SAS®
GLOBAL FORUM
2019

USERS PROGRAM

APRIL 28 - MAY 1, 2019 | DALLAS, TX
The digital economy is showing tremendous growth in the 21st century, and it is having a massive impact on the current society. E-commerce is one element of the Internet of Things (IoT), and its worldwide retail sales amounted to 2.3 trillion US dollars. This amount shows the popularity of online shopping, and it indicates an evolution of retailers in this industry. A recent study conducted by GE Capital Retail Bank found that 81% of consumers perform online research before buying products. This indicates that consumers rely heavily on others' opinions and experiences in order to buy a product. Businesses need to understand customers' views of their products and of competitors' products for strategic marketing. E-commerce businesses provide a platform to generate user-experience content through customer reviews, which are vital for a buyer to choose the best product out of numerous similar products available in the market. Companies need to analyze the customers' perspectives through reviews for better business, evaluate customer engagement, and devise strategies for the launch of their products. This paper focuses on analyzing customer reviews primarily on Amazon using Python and SAS® Text Miner. This project determined which product features are given high-ratings or low-ratings, and how the high-rating features of a best-selling product perform compared to a similar product that is sold by a different vendor.
Introduction

The rapid evolution of online market sale helps customers to find the good product, based on online reviews in e-commerce websites like Amazon. These reviews suggest to determine high and low rating features of any individual product and thereby provide access to compare high rating features of a bestselling product with that of other similar products available in the market. It is often important for any business to understand the customer opinion about a product. The reviews from e-commerce websites help companies to determine the customer engagement and plan strategies for future products to gain better business value. Textual data in the form of reviews from Amazon have been gathered to extract the features of a product.

Objective

- Understand customer opinion about a product
- Help companies to determine the customer engagement
- Plan strategies for future products to gain better business value
- Help customers to find best product out of numerous products
Analyzing Amazon’s Customer Reviews using SAS Test Miner for Devising Product Launch Strategies

Manideep Mellachervu
Oklahoma State University

Anvesh Reddy Minukuri
Comcast Corporation

Abstract

Introduction

Methods

Results 1

Results 2

Conclusion

Data Collection

- ASIN code is used for extracting reviews from Amazon
- Data Dictionary has Author, Header, Posted Date, User Rating, Comment
- Filter Node helps to separate the comments
Analyzing Amazon’s Customer Reviews using SAS Test Miner for Devising Product Launch Strategies

Manideep Mellachervu  Anvesh Reddy Minukuri
Oklahoma State University  Comcast Corporation

Text Parsing

- ‘Num’, ‘Punct’ types of attributes have been ignored.

Text Filter

- ‘Num’, ‘Punct’ types of attributes have been ignored.

Text Cluster

40 as the maximum number of clusters and 15 as the number of descriptive terms to describe clusters using Expectation-Maximization Cluster Algorithm

<table>
<thead>
<tr>
<th>Cluster ID</th>
<th>Descriptive Terms</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pill +big +charge +device +music +phone portable +little +size bass battery +speaker +volume +listen +easy</td>
<td>Easy to carry-Portability</td>
</tr>
<tr>
<td>2</td>
<td>perfect +pair +amaze +play +song +device +waterproof awesome +pool easily highly +small +boy +recommend +easy</td>
<td>Awesome Waterproof</td>
</tr>
<tr>
<td>3</td>
<td>+good, good sound, good product, good quality, +sound</td>
<td>Sound Quality is good</td>
</tr>
</tbody>
</table>

Text Topic

Number of Multi-Term Topics has been set to 20

### Abstract

Introduction

Methods

Results 1

Results 2

Conclusion

Text Parsing

- ‘Num’, ‘Punct’ types of attributes have been ignored.

Text Filter

- ‘Num’, ‘Punct’ types of attributes have been ignored.

Text Cluster

40 as the maximum number of clusters and 15 as the number of descriptive terms to describe clusters using Expectation-Maximization Cluster Algorithm

<table>
<thead>
<tr>
<th>Cluster ID</th>
<th>Descriptive Terms</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pill +big +charge +device +music +phone portable +little +size bass battery +speaker +volume +listen +easy</td>
<td>Easy to carry-Portability</td>
</tr>
<tr>
<td>2</td>
<td>perfect +pair +amaze +play +song +device +waterproof awesome +pool easily highly +small +boy +recommend +easy</td>
<td>Awesome Waterproof</td>
</tr>
<tr>
<td>3</td>
<td>+good, good sound, good product, good quality, +sound</td>
<td>Sound Quality is good</td>
</tr>
</tbody>
</table>

Text Topic

Number of Multi-Term Topics has been set to 20

<table>
<thead>
<tr>
<th>Topic ID</th>
<th>Topic</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+Charge, +Charger, +Purchase, +Port, +Disappont</td>
<td>Charge Port is not good</td>
</tr>
<tr>
<td>2</td>
<td>+nice, +sound, +bass, bass, +clear</td>
<td>Good Bass</td>
</tr>
<tr>
<td>3</td>
<td>+good, good sound, good product, good quality, +sound</td>
<td>Sound Quality is good</td>
</tr>
</tbody>
</table>
Analyzing Amazon’s Customer Reviews using SAS Test Miner for Devising Product Launch Strategies

Concept Link Diagram for word PRODUCT

Concept link diagram in adjacent figure shows word PRODUCT is highly associated with words GREAT, RECOMMENDED, EXCELLENT, AWESOME. It tells that particular speaker is a good product and reviewers highly recommend this product to other customers.

Concept Link Diagram for word SOUND

Concept link diagram in adjacent figure shows word SOUND is strongly associated with word GREAT, indicates that particular speaker is a great product with the good sound quality feature.
# Conclusion

<table>
<thead>
<tr>
<th>Product</th>
<th>High Rating Features</th>
<th>Low Rating Features</th>
<th>Average Rating</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beats</td>
<td>Portability, Sound Quality</td>
<td>Charge Port, Battery</td>
<td>3.6</td>
<td>170</td>
</tr>
<tr>
<td>JBL</td>
<td>Waterproof, Sound Quality, Battery life, Charger</td>
<td>Bluetooth</td>
<td>4.5</td>
<td>149</td>
</tr>
<tr>
<td>Ultimate Ears</td>
<td>Waterproof, Sound Quality, Bass, Battery life</td>
<td>Charger, Bluetooth, Product Component</td>
<td>4</td>
<td>146</td>
</tr>
<tr>
<td>Bose</td>
<td>Sound Quality, Bass Battery life</td>
<td>Battery, Bluetooth</td>
<td>4.4</td>
<td>160</td>
</tr>
<tr>
<td>Altec Lansing</td>
<td>Sound Quality, Battery life</td>
<td>Noise, Product Component</td>
<td>3.9</td>
<td>160</td>
</tr>
</tbody>
</table>

This feature analysis helps retailers and companies understand customer expectations to create future products that meet the needs of their customers and gain business value. Additionally, this analysis prototype also benefits the Amazon retailers to constantly view and evaluate the customer's reviews and increase their brand loyalty by following the above recommendations.

## References


## Acknowledgements

I sincerely thank my professors Dr. Goutam Chakraborty and Dr. Miriam McGaugh for their constant support and guidance through every phase of this project.