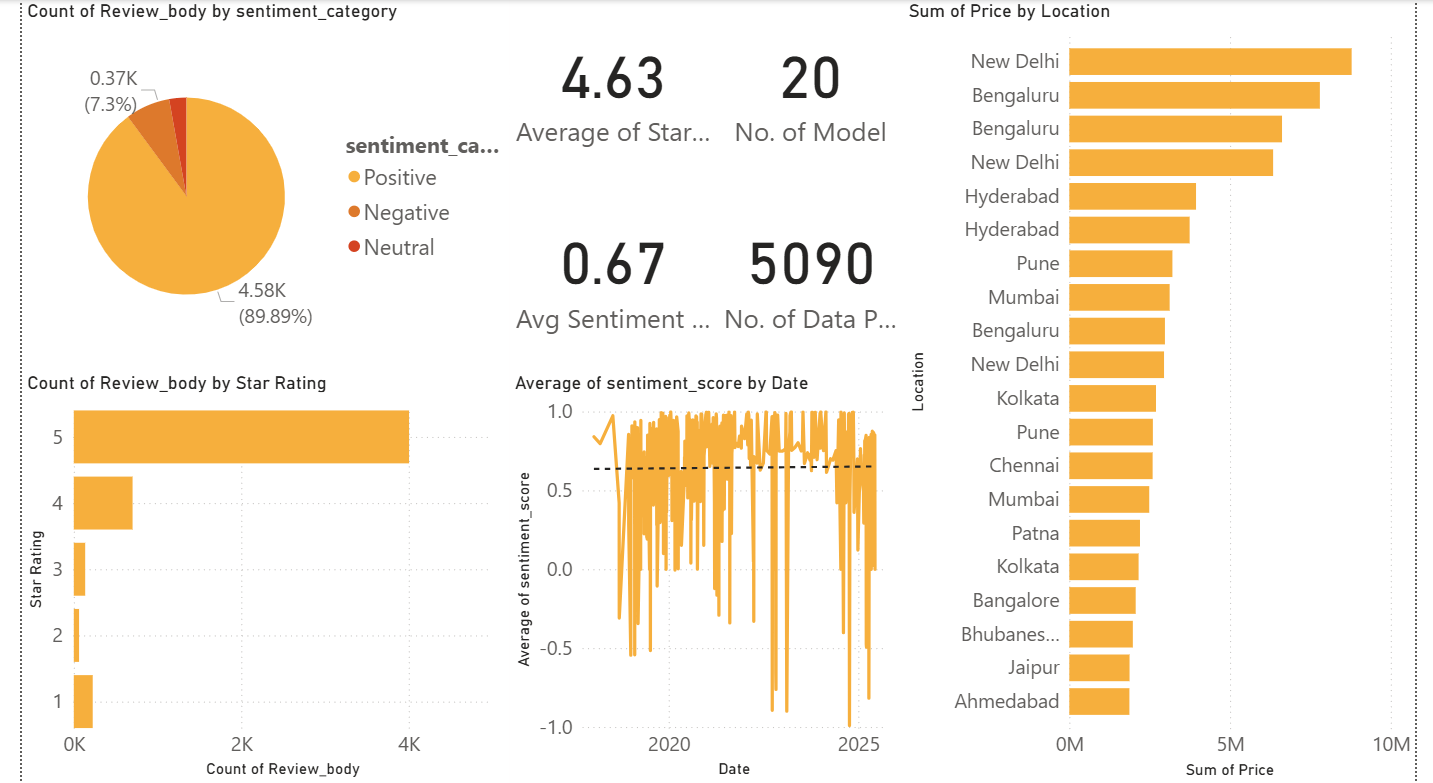
**Codes Used In Sentiment Analysis:**

****

****

****

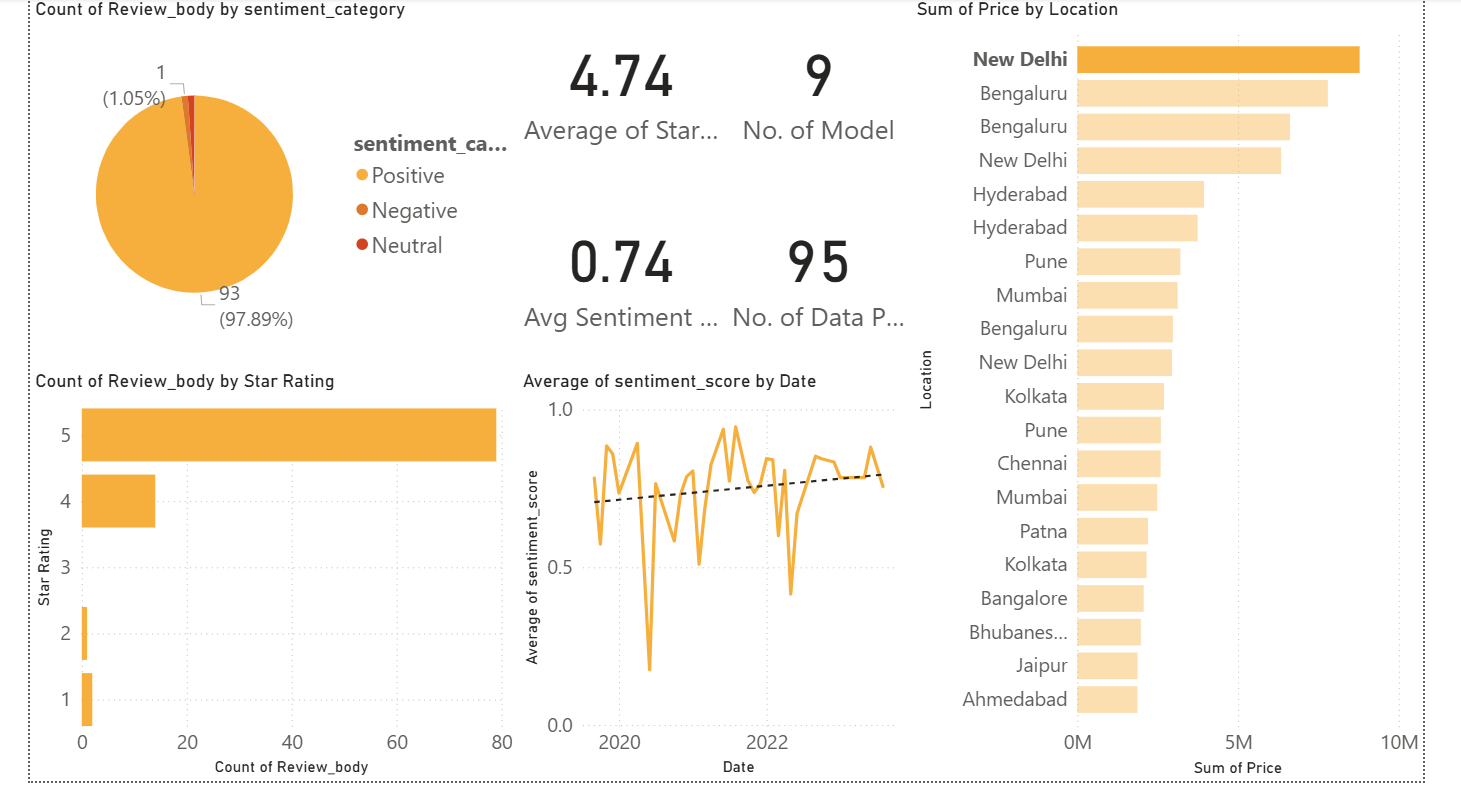
**Data Analysis Result and Insights:**



From the above graph we get the following:

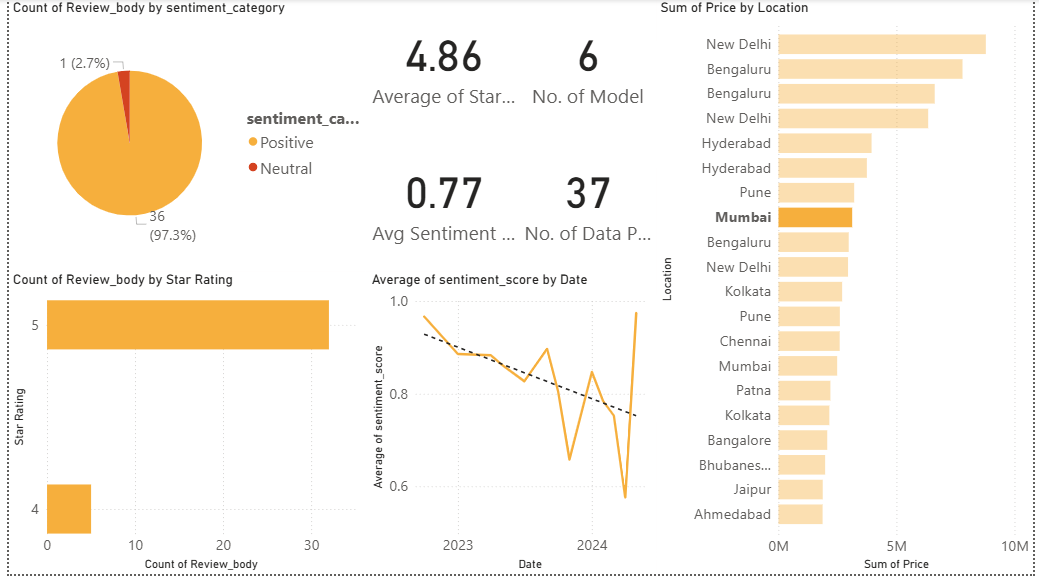
* + - 1. The majority of reviews were positive, with about 90% of feedback falling into the positive category, while only a small portion was negative or neutral.
      2. The average star rating across all iPhone models all over India was 4.63, indicating strong overall customer satisfaction. The average sentiment score was 0.67, further confirming a generally favourable perception among users.
      3. New Delhi and Bengaluru emerged as the major markets with the highest price contribution, followed by cities like Hyderabad, Pune, and Mumbai.
      4. The time series sentiment trend shows that positive sentiment has remained stable overall.
* **Specific location analysis:**

1. **New Delhi:**



From the above graph we get the following:

1. High Positive Sentiment: Around 98% of the reviews analyzed for New Delhi fall under the positive sentiment category, indicating strong overall satisfaction with iPhone models.
2. Star Ratings: The average star rating for New Delhi is 4.74,with the majority of reviews clustered around 5 stars.
3. Above average sentiment score: The average sentiment score calculated is 0.74.
4. Significant market contribution: New Delhi tops the price by location chart, implying that it contributes substantially to the total revenue generated. It emerges as a major urban market for flagship iPhone models.
5. Positive trend: The time series graph shows that sentiment scores have remained consistently positive over the recent years. A dip in the trend is observed in the post-lockdown period in India
6. **Mumbai :**



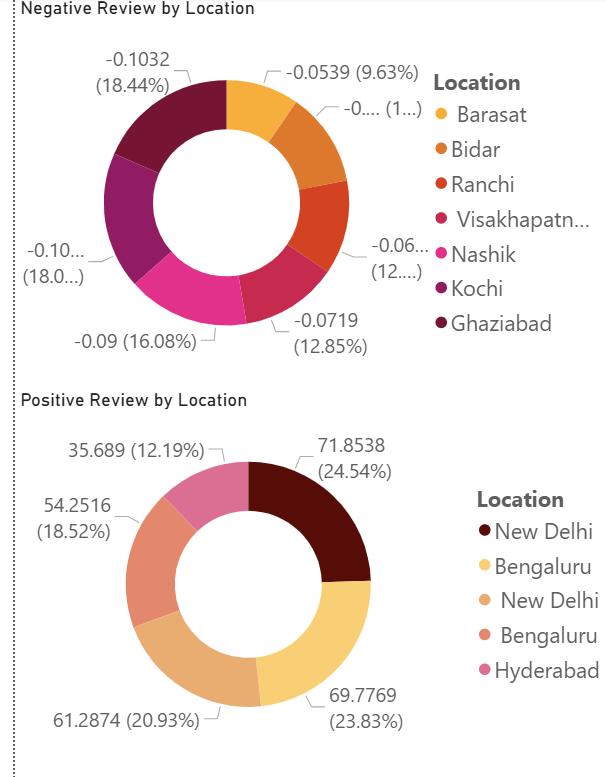
From the above graph we get the following:

1. Predominantly Positive Sentiment: 97.3% of the reviews from Mumbai fall into the positive category. 2.7% are marked as neutral, indicating very high customer satisfaction.
2. Excellent star rating: The average star rating is 4.86, one of the highest across regions.
3. Strong sentiment score: The average sentiment score is 0.77, aligning well with high star ratings. Both structured and unstructured (review text) data reflect a positive perception.
4. Slight declining trend: The time series trend shows that while overall sentiment remains high, there is a downward drift in reviews. The lowest sentiment score was obtained in April of 2024. This may suggest shifting expectations or minor issues in the product.
5. **Kolkata :**



From the above graph we get the following:

1. Highly Positive Feedback:About 97% of reviews are positive,while only around 3% are negative, reflecting strong customer trust and satisfaction .
2. Excellent Ratings: Impressive average star rating of 4.89.
3. Robust Sentiment Score: Average sentiment score is 0.76, showing textual sentiment aligns with star ratings given.
4. Upward trend: Unlike Mumbai, Kolkata’s sentiment score over time shows a positive trend. This is a healthy sign for brand loyalty and product quality.
5. Moderate Revenue Position: While Kolkata sits mid-tier in total price contribution, its strong sentiment trend suggests its potential to expand sales through local marketing and product positioning.
6. While Kolkata shows high customer sentiment and star ratings, its current moderate sales value and lower concentration of high-income tech clusters compared to cities like Bengaluru, Delhi, and Mumbai may limit its immediate capacity to become a flagship retail hub



* Negative reviews by location:

The analysis of negative sentiment by location reveals that certain smaller cities, including Barasat, Bidar, Ranchi, Visakhapatnam, Nashik, Kochi, and Ghaziabad, showed a relatively higher proportion of negative reviews. This indicates possible gaps in service quality, product performance, or delivery experience in these regions.

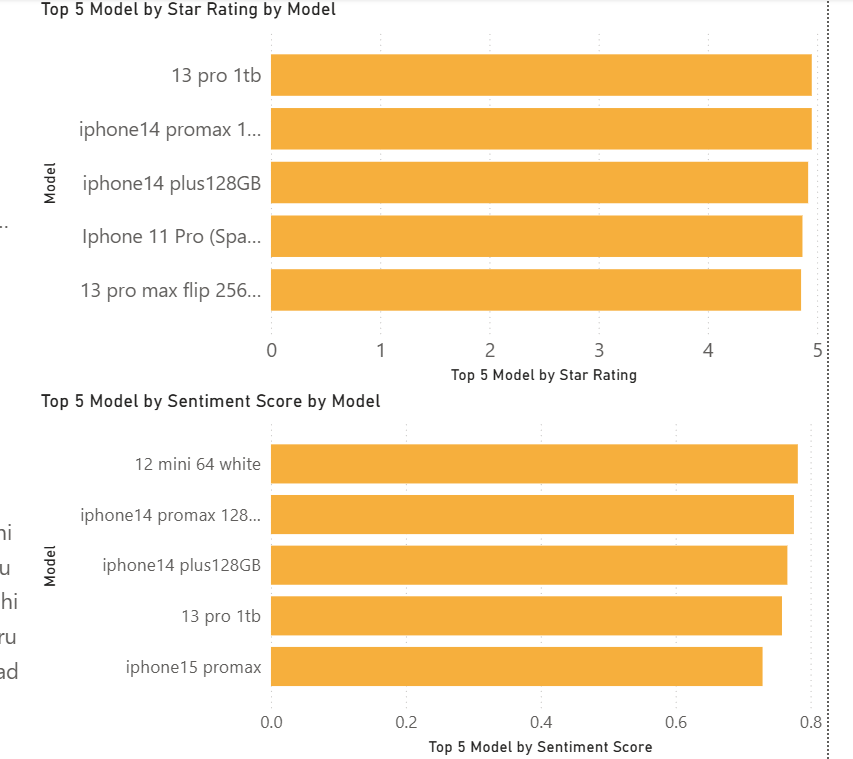
* Positive Reviews by location:

Positive reviews were predominantly concentrated in major metropolitan cities such as New Delhi, Bengaluru, and Hyderabad, which contribute significantly to sales. This suggests that in these cities, customers are generally satisfied with the overall experience, reinforcing these locations for brand loyalty and repeat purchases.

**Insights(from location specific analysis):**

1. Strong Demands and Affluent Markets: Metropolitan cities like Mumbai, Delhi, Bengaluru, and Pune are home to wealthy, urban populations of professionals, IT workers and students whose purchasing power fuels demand for premium products
2. Strategic Retail Locations: Apple’s Choice of prime retail spots in malls and iconic areas, such as Jio World Drive in Mumbai and Select CITYWALK in Delhi, strengthens brand visibility and attracts affluent shoppers, aligning with global retail strategies.

* **Star Rating vs Sentiment Score:**



From the above graph we get the following:

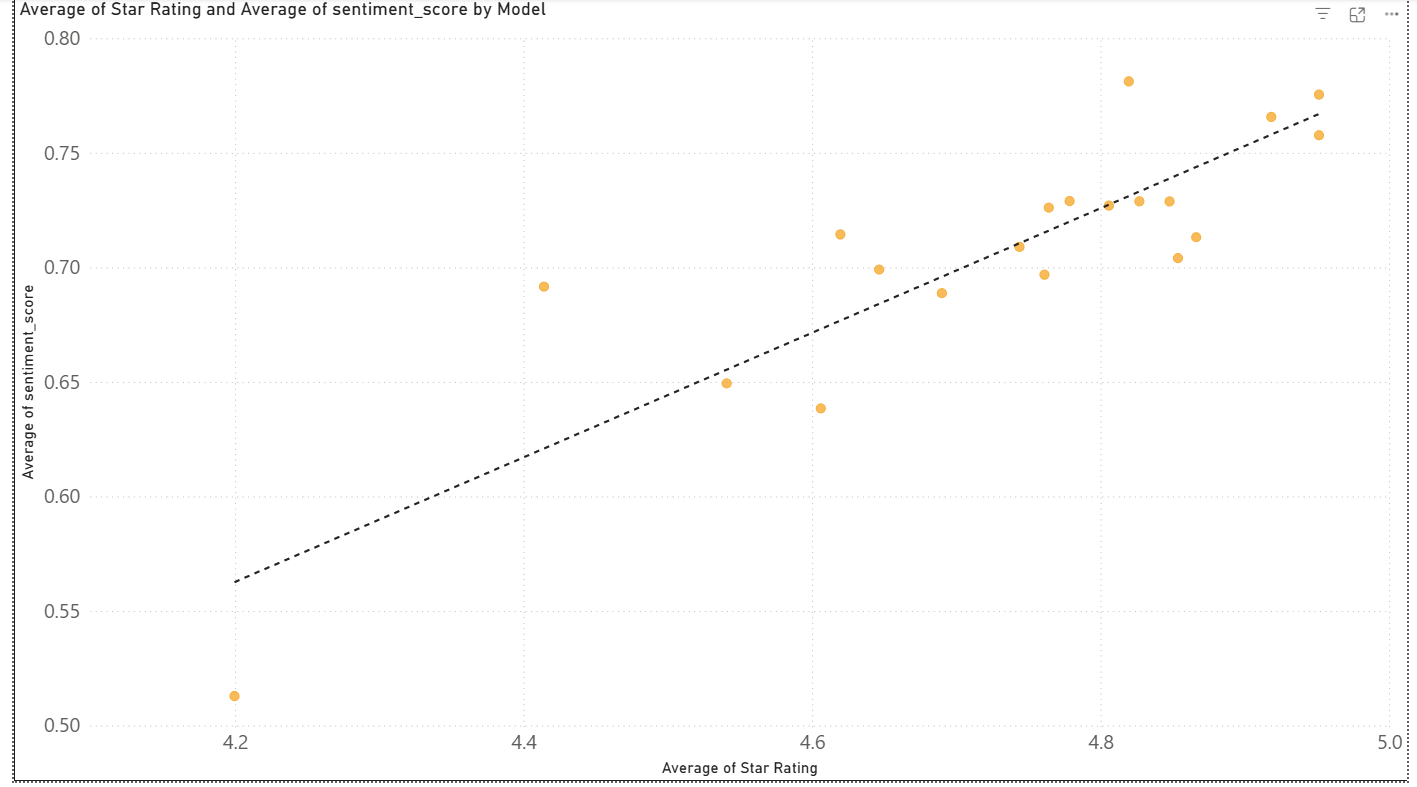
1. Top models by star rating:

The top five iPhone models with the highest average star rating were iPhone 13 Pro 1TB,iPhone 14 Plus 128GB, iPhone 11 Pro, and iPhone 13 Pro Max. The consistently high ratings indicate that premium storage variants and pro models deliver strong customer satisfaction, aligning with their positioning as flagship devices.

1. Top models by sentiment score :

When ranked by sentiment score, the iPhone 12 Mini 64GB (white), iPhone 14 Pro Max 128GB, iPhone 14 Plus 128GB, iPhone 13 Pro 1TB, and iPhone 15 Pro Max were the standout performers.

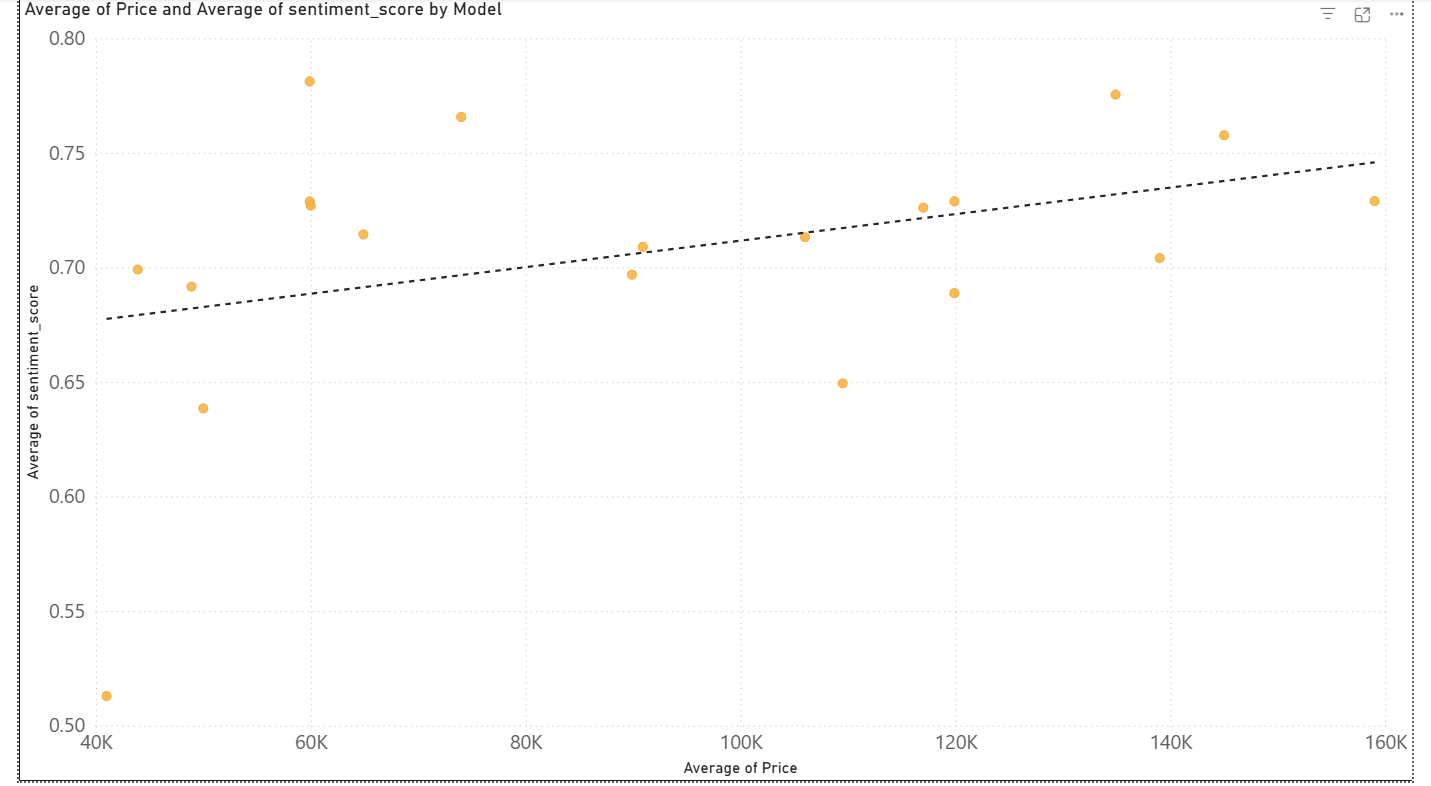
**Comparison between star ratings and sentiment score analysis:**



From the above graph we get the following:

1. A star rating is a direct, structured input where the customer assigns numerical values based on their initial experience. Many customers quickly give ratings out of habit or to meet the minimum requirement of initial ratings to post a review. These ratings may not represent the customer's opinion accurately.
2. Sentiment scores are implicit. They are calculated using natural language processing(NLP). It interprets words, phrases and other things which the star ratings cannot capture. The sentiment model used in our case is VADER . VADER is good at interpreting language in a general tone but may not fully detect context-specific meanings.
3. From the graph above, we can see an approx. linear relation between the Average star ratings and average sentiment scores for different models of iPhone.

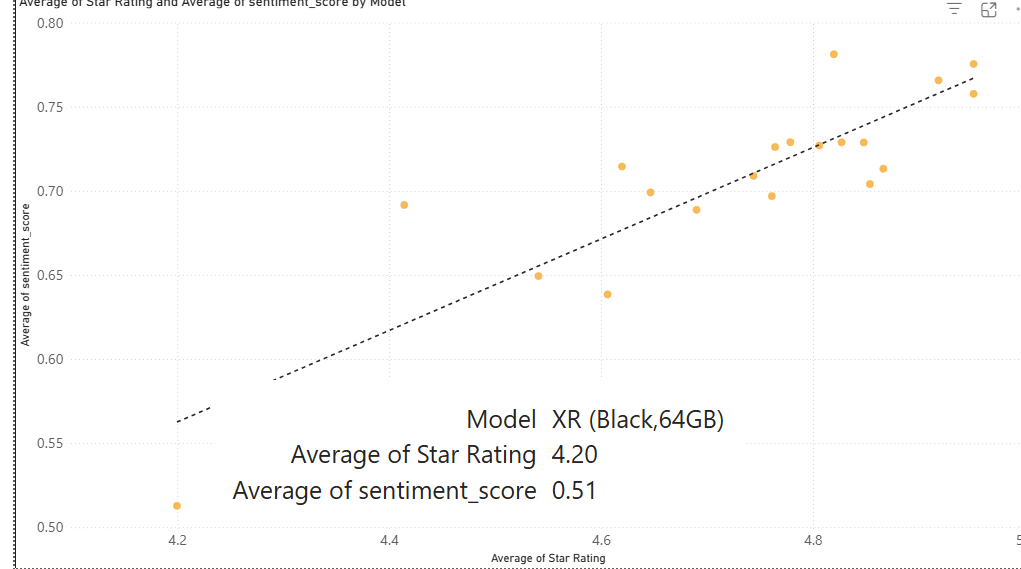
* **Sentiment Score vs Price Analysis :**



From the above graph we get the following:

1. The trend shows a slightly positive correlation between price and sentiment score,indicating that expensive iPhone models slightly receive more positive sentiment in customer feedback.
2. This suggests that customers who purchase premium flagship models are often satisfied with their overall experience, possibly due to better performance ,advanced features or perceived value for money.
3. However, price is not the only metric for customer satisfaction . The spread of points in the scatter plot also represent a strong positive sentiment in midrange models of iPhones. iPhone 12 Mini stands out with the highest customer sentiment score.
4. This insight can help guide product positioning and pricing strategies.

**Outliers**

****

A noticeable outlier with the lowest sentiment score was observed across the data. This outlier is the iPhone model XR. This point sits significantly below the trend line, suggesting a significant mismatch between the price and the perceived value.

**Actionable Insights:**

❖ Investigation on Outlier Models

We looked into the iPhone model with the lowest sentiment score and found that most of the negative feedback revolves around common issues like battery performance, heating, and delivery delays. These recurring complaints should be taken seriously by the company. Resolving these specific pain points could not only reduce negative reviews but also improve trust in future product launches.

❖ Reinforce Flagship Model Strengths

The high-end iPhone models are clearly performing well in terms of customer satisfaction. They consistently receive strong positive sentiment, especially around features like camera quality and display. These models should be marketed more aggressively in premium segments, as they are already attracting good sentiment and show strong brand loyalty.

❖ Enhance Mid-Range Value Propositions

Our analysis shows that mid-range and older models are still quite popular but are sometimes criticized when compared to the newer ones. To keep these models attractive, Apple should focus on maintaining key features like performance, build quality, and camera, even in lower-priced phones. This will help keep customer satisfaction high, even at lower price points.

❖ Region-Specific Gaps

Some regions, especially Tier-2 cities, show lower sentiment scores, mostly due to poor after-sales service, slow delivery, or lack of local support. Apple could work on improving its regional service experience in these areas — for example, by improving service center response times or making support more accessible in local languages.

❖ Next Best Location

Based on the data, Bengaluru and Hyderabad stand out as the next best locations to open a new Apple Store. Both cities have high review volumes, strong sales trends, and consistently positive sentiment. Expanding offline presence in these regions could help boost sales and create stronger customer engagement.

❖ Insights on Kolkata

Interestingly, Kolkata has one of the highest sentiment scores (around 0.81), even higher than the overall average. However, its sales numbers are relatively low. This suggests that people who do use iPhones there are very satisfied, but the reach is limited. To tackle this, Apple could focus on regional marketing efforts in Bengali, including online ads, influencer tie-ups, and localized promotions to tap into a larger user base.

**Conclusion:**

This project shows how sentiment analysis can produce useful insights for strategic business decisions when paired with interactive data visualization and structured star ratings. We discovered that Apple consistently enjoys high customer satisfaction across India’s major urban markets by methodically scraping ,cleaning and analyzing customer reviews of various iPhone models from major e-commerce platforms.

The region-by-region analysis demonstrates that cities such as New Delhi, Mumbai are not only technologically adept but also exhibit a strong correlation between premium pricing and favourable customer sentiment, thereby solidifying their position as perfect markets for Apple’s flagship retail strategy.

As indicated by outlier models with lower sentiment scores , the results show that while premium models typically have higher sentiment scores ,price by itself does not ensure satisfaction. In order to preserve Apple’s premium positioning this highlights the necessity of proactive problem solving,careful after-sales support and ongoing product improvement .

In summary, this study demonstrates that customer sentiment is a valuable strategic asset that guides more intelligent marketing, focused retail investments and a sustained competitive edge in India’s rapidly growing premium smartphone market .

**References**:

1. *Raw Data scraped (Iphone11)*

<https://drive.google.com/file/d/1HZMeKw00Clja9K0QU1rvn7OprZxibaRw/view?usp=drive_link>

1. *Final Excel data after cleaning (contains all the models)*

<https://docs.google.com/spreadsheets/d/1nbwZpdy2aR6h_7sK808ylwYhMTJX4dBe/edit?usp=drive_link&ouid=114193282698810927708&rtpof=true&sd=true>