Problem Chosen	2020 MCM/ICM Summary Sheet	Team Control Number
C		2009116

Market data tracking measurement based on sentiment analysis model and ARIMA model

Summary

Amazon offers customers the opportunity to score and rate. Other customers can use these reviews to help them make purchasing decisions. Based on information such as customer reviews, companies can make decisions that increase product appeal. If the company masters the future development trend of the product, it can provide a basis for business decisions.

As for task I, in order to accurately capture the effective information in the data set, a series of data processing methods to process the missing values, abnormal values and repeated observations in the original data set, and uses the combination of qualitative and quantitative methods to screen the unrelated variables. Besides, for making a preliminary understanding of the product market satisfaction, descriptive statistics is applied. And then this paper employs ***Lexicon-based approach*** to analyze the text data, and establishes an ***emotion score evaluation system***, which can quickly capture customer satisfaction from each comment.

The task II is divided into three subsections depending on the relevance of the issues. In subsection 1, a three-level evaluation system for star rating has been established. After using correlation test to judge the impact of comments and ratings on stars, a comprehensive evaluation system including voting information, comment content and customer identity was established by principal component analysis. On this basis, by referring to the Amazon platform crude oil star algorithm, the time span index is introduced to update the original three-level system, which makes the system more comprehensive.

For subsection 2, a ***time series model*** is introduced to study the influence of time on comment orientation. First, we visualize the time series to get the emotional trend chart of products and the time trend chart of other variables. Through image analysis, the trend of emotional score of different products is grasped, and the time trend of star rating, cumulative number of comments and cumulative number of likes of three products are obtained. Next, on the basis of the time series model, we introduce the difference ***autocorrelation moving average model (ARIMA)***, which can accurately predict the star rating situation and comment emotion trend in September and October 2015.

Finally, we discuss the influence of ***extreme reviews*** on other variables and the interaction between extreme reviews and ***extreme stars***. In the study of the impact of the follow-up comments on the number and emotional direction of extreme stars, because the data does not conform to the characteristics of normal distribution, the ***non parametric test*** becomes the main method of this paper. The influence of extreme emotion words in the comments on the follow-up stars is analyzed in two directions: positive and negative.

In summary, the empirical analysis and models about three productions developed in this paper are also a powerful tool for other electronic business platform, showing an excellent performance.

Keywords: ***Lexicon-based approach emotion score evaluation system principal component analysis time series model ARIMA non parametric test***

A Letter to the Marketing Director of the Sunshine company

Dear Sir or Madam,

According to your online sales needs, our team provides some suggestions for you to solve the related problem. We propose the mathematical model to analyze the sentiment polarity and figure out the relationship between three kinds of customer feedback. We also consider how the time factor affect the rating changes and offer a simplistic model can be used to forecast rating and sentiment trend in the short term. And at last we form a COM_S parameter to define the recommended products.

More specifically, Lexicon-based approach which is based on the sentiment dictionary is used to make the text sentiment analysis in the customer-supplied review text. And an important sentiment score is developed by the approach to find out the internal relationship between review and rating variables. Then, we establish two parameters Customer Reviews (CR) and Customer Satisfaction (CS) in order to help present the specific relationship among three variables. One of these two parameters, Customer Reviews, is analyzed by the principal component analysis method. Three principal components respectively are voting information, review content and the customer. The other one of two parameter, Customer Satisfaction (CS) can also be determined by a detailed calculation formula. Finally, Comprehensive Satisfaction (COM_S) can be derived from the CS summary of the same product.

And when we consider the time factor, the time series model can be utilized to depict the trend of different variables. Moreover, we also focus on how the extreme sentiment which is manifested as one star or five stars changes itself or affects other rating variables. The result shows the extreme attitude would influence other customers, especially in a long period.

Then, we give some advices on the product current problems which is identified as the important points to modify in the future. In the process of analyzing the reviews of each products, some common product problems need paying more attention to. For microwave oven, you should pay attention to the performance and price of the product itself, and make timely improvements to the product. For the hair dryer, focus more on product performance and appearance. For pacifier, the evaluation is generally high, and the company should pay attention to the issue of product working life.We believe that our method can help your company select most possibly successful products and avoid the risky products. Followings are the recommended list:

	Successful	**Unsuccessful**
Pacifier	cosco alpha, Fisher, ameda, and Sealy	chicco keyfit, ikea kladd, and batman
Microwave	sharp, lg, magic chef, whirlpool, profile, and Panda Small	Samsung, Danby, and jx7227sfss
Hair dryer	Salon, Waterpik, Micro Tweeze, DuWop, ISLAND HAWAII, Oral-B	Andis, conair, xtava, mhd, remington

Yours Sincerely,

Team#2009116

Contents

1 Introduction

1.1 Statement of the problem

Online marketplace where Amazon takes the leading place has become the most important platform to trade for most companies. As the mechanism for customers to communicate shopping experience with each other, various kinds of reviews and ratings provide potential consumers with subjective evidence. It can also provide the guidance for new entrants in choosing a target market and target product. In addition, in the social media era, consumers have louder and stronger voices than ever before, which is shown as over 50% of consumers frequently factor in online reviews before buying a product.[1] Because this process can be seen as a direct link to the former customer, getting the first-hand information of product in order to avoid the exaggeration or of the seller.

However, at this stage, it is not clear that their interaction, and how they influence the corporate strategy. Thus, we aim to explore the internal relationship of three variables and design a composite data measure to assist companies in the selection of potential successful products. Meanwhile, we discuss how the customer group reviews influence on individual review attitude orientation. At last, we rank the alternative products according to our method and offer a recommendation list to Sunshine company.

1.2 Literature Review

At present, there are mainly two methods for sentiment analysis, one is based on machine learning, and the other is based on the lexicon. Several approaches based on the machine learning have come into the spotlight in recent years and most of them are based on the big amount of personalized reviews from social media. Due to the reviews in the social media are not only English, but other languages such as Arabic. [2] Many researchers from non-English regions also apply this method to solve the issue, showing that this method is quite mature.

And the other method is the Lexicon based approach. This method also utilizes the text which expresses people’s sentiment or emotion in their social media, such as the Twitter posts and Micro-blog text. [3,4] Latest researches on the sentiment polarity lexicon also show the resource of the lexicon is extended to a particular domain, the stock market, without human intervention and addressing the scaling and thresholding problem. [5] Therefore, we believe that this method is potential to be widely used.

After we compared the two methods, we determine to choose the Lexicon-based approach. We believe that this method is most applicable when time and resources are limited, that is, similar to the conclusions obtained in the study of Urdu's Sentiment Analysis. [6]

After introducing the time factor, we propose a time series model to discuss the changes of three variables in the time dimension. After comparing several applicable models, we choose to use the traditional time series models and the Autoregressive Integrated Moving Average (ARIMA) model to observe the influence of time factors on different reviews and to predict the future indicator trend. We think that during the forecasting phase, the ARIMA model can achieve a quite accurate forecasting as theories of Matyjaszek, M. et al. [7] At the same time, we also conducted product forecasting of different products by product. This is in line with Nguyen, H. et al. The model's predictions for different product prices are similar [8].

1.3 Overview of Our Work

Our main goals are to address three issues that need to be explored: (1) the sentiment analysis of text reviews; (2) the relationship between reviews and ratings indicators and (3) the impact of time. Then we propose the product sequencing method which these three contents combined with.

In order to solve these problems, we first analyze the data characteristics of each digital parameter, and then apply the sentiment lexicon-based method to analyze the emotional attitude of the text of the text review, extract the emotional words in the

sentence through the inner join, and then calculate the sentence sentiment scores. This process of sentiment scoring is used to distinguish between the positive and negative attitudes of a sentence. Then by comparing text reviews and two rating variables, we get Customer Review (CR) to discuss how star ratings are affected by review factors, and then we discuss the influencing factors of Customer Satisfaction (CS) and draw a quantitative relationship. After that, we formed Comprehensive Satisfaction (COM_S) according to the product satisfaction based on the satisfaction of each customer. And we use this data measure to screen our candidate products and provide Sunshine Company with a list of alternative product suitable for online sales.

1.4 Notations

Symbol	Definition
P	Pacifier
H	Hair dryer
M	Microwave
hv	helpful_votes
tv	total_votes
sq	SentimentQDAP
sr	star_rating
n	number
v	vine
vp	verified_purchase
ci	customer_id
pp	product_parent
D	day
F_1	First principal component
F_2	Second principal component
F_3	Third principal component
CR	Customer Review
CS	Customer Satisfaction
COM_S	Comprehensive Satisfaction

1.5 Assumptions

- We assume that past trends of things will extend to the future. The reality of things is the result of historical development, and the future of things is an extension of reality. The past and future of things are connected.
- We assume that the data which forecast based on is irregular.
- We assume that there is no causality between market developments.
- We assume that there are no other relevant random variable changes that affect the time series analysis.
- We assume that the variables contained in the given data set can fully reflect the comprehensive score of the product.
- We assume that the dictionary selected by Lexicon-based method sentiment analysis contains all the required words.
- We assume when there are multiple sentences in a customer review, the average of the sentiment scores of all sentences can represent the emotional score of the customer.
- We assume that customer reviews are simple logical relationship without the use of derogatory words to indicate positive meanings, nor the complex combination of negative words, degree adverbs, and emotional words.
- We assume that the definition of affective words in the affective dictionary is not controversial.

2.Task 1

2.1 Problem Analysis

Because of the large amount of data provided by this topic, and mostly in the form of text, the focus of this topic is the preprocessing of original data and the extraction of

text information. In the processing of solving task 1, this paper first uses **a series of data processing methods** to process the missing values, abnormal values and repeated observations in the original data set, and uses t**he combination of qualitative and quantitative methods** to screen the unrelated variables. This step can not only remove the interference of miscellaneous information, but also accurately capture the effective information in the data set. Next, this paper makes **descriptive statistics** on the different variables of three kinds of commodities to make a preliminary understanding of the product market satisfaction. Finally, this paper uses the **Lexicon-based approach** to analyze the text data, and establishes an emotion score evaluation system, which can quickly capture customer satisfaction from each comment.

2.2 Methodology

The section is parted into three subsections, data preprocessing, descriptive statistics and Lexicon-based approach. In these subsections, the ideology of modelling in task 1 are introduced.

2.2.1 Data preprocessing

Due to the huge amount of data in the design of this problem, in order to remove the influence of interference information and extract the effective information in the data, the pretreatment of the original data is a crucial step in solving the problem. The data preprocessing in this paper can be divided into the following steps:

Step 1. Check for duplicate observations in the original data set
Step 2. Handling of missing value and abnormal value
Step 3. Standardization of observation data
Step 4. Delete useless variables

2.2.2 Lexicon-based approach

At present, there are two methods to solve the problem of emotion analysis: one is based on machine learning, the other is based on emotion dictionary.

For machine learning based methods, a large number of manually labeled corpus is input as training set, and emotion classification is realized by extracting text features and constructing classifier. It's difficult for us to do this because of the limitation of our personal computers and the length of the contest。 At the same time, because we are not clear about the internal relationship between the known classification and the comment text, this method is not suitable for the solution of this problem.

Therefore, we choose the method based on emotion dictionary, which only uses the known text and emotion dictionary first, and does not involve other rating indicators. The three dictionaries cited in this paper are:

(1) Dictionary with opinionated words from the Harvard-IV dictionary asused in the General Inquirer software

Dictionary with a list of positive and negative words according to the psychological Harvard-IV dictionary as used in the General Inquirer software. This is a general-purpose dictionary developed by the Harvard University.

(2) Dictionary with opinionated words from Henry's Financial dictionary

Dictionary with a list of positive and negative words according to the Henry's finance-specific dictionary. This dictionary was first presented in the Journal of Business Communication among one of the early adopters of text analysis in the finance discipline.

(3) Dictionary with opinionated words from Loughran-McDonald Financial dictionary

Dictionary with a list of positive, negative and uncertainty words according to the Loughran-McDonald finance-specific dictionary. This dictionary was first presented in the Journal of Finance and has been widely used in the finance domain ever since.

General steps are shown as following. First, by collecting the related text information, they created the specific dataset with text, user, emotion, sentiment information. Then, the sentiment dictionary can be extended by extraction and construction of a series of related dictionaries. Next, the sentiment value of a text can be obtained through the calculation of the weight. Eventually, the researchers got influence score to achieve their final goal.

2.3 Problem Solving

For the convenience of follow-up work, copy the contents of original hair dryer.tsv, microwave.tsv and pacifier.tsv files to excel and save them as CSV files. Data set pacifier.csv contains 18939 observations and 16 variables; microwave.csv contains 1615 observations and 16 variables; hair_dryer.csv contains 11470 observations and 16 variables. The 16 variables in the three files are the same, and the variables are divided into character type variables and numerical type variables. Among the variables, star rating, helpful voices and review body are the most important research objects; review ID is the comment number, which is the identification variable and unique. Variables also include product information, market information, customer information and comment information.

2.3.1 Detail of data preprocessing

Step 1. Check for duplicate observations in the original data set

Import the file into the SAS system, use the nodupkey option in the proc sort process to check whether there are duplicate observations in the data set, and use the review [ID] as the identification variable. If there are duplicate observations, store them in the work.dup data set. The inspection results of the three files show that there is no observation in the data set work.dup, so it shows that there is no repeated observation in the original data set.

Step 2. Handling of missing value and abnormal value

(1) Missing values

In Excel, you can simply check the missing value and fill in the missing value. The product "category" in pacifier.csv is missing. Pacifier is a baby product and can be filled directly. In addition, 8 observations with missing important information were found. These eight observations were deleted, leaving 18931 observations. There are 7 missing values in the review date variable of hair-dryer.csv, which can be supplemented manually. There is no missing value for microwave.csv.

(2) Abnormal values

First, we use the proc contents process to generate the variable description table work.var, then we use the proc SQL process to define the character type variable as the macro variable var_char, and the numerical type variable as the macro variable var_num, In order to call character type variable and numerical type variable more efficiently in the later process; then use the nmiss option of proc means process to check the missing value in numerical type variable, use the missing option of tables statement in proc freq process to check the missing value in character type variable, and check the frequency distribution of character type variable. The results show that there are no missing values in character type variables, 14 missing values in review [data] in packer.csv file and no missing values in other variables. There are no obvious error values. Delete 14 observations with missing values in the pacifier file, 18917 observations and 1615 observations in the microwave file, including 11470 observations.

Step 3. Standardization of observation data

Amazon is a platform with a strong freedom of comment, so there are a lot of junk comments that need to be removed. According to the observation, there are many junk comments or useless comments in the comments, and we need to filter them out. In addition, we need to analyze the emotion and attitude through the comment content, so useless comments with too short length also need to be removed. In order to facilitate the follow-up work, it is necessary to standardize the comment text.

(1) Delete too short comment

Use the length statement to check the string length of the review body observed in each of the three files. If the string length is less than 20, delete the comment. There are 1674 observation review bodies in the pacifier file with a string length less than 20, 738 observation review bodies in the hair dryer file with a string length less than 20, and 79 observation review bodies in the microwave file with a string length less than 20.

(2) Standardize data.

In order to facilitate the subsequent extraction of important information in review body, it is necessary to delete the punctuation in the review, modify all uppercase to lowercase, use tranwrd function to delete punctuation, and use translate function to

modify uppercase to lowercase in the review. In addition, use the translate function to modify the upper case observations of the variable vine and the variable verified_.

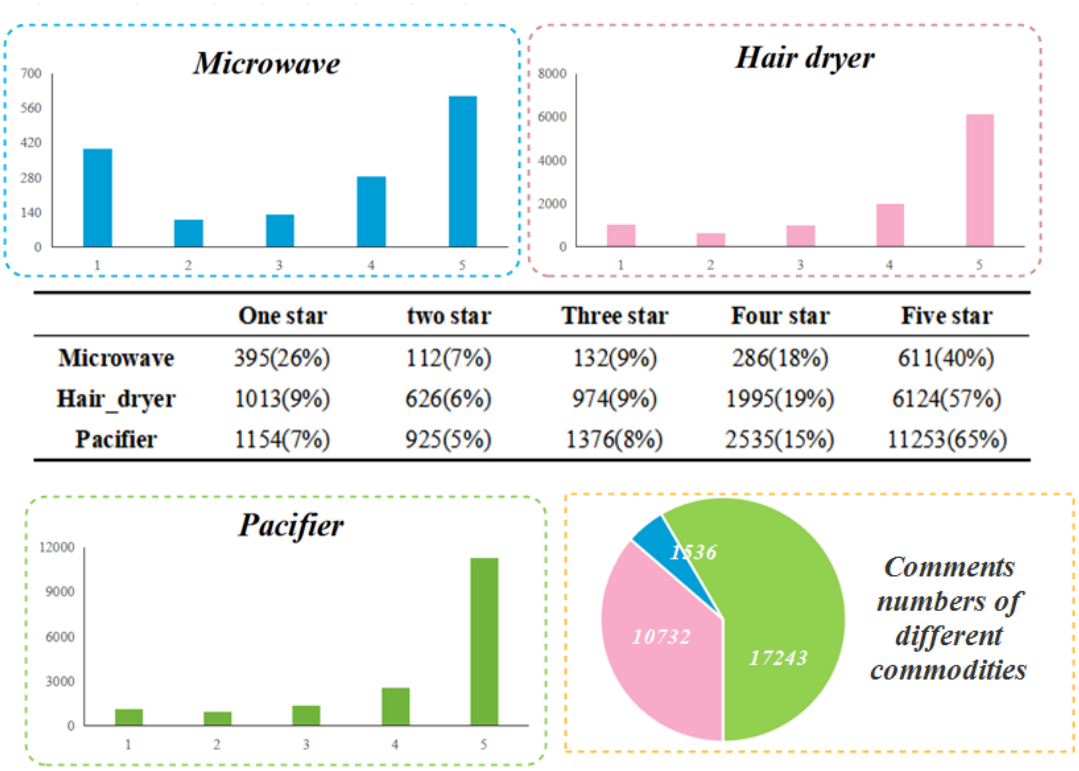
Step 4. Delete useless variables

There are some useless variables in the file that can be deleted. The three files market place are the same, which can be deleted. Product ID, product title and product parent can uniquely identify the product. Just keep one of the three, and keep product parent. Product category and review headline are useless variables and can be deleted. Use drop statement to delete useless variables.

Finally, 1674 observations are deleted from the pacifier file, 17243 observations and 10 variables are reserved. 738 observations were deleted from the hair guy file, 10732 observations and 10 variables were reserved. In the microwave file, 79 observations are deleted, 1536 observations and 10 variables are reserved.

2.3.2 Results of descriptive statistics

(1) Star_ratings



	One star	two star	Three star	Four star	Five star
Microwave	395(26%)	112(7%)	132(9%)	286(18%)	611(40%)
Hair_dryer	1013(9%)	626(6%)	974(9%)	1995(19%)	6124(57%)
Pacifier	1154(7%)	925(5%)	1376(8%)	2535(15%)	11253(65%)

Figure 1 The satr_rating of product analysis chart

As shown in the table above, it is the star ratings frequency table and proportion of three products. According to the table data, we can preliminarily know that five-star rating accounts for the largest proportion of the three products, and two-star rating accounts for the smallest proportion. For microwave, star ratings are concentrated on one and five stars, while star ratings of hair player and pacifier are concentrated on five stars, and the proportion of other stars is very low.

The above figure is the star ratings frequency distribution diagram of the three products. According to the distribution diagram, one conclusion can be drawn that people buyers of pacifier and hair player are more likely to give favorable comments. Pacifier and hair player are relatively successful products made by Amazon. However, buyers of microwave are more likely to give a low rating, indicating that Amazon's microwave is not very successful.

(2) Reviews

Customers' comments on products directly reflect the success of products. According to product reviews, we can explore customers' satisfaction with the use of the product, so as to find the popularity of the product and the existing problems. According to the top 50 words in the comments, we drew a word cloud map. As shown

in the figure below, it is a word cloud map drawn according to the reviews of three products.

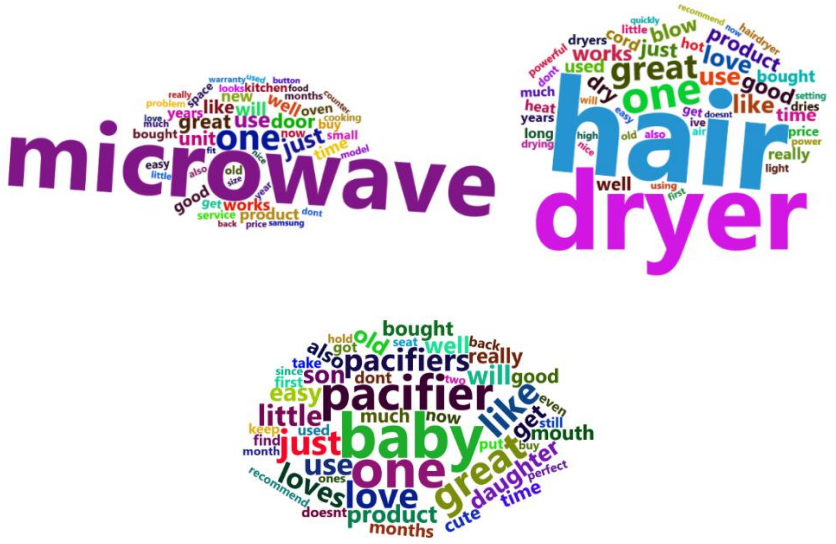


Figure 2 Word cloud diagram based on three product reviews

For microave, the products discussed by customers are button, space, month, price, year, size, works, counter, small. The words reflecting customers' emotional attitude include great, well, problem, good, small, sampling, o ld, etc. It shows that customers' comments on microwave are good or bad. Sunshine company should pay attention to the performance and price of the product itself, and make timely improvement on the product.

Customers talk more about hair dryer products themselves, such as cord, long, setting, hot, air, power, price, nice, etc. The words that reflect customers' emotional attitude include great, love, right, like, dry, etc. It shows that the customer's evaluation of hair dryer is generally better. Sunshine company should pay attention to product performance and appearance, and find out the factors of product success.

For pacifier, the product itself is often discussed by customers in terms of month, time and cute. The words that reflect customers' emotional attitude include great, love, etc. It shows that customers generally have a high evaluation of pacifier, and sunshine company has done well in the baby market. The company should pay attention to the service life of the products.

(3) Helpfulness ratings

Table 1 General analysis table of helpfulness ratings

product	helpful_votes	total_votes	ratio
microwave	8996	10635	0.846
hair dryer	24938	29202	0.854
pacifier	15244	20708	0.736

Table 2 Relationship table of star_ratings and helpfulness ratings

ratings	helpful_votes	total_votes	helpful/total
1	9418	12798	0.735896234
2	3052	4284	0.712418301
3	4362	5778	0.754932503
4	7948	9539	0.833211028
5	24398	28146	0.866837206

As shown in the table above, the total number of useful votes, the total number of votes and the proportion of useful votes in the total votes of the sub products are depicted. According to the statistical results, the proportion of effective reviews of hair dryer customers is the largest, while that of pacifier is the smallest.

In addition, we think about whether there is a relationship between star ratings and helpfulness ratings. Therefore, we calculated the percentage of useful comments by star rating.

Table 3 Analysis table of helpful reviews proportion by star ratings

proportion	1	2	3	4	5
pacifier	0.61	0.66	0.65	0.78	0.805
microwave	0.806	0.70	0.80	0.82	0.91
hair dryer	0.78	0.78	0.801	0.88	0.90

According to the analysis of the effective voting percentage of the stars, for pacifier and hair dryer, it can provide more effective information to give high scoring customer reviews. However, for microave, whether it's high score or low score customer reviews, it can provide more effective information.

2.3.3 Emotional analysis of reviews

Step 1. Divide sentences

Because the comments and other variables in the design of this topic all exist in the text situation, the ontology first divides the text to be analyzed according to the grammar. In this paper, inner join is used to extract the emotional words in the text

Step 2. Classification of emotional words

According to the emotional dictionary, the words extracted in the previous step are classified into commendatory words list, derogatory words list, neuter words list, negative words and degree adverbs, and then the emotional classifiers list and its emotional intensity are obtained,

Step 3. Remove function words

Function words refer to a kind of words that have no actual meaning. Compared with other words, function words are frequently used, such as' the ',' is', 'at', 'which', 'on', etc. In addition, there is also a kind of words, such as the word "want". In search engines, such words do not help search results much. Their emergence is not only difficult to help users narrow the scope of search, but also reduce search efficiency. Therefore, this paper removes these words from the problem, so as to improve the efficiency of text utilization.

Step 4. Calculate the sentiment score of each review

For example, the review title of the review_id in hair_dryer is R100TKSFK51G4K and the title is: "great blow dryer great price and very light weight". The number of active words after segmentation and removal of function words is 7, the negative emotion score is 0, the positive emotion score is 0.2857, and the final emotional score is 0.2857.

3. Task 2

3.1 Problem Analysis

In order to explore the impact of reviews and ratings on the Customer Reviews （CR）, we sorted out seven variables and tested their correlations. Due to the correlation between the variables, we use the ***principal component method*** to divide the variables of the period into three principal components. Based on this, we obtain ***the calculation formula for Customer Reviews*** (CR). Combining CR and review time, we ***calculate the Customer Satisfaction*** (CS) for each review and standardize CS. Then, according to the product classification, the ***Comprehensive Satisfaction*** （COM_S） of the buyer for the specific product can be obtained. Finally, we sorted each product of pacifier, hair dryer, and Microwave oven according to the comprehensive satisfaction, and ***analyzed the information of successful products and information of failed products***.

3.2 Methodology

Based on the study of the above models, we introduce a time series model to study the influence of time factors on comment orientation.

(1) Autoregressive model: An autoregressive model is a model that describes the relationship between current values and historical values. It is a method of predicting itself using the historical event data of the variable itself.

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-i} + \epsilon_t$$

y_t is the current value; μ is the constant term; p is the order; γ_i is the autocorrelation coefficient, and ϵ_t is the error value

(2) Integrated: The most important part of the ARIMA model is the smoothness of the time series data. Stationarity is the requirement that the fitted curve obtained through the sample time series can continue inertia along the existing morphology in the short time in the future, that is, the mean and variance of the data should not theoretically change too much.

(3) Moving average: The moving average model focuses on the accumulation of error terms in the autoregressive model. It can effectively eliminate random fluctuations in predictions.

$$y_t = \mu + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-i} + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

3.3 Problem Solving

3.2.1 Determination of star rating index CR

The number of characters in review_body that contain emotional information and comments. We choose the method based on sentiment lexicon to extract emotion words by using inner join, and calculate the emotion score of each text according to the score of emotion words. SentimentQDAP is used to express emotion score. In addition, the number of characters in the comments can also reflect the star rating results of customers. The number of characters for a comment is number.

Step 1. Impact indicators of star ratings

In the discussion of the final result of star rating, if the time factor is not considered first, only the impact of comments and ratings on star rating is discussed. We call this star rating indicator customer reviews. The main factors influencing customer reviews are helpful_votes (hv), total_votes (tv), sentimentqdap (sq), star_rating (sr), number (n), vine (v), verified_purchase (vp), customer_id (ci) and product_parent (pp). There may be correlation between variables, so we test the correlation of variables first. If the correlation coefficient is greater than or equal to 0.3, we think that there is a good linear correlation between variables. According to the correlation test, we found that correlation exists in some of the variables.

Table 1 Test of correlation

	star_rating	helpful_votes	total_votes	vine	verified_purchase	SentimentQDAP	number
star_rating	1.000	-.048	-.069	.020	.142	.335	-.132
helpful_votes	-.048	1.000	.994	.040	-.093	-.062	.153
total_votes	-.069	.994	1.000	.042	-.100	-.070	.163
vine	.020	.040	.042	1.000	-.244	-.028	.117
verified_purchase	.142	-.093	-.100	-.244	1.000	.117	-.273
SentimentQDAP	.335	-.062	-.070	-.028	.117	1.000	-.322
number	-.132	.153	.163	.117	-.273	-.322	1.000

Step 2. Principal component analysis

Principal component analysis is to use the idea of dimensionality reduction to explain most of the variables in the original data with fewer variables, and transform many variables we have with high correlation into independent or uncorrelated variables. It will usually select several new variables that can explain most of the

variables, namely the so-called principal components, and use them to explain the comprehensive indicators of the data.

- Justification of Our Approach

According to the results of Bartlett's test, the p value of Bartlett's test is less than 0.001, and the hypothesis of zero is rejected, that is to say, the research data can be extracted by principal components.

Table 2 KMO and Bartlett Test

Kaiser-Meyer-Olkin test sampling adequacy		**.524**
Bartlett Sphericity Test	Chi-square value	145929.849
	df	21
	Significance	.000

- Dividing principal components

Previous scholars also believed that the extracted principal components should account for 60-70% of data variation. The results of principal component extraction showed that the eigenvalues of the first three principal components are greater than 1, which account for 31.025%, 22.117% and 16.597% of the total data variation, and 69.738% of the total data. The number of principal components extracted is 3.

Table 3 Illustrated Variation Statistics

	Initial eigenvalue			Intercept and load square			Loop and load square		
	Total	Vara-tion%	Cumu-lative%	Total	Vara-tion%	Cumu-lative%	Total	Vara-tion%	Cumu-lative%
1	2.172	31.025	31.025	2.172	31.025	31.025	2.006	28.663	28.663
2	1.548	22.117	53.141	1.548	22.117	53.141	1.518	21.687	50.350
3	1.162	16.597	69.738	1.162	16.597	69.738	1.357	19.389	69.738
4	0.813	11.620	81.359						
5	0.742	10.605	91.964						
6	0.557	7.959	99.923						
7	0.005	0.077	100.000						

- Judgment of interpretation ability

The extracted principal components should have certain significance, that is, the ability to interpret the research content. The ability of interpretation suggests that the first three principal components are in line with the actual needs of the study. The ability of each principal component to interpret the corresponding variables (the data with correlation coefficient less than 0.3 has been eliminated), as shown in the following table:

Interpretation ability of corresponding variables of principal components

Table 4 Interpretation ability of corresponding variables of principal components

	1	2	3
helpful_votes	.996		
total_votes	.995		
SentimentQDAP		.801	
star_rating		.743	
number		.503	
vine			.786
verified_purchase			.726

It can be seen from the table above that the variable information interpreted by the first three principal components is basically the same as that of the classification. The first principal components (F1) are helpful_votes and total_votes. The second principal components (F2) are SentimentQDAP, star_rating and number. The third principal component (F3) is vine, verified purchase and number. Correspondingly, the first principal component mainly reflects the voting information of the comment, the second principal component mainly reflects the content of the comment, and the third principal

component mainly reflects the customer identity information. The extraction of the first three principal components has a good result interpretation ability.

Based on this, we can deduce the basic expressions of three principal components:

$$F_1 = 0.996 * hv + 0.995 * tv$$

$$F_2 = 0.801 * sq + 0.743 * sr + 0.503 * n$$

$$F_3 = 0.786 * v + 0.726 * vp$$

Step 3. Data element of Customer Reviews

According to the results of principal component analysis, the first principal component reflects the voting information of the comment, the second principal component reflects the content of the comment, and the third principal component reflects the customer identity information. It shows that the voting information is the main factor determining the final score of customer reviews, followed by the content of reviews, and finally the customer identity. The level of customer reviews directly reflects the level of customer satisfaction with products. The customer reviews (CR) score can be obtained by giving different weights to different variables. The CR calculation formula is as follows:

$$CR = \alpha_1 * F_1 + \alpha_2 * F_2 + \alpha_3 * F3$$

3.2.2 Analysis of Comprehensive Satisfaction

Step 1. Customer satisfaction Influence factor

As the most experienced e-commerce company, Amazon will surely fully understand the psychology of customers, and in particular will extract the reviews that have been recognized by the most customers. It is understood that Amazon has a specialized intelligent machine (machine learned model) that uses a complex star algorithm to give products a star rating. The parameters with the highest weight in the star algorithm are:

(1) The date of the review. The longer the date is, the more valuable it is;
(2) Helpful clicks. When a customer clicks "helpful", it means that this review is useful.
(3) Whether the review comes from Verified Purchases (VP).

In addition to the above three parameters with larger weights of Amazon rolls, the star algorithm also refers to other indicators with smaller weights:

(1) The number of characters in the review, that is, the length of the sentence;
(2) Review original score, that is, Review ratings given by each buyer;
(3) The contents of the Review. When building the model, we chose a sentiment dictionary-based method to extract the sentiment words by using inner join, and calculated the sentiment score of each review according to the scores of the sentiment words.

In the above, we used the principal component analysis method to divide the factors that determine the Customer Reviews (CR) score into three categories: vote information for reviews, content of reviews, and customer identity information.

The three data sets were reviewed from 03/02/2002 to 08/31/2015. We select 01/01/2016 as the standard time and use the intck function of SAS software to calculate the number of days between the comment time and 01/01/2016. The number of days in this interval is the "age" of the Review. The longer the "age" is, the more valuable the comment is. The interval between days is represented by day (D).

Step 2. Calculation of Customer satisfaction

According to analysis, we can introduce the factors that affect customer satisfaction (CS): helpful_votes (hv), total_votes (tv), SentimentQDAP (sq), star_rating (sr), number (n), vine (v), verified_purchase (Vp) and day (D). So the basic expression of Customer satisfaction (CS) is:

$$CS = \alpha * D + \alpha_1 * F_1 + \alpha_2 * F_2 + \alpha_3 * F3$$

Substituting the above formulas (1), (2), (3) into

$$CS = 0.3 * D + 0.3 * (0.996 * \text{hv} + 0.995 * \text{tv}) + 0.1 * (0.801 * \text{sq} + 0.743 * \text{sr} + 0.503 * \text{n}) + 0.3 * (0.786 * \text{v} + 0.726 * \text{vp})$$

$$CS = 0.3 * D + 0.2988 * \text{hv} + 0.2985 * \text{tv} + 0.0801 * \text{sq} + 0.0743 * \text{sr} + 0.0503 * \text{n} + 0.2358 * \text{v} + 0.2178 * \text{vp}$$

We use 0-1MinMax normalization, and the Customer satisfaction score is scaled to the [0,1] interval. The normalized Customer satisfaction is represented by CS *.

Step 3. Analysis of Comprehensive Satisfaction

After obtaining the satisfaction of each customer, the overall satisfaction of the product buyer can be obtained based on the product classification. Comprehensive Satisfaction (COM_S) is equal to the average value of buyer satisfaction for the same product,

$$\text{COM\_S} = \sum_{1}^{n} CS*_i/n$$

According to the analysis of the comprehensive product score, the highest product score of Pacifier is 0.713, and the lowest score is 0. Microwave oven's overall product score is highest at 0.533 and lowest at 0.01. The overall score of the hair dryer is 0.557, and the lowest is 0.013. Words show that for the three products of pacifier, Microwave and hair dryer, the company's best product is pacifier.

Step 4. Analysis of the successful products

Analyze the top ten products of Comprehensive Satisfaction of the three products respectively. We found that parafierde has a long history of successful product reviews. The number of review characters is around 150 words. Positive reviews account for a large proportion of valid votes. Star ratings are concentrated at 3-5. Customers are hardly members of verified_purchase or Amazon Vine. . Microwave has a long history of successful product reviews. The number of review characters is around 100 words. The proportion of negative reviews has increased, the number of valid votes has decreased, and the star rating has no centralized trend. Most buyers are verified_purchase, and there are almost no Amazon Vine members. Hair dryer has a long history, the number of comment characters is about 50 words, the proportion of positive comments is large, the number of valid votes is very small, the star rating is concentrated on 3-5, half of the buyers are verified_purchase, almost no.

This shows that there are more positive emotional reviews of successful products in pacifier and hair dryer, and the star rating given by customers is also high. Microwave has both positive and negative sentiment reviews, and star ratings are scattered. This is in line with the conclusions we initially obtained when analyzing stars and valid votes.

For Microwave's successful products, there are more buyers with verified_purchase identity, followed by hair dryer, and finally pacifier. The reviews of Microwave's successful products are the most reliable, followed by the hair dryer, and the reviews of Pacifier's successful products are less reliable. And whether a reviewer is an Amazon Vine member has nothing to do with the success of the product.

It can also be seen that the reviews of verified_purchase or Amazon Vine members are not significantly related to the number of valid votes. Explain that customers judge whether reviews are useful and will not be affected by the identity of the reviewer.

In addition, all products with a higher rating account for a larger proportion of positive emotional reviews. However, products with lower ratings among all products have a larger positive emotional review than before. We cannot directly judge the relationship between review sentiment and comprehensive score.

Table 5 Pacifier Top 10 Products with highest COM_S

COM_S	PRODUCT_TITLE	PRODUCT_PARENT
0.713	cosco alpha omega elite convertible car seat	485572042
0.703	Sealy Baby Soft Ultra MattressSealy	409107492
0.696	Fisher45;Price Aquarium Monitor	61825188
0.693	Velour Boppy Nursing Pillow - green	108369347
0.691	fisher-price infant-to-toddler rocker in blue/red	32051425
0.691	the ultimate baby wrap in navy	234474039
0.691	ameda purely yours breast pump with carry-all	985518895
0.690	baby jogger performance series double 20/navy	242140507
0.689	Vintage Teaberry 4 Piece Crib Set Vintage Teaberry	460235866

Among pacifier products, brands such as cosco alpha, Fisher, ameda, Sealy has the better performance. In the same way, we can get the following conclusions.Among microwave oven products, brands such as sharp, lg, magic chef, whirlpool, profile, and Panda Small are doing well. Among the hair dryer products, Salon, Waterpik, Micro Tweeze, DuWop, ISLAND HAWAII, Oral-B and other brands do a better job. Among Pacifier products, brands such as chicco keyfit, ikea kladd, and batman perform bad. Among microwave oven products, brands such as Samsung, Danby, and jx7227sfss are not doing well. Among the hair dryer products, brands such as andis, conair, xtava, mhd, remington do a worse job.

4. Task 3

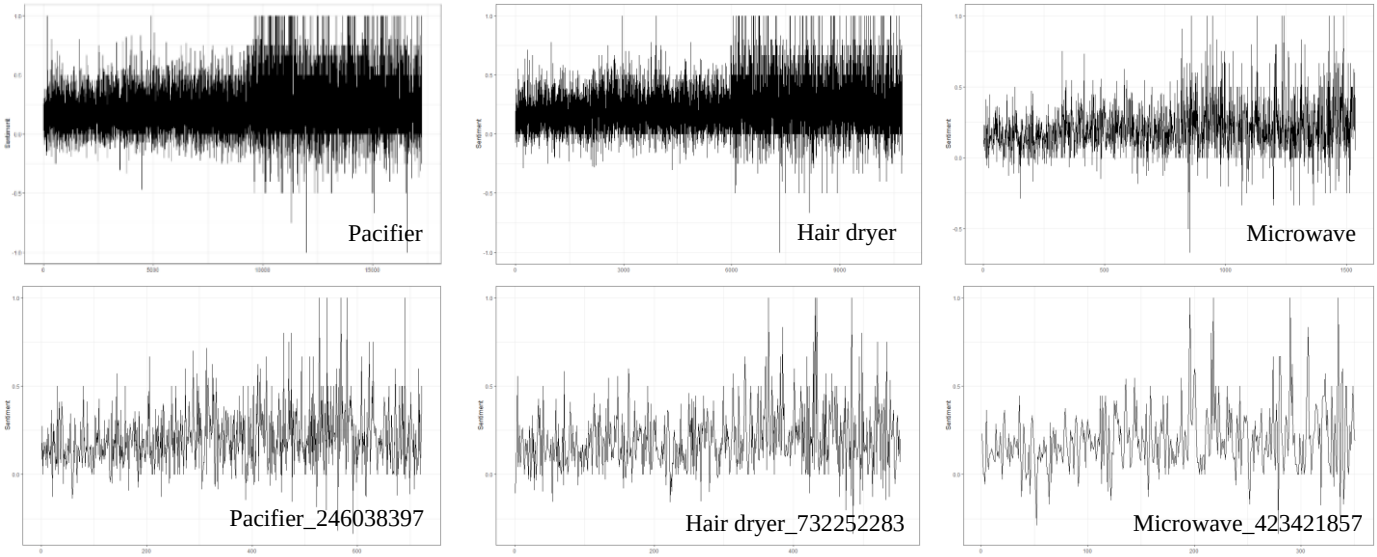
4.1 Problem Analysis

In order to explore the changes in customer reviews over time, and to give the trend of product reputation in the market. When solving task2, first of all, the original time series data with breakpoints in units of days is used to calculate the cumulative dynamic mean to obtain continuous time series data in units of months, and the data is modeled in conjunction with the time series ARMA model Fit and predict future trends. Next, calculate the change in the number of reviews and the change in review sentiment before and after each rating in each specific product, and use non-parametric tests to determine the impact of star rating on reviews. Finally, find out all reviews with extreme emotions, use non-parametric tests to determine whether they have an impact on star ratings, and combine KMEANS clustering to give the characteristics of customer reviews and discuss potential false reviews.

4.2 Problem Solving

4.2.1 Time series visualization and prediction of three variables

Drawing a line graph can intuitively show the evolution of emotional scores, and you can see the trend of emotional scores of product reviews in the time dimension. The data in the following time series models are processed by month, because the original data is from day to day, but there are cases where there are no comments for many consecutive days, so we analyze all data after integrating them monthly, and then In order to ensure the validity of the data, we filter out the first few months with a small number of valid comments that affect the quality of the modeling. We find that the probability of extreme emotion reviews is low, but with the increase of time and the increase in sales of goods, the number of extreme emotions will increase accordingly. The figure below shows the sentiment trends of the three product reviews, Pacifier, hair dryer, and microwave. It can be seen that the emotional score in product reviews increases with time and the volatility increases. Generally speaking, positive (positive) emotions are higher than negative (negative) emotions in comments. The product model



with the most number of reviews is extracted from the three products, and their product_parent codes are 246038397 (philips avent bpa free soothie pacifier, 0-3 months, 2 pack, packaging may vary), and 732252283 (remington ac2015 t | studio salon collection pearl ceramic hair dryer, deep purple), 423421857 (danby 0.7 cu.ft. countertop microwave), the emotional scores of the three products are all positive, but the emotional scores of Pacifier and Hair dryer show a downward trend, and Microwave products are in the rising stage later.

Figure 3 Emotional chart of three products

The figure shows a line chart of the three product ratings of Pacifier, Hair dryer and Microwave Oven over time. Observation from the vertical line (April 2005) shows that the stars have an upward trend, and the final stars are divided into 4.26. The number of likes rises faster, showing an exponential upward trend; H products The previous period fluctuated greatly. From 2005 to 2006, the star rating experienced a sharp decline, and then slowly rebounded. The final rating in August 2015 was 4.08. As a durable product, the number of likes is much higher than the number of reviews, that is, "watch and see" is more To the purchaser. After experiencing a sharp decline in the second half of 2004, the star rating of M products gradually rebounded. After reaching a peak in early 2008, it gradually decreased. After falling to a trough in mid-2012, it gradually rebounded. The final score in March 2015 was 3.39 points. It will continue to rise slowly. It is worth noting that M as a durable product sales and reviews are much lower than the two categories of products, but its number of likes is much higher than the number of reviews, the cumulative number of likes is 8996, the cumulative number of reviews is 1536 , The difference is almost 5 times, and in some time periods, the number of likes rose in a stepwise manner. At the end of 2014, with the sudden increase of vine comments, the number of likes rose sharply.

Figure 4 Pacifier Trend graph of cumulative number of star ratings, reviews and helpfulness ratings

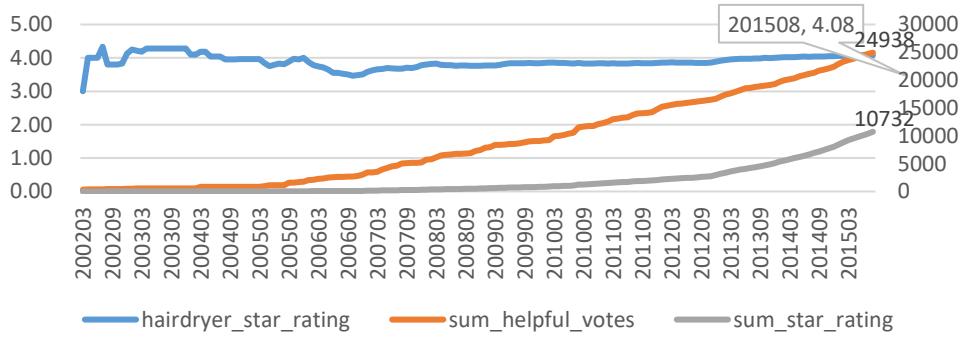


Figure 5 Hair dryer Trend graph of cumulative number of star ratings, reviews and helpfulness ratings

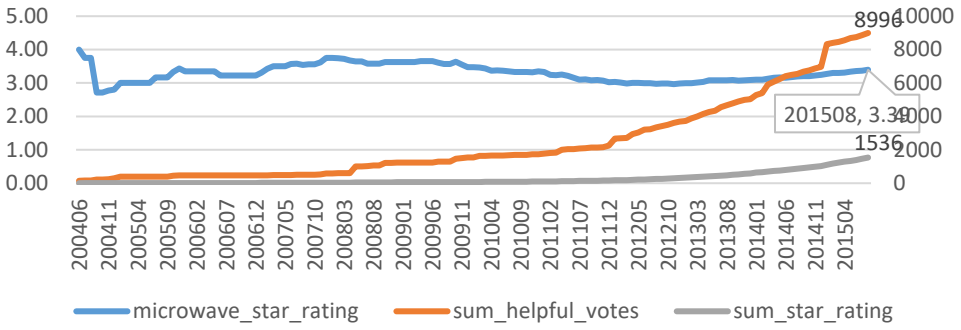


Figure 6 Microwave Trend graph of cumulative number of star ratings, reviews and helpfulness ratings

4.2.2 Prediction of star ratings

We then use a time series model to predict star ratings. First, we use the time series visualization described above to obtain the main trends of the three products. Then we check the stability of the time series and draw the ACF / PACF chart to find the optimal parameters. Then, a differential autocorrelation moving average model (ARIMA) model is established, and the star ratings in September and October 2015 can be predicted as follows:

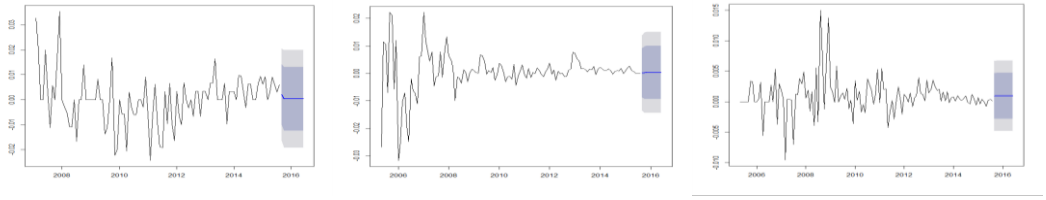


Figure 7 Star ratings Forecasting Graph

Microwave oven star ratings prediction for September: 3.401379; October: 3.395562.Hair dryer star ratings prediction for September: 4.080642; October: 4.081117.Pacifier star ratings prediction for September: 4.268717; October: 4.268825.

4.2.3 Prediction of reviews sentiment

We can use the same method to predict the sentiment trends of the three products in the next three months. The results are shown in the table below. Similar to the results obtained from the emotional trend chart, the sentiment scores of future reviews of P and H products will decline. The negative and positive emotions of P products will first increase and then decrease, while the negative emotions of H products will not change much , The decline in positive emotions is more serious, reflecting the possible problem of customer enthusiasm; M as a durable product, sales and reviews are very few compared to the other two products, relatively more data more consistent with the normal distribution, and relatively Less data is more volatile. Therefore, the emotional trend of future reviews of M products is rising, falling, and rising fluctuation trends, but generally it is rising.

table 6 Emotional charts of three product reviews in the next three months

date	sentiment	Aug-15	Sep-15	Oct-15	Nov-15	mini-Line chart
pacifier	negative	0.0556	0.0759	0.0626	0.0595	
	positive	0.2552	0.3322	0.2838	0.2641	
	sentiment	0.1996	0.2564	0.2212	0.2047	
hair dryer	negative	0.0679	0.0679	0.0679	0.0679	
	positive	0.2654	0.2714	0.2666	0.2652	
	sentiment	0.1975	0.2035	0.1987	0.1973	
microwave	negative	0.0606	0.0893	0.0653	0.0636	
	positive	0.2501	0.2760	0.2415	0.2517	
	sentiment	0.1896	0.1867	0.1762	0.1881	

4.2.4 Time series analysis

First, select star ratings, next_past, next_past_sen, and use multiple independent sample nonparametric tests. The test method is J-T test (Jonckheere-Terpstra):

It can be known by the results that p_1 is approximately 0, and $p_2 = 0.531 > 0.05$. At the confidence level of 0.05, there is a significant difference in the number of reviews before and after the occurrence of different star reviews, but there is no significant difference in the sentiment between the reviews before and after.

Check whether the one-star evaluation will cause significant changes in comments and comment sentiment. Select the number of front and rear comments and comment sentiment corresponding to the one-star sample, and use Wilcoxon to perform a non-parametric test on two paired samples.

Table 7 Test Statistics

	past_sen - next_sen	past - next
Z	-39.067[b]	-37.732[b]
Asymp. Sig. (2-tailed)	.000	.000

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

P is approximately 0, and there is a significant difference at a significance level of 0.001, that is, a one-star review will cause a significant change in the number of reviews and a significant change in review emotions.

Similarly, check whether the five-star evaluation will cause a significant change in the review:

P is close to 0, and there is a significant difference at a significance level of 0.001, that is, a five-star review will cause a significant change in the number of reviews and review sentiment.

In combination with the above, there is a difference in the number of reviews before and after different star reviews, so it cannot be judged that special star reviews will affect the number of reviews, but special star reviews, such as one-star and five-star, will cause changes in review emotions.

Table 8 Final Cluster Centers

	Cluster		
	1	2	3
star	3.55	3.90	4.17
next_past_sen	.56	-1.44	6.66

K-MEANS clustering is based on the star rating of each item and the emotional changes before and after the one-star review. It can be found that 1 group of products in the table has the lowest star rating and is least affected by the one-star review; 2 groups publish more with one-star review Negative reviews. The star rating of these products is 3.9, which is relatively high. It may be that the customer 's star rating is too high and does not match the product, resulting in a large number of negative reviews and one-star ratings. High, 1-star reviews in such products will cause a large number of positive reviews from customers. Possible situations are: first, when a good product is discredited, customers will praise it more positively; Influence, the phenomenon of good comments appears.

4.2.5 Interaction between star_rating and review

The data is summarized according to product_parent and date, and the number of reviews and the sentiment of reviews before and after a review of a certain type of product (excluding this review) are obtained. In order to eliminate the influence of time, the longer the time between the two comments, the smaller the mutual influence, so the time weight is 1 / [(date_after-date_before) +1], where +1 is for processing Comments occurred when the denominator was zero on the same day.

The pseudo code for the variable processing of comment number next and comment emotion next_sen after the comment has occurred is:

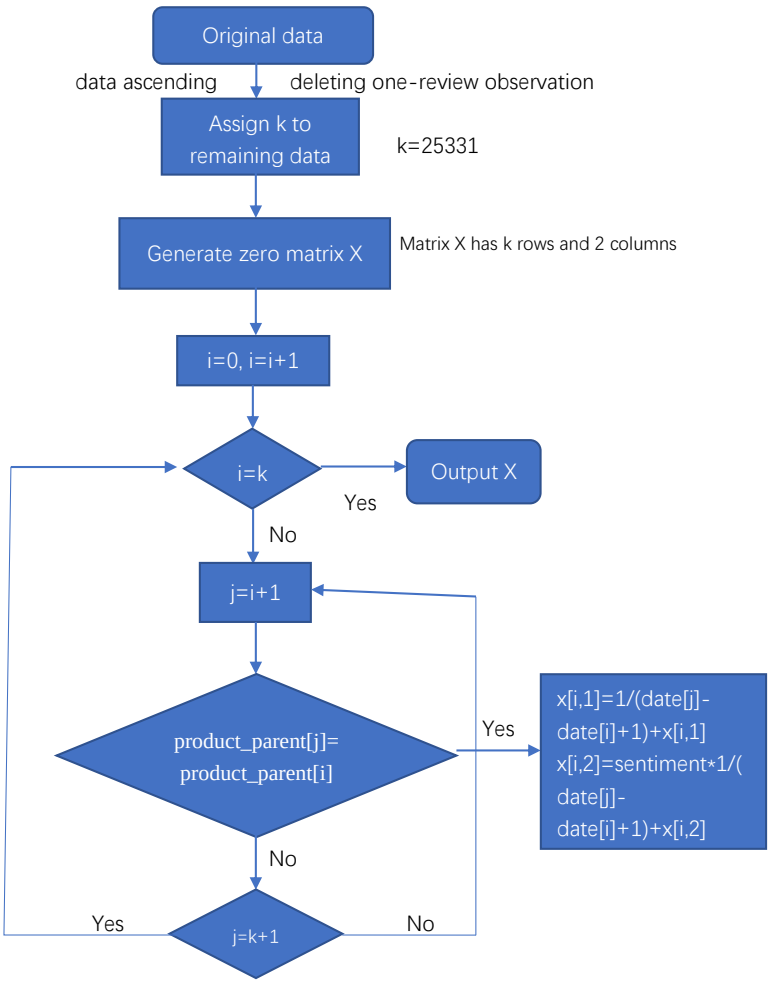


Figure 8 Process of variable handling pseudo code

The original data set is sorted by date in descending order, and according to the above pseudo-code calculation process, you can get the number of reviews before the comment occurs, and the change in comment sentiment past_sen, which is not repeated, and then you can get the change in the number of reviews before and after a certain comment next_past and comment sentiment Changes next_past_sen. After testing, both do not conform to the normal distribution, so non-parametric tests are selected for subsequent tests.

We selected 17 extreme negative words including disappointed and 17 extremely positive words including enthusiastic to obtain extreme emotional values NEG and POS.

Table 9 Final Cluster Centers

Cluster	1	2	3	4	5	6	7	8
COM_S	0.1283	0.1321	0.1699	0.1334	0.0399	0.0909	0.0982	0.1235
star_rating	4	4	4	4	**5**	4	4	4
helpful_votes	1	156	23	331	814	0	17	1
vine	0	0	0	0	**1**	0	0	0
verified_purchase	**1**	**1**	0	**1**	0	**1**	**1**	**1**
reviewzishu	51	115	291	67	233	38	254	44
NEG	0	0	0	0	**1**	0	0	0
POS	0	0	**1**	0	0	0	**1**	0
reviews	46	205	17	518	43	722	619	379

The number of classifications is determined through the stone map, and the entire sample is divided into 8 categories in the table using Kmeans clustering. You can see that extreme negative words appear in category 5, which is characterized by low

product satisfaction, few reviews, but star ratings. High ratings, high number of valid points and VINE reviews, long word reviews, and false information. Such products are either good products that have been maliciously smeared or bad products that have been touted;

Positive vocabulary appears in categories 3 and 7, of which the third category has high product satisfaction, fewer reviews and likes, longer reviews, fewer vine reviews, and fewer verified_purchase certifications, and doubts the validity and credibility of reviews , Or there is a phenomenon of praise; the satisfaction of the seventh category of products is acceptable, the number of reviews is high, and the number of likes is small. With the increase of the number of reviews, the probability of occurrence of extreme words will increase accordingly, so there is no abnormal review behavior in this group.

Table 10 Table of Negative words Correlations

		star_rating	NEG
star_rating	Pearson Correlation	1	-.209**
	Sig. (2-tailed)		.000
	N	29511	29511
NEG	Pearson Correlation	-.209**	1
	Sig. (2-tailed)	.000	
	N	29511	29511

**. Correlation is significant at the 0.01 level (2-tailed).

Table 11 Table of Positive words Correlations

		star_rating	POS
star_rating	Pearson Correlation	1	.240**
	Sig. (2-tailed)		.000
	N	29511	29511
POS	Pearson Correlation	.240**	1
	Sig. (2-tailed)	.000	
	N	29511	29511

**. Correlation is significant at the 0.01 level (2-tailed).

It can be seen that at a confidence level of 0.01, there is a correlation between extreme emotions and star ratings. There is a negative correlation between extreme negative emotions and star ratings. There is a positive correlation between extreme positive emotions and star ratings. The coefficients are all less than 0.3, and the correlation is low. Considering the frequency of extreme emotional words, non-parametric tests are performed. The results are as follows:

Table 12 Test Statistics

	star_rating
Mann-Whitney U	5310158.500
Wilcoxon W	5655854.500
Z	-31.131
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: NEG

	star_rating
Mann-Whitney U	63602803.000
Wilcoxon W	266080429.000
Z	-43.030
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: POS

It can be seen that there is a significant difference between star ratings and other star ratings for comments with extreme emotional words, which can explain that there is indeed a relationship between extreme emotions and star ratings. The relationship between extreme negative emotions and star ratings is Negative correlation, positive correlation between extreme positive emotion and star rating.

5. Model Evaluation

5.1 Model Limitations

Time series analysis:

(1) The model has the defect of prediction error and ignores the changes of other random variables. When there are large changes in the outside world, there are often large deviations.

(2) The time series prediction method is effective to predict the effect of changes in the short term. Because objective things, especially economic phenomena, are more likely to change external factors over a longer period of time, they must have a significant impact on market economic phenomena.

Lexicon-based method:

(1) The sentiment score of a review is the average of all its sentence scores. This method is not consistent with the actual situation. Just as the paragraphs in the article are of different importance. In a sentence, the sentence before and after the sentence also have important differences.

(2) There is a type of text that uses a derogatory word to indicate positive meaning. Lexicon-based method cannot recognize this situation. For this problem, deep learning methods are needed to effectively solve it. It is difficult to solve this problem with ordinary machine learning methods.

(3) For the judgment of positive and negative comments, the algorithm ignores many other negative words, degree adverbs, and emotion words; it is too simple to judge the strength of emotions.

(4) The accuracy rate of Lexicon-based method is high (more than 80%), but the manual workload is relatively large.

(5) Ignore the existence of emoticons and the emotions expressed in the comments. Emoticons often carry some emotional polarity, which expresses the emotions of customers more clearly. But the emoticons were ignored in the emotional score.

Principal component analysis disadvantages:

(1) Principal component analysis is suitable for data that has a strong correlation between variables. If the original data is weakly correlated, it will not play a good role in reducing dimensions.

(2) After the dimensionality reduction, there is a small amount of information loss and it is impossible to contain 100% of the original data.

(3) After the original data is standardized, the meaning will change, and the interpretation meaning of the principal components will be more vague than the original data.

5.2 Model Strengths

Time series analysis:

The time series analysis and prediction method highlights the role of time factors in forecasting, and does not consider the impact of specific external factors for the time being. Time series analysis carries out volume prediction. In fact, all the influencing factors are attributed to the time factor, and only the comprehensive effect of all influencing factors is recognized, and it will still play a role in predicting objects in the future. The input variables can be adjusted according to the time series model to keep the system development process at the target value, that is, the necessary control can be performed when the process is predicted to deviate from the target. By analyzing time series, we can make reasonable predictions and grasp the future development trends in advance to provide a basis for business decisions.

Lexicon-based method:

Lexicon-based approach outperforms Supervised Machine Learning approach not only in terms of Accuracy, Precision, Recall and F-measure but also in terms of economy of time and efforts used.

The ARIMA model achieves a more accurate forecast and it is a simplistic modeling technique.

Advantages of principal component analysis:
(1) The data is not required to be normally distributed. The principal component is to rotate the base group in the direction in which the data is most discrete.
(2) Through the synthesis and simplification of the original variables, the weight of each indicator can be objectively determined, and the randomness of subjective judgment can be avoided. The final result is only related to the data, and has nothing to do with model selection.
(3) PCA transforms the original indicators into independent principal components. Practice has shown that PCA is a linear dimensionality reduction method with the least amount of lost original data information.

6. Conclusion

1. By analyzing star_ratings, reviews, helpfulness ratings, and SentimentQDAP, the overall characteristics of three products are initially obtained.
(1) For microwave, star ratings are concentrated on 1 star and 5 stars, while hair dryer and pacifier are concentrated on 5 stars. In addition, high-rated customer reviews provided by hair dryer and pacifier can provide more effective information. It can be assumed that pacifier and hair dryer are Amazon's more successful products.
(2) For microwave, the company should pay attention to the performance and price of the product, and make timely improvements to the product. For the hair dryer, the company should focus on product performance and appearance issues to find out the factors of product success. For pacifier, the evaluation is generally high, and the company should pay attention to the product life.
2. Use time series visualization to depict the trend of comment sentiment and star rating, cumulative number of comments, and cumulative number of likes in the time dimension.
(1) With time flowing and the increase in sales of goods, the number of extreme attitudes will increase accordingly.
(2) Positive (positive) emotions are higher than negative (negative) emotions in the three product reviews. However, the emotional scores of Pacifier and Hair dryer showed a downward trend, and Microwave products were in a rising stage in the later period.
3.The ARIMA model is used to predict the emotional score. Once the product is sold on the market, based on customer ratings and reviews, it can be inferred that the product's score will be within a certain period of time. The company's grasp of future development trends can provide a basis for business decisions. Here are the prediction results for the sentiment scores of the three commodities over the next three months:
(1) Both negative and positive emotions of Pacifier show a trend of rising first and then falling.
(2) The negative emotion of the Hair dryer changes little, and the decline of the positive emotion is relatively serious, reflecting the possible problem of customer enthusiasm.
(3) As a durable product, Microwave has fewer sales and reviews than the other two products. The emotional trend of future reviews is a rising, falling, and rising trend, but it is generally on the rise.
4. The decision indicators for Customer Reviews were determined using the principal component analysis method. The first principal component is voting information, the second principal component is the review content, and the third principal component is the customer. Identity information. With the influence of the “age” of the Review, the calculation formula for determining the Customer Satisfaction (CS) was obtained. Comprehensive Satisfaction (COM_S) is derived from the CS summary of the same product.
5. Using non-parametric test methods, we conclude that special star reviews (one star, five stars) can cause emotional changes in reviews.
6. Using the non-parametric test method, there is a correlation between extreme attitude reviews and ratings: there is a relationship between extreme attitudes and star ratings, there is a negative correlation between extreme negative attitudes and ratings, and extreme positive attitudes and ratings are: Positive relationship.

7. References

[1] Carolanne, M. (2017). Understanding the Value of Online Customer Reviews

[2] Gamal, D., Alfonse, M., EI-Horbaty, E., & Salem, A. (2019). Implementation of Machine Learning Algorithms in Arabic Sentiment Analysis Using N-Gram Features. *Procedia Computer Science*, Volume 154, 2019, Pages 332-340

[3] Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, Volume 36, September 2019, 101003

[4] Zhang, S., Wei, Z., Wang, Y., & Liao, T. (2018). Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary. *Future Generation Computer Systems*, Volume 81, April 2018, Pages 395-403

[5] Bernabé-Moreno, J., Tejeda-Lorente, A., Herce-Zelaya, J., Porcel, C., & Herrera-Viedma, E. (2020). A context-aware embedding supported method to extract a fuzzy sentiment polarity dictionary. *Knowledge-Based Systems*, Volume 190, 29 February 2020, 105236

[6] Mukhtar, N., Khan, M., & Chiragh, N. (2018). Lexicon-based approach outperforms Supervised Machine Learning approach for Urdu Sentiment Analysis in multiple domains. *Telematics and Informatics*, Volume 35, Issue 8, December 2018, Pages 2173-2183

[7] Matyjaszek, M., Fernández, P., Krzemień, A., Wodarski, K., & Valverde, G. (2019). Forecasting coking coal prices by means of ARIMA models and neural networks, considering the transgenic time series theory. *Resources Policy*, Volume 61, June 2019, Pages 283-292

[8] Nguyen, H., AsifNaeem, M., Wichitaksorn, N., & Pears, R. (2019) A smart system for short-term price prediction using time series models. *Computers & Electrical Engineering*, Volume 76, June 2019, Pages 339-352