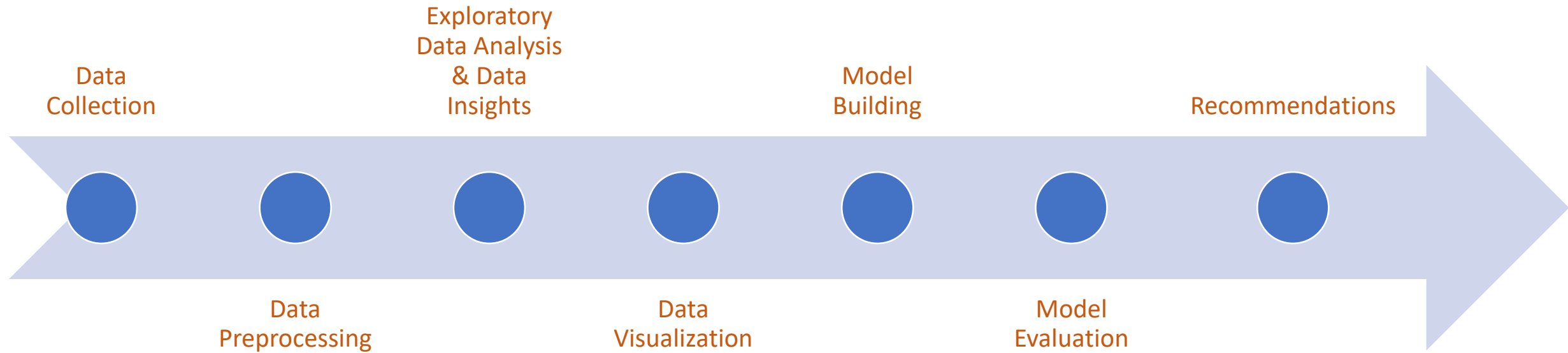# Twitter Sentiment Analysis-
Biggest Indian Online Retailers
(2019 to Present)

Project Overview :

Sentiment Analysis helps us to identifies and extracts subjective information from the user experience. In this project we are Collecting India's biggest Online retailers tweets from twitter and analysing the positive and negative tweets. This model help to find out the positive and negative tweets and business developments. This model has the roc-auc-score of 97%.

Project Flow

Data
Collection

Data
Preprocessing

Exploratory
Data Analysis
& Data
Insights

Data
Visualization

Model
Building

Model
Evaluation

Recommendations

Approaches for twitter sentiment analysis:

- Data Scrapping :
    1. From twitter we scrapped the tweets of amazon, flipkart, lenskart, ajio, bigbazaar, Zomato, snapdeal from 2019 to at present.
    2. After extracting all the tweets from the twitter we convert it into DataFrame.
    3. After converting the DataFrame we removed the Duplicates from the Data.
    4. we removed the duplicates and convert it into Comm separated value(csv)file.

- Data Preprocessing:
    1. We are loading the Data into new ipnby file and taking the tweets feature from the Data for the cleaning and model building purpose.
    2. This tweets are text in nature. We create some function to clean the text. And lemmatizing the words.
    3. We are using textblob library to finding the subjectivity and polarity. After finding the polarity we find the positive and negative tweets using the previous polarity

EDA & modelling & evaluation

- ## Exploratory Data Analysis:

  1. We are performing some exploratory analysis on whole tweets and after that we are visualizing the tweets of each and every companies.

  2. We are visuazling the top most frequent positive and negative words on whole tweets and after that visualizing the tweets of each and every companies.

- ## Model Building:

  - We are converting all words into vectors using TF-IDF.

  1. We are splitting our Dataset into Train and test Data with 70 30 ratio. After splitting we are importing the Random Forest Classifier and fit the x and y train data and after fitting predicting the results.

  2. we evaluate our model using sklearn metrics like accuracy score auc-roc-score confusion matrix.

  3. this model has the roc-auc-score of 97%. Precision recall f1 score also more than 92%.

  4. This model will able to predict if the new user tweets it will able to predict whether tweet is positive or negative.

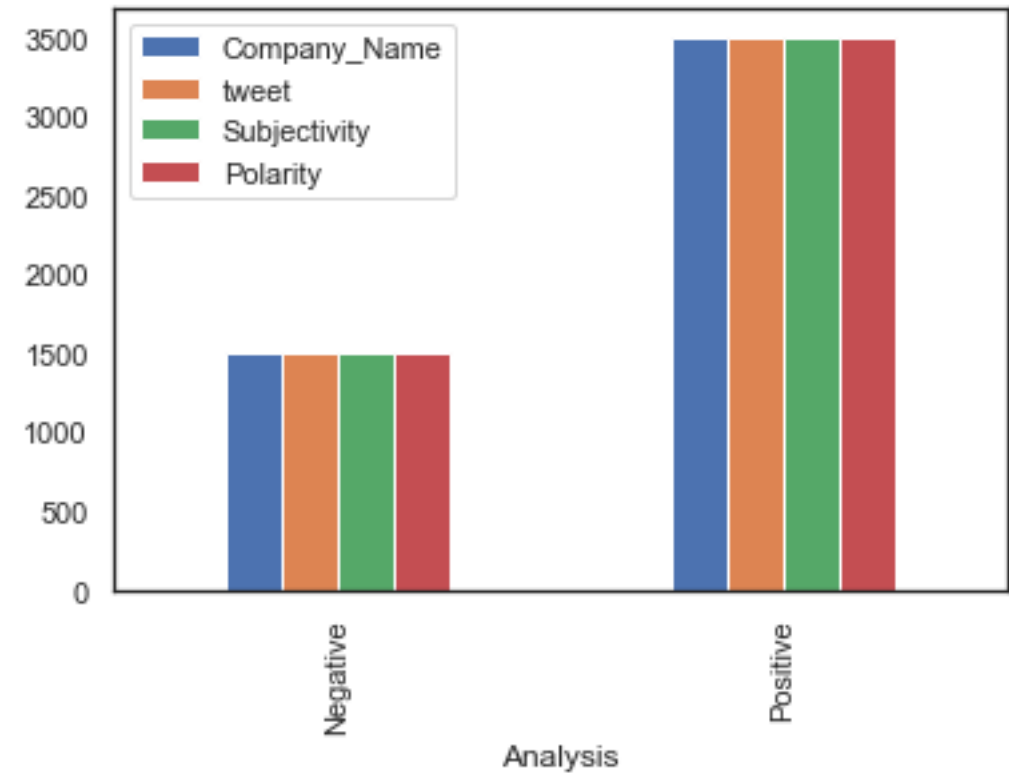# India's Biggest Online retailers Positive and negative tweets proportion
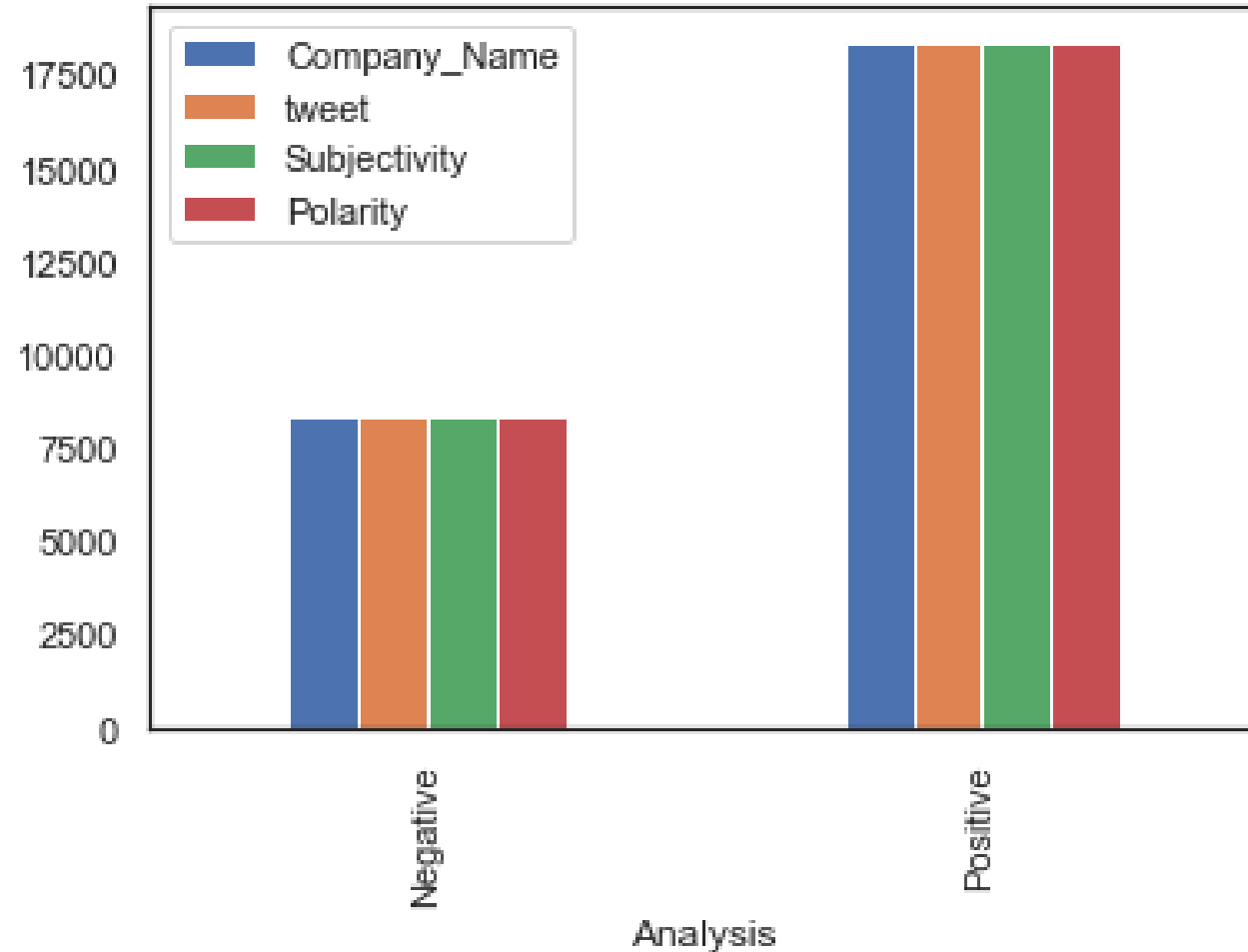
Positive and negative proportion

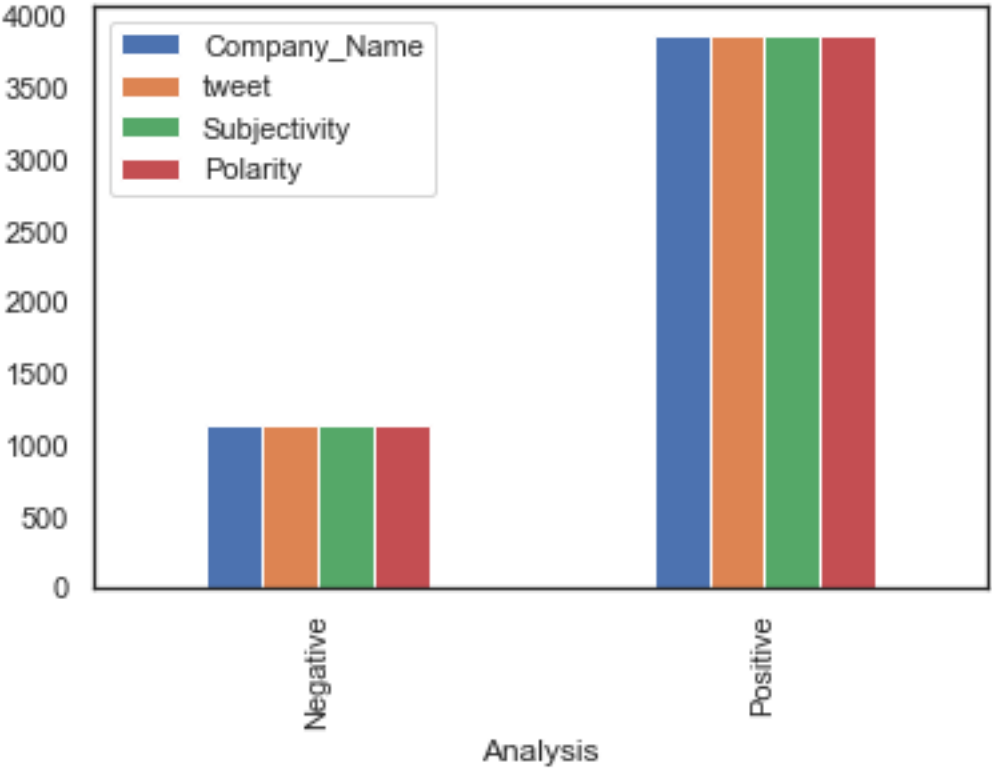# AmazonIndia and Flipkart positive negative tweet proportion
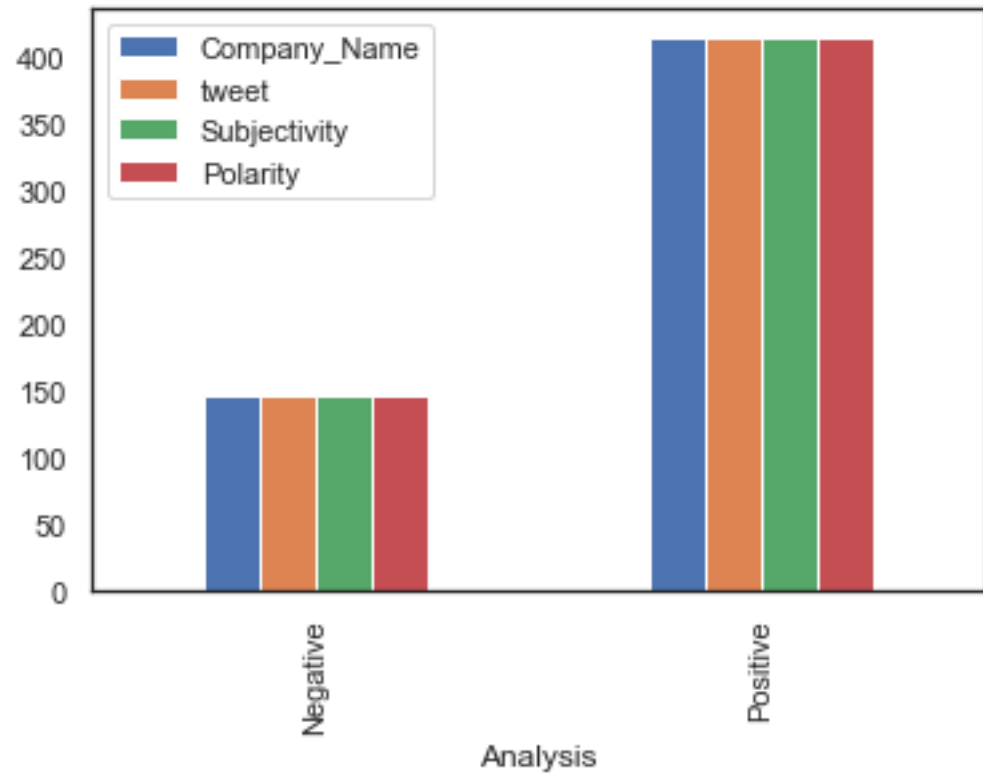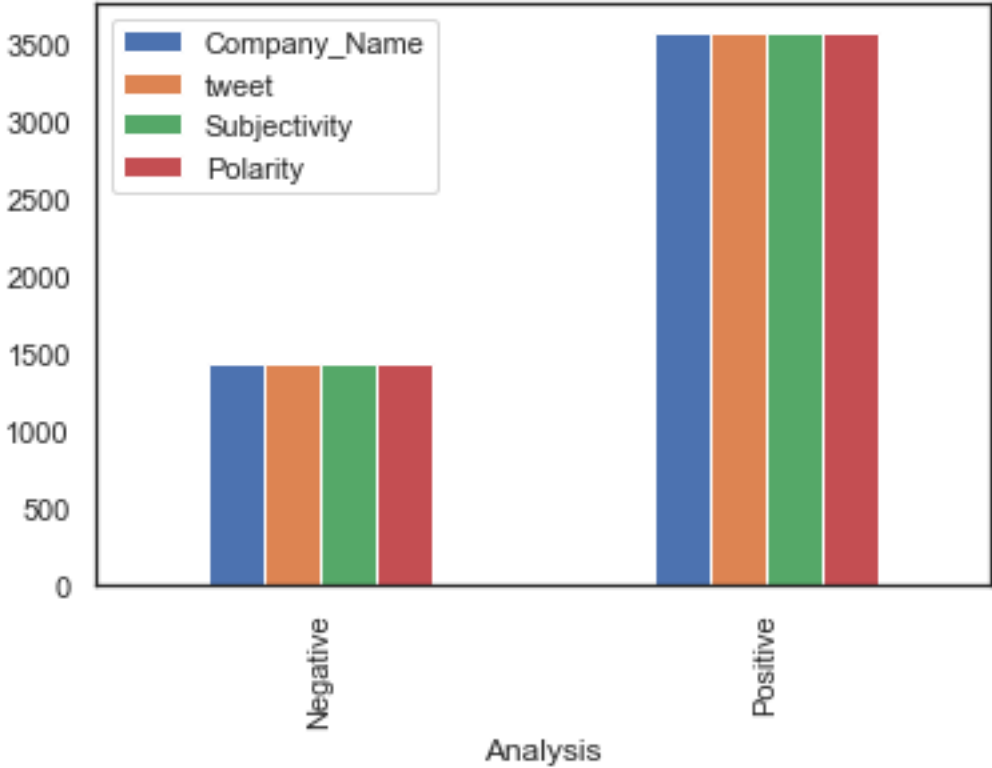

AmazonIndia

Flipkart

SnapDeal and Myntra
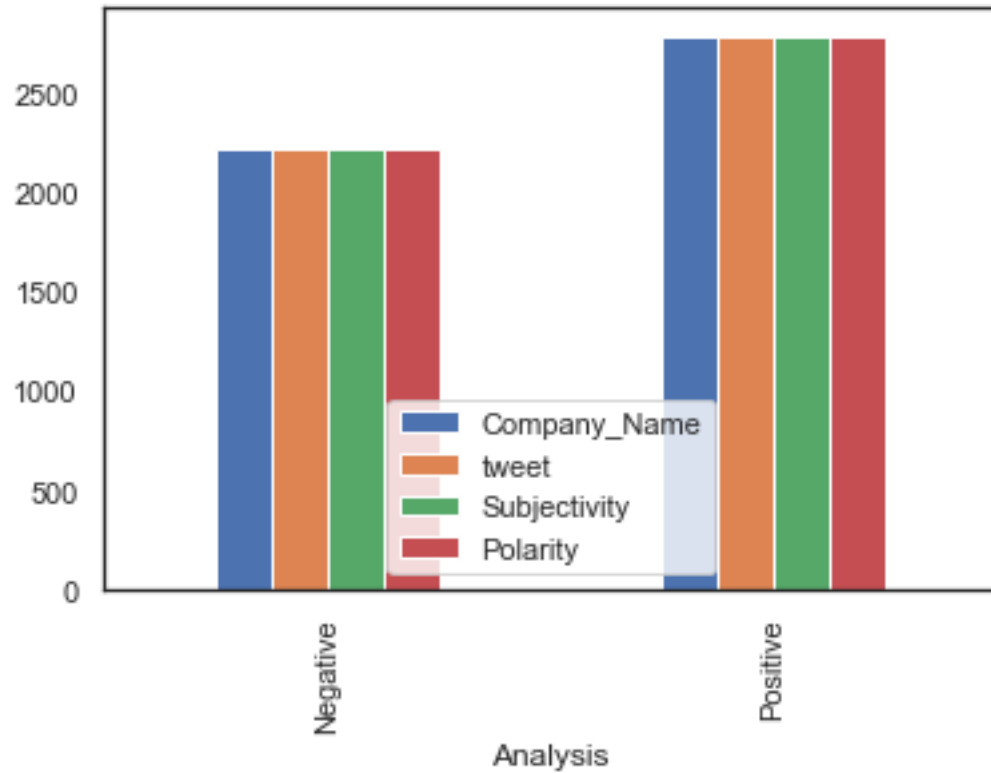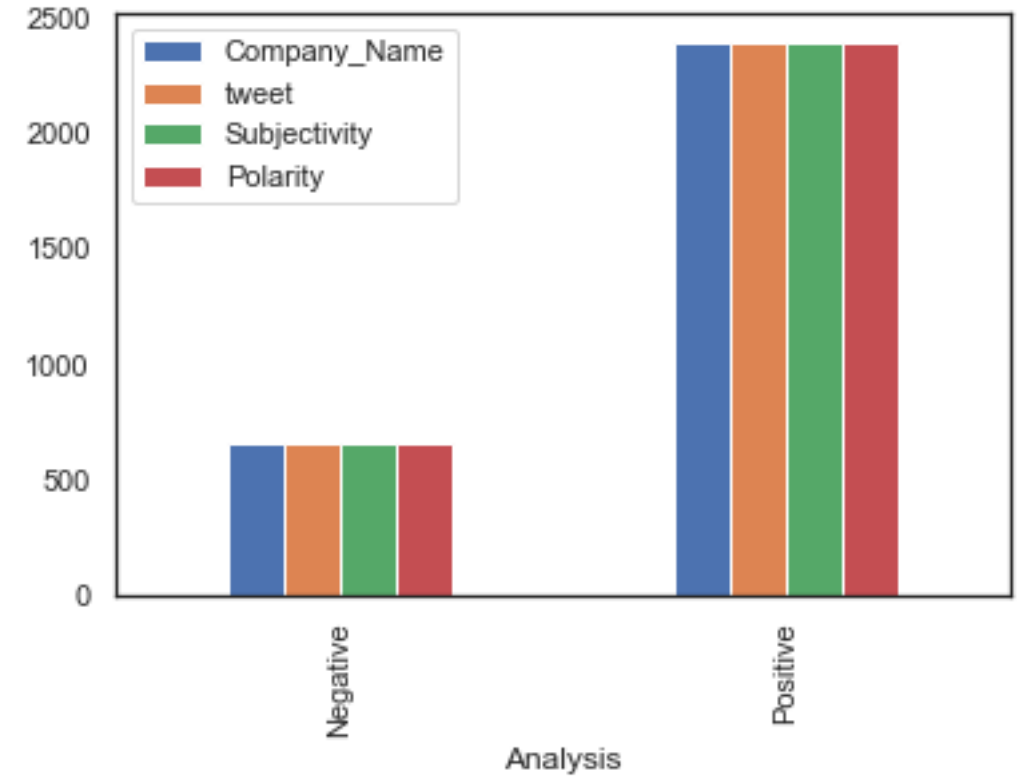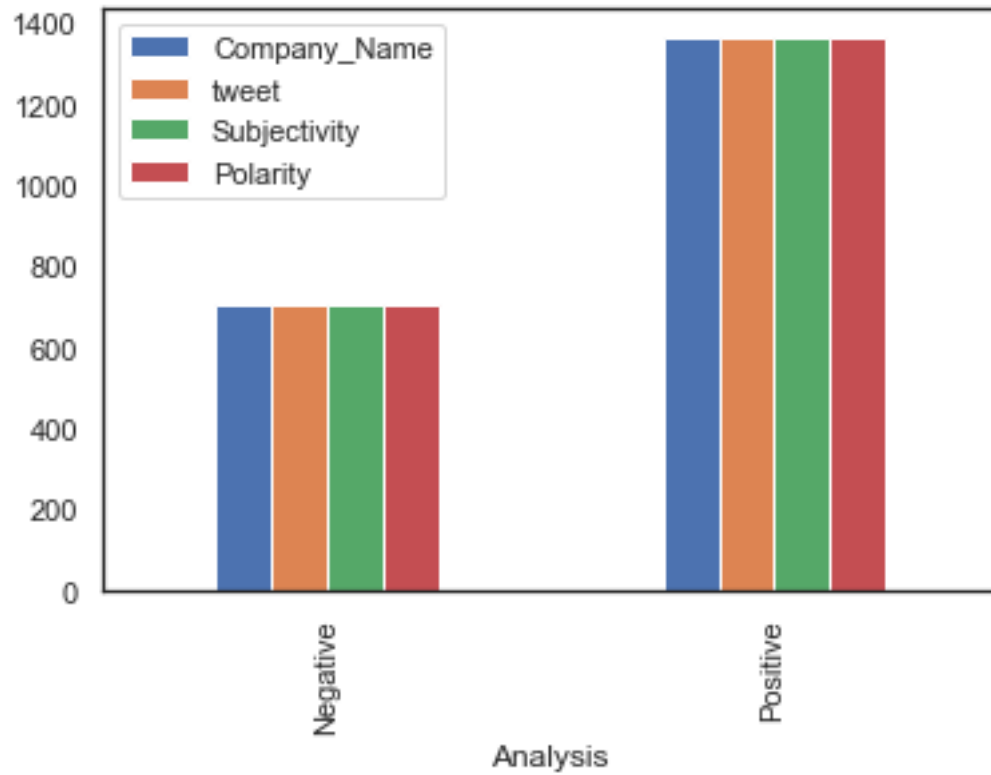
Zomato and BigBazaar
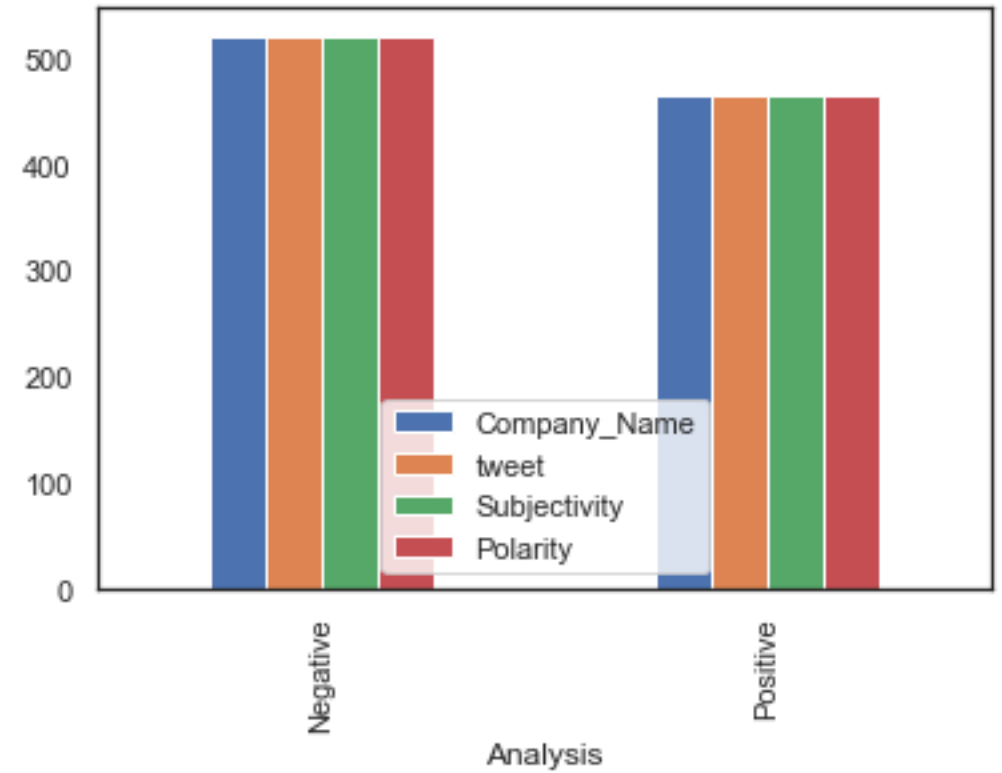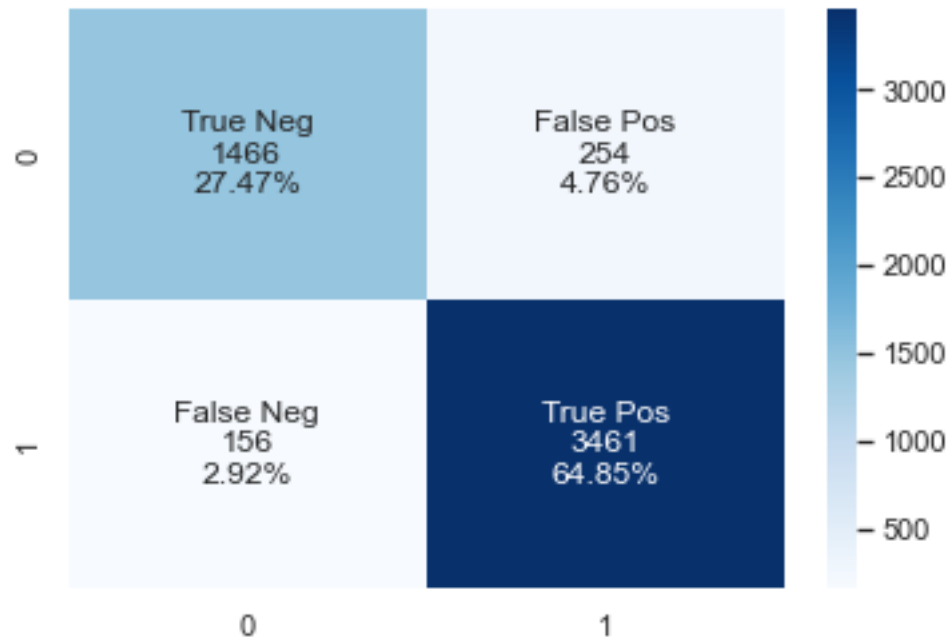

Zomato

BigBazaar

Lenskart and Ajio

Lenskart

Ajio

Confusion matrix and Classification report

Confusion matrix

	precision	recall	f1-score	support
0	0.90	0.85	0.88	1720
1	0.93	0.96	0.94	3617
accuracy			0.92	5337
macro avg	0.92	0.90	0.91	5337
weighted avg	0.92	0.92	0.92	5337

Classification Report

Roc Curve
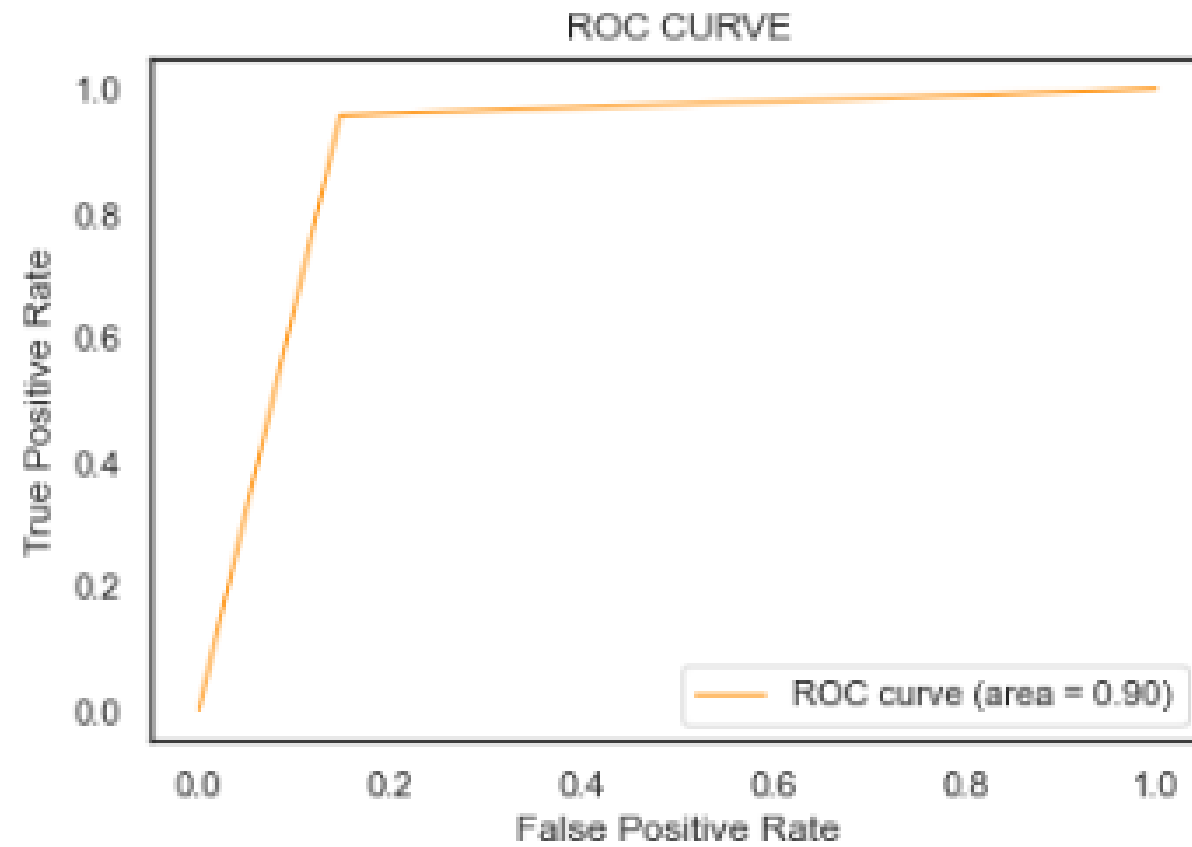
Conclusion - Recommendations

❖ From this Diagrams we come to know that positive tweets is higher than negative tweets. This tweets are from india's topmost online retailers.

❖ while analyzing all company tweets individually, Negative tweets of Ajio and Lenskart is more than 50%.

❖ from the analysis of most frequent negative words we can know that delivery and quality of product is the main reason for the negative tweets.

Recommendations :

1. While packing the product make sure to verify whether the product ordered and packed product is same and accurate.

2. Make sure that all the products are delivered on time.

3. If the user applied for the return or exchange take immediate steps to resolve it.

4. when the customer doing online payments , if there is any troble in transaction try to solve it immediately. From this steps users will surely trust the company