

PHONE REVIEW SENTIMENT ANALYTICS

BY

VISHAL GADDAM

OBJECTIVE

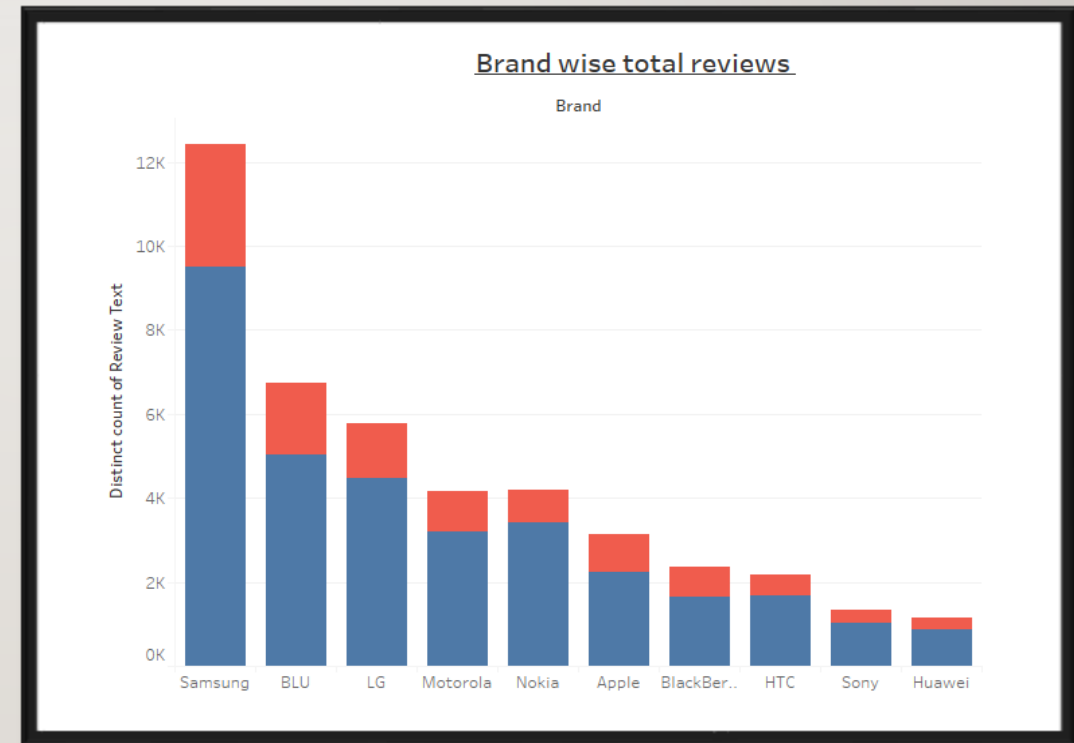
- Understanding the market competitors and customer preferences.
- Analysing the most sought out features by the customers.
- Finding out the user experiences with current devices.
- Pricing strategies.
- Recommendations to increase the market share and brand value.

BACKGROUND

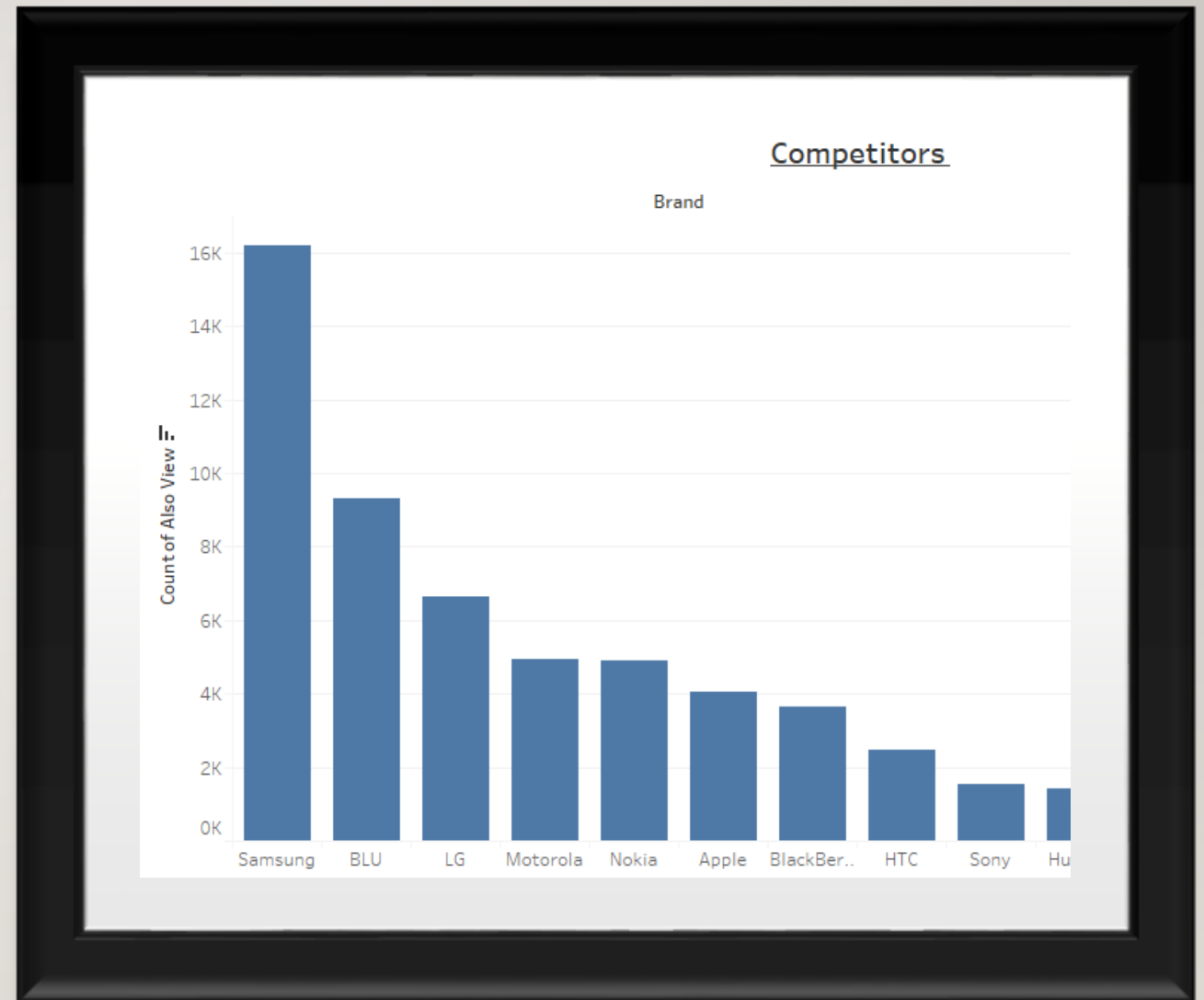
- The user reviews from the amazon website are given as the company is new entrant into the market.
- Competitor market share and the reviews received for their products.

INSIGHTS

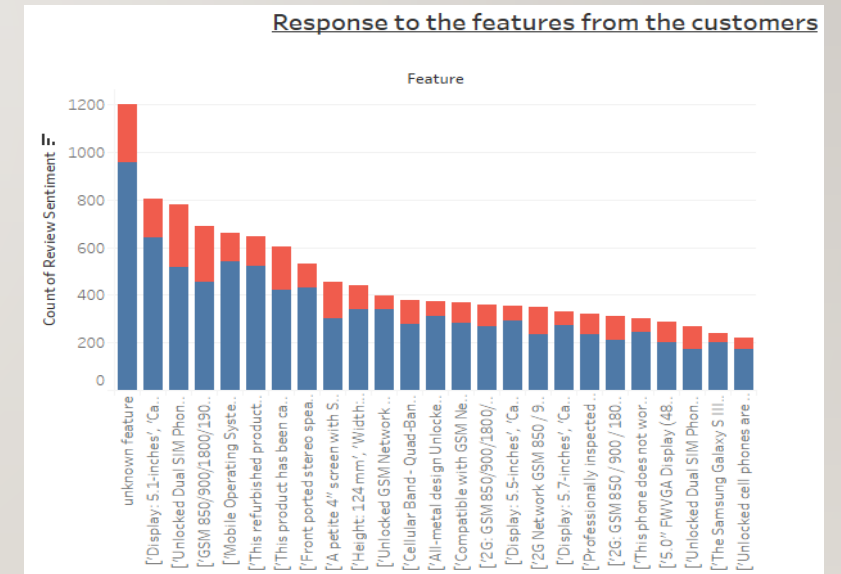
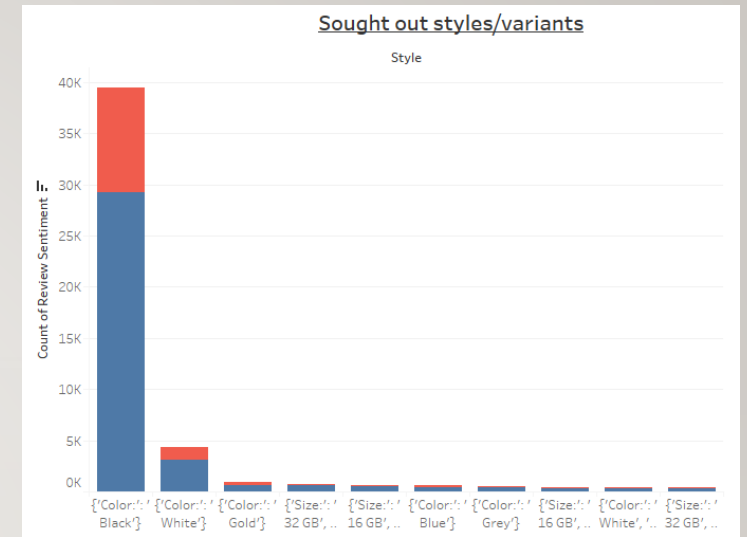
- Samsung is the top most brand with highest number of reviews followed by BLU, LG.



-
- Samsung , Blu , LG are the top most competitors. These brands are most viewed when a customer decides to buy a mobile.



-
- Black, white and gold are the most sought out styles followed by the storage sizes.
 - Screen size of 5.1 inch, unlocked phones garnered high positive response from the customers.



RECOMMENDATIONS

- Keeping an eye on the major competitors
- Targeting the price ranges of 40\$, 80\$, 120\$.
- Giving multiple options in colors and specifications.
- Offering more unlocked phones.

APPENDIX – DATA SOURCES

- Here is a snapshot of our data dictionary
 - Phone data such as asin, review text, review_sentiment, reviewer Id.
 - Phone metadata such as category, brand , feature. Rank, also_view etc
- Positive words, negative words, stop words text files for text analytics.
- Data Source used is Amazon users phone orders dataset.

APPENDIX – DATA METHODOLOGY

- We conducted a thorough analysis of the phone orders dataset and predicted the review sentiments. The process included:
 - Data importing and cleaning the datasets for null values and irregularities using pandas.
 - Merging the datasets and creating train and test datasets for the model.
 - Text analytics using Naïve Bayes for predicting the review sentiments.
 - Used tableau for the visualizations and deriving insights

APPENDIX – DATA ASSUMPTIONS

- Removed some fields present in the data which are statistically insignificant.
- There were some null values in certain columns.