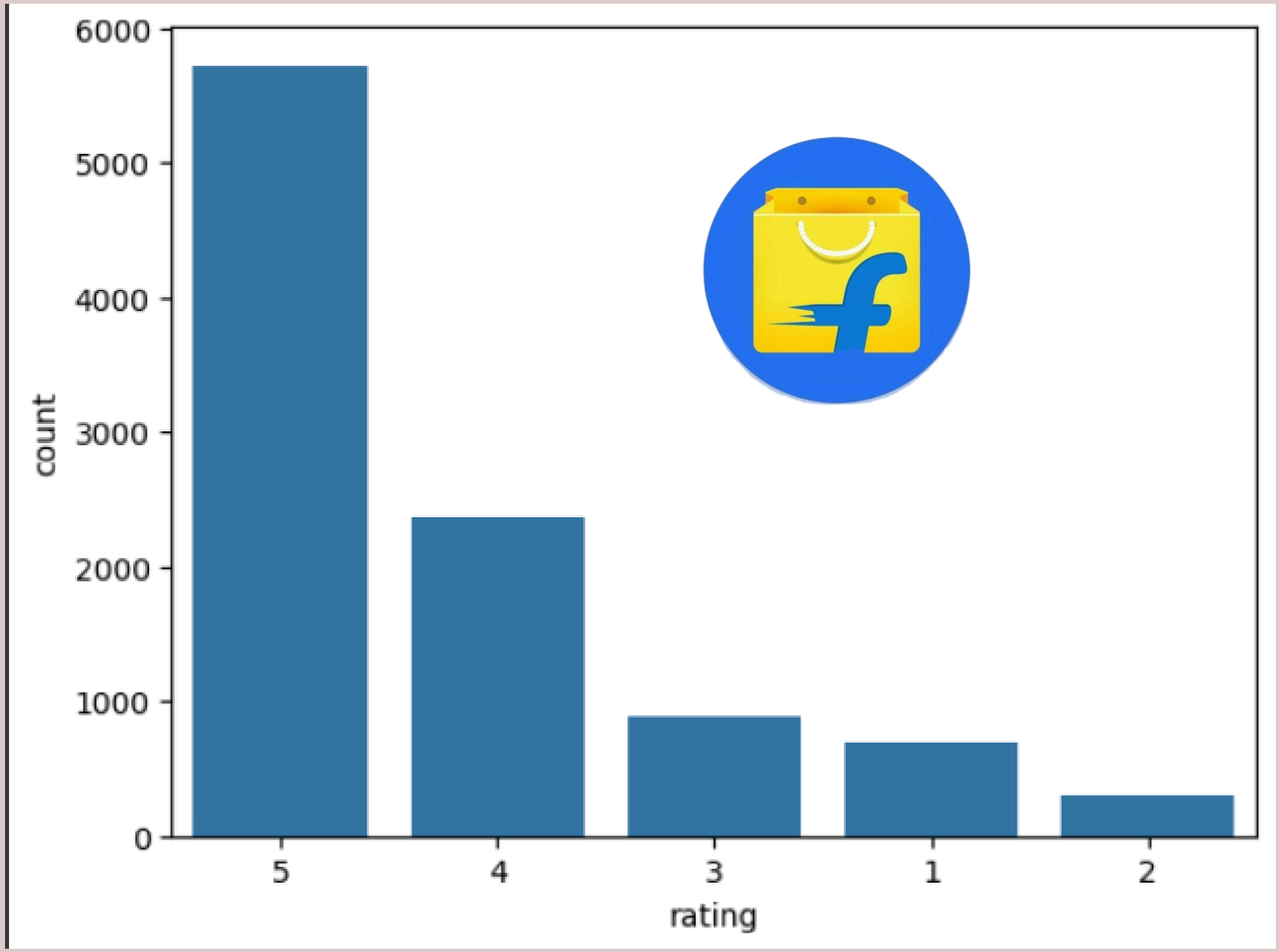
Flipkart

ENHANCING E-COMMERCE INSIGHTS: A COMPREHENSIVE SENTIMENT ANALYSIS OF FLIPKART REVIEWS

By Pranav Mishra
(A Data Analyst Fellow)

LIBRARIES USED FOR FLIPKART SENTIMENT ANALYSIS OF REVIEWS

- Pandas
- Scikit-Learn
- Matplotlib
- Seaborn
- Natural-Language-Toolkit (NLKT)
- TfidfVectorizer
- WordCloud
- Warnings



INTRODUCTION TO SENTIMENT ANALYSIS

Sentiment Analysis is a powerful tool that helps businesses understand customer emotions and opinions. In this presentation, we will explore the **insights** gained from analyzing **Flipkart reviews**, focusing on customer satisfaction and areas for improvement.

INTRO TO DATASET OF SENTIMENT ANALYSIS

```
data = pd.read_csv('/content/flipkart_data.csv')
data.head()
```

	review	rating
0	It was nice produt. I like it's design a lot. ...	5
1	awesome sound....very pretty to see this nd th...	5
2	awesome sound quality. pros 7-8 hrs of battery...	4
3	I think it is such a good product not only as ...	5
4	awesome bass sound quality very good bettary l...	5

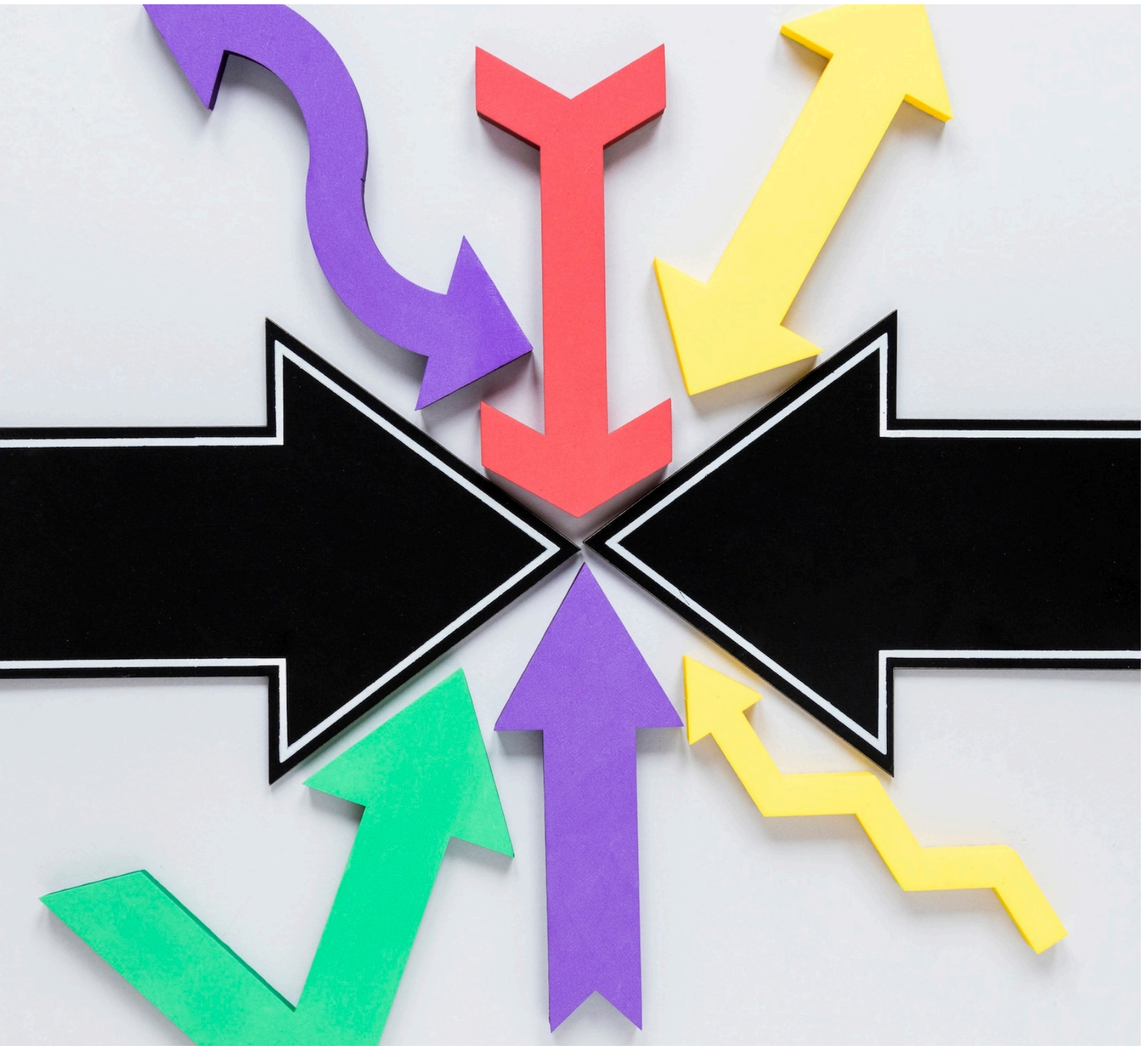
```
pd.unique(data['rating'])
```

```
array([5, 4, 1, 3, 2])
```

IMPORTANCE OF E-COMMERCE INSIGHTS

Flipkart

In today's competitive landscape, **e-commerce insights** are essential for driving sales and enhancing customer experience. By analyzing reviews, companies can better understand **customer preferences** and make data-driven decisions for product development and marketing strategies.



METHODOLOGY OF REVIEW ANALYSIS

We employed a systematic approach to analyze **Flipkart reviews** using natural language processing techniques. The process includes data collection, **preprocessing**, sentiment classification, and extracting key themes to provide a comprehensive overview of customer feedback.

PREPROCESSING

```
# For Label
pos_neg = []
for i in range(len(data['rating'])):
  if data['rating'][i] >= 5:
    pos_neg.append(1)
  else:
    pos_neg.append(0)

data['label'] = pos_neg
```

```
nltk.download('punkt')
from tqdm import tqdm

def preprocess_text(text):
  preprocessed_text = []
  for sentence in tqdm(text):
    # It will remove punctuations mark
    sent = re.sub(r"[^\w\s]", "", sentence)

    # Converting lowecase and removing stopwords
    preprocessed_text.append(" ".join(token.lower() for token in nltk.word_tokenize(sentence)
     if token.lower() not in set(stopwords.words('english'))))
  return preprocessed_text
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
```

'punkt' is a tokenizer divides a text into a list of sentences by using an unsupervised algorithm to build a model for abbreviation words, collocations, and words that start sentences.

```
preprocessed_review = preprocess_text(data['review'].values)
data['review'] = preprocessed_review
```

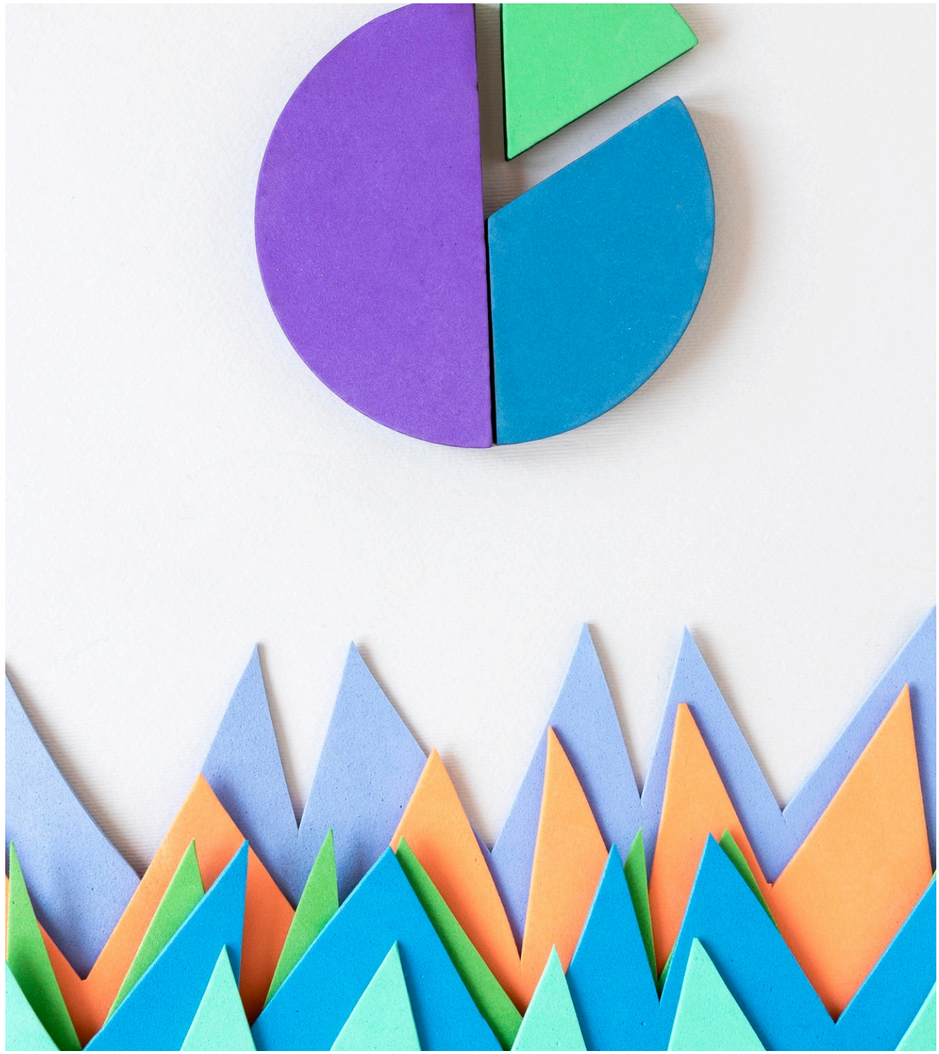
```
100%|██████████| 9976/9976 [00:15<00:00, 648.57it/s]
```

```
data.head()
```

	review	rating	label
0	nice produt like design lot easy carry looked ...	5	1
1	awesome soundvery pretty see nd sound quality ...	5	1
2	awesome sound quality pros 78 hrs battery life...	4	0
3	think good product per quality also design qui...	5	1
4	awesome bass sound quality good bettary long l...	5	1

KEY FINDINGS FROM REVIEWS

Our analysis revealed several crucial insights. The majority of customers expressed **satisfaction** with product quality, but common complaints included **delivery delays** and customer service issues. Understanding these aspects can guide Flipkart in enhancing their overall service.



By this code I have seen that words like awesome, product, good, quality, etc. have high frequency in positive label

```
consolidated = ' '.join(
    word for word in data['review'][data['label']==1].astype(str)
)
wordCloud = WordCloud(width=1500,height=800,random_state=21,max_font_size=110)
wordCloud.generate(consolidated)
plt.figure(figsize=(12,10))
plt.imshow(wordCloud,interpolation='bilinear')
plt.axis('off')
plt.show()
```

Output is on next slide


good battery
boatread
totally
clear sound
awesome product
good bass
bestread
sound
awesomeread
headphoneread
big
product price
one
must buy
excellent product
great
nice
bought
love
bass good
bluetooth
feel
amazingread
productsread
great product
sound qualityread
u
got
read good
best headphone
good productread
clear
long time
month
deal
used
battery backup
time
price range
hour
flipkartread
call
must buyread
better
thing
product read
look
bass
product boat
head
awesome bass
boat
low
go itread
value moneyread
go
happy
music
superb
want
cant
really good
connectivity
headphone
everything
goodread nice
priceread
good sound
dont
base
comfortable
cool
aux cable
overall
use
awesome productread
nice product
super
buy
year
quality good
awesome read
value money
build quality
ear
quality
thanks flipkart
super bass
build
little
read awesome
superbread
quality awesome
superread
work
battery
productread awesome
still
pretty
wow
feature
first
using
price
headset
good product
bass sound
flipkart
design
excellentread
read best
song
read nice
best product
productread good
perfect
delivery
experience
give
audio
range
amazing
worth
loved
goodread
excellent
day
purchase
loved itread
soundread
sound bass
bassread
mic
hear
awesome
well
niceread
boat rockerz
wireless
budget
looking
nd
wire
productread best
really
product
much
nice productread
amazing sound
fantastic
awesome sound
love itread
amazing product
problem
productread nice
sale

```python
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score

model = DecisionTreeClassifier(random_state=0)
model.fit(X_train,y_train)

#Model Test
pred = model.predict(X_train)
print(accuracy_score(y_train,pred))
```

```
0.9209273182957394
```

RECOMMENDATIONS FOR FLIPKART

Based on our findings, we recommend that Flipkart focus on improving **logistics** and customer support. Implementing a more efficient delivery system and training staff can significantly enhance the **customer experience** and boost overall satisfaction.

CONCLUSION AND FUTURE WORK

In conclusion, sentiment analysis of Flipkart reviews provides valuable insights into customer opinions. Future work should involve continuous monitoring of reviews and adapting strategies based on evolving customer needs to maintain a competitive edge in e-commerce.

Thanks!

By Pranav Mishra
(A Data Analyst Fellow)