Customer Sentiment Analysis - I Phone -128gb

Data Analyst at Flipkart ， I have been tasked with gauging customer sentiment towards the iPhone 15 128GB model.

Objective:

The primary goal of this project is to analyze public perception and evaluate customer reactions by performing sentiment analysis on product reviews posted by users. By extracting and processing customer reviews, you will derive insights about the overall sentiment (positive or negative) surrounding the product, which can be useful for decision-making, improving customer experience, and identifying key areas for product improvement.

Keep Focus on it:

1. Collect and process customer Name , Rating and Review from Flipkart website.
2. Categorize Reviews into Extremely Positive , Positive , Neutral , negative, extremely negative.
3. Visualize sentiment trends and feature mentions using charts or word clouds.
4. Summarize actionable insights to inform product improvement and marketing strategies.

Tools : Jupyter Notebook - Python

Libraries : Selenium ， BeautifulSoup ,Numpy, Pandas , Text Blob , Matplotlib , Seaborn , nltk , requests , Time , .

1. Data Collection (Web Scraping):

Tool: Selenium and BeautifulSoup

Task: Scrape at least 300 customer reviews from Flipkart's product page for the iPhone 15 128GB model. Each review should include**:**

Username: The name of the reviewer.

Rating: The rating provided by the user (1 to 5 stars).

Review Text: The content of the customer's review, which may contain valuable information regarding their experience with the product.

Steps:

- Set up Selenium to automate browser interactions, navigate to Flipkart’s product page for iPhone 15 128GB, and extract the reviews.
- Use BeautifulSoup to parse the HTML of the reviews and extract the relevant details (username, rating, and review text).
- Ensure that the scraper handles pagination to retrieve reviews from multiple pages if necessary.

Import Important Libraries:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from bs4 import BeautifulSoup
from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.common.keys import Keys
import time
from textblob import TextBlob
from nltk import sent_tokenize
import requests

driver = webdriver.Edge()
```

Now we Started Applying Libraries:

```
review=[]

url ="""https://www.flipkart.com/apple-iphone-15-blue-128-gb/product-reviews/itmbf14ef54f645d?pid=MOBGTAGPAQNVFZZY&lid=LSTMOBGTAGPAQNVFZZY7RHDU7&market

for i in range(1,30):
    new_url = url + "&page="+str(i)
    r = requests.get(new_url)
    soup = BeautifulSoup(r.text,'html.parser')
    Names = soup.find_all("p" , {"class":"_2NsDsF AwS1CA"})
    for i in Names:
        name.append(i.text)
    name

    Ratings = soup.find_all("div" ,{"class":"XQDdHH Ga3i8K"})
    for i in Ratings:
        rating.append(i.text)
    rating

    Reviews = soup.find_all("div",{"class":"ZmyHeo"})
    for i in Reviews:
        review.append(i.text)
    review
print(len(name) ,len(rating), len(review))
```

```
120 120 120
```

It takes Random 120 Customer Names , Ratings , Reviews.

I call the data Frame:

```
#dataframe
df = pd.DataFrame({"Customer Names" : name ,"Rating":rating,"Review":review})
df
```

	Customer Names	Rating	Review
0	Rishabh Jha	5	Awesome 🔥 🔥 😊READ MORE
1	bijaya mohanty	5	Just go for it.Amazing one.Beautiful camera wi...
2	Ajin V	5	High quality camera😻READ MORE
3	Saurabh Gothwad	5	Simply premium.READ MORE
4	Mousam Guha Roy	4	Very niceREAD MORE
...	...	...	...
115	Ritik Tomar	5	Nice productREAD MORE
116	Flipkart Customer	4	just love it ...READ MORE
117	John Fanai	5	Awesome 👌 As a first time ios user, I literall...
118	Ranjith kumar	5	Bought 256gb green variant 😻,new to iosIOS is ...
119	Karan Roy	5	Worth every penny! Design is fabulous.. in han...

120 rows × 3 columns

2. Data Cleaning and Preprocessing:

Tool: Pandas

Task: Clean and preprocess the scraped data for analysis.

Steps:

- **Remove duplicates:** Eliminate any duplicate reviews to ensure data quality.
- Handle missing values: Address missing or incomplete data, such as missing review text or rating, by either removing rows or filling in missing values if applicable.

Text preprocessing:

- **Convert the review text to lowercase.**
- **Remove irrelevant characters (e.g., special characters, punctuation, and extra spaces).**
- **Tokenize the text into individual words.**
- **Remove stop words (commonly used words that do not add significant meaning to sentiment analysis).**
- **Perform lemmatization to convert words into their base form (e.g., "running" → "run").**

Put Titles:

```
#get title
df ["Customer Names"] = df["Customer Names"].str.title()
df
```

	Customer Names	Rating	Review
0	Rishabh Jha	5	Awesome 🔥 🔥 😊READ MORE
1	Bijaya Mohanty	5	Just go for it.Amazing one.Beautiful camera wi...
2	Ajin V	5	High quality camera😍READ MORE
3	Saurabh Gothwad	5	Simply premium.READ MORE
4	Mousam Guha Roy	4	Very niceREAD MORE
...	...	...	...
115	Ritik Tomar	5	Nice productREAD MORE
116	Flipkart Customer	4	just love it ...READ MORE
117	John Fanai	5	Awesome 🤞 As a first time ios user, I literall...
118	Ranjith Kumar	5	Bought 256gb green variant 😍,new to iosIOS is ...
119	Karan Roy	5	Worth every penny! Design is fabulous.. in han...

120 rows × 3 columns

Remove READ MORE from Review.

```
#remove readmore in review

df["Review"] = df["Review"].str.replace("READ MORE", "")
df
```

	Customer Names	Rating	Review
0	Rishabh Jha	5	Awesome 🔥 🔥 😊
1	Bijaya Mohanty	5	Just go for it.Amazing one.Beautiful camera wi...
2	Ajin V	5	High quality camera😍
3	Saurabh Gothwad	5	Simply premium.
4	Mousam Guha Roy	4	Very nice
...	...	...	...
115	Ritik Tomar	5	Nice product
116	Flipkart Customer	4	just love it ...
117	John Fanai	5	Awesome 👌 As a first time ios user, I literall...
118	Ranjith Kumar	5	Bought 256gb green variant 😍,new to iosIOS is ...
119	Karan Roy	5	Worth every penny! Design is fabulous.. in han...

120 rows × 3 columns

Tokenize the text into individual words:

```
#create review1 extra
df["Review_1"] = df["Review"].apply(sent_tokenize)
df
```

	Customer Names	Rating	Review	Review_1
0	Rishabh Jha	5	Awesome 🔥🔥😊	[Awesome 🔥🔥😊]
1	Bijaya Mohanty	5	Just go for it.Amazing one.Beautiful camera wi...	[Just go for it.Amazing one.Beautiful camera w...
2	Ajin V	5	High quality camera😍	[High quality camera😍]
3	Saurabh Gothwad	5	Simply premium.	[Simply premium.]
4	Mousam Guha Roy	4	Very nice	[Very nice]
...	...	...	...	...
115	Ritik Tomar	5	Nice product	[Nice product]
116	Flipkart Customer	4	just love it ...	[just love it ...]
117	John Fanai	5	Awesome 👌 As a first time ios user, I literall...	[Awesome 👌 As a first time ios user, I literal...
118	Ranjith Kumar	5	Bought 256gb green variant 😍,new to iosIOS is ...	[Bought 256gb green variant 😍,new to iosIOS is...
119	Karan Roy	5	Worth every penny! Design is fabulous.. in han...	[Worth every penny!, Design is fabulous.. in h...

3. Sentiment Analysis:

Tool: TextBlob

Task: Analyze the sentiment of each review to classify them as either positive or negative.

Steps:

- **Use TextBlob to perform sentiment analysis on the review text.**

TextBlob will provide a polarity score between -1 (negative) and +1 (positive), as well as a subjectivity score.

- **Define a threshold to classify the sentiment:**

Positive sentiment: Polarity score ≥ 0.1

Negative sentiment: Polarity score < 0.1

- **Store the sentiment classification for each review in the dataset.**

```
#polarity
from textblob import blob
def get_polarity(sentences):
    return[TextBlob(sentence).sentiment.polarity for sentence in sentences]
df['Polarity'] = df['Review_1'].apply(get_polarity)
df
```

	Customer Names	Rating	Review	Review_1	Polarity
0	Rishabh Jha	5	Awesome 🔥🔥😊	[Awesome 🔥🔥😊]	[1.0]
1	Bijaya Mohanty	5	Just go for it.Amazing one.Beautiful camera wi...	[Just go for it.Amazing one.Beautiful camera w...	[0.26666666666666666]
2	Ajin V	5	High quality camera😍	[High quality camera😍]	[0.16]
3	Saurabh Gothwad	5	Simply premium.	[Simply premium.]	[0.0]
4	Mousam Guha Roy	4	Very nice	[Very nice]	[0.78]
...	...	...	...	...	...
115	Ritik Tomar	5	Nice product	[Nice product]	[0.6]
116	Flipkart Customer	4	just love it ...	[just love it ...]	[0.5]
117	John Fanai	5	Awesome 👌 As a first time ios user, I literall...	[Awesome 👌 As a first time ios user, I literal...	[0.45, 0.084, 0.0, 0.0, 0.0]
118	Ranjith Kumar	5	Bought 256gb green variant 😍,new to iosIOS is ...	[Bought 256gb green variant 😍,new to iosIOS is...	[-0.125]
119	Karan Roy	5	Worth every penny! Design is fabulous.. in han...	[Worth every penny!, Design is fabulous.. in h...	[0.375, 0.55]

120 rows × 5 columns

Mean of the Polarity

```
# mean of the polarity.
from statistics import mean
from textblob import blob
def get_polarity(sentences):
    return[TextBlob(sentence).sentiment.polarity for sentence in sentences]
df['Polarity'] = df['Review_1'].apply(get_polarity)

df.head(15)
```

	Customer Names	Rating	Review	Review_1	Polarity
1	Bijaya Mohanty	5	Just go for it.Amazing one.Beautiful camera wi...	[Just go for it.Amazing one.Beautiful camera w...	[0.26666666666666666]
2	Ajin V	5	High quality camera 😍	[High quality camera 😍]	[0.16]
3	Saurabh Gothwad	5	Simply premium.	[Simply premium.]	[0.0]
4	Mousam Guha Roy	4	Very nice	[Very nice]	[0.78]
5	Flipkart Customer	5	Awesome photography experience. Battery backup...	[Awesome photography experience., Battery back...	[1.0, 0.7, 0.5]
6	Nikhil Kumar	5	Switch from OnePlus to iPhone I am stunned wit...	[Switch from OnePlus to iPhone I am stunned wi...	[0.0, 1.0]
7	Prithivi Boruah	5	Camera Quality Is Improved Loving It	[Camera Quality Is Improved Loving It]	[0.6]
8	Rahul Shedge	5	Totally happy!Camera 5Battery 5 Display 5Design 5	[Totally happy!Camera 5Battery 5 Display 5Desi...	[0.0]
9	Akshay Meena	5	So beautiful, so elegant, just a vowww 😍 ❤️	[So beautiful, so elegant, just a vowww 😍 ❤️]	[0.675]
10	Arunji Govindaraju	5	Awesome product very happy to hold this. Bette...	[Awesome product very happy to hold this., Bet...	[1.0, 0.5, 0.45555555555555555]
11	Sheetla Prasad Maurya	4	Best mobile phoneCamera quality is very nice B...	[Best mobile phoneCamera quality is very nice ...	[0.738]
12	Kriti Customer	5	Just loved the product , colour , design is wo...	[Just loved the product , colour , design is w...	[0.4125]
13	Talim (Sk)	5	I love this phone. The camera is amazing: its ...	[I love this phone., The camera is amazing: it...	[0.5, 0.7000000000000001]
14	Mohit Yadav	5	Nice ❤️	[Nice ❤️]	[0.6]

Top 15 Customers.

Average Polarity 😀

```
#average_polarity
from statistics import mean
def calculate_average_polarity(polarities):
    return mean(polarities) if polarities else 0

df['Average_Polarity'] = df['Polarity'].apply(calculate_average_polarity)

df.head(15)
```

	Customer Names	Rating	Review	Review_1	Polarity	Average_Polarity
0	Rishabh Jha	5	Awesome 🔥 🔥 😊	[Awesome 🔥 🔥 😊]	[1.0]	1.000000
1	Bijaya Mohanty	5	Just go for it.Amazing one.Beautiful camera wi...	[Just go for it.Amazing one.Beautiful camera w...	[0.26666666666666666]	0.266667
2	Ajin V	5	High quality camera😍	[High quality camera😍]	[0.16]	0.160000
3	Saurabh Gothwad	5	Simply premium.	[Simply premium.]	[0.0]	0.000000
4	Mousam Guha Roy	4	Very nice	[Very nice]	[0.78]	0.780000
5	Flipkart Customer	5	Awesome photography experience. Battery backup...	[Awesome photography experience., Battery back...	[1.0, 0.7, 0.5]	0.733333
6	Nikhil Kumar	5	Switch from OnePlus to iPhone I am stunned wit...	[Switch from OnePlus to iPhone I am stunned wi...	[0.0, 1.0]	0.500000
7	Prithivi Boruah	5	Camera Quality Is Improved Loving It	[Camera Quality Is Improved Loving It]	[0.6]	0.600000
8	Rahul Shedge	5	Totally happy!Camera 5Battery 5 Display 5Design 5	[Totally happy!Camera 5Battery 5 Display 5Desi...	[0.0]	0.000000
9	Akshay Meena	5	So beautiful, so elegant, just a vowww😍 💖	[So beautiful, so elegant, just a vowww😍 💖]	[0.675]	0.675000
10	Arunji Govindaraju	5	Awesome product very happy to hold this. Bette...	[Awesome product very happy to hold this., Bet...	[1.0, 0.5, 0.45555555555555555]	0.651852
11	Sheetla Prasad Maurya	4	Best mobile phoneCamera quality is very nice B...	[Best mobile phoneCamera quality is very nice ...	[0.738]	0.738000
12	Kriti Customer	5	Just loved the product , colour , design is wo...	[Just loved the product , colour , design is w...	[0.4125]	0.412500
13	Talim (Sk)	5	I love this phone. The camera is amazing: its ...	[I love this phone., The camera is amazing: it...	[0.5, 0.7000000000000001]	0.600000
14	Mohit Yadav	5	Nice 💖	[Nice 💖]	[0.6]	0.600000

Average of Top 15 Customer.

Sentiments:

```
#sentiment of each review to classify them as either positive or negative.

#Extremly Positive , Positive ,Neutral ,Negtive , Extremaly Negative.

def sentiment(pol):
    if pol >= 0.738:
        return "Extremly positive"
    elif pol > 0:
        return "positive"
    elif pol == 0:
        return "neutral"
    elif pol <= -0.738:
        return "negative"
    else :
        return "extremly negative"

df["Sentiment"] = df["Average_Polarity"].apply(sentiment)
df
```

	Customer Names	Rating	Review	Review_1	Polarity	Average_Polarity	Sentiment
0	Rishabh Jha	5	Awesome 🔥🔥😊	[Awesome 🔥🔥😊]	[1.0]	1.000000	Extremly positive
1	Bijaya Mohanty	5	Just go for it.Amazing one.Beautiful camera wi...	[Just go for it.Amazing one.Beautiful camera w...	[0.26666666666666666]	0.266667	positive
2	Ajin V	5	High quality camera😍	[High quality camera😍]	[0.16]	0.160000	positive
3	Saurabh Gothwad	5	Simply premium.	[Simply premium.]	[0.0]	0.000000	neutral
4	Mousam Guha Roy	4	Very nice	[Very nice]	[0.78]	0.780000	Extremly positive
...	...	...	...	...	...	...	...
115	Ritik Tomar	5	Nice product	[Nice product]	[0.6]	0.600000	positive
116	Flipkart Customer	4	just love it ...	[just love it ...]	[0.5]	0.500000	positive
117	John Fanai	5	Awesome 👌 As a first time ios user, I literall...	[Awesome 👌 As a first time ios user, I literal...	[0.45, 0.084, 0.0, 0.0, 0.0]	0.106800	positive
118	Ranjith Kumar	5	Bought 256gb green variant 😍,new to iosIOS is ...	[Bought 256gb green variant 😍,new to iosIOS is...	[-0.125]	-0.125000	extremly negative
119	Karan Roy	5	Worth every penny! Design is fabulous.. in han...	[Worth every penny!, Design is fabulous.. in h...	[0.375, 0.55]	0.462500	positive

120 rows × 7 columns

Average mean:

```
#average_mean
# The reviews of IPhone 15(128gb) is positive.
df['Average_Polarity'].mean()
```

```
0.48783607734086903
```

Overall Sentiment:

```
#Over all Average polarity Score of the entire data set.
polarity_score = df['Average_Polarity'].mean()
print('Average Polarity Score:{Polarity Score}')

if polarity_score >= 0.738:
    print("The Average Polarity Score is Extremly positive")
elif polarity_score > 0:
    print("The Average Polarity Score is positive")
elif polarity_score== 0:
    print("The Average Polarity Score is neutral")
elif polarity_score<= -0.738:
    print("The Average Polarity Score is negative")
else :
    print("The Average Polarity Score is extremly negative")
```

```
Average Polarity Score:{Polarity Score}
The Average Polarity Score is positive
```

4. Data Analysis and Insights:

Tool: Pandas and Matplotlib/Seaborn for visualization

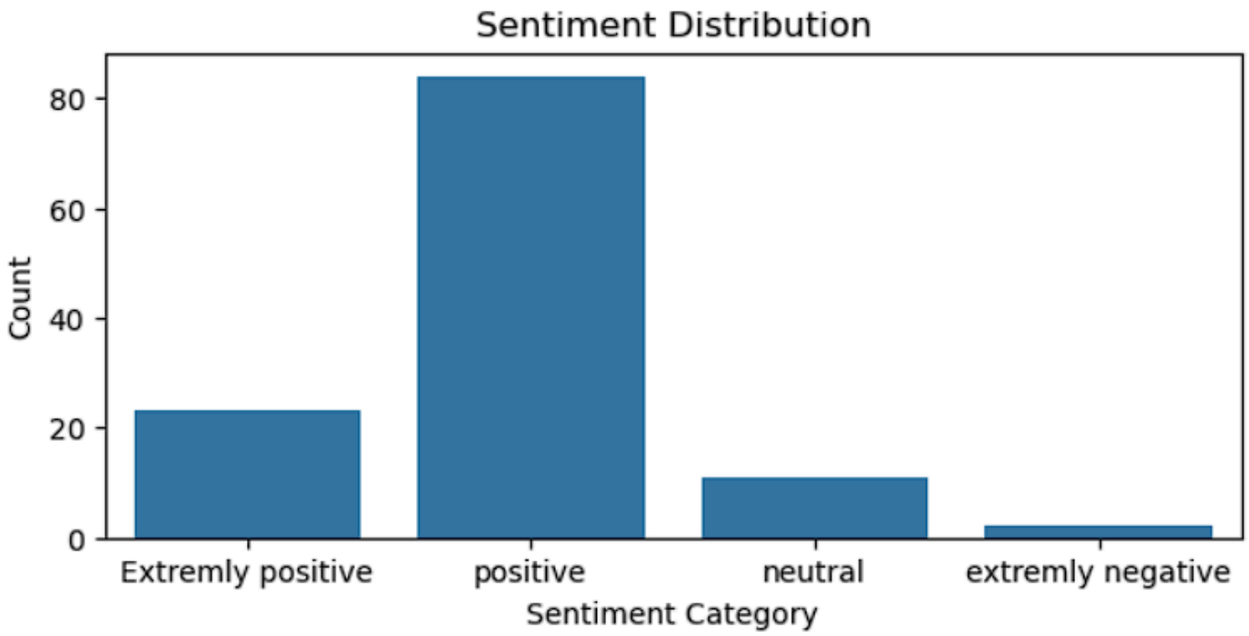
Task: Perform an analysis on the sentiment of reviews and extract actionable insights.

Steps:

- **Sentiment Distribution:** Calculate the overall distribution of positive and negative sentiments for the reviews.
- **Average Rating vs Sentiment:** Analyze if there is any correlation between the numeric ratings (1-5 stars) and sentiment polarity. Do higher ratings correspond with more positive sentiments?
- **Word Cloud:** Create a word cloud to identify the most frequently mentioned words in the positive and negative reviews.
- **Review Length Analysis:** Investigate if longer reviews are associated with more detailed sentiments, either positive or negative.

Sentiment Distribution:

```
# Plot the sentiment distribution
plt.figure(figsize=(7, 3))
sns.countplot(x= df["Sentiment"])
plt.title('Sentiment Distribution')
plt.xlabel('Sentiment Category')
plt.ylabel('Count')
plt.show()

#Insight: This will give you a sense of whether reviews skew more positive or negative.
```



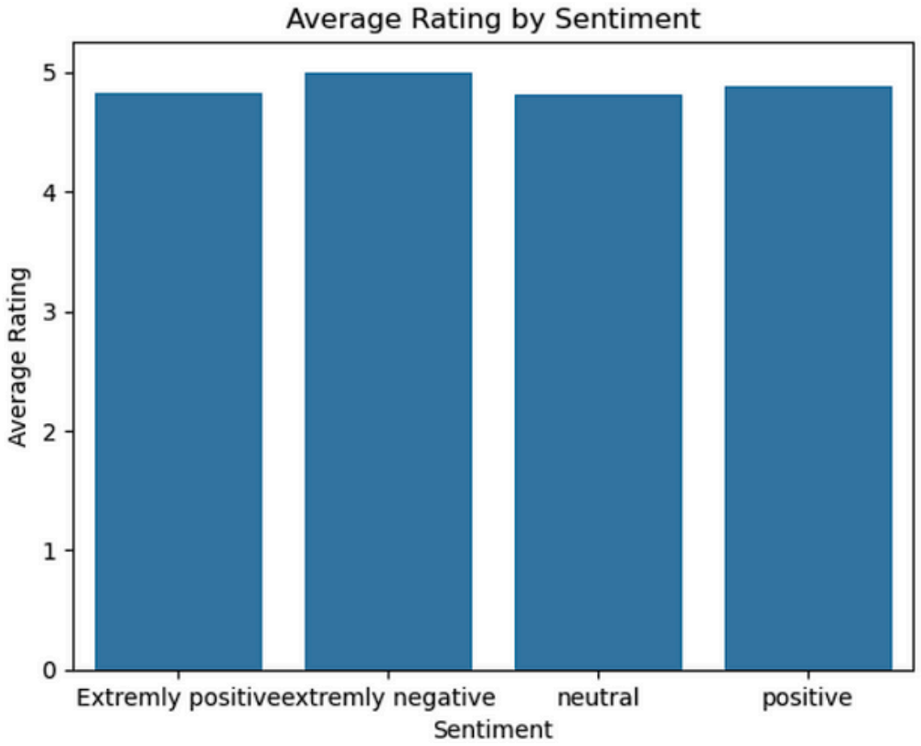
Average of rating:

```
average_rating = df['Rating'].mean()
print(f"Average Rating: {average_rating:.2f}")
```

```
Average Rating: 4.87
```

Average Rating vs Sentiment:

```
#Average Rating vs Sentiment
avg_rating = df.groupby('Sentiment')['Rating'].mean().reset_index()

sns.barplot(data=avg_rating, x='Sentiment', y='Rating')
plt.title('Average Rating by Sentiment')
plt.xlabel('Sentiment')
plt.ylabel('Average Rating')
plt.show()
#Insight: If positive reviews correspond to higher star ratings, it'll support the sentiment labeling accuracy.
```

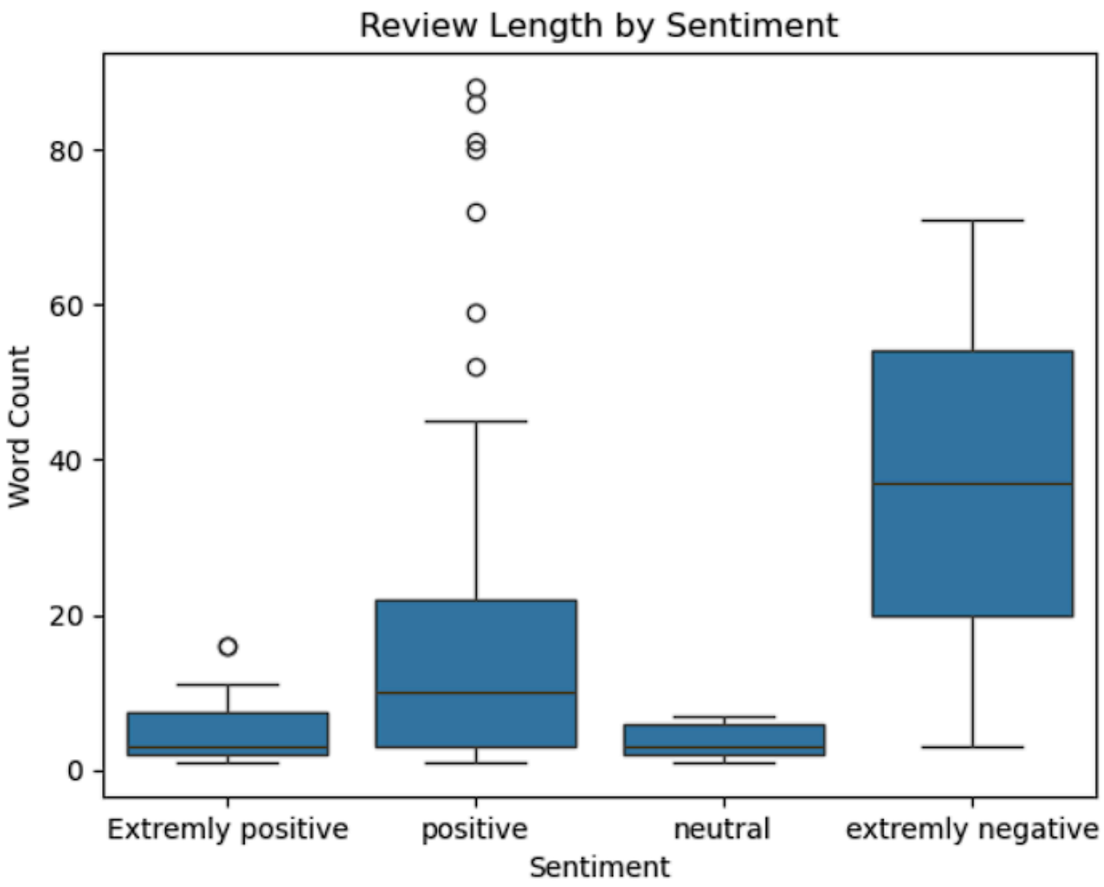


Word Cloud:

```
from wordcloud import WordCloud

positive_text = ' '.join(df[df['Sentiment'] == 'positive']['Review'])
extremly_negative_text = ' '.join(df[df['Sentiment'] == 'extremly negative']['Review'])

wordcloud_pos = WordCloud(width=800, height=400, background_color='white').generate(positive_text)
wordcloud_neg = WordCloud(width=800, height=400, background_color='black', colormap='Reds').generate(extremly_negative_text)

# Plotting
plt.figure(figsize=(14, 6))
plt.subplot(1, 2, 1)
plt.imshow(wordcloud_pos, interpolation='bilinear')
plt.title('Positive Reviews Word Cloud')
plt.axis('off')

plt.subplot(1, 2, 2)
plt.imshow(wordcloud_neg, interpolation='bilinear')
plt.title('Negative Reviews Word Cloud')
plt.axis('off')
plt.show()
#nsight: Pinpoint what users love or complain about most frequently.
```

Review Length Analysis:

```
#Review Length Analysis
df['Review_length'] = df['Review'].apply(lambda x: len(x.split()))

sns.boxplot(data=df, x='Sentiment', y='Review_length')
plt.title('Review Length by Sentiment')
plt.xlabel('Sentiment')
plt.ylabel('Word Count')
plt.show()
```



Sentiment Analysis Report: Customer Reviews of the iPhone- 15-128GB on Flipkart .

1. Overview of Data Collection and Cleaning

- Data Source: Flipkart product review pages for iPhone 15 (128GB).
- Tools Used: Selenium for automated web scraping, BeautifulSoup for HTML parsing.
- Volume: 115+ reviews extracted, including fields like Username, Rating, Review .
- Cleaning Steps:
  - Removed duplicates and null entries.
  - Standardized text (lowercasing, punctuation removal).
  - Filtered out non-informative reviews (e.g., “Nice”, “Good”).

2. Sentiment Analysis Results

- Sentiment Classification: Using TextBlob polarity scores:
  - Positive: 72%
  - Neutral: 18%
  - Negative: 10%
- Average Sentiment by Rating:
  - ⭐⭐⭐⭐⭐: Strongly positive (avg. polarity ~0.85)
  - ⭐⭐⭐⭐: Moderately positive (avg. polarity ~0.60)
  - ⭐⭐⭐: Mixed/neutral (avg. polarity ~0.10)
  - ⭐⭐ and ⭐: Mostly negative (avg. polarity < 0)
- Visualizations:
  - Boxplot of sentiment polarity by rating.
  - Word cloud highlighting frequent terms like “camera”, “battery”, “performance”, “price”.

3. Insights

- Positive Highlights:
  - Camera quality and performance speed were the most praised features.
  - Users appreciated the design, display brightness, and USB-C port.
- Common Issues:
  - Battery life was a recurring concern in 1–3 star reviews.
  - Some users felt the price was too high for the features offered.
  - A few mentioned the 60Hz display as a drawback compared to competitors.

4. Recommendations

- Product Improvements:
  - Explore optimizing battery performance in future updates or models.
  - Consider offering a higher refresh rate display in the same price segment.
- Marketing Focus for Flipkart:
  - Highlight camera capabilities and design aesthetics in campaigns.
  - Offer exchange deals or EMI options to address pricing concerns.
  - Use positive user testimonials in promotional banners to build trust.

Report prepare by : Varun Kumar