Unveiling the Unlocked Phone Market: Insights from Amazon Reviews and Sentiments

Alissa Dao, Long Bui
ECON 314: Applied Empirical Data Analysis

Introduction

Background: The market for unlocked phones has grown and changed significantly in recent years, partly due to the growing appeal of internet retailers like Amazon.com. Users using unlocked phones can select their preferred network provider since they are not restricted to any particular carrier. This adaptability has increased consumer demand for unlocked phones.

About the dataset:

- Our dataset comprises 400 thousand reviews of unlocked mobile phones from Amazon.com with different types and brands in 2016.
- Originally, there were six variables, which were product name, brand name, price, rating, review, and review votes.
- For the research, we created two other columns: the review length, price range, and sentiment score of the reviews.
  - The price range: low: [11.11; 71.05), Low-Mid: [71.05;116.99), Mid: [116.99,325.16), High-Mid: [325.16;622.9), High: [622.9;2598]

Objectives:

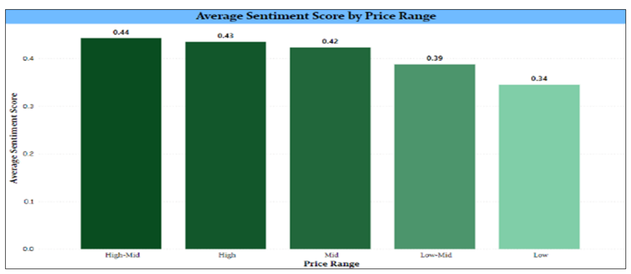
- Understanding people's opinions and sentiments about unlocked phones.
- Understanding the relationship between price, review length, and review votes.
- Understanding the factors that influence the sentiment score of the products.

Sentiment Analysis

Method:

- TextBlob is a Python package that is used to identify a text's sentiment including positive, negative, or neutral emotions.
- TextBlob analyzes every word in a passage to determine its sentiment and meaning ("love": positive, "hate": negative).
- Score nearer +1 indicates a highly positive text while -1 indicates negative sentiment. If it's close to 0, it's neutral.
- After getting the score, we conducted a t-test to see if the average sentiment score is significantly different between each price range.

Results:



	Comparison	P-value
High vs High-mid	High-mid > High	0.101
High vs Mid	High > Mid	0.025*
Low-mid vs Low	Low-mid > Low	< 0.001*
Mid vs Low	Mid > Low	< 0.001*
High-mid vs Low	High-mid > Low	< 0.001*
High vs Low	High > Low	< 0.001*
Mid vs Low-mid	Mid > Low-mid	< 0.001*
High-mid vs Low-mid	High-mid > Low-mid	< 0.001*
High vs Low-mid	High > Low - mid	< 0.001*
High-mid vs Mid	High-mid > Mid	< 0.001*

* statistically significant at 0.05

Interpretations:

- Every pair of pricing ranges is statistically significant, **except for "High vs. High-mid**. For "High vs. High-mid," the p-value is higher than the cutoff point of 0.05, which means that there is no statistically significant difference in the average sentiment scores between the "High" and "High-mid" pricing range.

Key Word Analysis

Method: Implemented log transformation to normalize the data, applied Apriori algorithm to extract key phrases to measure their frequencies of appearance on both negative and positive comments.

Results:

Negative Keywords					
Attributes	Low	Low-Mid	Mid	High-Mid	High
Phone stopped working	Most Freq	Middle Freq	Most Freq	Most Freq	Most Freq
Waste money	Top Freq	Most Freq	Top Freq	Most Freq	Most Freq
Battery Sucks	Top Freq	Top Freq	Top Freq	Top Freq	Top Freq
Phone Screen	Top Freq	Middle Freq	Top Freq	Low Freq	No Mention
Speed	Low Freq	Low Freq	Low Freq	No Mention	No Mention
Customer Service	No Mention	Low Freq	Middle Freq	Middle Freq	Top Freq
International Warranty	Low Freq	No Mention	No Mention	Low Freq	Low Freq
Phone Camera	Middle Freq	No Mention	No Mention	Low Freq	Low Freq
Bluetooth	Low Freq	Low Freq	Low Freq	No Mention	Low Freq
App Store	No Mention	Low Freq	No Mention	No Mention	Low Freq
Uniqueness	Text Message Sucks		SmartWatch		Scammers (20)

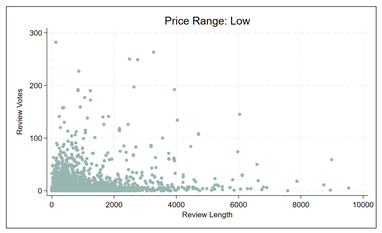
Positive Keywords					
Attributes	Low	Low-Mid	Mid	High-Mid	High
Good Condition	Most Freq	Most Freq	Most Freq	Most Freq	Most Freq
Good Battery	Middle Freq	Middle Freq	Top Freq	Middle Freq	Top Freq
Good Price	Top Freq	Top Freq	Middle Freq	Top Freq	Low Freq
Phone Screen	No Mention	Low Freq	No Mention	No Mention	No Mention
Work perfectly	Middle Freq	Middle Freq	Middle Freq	Top Freq	Most Freq
Uniquesness				Fast-phone	Fast shipping

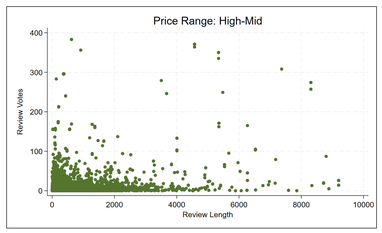
Interpretations: Battery life is a key concern for customers across all price ranges. Mid-range devices are also criticized for customer service and international warranty issues. Lower-priced phones often experience problems with text messaging, while mid-range phones may have screen issues. Scammers are a major concern for high-end products. Smartwatches and app store quality receive sporadic mentions in the mid-range. Price is scrutinized across the board, with high-end products receiving comments about being overpriced.

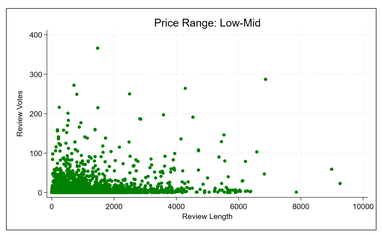
Relationship: Review Vote, Review Length & Price Ranges

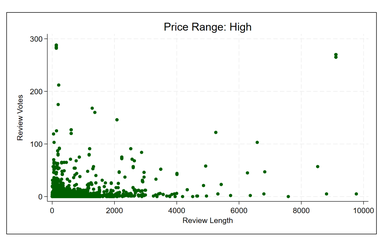
Methods: We transformed the price range variable into four dummy variables with "Low" as a reference variable and utilized the OLS regression model to understand the relationship between each price range, review length, and review votes.

Results:









Correlation: Review Votes & Review Length	
Review Votes	Review Length
General	0.32
Price = High	0.40
Price = High-Mid	0.33
Price = Low	0.32
Price = Mid	0.32
Price = Low-Mid	0.30

Linear Regression

Number of obs	385,057	
F(5,385051)	216.34	
p-value	< 0.001	
R-squared	0.10	
Root MSE	7.78	

review votes	Coefficient	std.err	p-value
priceRange			
Low-mid	-0.17	0.03	< 0.001*
Mid	0.07	0.03	0.016*
High-mid	0.66	0.05	< 0.001*
High	0.94	0.08	< 0.001*
reviewlength	0.01	< 0.001	< 0.001*
_cons	0.06	0.04	0.187

*statistically significant at 0.05

Interpretations:

- These differences in correlation values across different price ranges indicate that the relationship between review length and review upvotes may vary depending on the price range of the products.
- Higher-priced phones tend to receive more review votes, while longer reviews also contribute positively to review votes.

Sentiment Scores Multivariate Regression Analysis

Method: Optimized Python packages (sklearn, statsmodel) to conduct linear and polynominal regressions to measure how other attributes correlates with sentimentScore.

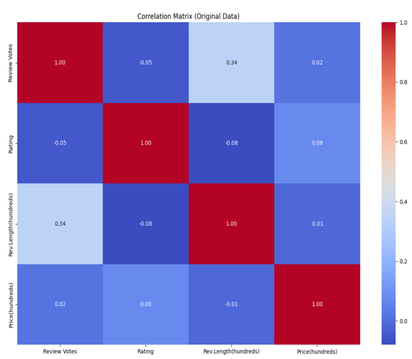
Results:

Multivariate Regression

Number of obs	385,057	
R-squared	0.52	

sentiment score	Coefficient	std.err	p-value
review length	0.01	<0.001	< 0.001*
review votes	-0.07	<0.001	< 0.001*
price	0.001	<0.001	< 0.001*
rating	0.1	<0.001	< 0.001*
_cons	0.00	<0.001	< 0.001*

*statistically significant at 0.05



Interpretations: The regression analysis suggests that there's a strong positive correlation between the length of reviews and the sentiment score, meaning that buyers tend to provide more detailed comments when they are satisfied with a product. Conversely, there's a notable negative correlation between the number of review votes and the sentiment score, indicating that shoppers often perceive negative comments as more helpful and therefore tend to seek them out more actively.