Problem Chosen	2020 MCM/ICM Summary Sheet	Team Control Number
C		2012358

Summary

As online marketplace on the rise, Sunshine Company plans to promote three new products in the online market including a microwave oven, a baby pacifier, and a hair dryer, and we're hired to search effective strategies about sales and product optimization. Hence we establish models based on previous review data sets of Amazon assisting our exploration of products' features, markets' condition and consumers' mentality and to gain some rules of these factors.

Firstly, we roughly estimate these data sets from three aspects: advantages, disadvantages and usage which provides a superb basis for the following analyses. Then they are classified according to their functions incarnating their characters for help Sunshine Company succeed. For the following section, as data has been managed, we establish models to analyze three closely related directions: product, market and consumer.

For identifying the dominant features of one product, we establish weight allocation model helping us give different weight to feature words in consumer's review text according to whether he is vine and how much helpful votes he gained. After that we calculate the sum weight of factors and find the features which appear most times for we deem consumers concern features mentioned in reviews. And then we build model using moving average method to investigate whether a product's reputation can be suggested by time for getting company's appropriate timing entering markets. By setting diverse weight to star rating according to vine and helpful votes since we think the more helpful votes it gained, the more effects on readers it has, it's obviously that these markets have various market tendency. Microwave oven market is fluctuate and hair dryer market and baby pacifier are stable.

In the third part, we select comprehensive scoring index analysis model to give review texts high, middle and low scores based on setting weight to diverse emotion words they contained according to their emotional intensity. Then we combine text system with star rating which is also classified into 3 categories getting 9 combinations. After that we use the flow chart of reviews' scores calculation graph to perceive the correlation between specific star ratings that they're highly relative and reviews whose scores are counted with the method above. At last, randomly picking 45 quality descriptors from texts and comparing sample's average star rating with total's give the relationship of them that specific negative and positive quality descriptors are more associated with rating levels(more details are shown in Appendix).

Last but not least, there are some suggestions concluded from analyses above: enhancing product performance in important feature areas, such as price of baby pacifiers; entering the microwave oven market at appropriate time point; finding the direction of future product optimization by calculating customer's review texts and star ratings after entering the market.

A Letter to the Marketing Director of Sunshine Company

Dear Marketing Director,

We are writing to you informing you our analysis result of the microwave oven, the baby pacifier and the hair dryer markets. We establish ideal models to analysis these three markets from two aspects, product's features and market operation mechanism, and our suggestions are put forward based on the properties reflected in models. Our models are made all on account of the data from Amazon, a huge product online shopping center, which means we can predict how your products perform after entering this market.

The features of products predominantly affect consumers' evaluations, so we establish the weighting model to figure out features mentioned the most times in reviews, since if consumers concern this feature, they will write their attitudes on them many times. After calculation we find features repeatedly appraised respectively are seven features in hair dryer market, setting, attachment, price, speed, weight, contour and durability, five features in microwave oven market, appearance, efficiency, function, volume and price, and six features in pacifier market, design, colour, safety, price, shape and fitting. This suggests that if your company desires to achieve commercial success, paying attention to these features are vitally important.

Considering long term marketing strategies, we offer an approach to test product features. As having given the weight to each feature, we build a way to measure which of these features need to be improved after entering the market. With the equations shown in the report details and average star rating gained in online markets, we can calculating rough scores of each features and make product optimization according to the result.

Timing of participation is also a key topic, hence we select moving average method to probe into the correlation between time and reviews' estimate. One month is regarded as a unit, and the consequence of exploration is: in microwave oven market, every ten months is a circle, and at the end of the circle consumers' attitudes will reach the highest level while at the intermediate nodes of circles; in hair dryer market and pacifier market, change is hardly visible but fluctuate in hair dryer market is bigger than pacifier's.

By means of discover above, we deem at the time point when reviews' attitudes are most negative is a decent business opportunity to enter microwave oven market, since then gaining benefits is apparent. And for baby pacifier and hair dryer markets, timing is not so specific because their customers' attitudes have little relevance to time.

After a series of model analysis, among these three products, a microwave oven, a baby pacifier, and a hair dryer, we conclude that choosing a appropriate time is vital and compute flaw of products' features assisting company in winning the market.

The recommended strategies obtained from models give Sunshine Company a new market-grabbing pattern may hopefully help the company promote products.

Sincerely,

MCM Team Mmbers

Contents

1 Restatement of the problem

In the era of online shopping, Amazon offers approaches including "star rating", expressing satisfaction level of the product, "reviews" , showing comments, and "helpfulness rating" , rating reviews whether helpful, for consumers to evaluate their purchases, in order to help the company choose market strategies of market participation, potential success and so on.

To estimate the online sale strategies and design characters of these three new productions, a microwave oven, a baby pacifier, and a hair dryer of Sunshine Company, it is essential to establish an objective and systematic analysis model using data to inform strategy. Identify and discuss practical strategies for the company by means of assessing the wealth data representing customer-supplied ratings and reviews for these three products sold in the Amazon, which performs how the product itself, market atmosphere and consumers' preference affect the sale.

In this paper, we have developed a complete evaluation analysis plan based on existing evaluation data on the Internet and focused on three parts, product evaluation of the present market to look for preponderance features, market analysis to learn the tendency of market changes, and consumer analysis which express consumer's appetite.

2 Planned Approach

Using the reviews data from Amazon websites, we are expecting to seek dominant features of three products, distribution of consumers' reviews attitudes according to time, the combination of text and star rating, how specific star ratings incite reviews, and the relation between specific quality descriptors and rating levels.

Firstly, we roughly analysis the data sets of three products with definition of advantages, disadvantages and potential usage. Then we evaluate these data from three aspects: product itself, product market and consumers.

In product part, the task is to search primary features of three products, which means, in other words, searching what products' features consumers are concerned most. A weight allocation model is built to calculate scores of reviews effectively and fairly taking vine and helpful votes into consideration and the weight setting mainly connect to key words describing consumers' opinions of product features. After summing the total score of three products individually, we use the linear correlation method to solve the proportion of each feature, and the more proportion feature occupies, the more significant it is in consumers' reviews.

For product market part, we weighted moving average method in time series model to establish tendency curve and perceive markets' trends. Every month is regarded as a unit, and the weight of month's score is related to reviews. Then we compare one time moving average method and two times moving average method in this condition and choose more suitable one. Accompany with the output of scatter plots, it is easily to see the tendency of markets.

Consumer's mentality is valuable to comprehend, so here we should judge from three aspects, reviews indicate a potentially successful or failing product, whether specific star ratings incite reviews and relationship between specific quality descriptors and rating levels. In the first aspect, we build a text mark system, giving different adjectives, verbs and adverbs weights to count sum-score of review texts, and then combine star rating and text score to divide into 9 combination types. In the second aspect, we choose stratified sampling to explore the correlation between seeing a series of specific star rating and writing reviews. In the third aspect, we make use of the proportion of the total average rating levels of samples' average rating levels whose texts including specific quality descriptors to seek the correlation.

Finally, we will give Sunshine Company our analysis result and proposals based on models.

3 Assumptions

To simplify the real life situation, we will make the following assumptions as a start of construction of our models.

- The consumers' reviews of the product is comprehensive which means consumers will consider all main factors and give it reviews.
- When consumers review a product, they look at recent reviews of the product.
- The review system offered by Amazon won't change and Sunshine Company can predict the future market from the given data.

4 Notations

Table 1: The List of Notation

Symbol	Meaning
ω_i	The weight of product quality's influencing factors
f_i	The main influencing factors' score
$score$	Star rating score
n	The number of all reviewers
n_0	The number of ordinary reviewers
k	The number of ordinary reviewers who get vote
j	The number of vine reviewers who get votes
HV	The number of helpful votes
TV	The number of total votes
$y_i$andY_i	The weight of review
$mark_t$	The synthetic mark of every month

5 Data Comprehension

In order to find a way to help Sunshine Company's online products succeed, we want to use the given data sets of reviews. By analyzing the characteristics and functions of star rating, review text and helpfulness ratings, and then analyzing the relationship between these factors that carries on the interpretation, and proposes are using these characteristics and the relation to guide the online sale strategy methods.

- Star Rating: The range of values is 1,2,3,4,5, which can most directly reflect the user's satisfaction with the product. Star rating is characterized by simplicity and clarity, with a large number of comments that users can easily access without knowing anything about the general quality of a product before. But it is precisely because of the rating mechanism, the stars often do not reflect very accurately what the user really thinks (maybe the user wants to score 4.5, but unfortunately he can only choose between 4 or 5 scores). For this reason, we can process the data more efficiently because of the simplicity of the score. In chapter 7, based on the data set providing time series, we use specific methods to discover how a product's user's star rating changes over time. On this basis, we can predict the future sales trend of a product within a certain margin of error. There is no doubt this approach can provide a great reference for the online product promotion strategy of Sunshine Company.

- Reviews: The amount of information contained in a consumer's review comment is often enormous. Unlike star rating alone, the description of the product can go into almost every aspect. Also, because users are able to reflect their true thoughts in their review texts, therefore, the user's comments are more indicative of the user's preference for a product than the star rating. However, how to make the abstract review texts translate into concrete math scores is a difficulty. In 8.1, we will propose a keyword - based scoring model quantifying the user's texts. With the means to measure user reviews, we can analyze the market and provide Sunshine Company sales strategies.

- Helpfulness Ratings: Helpfulness rating is the indicator based on the user's text and reflects the degree how other users agree with him. This is undoubtedly an important factor affecting the credibility of user reviews. We can do this based on the number of helpful votes a user gained to judge whether the user has provided comprehensive and effective information on the product so that we can make specific improvements or adjustment of the product's sales strategies, which helps the Sunshine Company's products to better adapt to the market demand and achieve success.

We believe that there is no doubt there is a close relationship between a user's star rating and a review text, when we put the user's star rating combined with a review we'll get a more realistic measure of what users like about the product. According to this index, we can analyze the degree of preference of Sunshine Company's products in consumers' mind, so as to make timely decisions and adjustments.

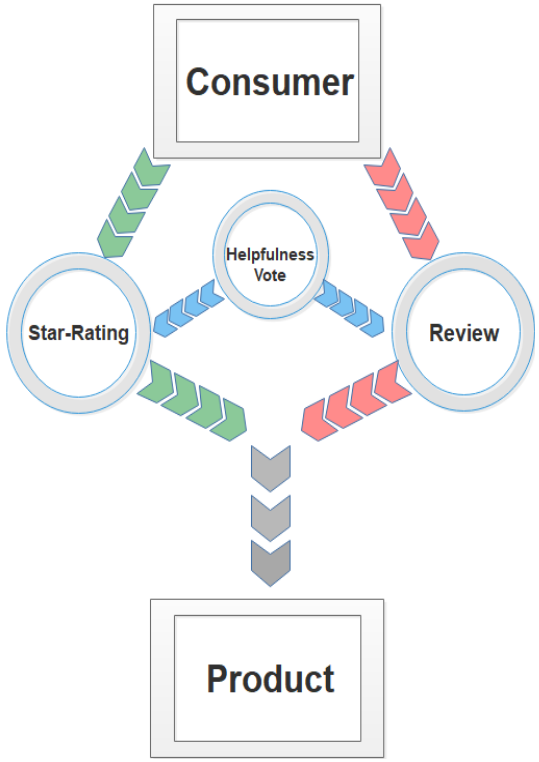


Figure 1: Relations Between Factors

6 Product Evaluation

Based on the data analysis above, we consider to establish a multiple regression analysis model to seek out what features of products are decisive factors that reviewers are concerned dominantly when criticizing the using experiences. The process of multiple regression analysis is consist of data filtration, building model and calculating weight.

6.1 Data Filtration

Firstly, We believe that each customers' evaluation of the product is comprehensive, besides the reviewer will consider each factor of a product, score each factor, and combine the reviewers' preferences(i.e. weights)for different factors of the product to finalize the product. Hence the unique comments gained a mass of helpful votes will be utilized to represent the majority's concern and attitudes to products, as numerous reviewers voted for their helpful use. So hypothesize the number of users who voted helpful to a comment is the number of users who had the same preferences as the comment,regardless of the fact that each user's rating on different factors might fluctuate. Concluding from portion of these valid reviews, comments on each products of a microwave oven, a baby pacifier, and a hair dryer have several primary elements.

Secondly, we suppose there are n factors to measure the quality of a product, and ϖ_i represents the weight of the i-th factor (the reviewer's preference for the factor). Assuming f_i is the reviewer's score for the i-th factor of the product (1 to 5), we judge the score from content of comment on behalf of customer preference for each factor. If a review mentions one feature, it'll be considered the reviewer pays attention to this feature, then we give the feature a score according to the feeling of the reviewer. However, if one reviewer doesn't mention one dominant feature, we'll give it an average score.

Thirdly, with combination of the star rating, which is presented as $score_i$, we could gain the weight of each factors via curving fit these review data based on ratings and reviews.

6.2 Building Model

After the analysis above, obviously we can get a synthetic score formula for linear fitting like this:

$$score = \sum_{i=1}^{n} \omega_i f_i, \omega_i \in (0, 1), f_i \in 1, 2, 3, 4, 5, i = 1, 2, \ldots n \tag{1}$$

Fitting the data points every reviewer into a regression line, then the weight of each factor can be figured out.

6.3 Calculation and Result

In Matlab fitting calculation, we can get the weight of each factor of these three products, which show the proportion of factors in consumers' concern for the products.

Product	Feature	Proportion
Hair Dryer	Setting	100/179
	Attachment	62/179
	Price	48/179
	Weight	41/179
	Contour	27/179
	Durability	42/179
Microwave Oven	Appearance	14/20
	Efficency	9/20
	Function	11/20
	Volume	7/20
	Price	6/20
Pacifier	Design	9/36
	Colour	9/36
	Safety	11/36
	Price	14/36
	Shape	13/36
	Fitting	9/36

Figure 2: Proportion of Features

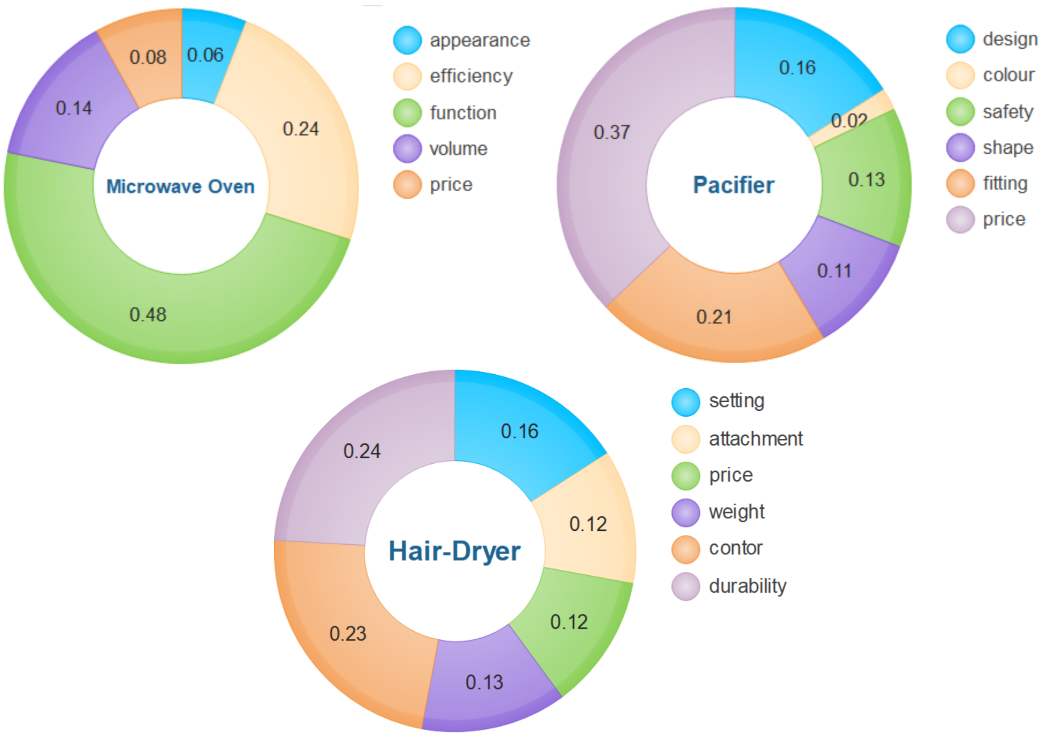


Figure 3: Weight Allocation

These three pie charts in figure3 reflect the majority's feature preference for hair dryer, microwave oven and pacifier, embodying average star rating score of every product market, which can be used for deciding where Sunshine Company should improve once their three products are placed on sale.

For Sunshine Company tracking their market status, comparing its star rating score to the market average which will show what factors' score below the market, and ameliorating its products defects build the company prosperous. For instance, if microwave oven of Sunshine Company gets 4.3 star rating average score, making use of the equation and the weight of product features of microwave oven market gained above, then the linear correlation line which is closest to star rating and feature is obtained by linear fitting that raises appearance score is 1, efficiency is 5, function score is 5, volume score is 4, price score is 1. Paralleling with the data in table 3, appearance score and price score are much lower than the average. Therefore Sunshine Company should pay more attention to these two aspects of its microwave oven.

7 Market Analysis

Another pivotal situation is what is the right timing of the participation, and this needs to take a product's reputation in the online market place into account concluding the market's prosperity tendency in order to seek a super excellent opportunity to enter the market. Hence forecasting the product's reputation over the next 1 to 2 years is crucial. We take advantage of weighted moving average method in time series model to establish tendency curve and predict the future trend.

The data of reviews are from August 2010 to August 2015, and we define one month as a unit of time measurement, whereupon there are 60 time measurement units. Giving each and every ordinary reviewer a base weight(1 point), but giving Amazon Vine members base weight(5 point)for they are invited to enjoy the product freely and submit their experience report who once have written accurate and insightful reviews. Whoever gets one helpful vote will increase an extra 0.05 point, gets one unhelpful vote will reduce 0.1 point, and we assume the number of all reviewers is n, the number of ordinary reviewers is n_0, the number of ordinary reviewers who get votes is k, the number of vine reviewers who get votes is j. In this way it's easily to learn everyone's weight(y_i /Y_i)in all situations:

- Ordinary Reviewe: has no vote: y_i= 1(i=1,2,...n_0-k)

  has votes: y_i =1+0.1(2HV -TV)(i=1,2,...k)

- Vine Reviewer: has no vote:Y_i =5(i=1,2,...$n - n_0 - j$)

  has votes: Y_i =5+0.1(2HV -TV)(i=1,2,...j)

Normalizing the obtained mark,

$$\frac{1}{(n_0 - k) + \sum_{i=1}^{k} [1 + 0.1(2HV - TV)] + \sum_{i=1}^{j} [5 + 0.1(2HV - TV)] + 5(n - n_0 - j)} \tag{2}$$

under this kind of data processing ,we ponder the influence of vine members and helpful votes and the influence of these two variables is shown in the weight calculation.

Next we consider the star rating effects, which reflect whether the sale status is healthy and flourishing or gloomy. Combining the weight calculated above and star rating in order to get a synthetic mark ,this shows consumers' attitude to the product, which means if the mark is high ,then consumers love this product and give it a good evaluation ,otherwise it's opposite. The mark equation is:

$$mark_t = y_i score_i + Y_i score_i, (t = 1, 2, \ldots 60) \tag{3}$$

Two conditions are affecting the product's reputation in the time measure unit, how many people buy this commodity and how these buyer appraise it, and these conditions can be converted to mark. The more mark in one month, the more reputation one product will get during this period, for if more consumers purchase it, n will be bigger, if more reviewers give it high star rating, $score$ will be greater, leading the mark to be higher.

7.1 Moving Average Method

In moving average method , giving high weights to the observation closed to the prediction period , while observations that are further away from the forecast period are given smaller weights accordingly. Adjusting the effect of each observation value on the predicted value with different weights ,so that the predicted value can more closely reflect the future development trend of the market. Since one month is considered as one unit, we can obtain the synthetic mark of reviewers and draw the fitting curve through weighted moving average method ,and the equation can be built as:

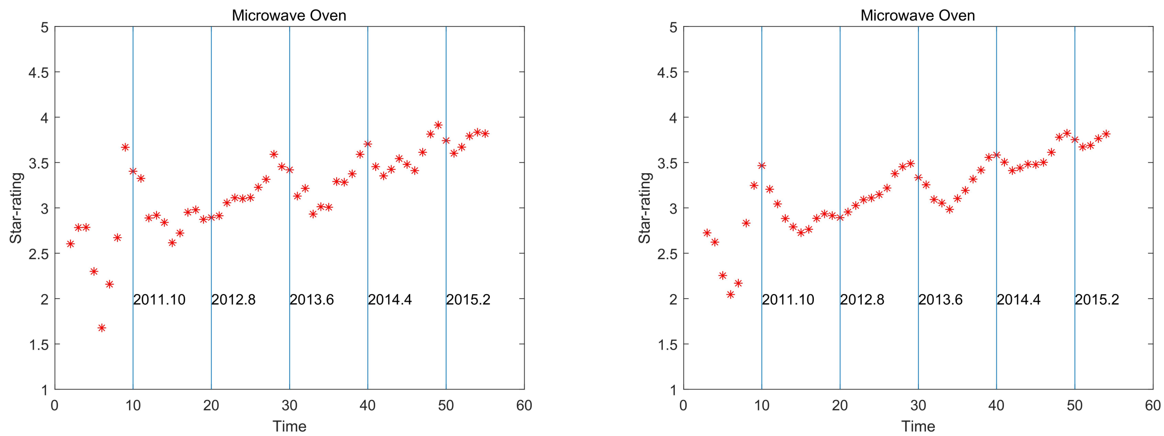
One time moving average method:

$$M_t^1 = \frac{1}{N}(y_t + y_{t-1} + \ldots y_{t-N+1}) \tag{4}$$

Two times moving average method:

$$M_t^2 = \frac{1}{N}(M_t^1 + M_{t-1}^1 + \ldots M_{t-N+1}^1) \tag{5}$$

These are two kinds of methods in time series model, we want to choose one suited method to draw the market tendency chart. We select microwave oven as an example and draw individual smooth scatter graph to compare which method works well.



(a) One time moving average method

(b) Two times moving average method

Figure 4: Compare between one and two time moving average method

From these two graphs we can see that two times moving average method works much better than the other one, trend of the mark obvious, and continuity is stronger, then we choose two time moving average method to analyse the market preference trend. Since then we can plot figures to describe the product's reputation with the normalized mark below.

7.2 Result and Advice

Firstly, the general mark trend is clearly upwards, which means reviews of microwave oven are better and better these years, and we can expect the mark must be higher as time goes on. What can easily perceive is that there is a cycle every 10 months during these five years and it will reach its peak in the 10th month of every cycle. On a 10-month cycle basis, in figure 1 (b), every ten-month cycle is labeled by a blue line, and at the beginning of each cycle, the star-rating will decrease to a low position compared to other time of the cycle, but it will grow steadily and finally get to the highest level. The regular means consumers' attitudes of microwave oven will decrease firstly then increase to the highest position during one cycle and imply product's reputation decreases at the start and increases then during the cycle, which gives Sunshine Company can seize the time when star-rating at lowest level to enter the microwave oven market ,enjoying advantages of star-rating growth and gaining interest.

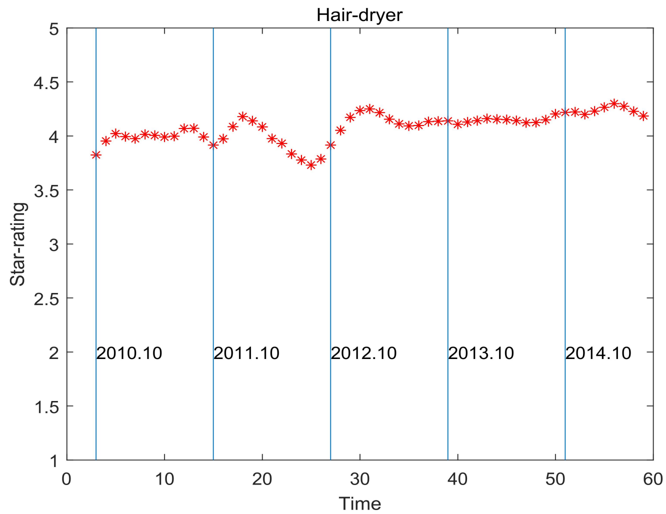


Figure 5: Scatter plot of Hair-dryer

Secondly, in hair-dryer market, the scattered points are almost connected into a smooth straight line from August 2010 to August 2015, but the star-rating just goes down around September 2012. So in the long term, mark of hair-dryer doesn't show a big change but a little fluctuation, which means reviews of hair-dryer hardly have any ups and downs, and consumers always give it an appropriate comment and star rating around 4. What strategy for Sunshine Company can summarize from this graph is entering the hair dryer market at any time is fine, since there no big variation of consumers reviews. Although the risk of this market is smaller than the microwave oven's, it also has weakness that companies may have no chance to acquire more benefits.

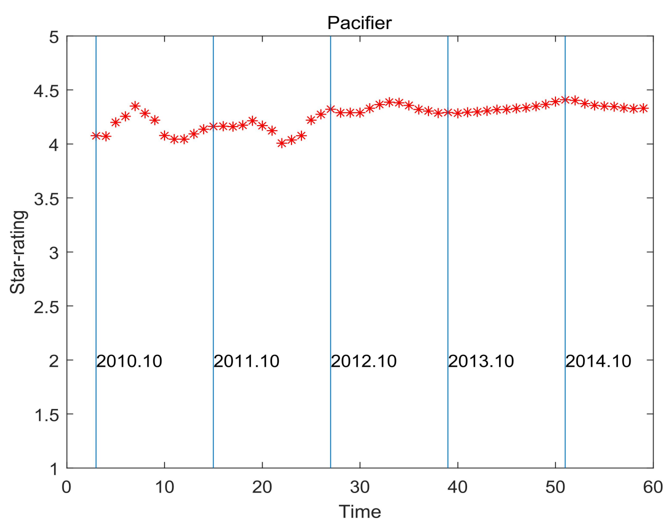


Figure 6: Scatter plot of Pacifier

Thirdly, according to the figure3, pacifier market is even smoother than hair-dryer market meaning purchasers hardly give extreme reviews and star rating. Nonetheless, the average star-rating of pacifier is higher than hair-dryer's staying close to 4.5 , showing entering pacifier market will get better consumers' responses and more numerous benefits. In this case , selecting hair-dryer market not as good as pacifier market to keep high yield.

In general, if Sunshine Company desires to obtain huge benefits, entering and being stationed in microwave oven market is a splendid choice but combining high threats. If the company want to choose a stable market , then hair-dryer market and pacifier market are both decent ,and pacifier market is more worth considering.

8 Consumer Dissect

8.1 Product Portent

Text and star rating of reviews that are measures directly represent consumers' taste can be used as one efficient norm. We extract the adjective description of product indicating the possibility of successful or failing product, and put adjective verbs into combination of star rating saying what product gets notice once it enters the market.

In order to explore relation and combinations, we build comprehensive scoring index analysis model to analysis. Having gained English Emotional Corpus[**?** , ref3]ontaining almost all emotional verbs helps to filtrate available text to calculate score of reviews.

8.1.1 Text Score

We set a system to estimate score of review texts by giving emotional verbs of review text weights judging the text's emotional trend of the reviewer which is positive or negative. Negative words are the same.

Above all, setting every word to evaluate a base unit from 1 to 5 (include 1 and 5), for example, "pretty" is a positive adjective and we give it 3 units as the base unit, whenever there is one pretty in the review text, the text will get 3 base units. Then considering adverbs around the adjective in sentences which will strengthen mood of texts, through multiplying an artificial coefficient of adverbs by the base unit of the adjective and get final point of one sentence in the text, and the coefficients of adverbs have 5 levels (including 0.5, 1.5, 2, -0.5, -1.5, -2), which means, for instance, one sentence including phrase "extremely pretty", the coefficient of "extremely" is 2, hence score of this sentence is 2 times 3 resulting in 6. Besides, verbs can also show the strong emotion, so giving each verb a coefficient as same as the setting of adjectives. If there is no adjective in the sentence, a middle score(3) will be given .To calculate total score of one review text, we need to get sum-score of sentences via the equation below:

$$score_{text} = 2\frac{3p}{(3p+1)}\frac{\sum\limits_{i=1}^{n}(\sum\limits_{j=1}^{t_{ai}})D_{aij}G_{ai} + \sum\limits_{i=1}^{m}(\sum\limits_{j=1}^{t_{vj}} D_{vij})G_{vi}}{\sum\limits_{i=1}^{n}(\sum\limits_{j=1}^{t_{ai}} D_{aij})|G_{ai}| + \sum\limits_{i=1}^{m}(\sum\limits_{j=1}^{t_{vj}} D_{vij})|G_{vi}|} + 3 \tag{6}$$

p is the number of all these words, n is the number of adjectives, m is the number of verbs, G_{ai} represents the base unit of the i-th adjective, G_{vi} represents the base unit of the i-th verb.

For i-th adjective, t_{ai} is the number of adverbs around the i-th adjective, and D_{aij} represents coefficient of i-th adjective's j-th adverb. For i-th verb, t_{vi} is the number of adverbs around the i-th verb, and D_{vij} represents coefficient of i-th verb's j-th adverb. And the equation part, $(3p/3p+1)$, is a correction factor, since the more text content is, the stronger his emotion is, in order to prevent the emotions calculated by the formula from being too extreme when the user evaluates too little content, we have added a correction factor.

Here we give an example to illustrate this equation:

Text content: *The heat did not last for long, just cold air blows. The dryer is stylish and the attachment does not fall off. It is sad the heat stopped within a week.*

Reviewer's star rating is 2.

What we need to extract are adjectives, verbs, adverbs to embellish adjectives and verbs.

Ensuring a word is positive or negative according to the content, words extracted are:

Adjectives: long(1) cold(-1) stylish(1) sad(-1)

Verbs: last(0) blows(0) fall off (-1) stop(-1)

Adverbs: not (embellishing "long" and "stop")(-0.5)

The total number of these words except words whose base unit is 0 is 10, the numbers in brackets represent the base unit of words, the total score of adjectives is the equation: not $\times$ long + cold + stylish + sad=-0.5-1+1-1+0.5, and the total absolute value score of adjectives is 3.5. The total score of verbs is the equation: not$\times$ fall + stop =0.5-1=-0.5, and the total absolute value score of verbs is 1.5.

Hence $score_{text}$ = [3 $\times$ 10/(3 $\times$ 10+1)] $\times$ (-0.5+0.5)/(1.5+3.5) $\times$ 2+3=3 is the result score of the text.

The value assigned to each text can be used to combine with star rating and potentially heralds success or failure of the product.

8.1.2 Comprehensive Scoring Index Analysis Model

As we have graded reviews above ,then endowing weights to score of reviews and star rating and the sum weights is 1. Dividing star rating score into three standards respectively representing high rating, middle rating and low rating, and high rating includes 4 and 5 stars, middle rating is 3 stars and low rating includes 1 and 2 stars. We also divided the review texts into three categories of low score, middle score and high score, and the low score category contains texts scored from 1 to 2.5(exclude 2.5), the middle score category contains texts scored from 2.5 to 3.5(include 2.5 and 3.5), and high score category contains texts scored from 3.5 to 5 (exclude3.5).

Then the evaluation can be divided into 9 categories according to their types: high rating & high score, middle rating & high score, low rating & high score, high rating & middle score, middle rating & middle score, low rating & middle score, high rating & low score, middle rating & low score, low rating & low score. According to the survey [1], sensitivity of consumers for negative comments is higher, so we deem details in texts are more convince. Hence giving the text score higher weights than score rating, and we could set weights for each type as below [2, 4]: The basis of weight setting is as follows:

1. For high rating & high score, middle rating & middle score, and low rating & low score, review

Type	high score high rating	high score middle rating	high score low rating	middle score high rating	middle score middle rating
Text Weight	0.55	0.65	0.85	0.65	0.55
Rating Weight	0.45	0.35	0.15	0.35	0.45
Type	middle score low rating	low score high rating	low score middle rating	low score low rating	
Text Weight	0.65	0.95	0.75	0.55	
Rating Weight	0.35	0.05	0.25	0.45	

texts are consistent with star rating, but review texts are more convinced, hence we give review texts a bit more weight.
2. When there is a contradiction between text's emotion and rating's, consumers have higher receptivity for the text, so we give text a bigger weight, such as low rating & high score. Meanwhile, negative reviews have a hugger impact on consumers, so we give 0.95 to the text weight in high rating & low score type.
3. For middle rating & high score and middle rating & low score, details will more convincing, so we give text higher weight.
Finally using $sumscore$ to express synthetic sum-score, and A is weight, then we can get equation:

$$sumscore = Arating + (1 - A)score_{text} \tag{7}$$

Deciding which product has potential success just calculates sum score of all reviewers, and if total sum-score is high then it means it's a potentially successful product.

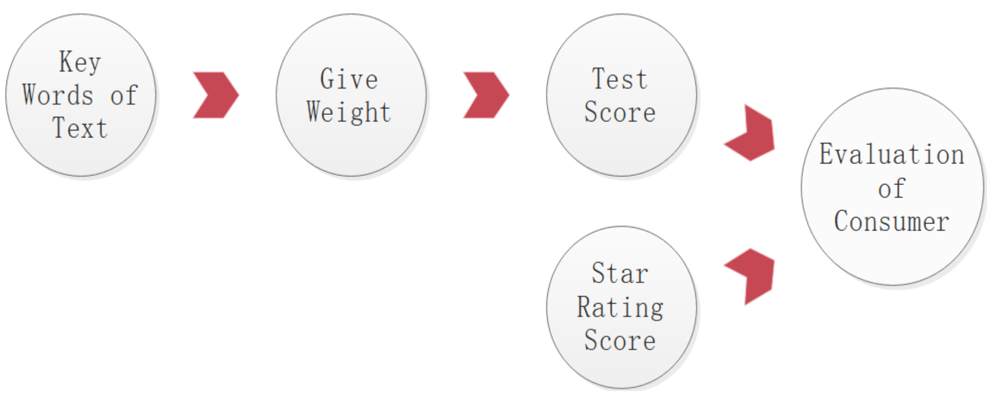


Figure 7: The Flow Chart of Reviews' Score Calculation

8.2 Star Ratings Effects

It is easy appearing copycat effect in reviews these days, so some reviewers may be affected by a series of low star rating in a period of time and also mark a low score. However this will lead to a biased result of the product's quality, so that reckoning the relation between specific star rating and reviews is a significant topic. We select stratified sampling to explore the correlation between seeing a series of low star rating and writing reviews.

Three products data sets including kinds of brands, and we assume consumers can only see and concern star rating of the brand he purchases, but not learning the reviews of the other brands of this product in markets. As not every brand gets the same attention which reflects on the number of reviews, we select several brands of each product whose reviews are more than 50, because if review number is too small, there will be a high probability of a large error, and separate brands into three sample sets. The big sample set including brands having more than 300 reviews, the middle sample set including brands having from 100 to 299 reviews, and the small sample set including brands from 50 to 99 reviews.

Considering the reality, in Amazon's product reviews websites page, there are two arrangement methods to list the comments of this product, according to the most recent time and the most helpful votes, which means time and helpful vote strongly compact on users. Now we'll combine the exploration from two aspects, time and helpful vote.

Users usually keep a watchful eye on the most recent reviews when preparing to write reviews affecting on writing reviews, and in this reason, we assume users observe 15 reviews at the time closest to the node where the comments are written and analysis the influence of these 15 reviews' star rating. Also, the more recent star rating is, the more possibility it has huge effect on watchers, since most people read reviews from top of the arrangement list and the nearest one must be read for the most time. So setting weights to these reviews' star rating according to time and make them normalized. The score gained after multiplying each of the 15 star rating scores by the normalized weight and summing them is defined as preliminary synthetic score (replacing it with PSS in the following).

We deem the abnormal extreme PSS embodies superior potential impact on users, and pick up respectively period of time from 30 brands when their PSS are at the extreme value (extreme high or extreme low), meanwhile one text score of a random user who writes reviews after seeing these extreme star rating is calculated based on last part. It's expected that after collecting and analyzing multiple sets of PSS and scores of texts the influence of specific star rating on more reviews.

8.2.1 Mapping and Analysis

Drawing these data sets on the rectangular coordinates, X-axis describes PSS numerical value and Y-axis describes the users' texts scores after seeing those reviews, and we judge whether these users have potential to be affected by star rating through observing the scatter plot distribution of data sets. In the scatter diagram above it is apparent that dots are focused on the two extremes of the graph which shows, when PSS is extremely low, the score of text also tend to be extremely low or high, and when PSS is extremely high, most score of text tend to be high. Obviously, extreme PSS has a strong effect on the texts' score of reviewer having seen a series of specific star rating. For the peculiar case that some users' text score present absolutely inverse consequence, for instance, when seeing a series of low star rating, they write texts containing high praise, we suspect this is due to the users' rebuttal psychology. What should be attentive is when encountering extreme PSS conditions, texts' scores almost never at the average position in this plot, from which we deem it's a battery of specific star rating make an impression on readers reflecting on content of review texts, and pull the customer's evaluation from the average value to the extreme value.

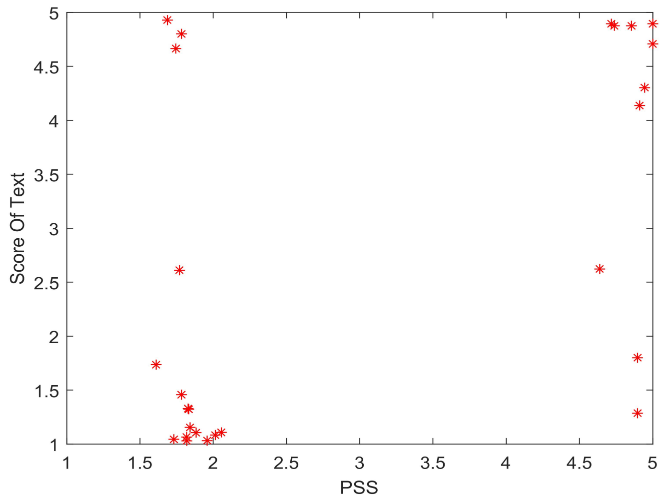


Figure 8: The Flow Chart of Reviews' Score Calculation

8.3 Connection between Star Rating and Manifestations

From the research above, it's obviously that consumers' star rating sometimes doesn't has close connection to reviews as we think of, and the condition one review equips with two extreme manifestations, such as high rating & low score type. To explore whether specific quality descriptors have relation with rating levels, we randomly select 45 descriptors from the English Emotional Corpus, such as terrible, and make a thorough inquiry connection.

In the first place, count how many review texts including these descriptors among three products' data sets and get their star ratings to figure up the average star rating. Then we put the average star rating of one product's total reviews into account, and contrast average star rating of samples with the total average to understand how does the relationship work. Here is the chart of two average of diverse descriptors in three products' markets. (entire chart is put at the appendix)

Specific Descriptor	Product	Number of Texts	Samples' Average	Total Average	Relative Error
garbage	hair-dryer	17	1.1176	4.1160	72.85%
garbage	microwave	10	1	3.4446	70.97%
junk	hair-dryer	50	1.56	4.116	62.10%
junk	microwave	20	1.4	3.4446	59.36%
horrible	microwave	13	1.6153	3.4446	53.11%
useless	pacifier	87	2.4023	4.3032	44.17%
horrible	hair-dryer	42	2.3333	4.1160	43.31%
useless	hair-dryer	41	2.3415	4.1160	43.11%
disappointed	pacifier	283	2.4841	4.3032	42.27%
junk	pacifier	29	2.759	4.303	35.88%
garbage	pacifier	17	2.6471	4.1160	35.69%
disappointed	hair-dryer	242	2.6570	4.1160	35.45%
useless	microwave	8	2.2500	3.4446	34.68%
disappointed	microwave	24	2.2500	3.4446	34.68%

The chart above is selected including only highly connected specific quality descriptors, in which we can learn not every quality descriptor is associated with rating levels. However, it's naturally to figure out that all these quality descriptors are negative. Therefore, we draw a conclusion that negative emotion words are more strongly correlated with ratings.

9 Conclusion and Discussion

The main missions we attempt to solve are creating online sales strategy and identify potentially important design features that would enhance product desirability for Sunshine Company. For this reason, we analysis from three most closely related to product sales factors: product, market and consumer according to reviews which is regarded as an important basis throughout the article.

To search what features one ideal hair dryer (or a microwave oven or a baby pacifier) need to be equipped with, we use weight allocation system and regression linear analysis to explore features appearing most times that we deem purchasers care for these features. As a result, the most significant feature of hair dryers, microwave ovens and baby pacifiers are setting, appearance and price. This declares if the setting of hair dryers, appearance of microwave ovens and price of baby pacifiers are wonderful then the company has a huge possibility to success.

Market atmosphere is complicated and changeable, but we still figure out some unique discoveries through moving average method. Firstly, microwave oven market shows drastic fluctuations and there is a rule that every ten months, microwave oven market will get to a peak value but go down at the beginning of the next ten months which called time circle. However, markets of hair dryer and baby pacifier don't have these characters, as they always keep on a stable score.

Consumers' reviews containing lots of messages reflect that content of review text may indicate a potentially successful or failing product, a battery of specific star rating make an impression on readers reflecting on content of review texts and negative quality descriptors are intense associated with rating levels. We individually gain these conclusions from the Comprehensive Scoring Index Analysis Model, the flow chart of reviews' score calculation and contrasting average star rating of samples with the total average.

In general, after a series of model analysis, the following Suggestions are made: enhancing product performance in important feature areas, such as price of baby pacifiers; entering the microwave oven market at appropriate time point; finding the direction of future product optimization by calculating customer's review texts and star ratings after entering the market.

10 Strengths and Weaknesses

10.1 Strengths

- Processing and Analysis of the Data Sets: We first process the data set, remove irrelevant content, and classify by brand, time, etc.
- Low Feature and High Accurate:We point out the characteristics and commonalities of each product. In Question 2 a. we refine specific 6 out of total feature according to each product for Sunshine Company to tract.

- Effective Combination of Review Scoring,Star Rating and Helpfulness Votes: We used almost all data sets and adopted a weight calculation method to consider the authority and high helpfulness votes to be more accurate.
- Review Text Analysis Effectiveness: We have built a thesaurus and based on this we can give specific and efficient scoring of review content.
- Comprehensive Scoring Mode: We use three text scoring modes and three star scoring to establish nine scoring weights, and based on this we get a comprehensive product score to help Sunshine Company's success.
- Make Full Use of Information Retrieval Technology: In Question 2 d, considering the real situation of the user browsing the product, we use MATLAB code inspection to select the most extreme data segment among the many data. Combined with the text analysis of the user, the scatter plot drawn has a very high accuracy.

10.2 Weaknesses

- No Verification of All The Data Provided: In Question 2 d, we have only selected some typical brands' products for analysis ,but not the whole data set.
- No Consideration of Business Cycle and Alternative Products: We don't consider the fluctuations in the economic cycle leading to changes in consumer purchasing power and the changes in alternative products. The overall market outlook is also important for companies
- Lack of Emotional Judgment on the Entire Sentence. In Question 2 c, we established a method model for weighting text scores based on user sentiment keywords. Although the model has a certain degree of accuracy, it does not consider the overall expression of the user's emotions in the sentence, so there is still a certain error in the text scoring.

References

[1] MIKOLOV T,SUTSKEVER I,CHEN K ,et al.Distributed representations of words and phrases and their conpositionality[J]. Advances in Neural Information Processing Systems,2013(26):31113119.

[2] WANG Qian, FU Ku, Calculation Method of Comprehensive Score of Commodity Evaluation BasedonLSTM-AENeuralNetwork,2018.4.

[3] Retrieved March 8, 2020, from https://github.com/aesuli/SentiWordNet

[4] WANG Xiao-yun, SHI Ling-ling, A Comprehensive Scoring Model of Product Based on Emotional Quantification of Web Reviews, 2015.9.

A Appendix

Specific Descriptor	Product	Number of Texts	Samples' Average	Total Average	Relative Error
perfect	pacifier	1475	4.718	4.3032	9.64%
	hair-dryer	687	4.6288	4.1160	12.46%
	microwave	214	4.4065	3.4446	27.92%
useless	pacifier	87	2.4023	4.3032	44.17%
	hair-dryer	41	2.3415	4.1160	43.11%
	microwave	8	2.2500	3.4446	34.68%
disappointed	pacifier	283	2.4841	4.3032	42.27%
	hair-dryer	242	2.6570	4.1160	35.45%
	microwave	24	2.2500	3.4446	34.68%
wonderful	pacifier	255	4.7373	4.3032	10.09%
	hair-dryer	152	4.6513	4.1160	13.01%
	microwave	18	3.8889	3.4446	12.90%
horrible	pacifier	68	2.7059	4.3032	31.12%
	hair-dryer	42	2.3333	4.1160	43.31%
	microwave	13	1.6153	3.4446	53.11%
garbage	pacifier	17	2.6471	4.1160	35.69%
	hair-dryer	17	1.1176	4.1160	72.85%
	microwave	10	1	3.4446	70.97%
love	pacifier	6333	4.714	4.303	9.55%
	hair-dryer	2436	4.656	4.116	13.12%
	microwave	172	4.24	3.4446	23.10%
great	pacifier	3773	4.572	4.303	6.25%
	hair-dryer	2809	4.509	4.116	9.55%
	microwave	368	4.22	3.4446	22.51%
hate	pacifier	83	3.66	4.303	14.94%
	hair-dryer	139	3.75	4.116	8.89%
	microwave	24	3.167	3.4446	8.06%
happy	pacifier	654	4.494	4.303	4.44%
	hair-dryer	565	4.304	4.116	4.57%
	microwave	91	3.95	3.4446	14.67%
amazing	pacifier	214	4.757	4.303	10.55%
	hair-dryer	247	4.7	4.116	14.19%
	microwave	12	3.8333	3.4446	11.28%
excellent	pacifier	301	4.811	4.303	11.81%
	hair-dryer	336	4.75	4.116	15.40%
	microwave	42	4.333	3.4446	25.79%
cheap	pacifier	390	3.513	4.303	18.36%
	hair-dryer	442	3.568	4.116	13.31%
	microwave	68	2.559	3.4446	25.71%
junk	pacifier	29	2.759	4.303	35.88%
	hair-dryer	50	1.56	4.116	62.10%
	microwave	20	1.4	3.4446	59.36%
careful	pacifier	83	3.795	4.303	11.81%
	hair-dryer	101	3.353	4.116	15.54%
	microwave	20	3.35	3.4446	2.75%