JSR Publications

Available Online

J. Sci. Res. **13** (2), 335-345 (2021)

JOURNAL OF SCIENTIFIC RESEARCH

www.banglajol.info/index.php/JSR

Brand Positioning and Segmentation of Sneakers through Multi-Dimensional Customer Experience Analysis

R. Senthilkumar*, B. RubanRaja, Monisha

Department of Statistics, Loyola College, Chennai – 600 034, India

Received 28 June 2020, accepted in final revised form 12 February 2021

Abstract

A huge corpus of valuable information on customer experience is available as unstructured form in customer reviews on e-commerce websites. Multivariate data analysis techniques are effective in uncovering hidden patterns and segments in structured data. A major challenge is to convert the unstructured data into a structured form for applying multivariate techniques. In this article, we have provided a text analysis based approach coupled with multivariate techniques to uncover the sentiment of various features associated with different brands and to determine the brand positions and segments through perceptual mapping and cluster analysis.

Keywords: Hierarchical clustering; Dendrogram; Multivariate analysis; Perceptual mapping.

doi: http://dx.doi.org/10.3329/jsr.v13i2.47841 J. Sci. Res. **13** (2), 335-345 (2021)

1. Introduction

People living in the 21st century have all experienced the ease of shopping through various e-commerce websites. Customers are loaded with piles of offers, discounts, and coupons when shopped online. When we look 20 years back, most of the businesses have grown by word of mouth. Recommending or suggesting their experience in using the product makes the particular brand create its own identity with its customers. It enables producers to pitch in customers by perceiving their choices and also enhance their products. In this era of online shopping, word of mouth and recommendation has gained their new name which is called **"reviews"** and "**ratings"**. Thousands of products are sold online every day. Reviews written by customers are segregated as positive and critical reviews by e-commerce algorithms. It helps other customers to differentiate between the same products but different brands. Online shopping has replaced the personal, human element with a digital conversion funnel. As Amazon approached four percent of net retail sales in 2017 and captured 44 % of e-commerce revenue, consumers are increasingly making purchase

* *Corresponding author:* senthilkumar0185@gmail.com

decisions with the best tool available. Once shoppers leave Amazon’s home page they never see an item without a star rating and review count displayed next to the name, yet consumers have very little knowledge on how Amazon solicits and shapes that information. Product reviews are not the sole deciding factor in where items appear in Amazon rankings; sales are still the champion when Amazon’s A9 algorithm is delivering search results. Amazon happily indicates ‘Best Seller’ and ‘Amazon’s Choice’ next to products but those only appear on two results. Every product gets a boost from its review summary right on the results page:the real answer is that reviews are designed to influence conversion; 22 % of shoppers won’t look anywhere else once they’ve identified an Amazon product they want to buy and reviews are the real deal when it comes to this decision. Research shows that 84 % of shoppers trust online reviews as much as a personal recommendation and 91% of shoppers occasionally or regularly read online reviews. Amazon strategically place reviews in higher location priority before the product specifications or detail bullets. Customer feedback appears before your marketing team’s carefully crafted messaging and even gets its rollover graphic right on the search page.A9 delivers awareness when consumers are searching; its reviews that push consumers from consideration to purchase. Approximately 70 % of consumers read reviews before making a purchase and a whopping 68 % of those shoppers form an opinion on a product after reading between one and six online reviews. With so many reviews available and so few being consumed it’s easy to see why companies are scrambling to gather positive feedback on Amazon.

In e-business, new value can be created by how transactions are enabled. Grounded in the rich data obtained from case study analyses and in the received theory in entrepreneurship and strategic management, we develop a model of the sources of value creation. The model suggests that the value creation potential of e-businesses hinges on four interdependent dimensions, namely: efficiency, complementarities, lock-in, and novelty.

Amit and Zott [1] suggests that no single entrepreneurship or strategic management theory can fully explain the value creation potential of e-business. Rather, integration of the received theoretical perspectives on value creation is needed. To enable such integration, we offer the business model construct as a unit of analysis for future research on value creation in e-business. A business model depicts the design of transaction content, structure, and governance to create value through the exploitation of business opportunities. We propose that a firm's business model is an important locus of innovation and a crucial source of value creation for the firm and its suppliers, partners, and customers. Bilotti [2] investigated that the marketing tactics utilized by Dove brand and Nike, Inc. Although the means were different, both companies successfully generated emotional brand attachment between their products and modern consumers. Charles *et al.* [3] pointed out that they act as various forms of model: to provide means to describe and classify businesses; to operate as sites for scientific investigation, and to act as recipes for creative managers. We argue that studying business models as models is rewarding in that it enables us to see how they embody multiple and mediating roles. We illustrate our ideas concerning practices in the real world and academic analyses, especially in this Long Range Planning Special Issue on Business Models.

In today's complex business world, strategic planning is indispensable to achieving superior management. Steiner [4] classic work, known as the bible of business planning, provides practical advice for organizing the planning system, acquiring and using information, and translating strategic plans into decisive action. An invaluable resource for top and middle-level executives, Strategic Planning continues to be the foremost guide to this vital area of business management. Hussain *et al.* [5] studied that the highlight Nike's strategies which focus on innovation and emphasis on its research and development department, provision of premium pricing for its customers, broad differentiation strategy, market Segmentation Strategy and Closed-Loop strategy. The Adidas strategies focus on broad differentiation, innovation, trying to produce new products, services and process to cope up with the competition. It embraces a multi-brand strategy, emphasis on expanding activities in the emerging markets, continuously improving infrastructure, processes and systems, foster a culture of challenging convention and embracing change, foster a corporate culture of performance, passion, integrity and diversity. These strategies coupled with its resources and unique capabilities form the basis of sustainable competitive advantage for both companies.In this paper, we have investigated that to identify the brand positions of various sneaker brands and the features influencing the sale of the particular brand. The given an excellent insight to each of the brands about their competitors. To provide a general idea to each brand about their areas that need improvement this would increase their customer base. To give a wider picture about vacant positions to be captured i.e. the position where there is a possibility to capture a larger number of customers.

2. Data Collection

2.1. *Text mining*

Text Mining is also known as Text Data Mining. The purpose is too unstructured information, extract meaningful numeric indices from the text. Thus, make the information contained in the text accessible to the various algorithms. Information can extract to derive summaries contained in the documents. Hence, you can analyze words, clusters of words used in documents. In the most general terms, text mining will "turn text into numbers". Such as predictive data mining projects the application of unsupervised learning methods.

2.2. *Natural language processing*

Natural Language Processing or NLP is a field of Artificial Intelligence that gives the machines the ability to read, understand and derive meaning from human languages.

Syntax and Semantic analysis are two main techniques used with natural language processing.

2.2.1. *Syntax*

Syntax refers to the arrangement of words in a sentence such that they make grammatical sense.In NLP, syntactic analysis is used to assess how the natural language aligns with the grammatical rules.Computer algorithms are used to apply grammatical rules to a group of words and derive meaning from them.
Here are some syntax techniques that can be used:

- *Lemmatization*: It entails reducing the various inflected forms of a word into a single form for easy analysis.
- *Morphological segmentation*: It involves dividing words into individual units called morphemes.
- *Word segmentation*: It involves dividing a large piece of continuous text into distinct units.
- *Part-of-speech tagging*: It involves identifying the part of speech for every word.
- *Parsing:* It involves undertaking grammatical analysis for the provided sentence.
- *Sentence breaking*: It involves placing sentence boundaries on a large piece of text.
- *Stemming*: It involves cutting the inflected words to their root form.

2.2.2. *Semantics*

Semantics refers to the meaning that is conveyed by a text. Semantic analysis is one of the difficult aspects of Natural Language Processing that has not been fully resolved yet.It involves applying computer algorithms to understand the meaning and interpretation of words and how sentences are structured.
Here are some techniques in semantic analysis:

- *Named entity recognition (NER)*: It involves determining the parts of a text that can be identified and categorized into preset groups. Examples of such groups include names of people and names of places.
- *Word sense disambiguation*: It involves giving meaning to a word based on the context.
- *Natural language generation*: It involves using databases to derive semantic intentions and convert them into human language.

2.3. ***Data preparation***

To segment the customers and to know the position of each brand in the market, the data must be collected for only base price models in each brand. With the help of R software using the product URL and by looping through reviews in all pages for a specific product, the data is finally extracted and preprocessed before analyzing it. The following steps involved in data preparation are:- Data extraction, Text Preprocessing, Tokenization, Topics identification, Topics mapping, Breakdown reviews to sentence, Topic flag variables and Sentence level sentiment identification.

2.3.1. *Data extraction*

The amazon reviews are scraped for twelve different base price sneaker brands. The product code has been used to scrape reviews for each brand. The **"rvest"** package in R has been used to scrape amazon reviews.

2.3.2. *Text preprocessing*

To preprocess your text simply means to bring your text into a form that is predictable and analyzable for your task. The following types of text preprocessing techniques are:-

- *Lowercasing:* Lowercasing all the text data, although commonly overlooked, is one of the simplest and most effective forms of text preprocessing.
- *Stemming:* Stemming is the process of reducing inflection in words (e.g. troubled, troubles) to their root form (e.g. trouble). The "root" in this case may not be a real root word, but just a canonical form of the original word.
- *Lemmatization:* Lemmatization on the surface is very similar to stemming, where the goal is to remove inflections and map a word to its root form. The only difference is that, lemmatization tries to do it the proper way. It doesn't just chop things off; it actually transforms words to the actual root. For example, the word "better" would map to "good". The original word to lemmatized word (Table1) is given below:-

Table 1. The Original word to Lemmatized word.

OriginalWord	Lemmatized Word
Trouble	Trouble
Troubling	Trouble
Troubled	Trouble
Troubles	Trouble
Goose	Goose
Geese	Goose

- *Stopword removal:* Stop words are a set of commonly used words in a language. Examples of stop words in English are "a", "the", "is", "are" and etc. The intuition behind using stop words is that, by removing low information words from text, we can focus on the important words instead.
- *Noise removal:* Noise removal is about removing characters digits and pieces of text that can interfere with your text analysis. Noise removal is one of the most essential text preprocessing steps. It is also highly domain dependent.For example, in reviews,

noise could be all special characters, emoji's, as it signifies concepts that can characterize a review.

2.3.3. *Tokenization*

The process of splitting an input text into meaningful chunks is called Tokenization, and that chunk is actually called token or given a character sequence and a defined document unit, tokenization is the task of chopping it up into pieces, called tokens, perhaps at the same time throwing away certain characters, such as punctuation. Tokens can be:

- A useful unit for further semantic processing
- Can be a word, sentence, paragraph, etc

2.3.4. *Topics identification*

There are different topics represented by reviews and at this step the frequency count of those topics are obtained. For example, the word durability, durable, long lasting is related to the topic durability. This count enables us to know how much of these words are present in a document and helps to decide a topic related to those words.

2.3.5. *Topic mapping*

There exists a sequence of words which are related to a particular topic. For example, the words original, fake, duplicate, etc…all comes under the topic called Authenticity of the product. In this step, all those words are mapped to their respective topics.

2.3.6. *Breakdown reviews into sentences*

A review can be a single word, a paragraph or a sentence. When it a para, it is split into sentences and a sentence id, review id is given to them. For example, a small chunk of original data after text preprocessing. The original data consists of a review_id and review (Table 2) is given below.

Table 2. The original data consists of a Review id and Review.

Review id	Review
1	Receivedproduct shown description pic. Regret closer look shoe looks cheap real adidas. Zoomed pics buying
2	Typical sneakers. These second pair shade blue camouflage pattern. Feels comfortable daily use. Goes well chinos looks best light blue jeans black shirt.

After topic mapping and breaking down reviews into sentences, the data is transformed to the following way as given in Table 3.

Table 3. The data consists of a Review id into Sentence id and Review column.

Review id	Sentence id	Review column
1	0	received product shown description pic
1	1	regret closer look shoe looks cheap real adidas
1	2	zoomed pics buying
2	0	typical sneakers
2	1	these second pair shade blue camouflage pattern
2	2	feels comfortable daily use
2	3	goes well chinoslooks best light blue jeans black shirt

2.3.7. *Topic flag variable*

To create a variable of each topic, the values of the variables are binary i.e. if the words under each topic are present in a sentence then the value of the variable is 1 else 0.The data after creating variables are given in Table 4.

Table 4. Each topic creating variables.

Review id	Sentence id	Review column	Value for money	Trait s	Color	Authenticity
1	0	received product shown description pic	0	0	0	0
1	1	regret closer look shoe looks cheap real Adidas	1	0	0	1
1	2	zoomed pics buying	0	0	0	0
2	0	typical sneakers	0	0	0	0
2	1	these second pair shade blue camouflage pattern	0	0	1	0
2	2	feels comfortable daily use	0	0	0	0
2	3	goes well chinos looks best light blue jeans black shirt	0	1	1	0

2.3.8. *Sentence level sentiment identification*

The sentiment of a piece of text is its positivity or negativity, and sentiment analysis gives an objective idea of whether the text uses mostly positive, negative, or neutral language.In order to calculate the sentiment of a piece of text, we split it into individual words.We have a database of words, each with a "score" to determine how positive or negative it is.For this data, sentiment scores are calculated on the basis of count of positive and negative words. If there are two positive words in a sentence, then the sentiment score for the sentence is 2. If it is negative, then it is -2. The data after assigning sentiment scores for each sentence is given in Table 5.

Table 5. The data after assigning sentiment scores for each sentence.

Review id	Sentence id	Review column	Value for money	Traits	Color	Authenticity	Score
1	0	received product shown description pic	0	0	0	0	0
1	1	regret closer look shoe looks cheap real Adidas	1	0	0	1	-2
1	2	zoomed pics buying	0	0	0	0	0
2	0	typical sneakers	0	0	0	0	0
2	1	these second pair shade blue camouflage pattern	0	0	1	0	0
2	2	feels comfortable daily use	0	0	0	0	1
2	3	goes well chinos looks best light blue jeans black shirt	0	1	1	0	2

2.3.9. *Sentence level topic modeling data*

The score has to be multiplied to each variable and a transformed data is given in Table 6.

Table 6. The score to be multiplied each variable and transformed data.

Review id	Sentence id	Review column	Value for money	Traits	Color	Authenticity	Score
1	0	received product shown description pic	0	0	0	0	0
1	1	regret closer look shoe looks cheap real Adidas	-2	0	0	-2	-2
1	2	zoomed pics buying	0	0	0	0	0
2	0	typical sneakers	0	0	0	0	0
2	1	these second pair shade blue camouflage pattern	0	0	0	0	0
2	2	feels comfortable daily use	0	0	0	0	1
2	3	goes well chinos looks best light blue jeans black shirt	0	2	2	0	2

3. Results and Discussion

In Table 7, various cluster solutions and their corresponding Pseudo F-statistic value is recorded and based on the results, three cluster solution which resulted in the largest F-statistic value is chosen as the optimal cluster solution.

Table 7. Pseudo F-statistic values for different clusters.

Clusters	Index
2	5.6245
3	6.2904
4	5.6828
5	5.7612
6	5.1759
7	5.3147
8	6.0296

Fig. 1 shows that, the horizontal axis in the dendrogram represents the clusters whereas the vertical axis represents the distance or dissimilarity between clusters. Cluster 3 brands are Adidas, Centrino, Vans, Sparx, Converse and Reebok classics; the cluster 2 brands are Reebok, Redtape, Afrojack and United Colors of Benetton and cluster 1 brands are Hushpuppies and Levi. The two outliers, Adidas and Centrino brands are fused in rather arbitrarily at much higher distances. The characteristics of each cluster can be determined by analyzing the variables separately for each cluster.

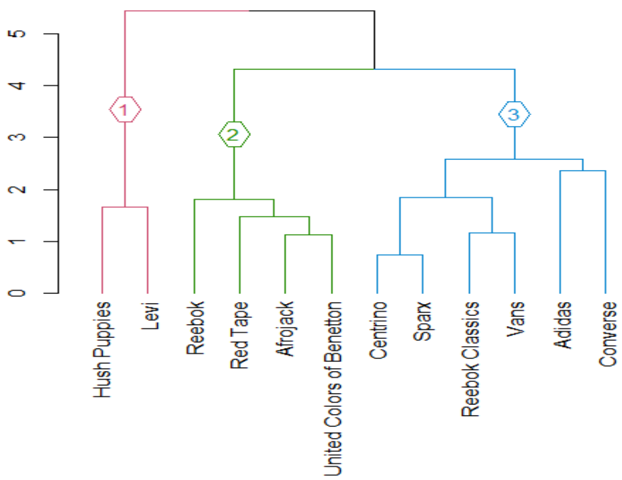

Fig. 1. Dissimilarity between clusters for various brands.

Perceptual mapping is a visual representation of where a brand, product, or service stands among competitors. It is also known as positional mapping. Perceptual mapping is a diagrammatic technique used by asset marketers that attempts to visually display the perceptions of customers or potential customers. The positioning of a brand is influenced by customer perceptions rather than by those of businesses.

From Fig. 2, the perceptual map reveals two major segments in which Levis brand has the strong hold in terms of style and colour, attracting customers interested in the looks of sneakers. Adidas, Vans and Reebok classic are major competitors in the durability aspect of sneaker brands.

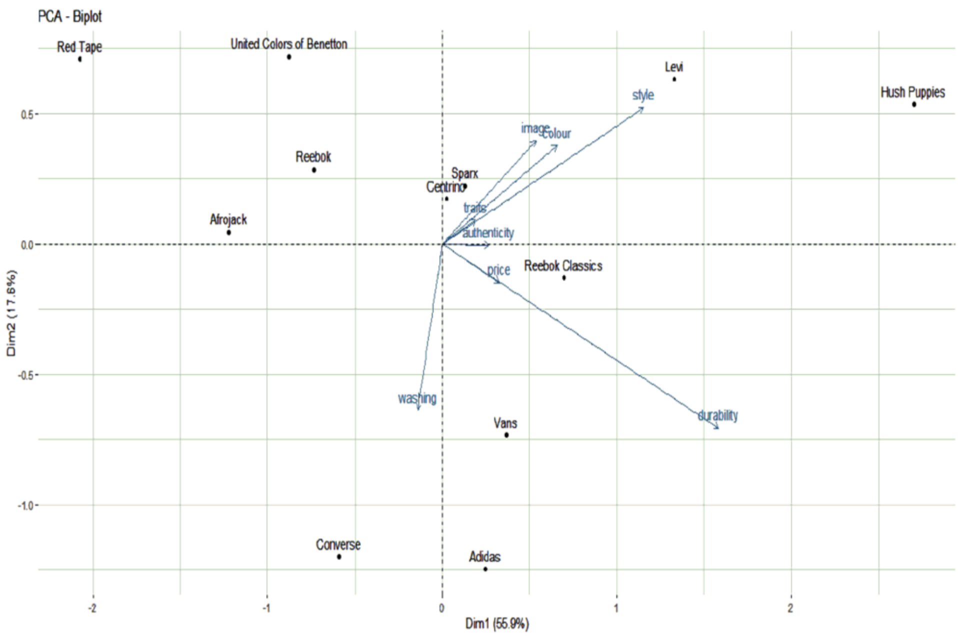

Fig. 2. Perceptual mapping of competing products.

It is more meaningful to look at the variables in their original scales when data is centered, negative values mean "lower than most" and positive values mean "higher than most". From Tables 8 and 9, it can be concluded that cluster 1 brand with relatively moderate durability and style. Cluster 2 brands with low durability and cluster 3 are brands with high color, durability and style features.

Table 8. The brand wise clusters membership.

Product	Value for money	Traits	Colour	Authenticity	Washing	Durability	Style	Image	Cluster
Adidas	-0.7070	-0.7953	-1.0603	-0.3978	0	1.3245	0.2645	-0.7953	1
Afrojack	-0.4873	-0.5303	-0.7953	-0.9824	0	-0.7953	0.2645	-0.7938	2
Centrino	0.0163	0.1005	-0.4044	-0.4583	0	0.0525	0.5613	-0.2105	1
Converse	0.4160	-0.3537	-0.4420	-0.7953	1.3245	0	0	-0.7958	1
Hush Pupies	0.1131	-0.0143	1.0595	-0.5598	-0.2654	1.6778	2.3844	0.2645	3
Levis	-0.1954	-0.4844	0.1131	-0.1799	-0.2654	0.6178	0	0.2645	3
Red Tape	-0.9278	-0.4420	-0.4420	-0.9278	0	-0.1855	0	-0.7953	2
Reebok	0.2645	-0.4099	0.2645	-0.1594	0	-0.7953	0	-0.4420	2
Reebok Cl	0.3662	0.1247	0.2256	-0.0351	0.2521	0.4505	0.8033	-0.2096	1
Sparx	-0.1329	0.0302	-0.5100	-0.4896	-0.9881	0.3609	0.3705	-0.1770	1
United Co	-0.4773	-0.1538	-0.3978	-0.5303	-0.7953	-0.7953	0.2645	-0.4420	2
Vans	-0.0635	0.2031	0.0802	-0.3328	0.0525	0.8702	0.1131	-0.7953	1

Table 9. Clusters mean vector.

Clusters	Value for money	Features	Colour	Authenticity	Washability	Durability	Style	Image
1	-0.0174	-0.1150	-0.3518	-0.4182	0.0436	0.5098	0.3521	-0.4972
2	-0.4055	-0.3840	-0.3426	-0.6500	-0.1988	-1.0603	0.1322	-0.6187
3	-0.0411	-0.2494	0.5863	-0.3698	-0.2654	1.1478	2.1195	0.2645

4. Conclusion

The perceptual mapping based on customer sentiment identifies Adidas, Vans and Reebok classic with a strong position in the market in terms of durability and Levis brand in the position of aesthetics. Cluster analysis identifies three major segments in which cluster 1 consists of relatively moderate durability and style, brands in cluster 2 consists of low durability and brands in cluster 3 comprises of high satisfaction in aesthetics. These insights in market positions can help brands understand the market competitors in terms of customer perception and to deploy changes in manufacturing and advertisement aspects to establish a better position in the market.

References

1. R. Amit and C. Zott, Strategic Manage. J. **22**, 493 (2001). https://doi.org/10.1002/smj.187
2. K. Bilotti, Emotional Brand Attachment: Marketing Strategies for Successful Generation, CMC Senior Theses (2011) Paper 273.
3. C. Baden-Fuller and M. S. Morgan, Long Range Planning **43**, 156 (2010). https://doi.org/10.1016/j.lrp.2010.02.005
4. G. A. Steiner, Strategic Planning: What Every Manager Must Know (1997).
5. H. A. A. Mahdi, M. Abbas, and T. I. Mazar, Int. J. Business Manage. Econ. Res. (IJBMER) **6**, 167 (2015).
6. M. E. Porter, The Competitive Advantage: Creating and Sustaining Superior Performance (Free Press, New York, 1998).
7. M. D. Shank and M. R. Lyberger, Sports Marketing: A Strategic Perspective, 5th Edition (Routledge, UK, 2014). https://doi.org/10.4324/9781315794082
8. C. Zott and R. Amit, Organization Sci. **18**, 181 (2007). https://doi.org/10.1287/orsc.1060.0232
9. C. Zott, R. Amit, and L. Massa, J. Manage. **37**, 1019 (2011). https://doi.org/10.1177/0149206311406265