Problem Chosen	2020 MCM/ICM Summary Sheet	Team Control Number
C		2011400

Product Evaluation Model Based on Entropy Weight Method

Summary

In order to provide marketing strategies for Sunshine Company, our team develops a TOPSIS (Technique for Order Preference by Similarity to an Ideal Solution) model based on entropy weight method. Our model has three parts:

In the first part, we process the data using Natural Language Processing (NLP) and quantify the results. Each column of data is analyzed and the redundancy is removed. By analyzing the column of product title in pacifier.tsv with NLP and extracting keywords from it, some irrelevant information in the data is founded and deleted.

In the second part, we build TOPSIS model based on Entropy Weight Method, sorting according to the proximity between evaluation object and idealized target, doing relative pros and cons assessment of existing objects. First, we use Grey Relational Analysis to calculate the correlation coefficient between variables, choosing variables with a high correlation with star ratings and reviews. Then use Entropy Weight Method to rate star ratings and reviews. Applying TOPSIS model to each purchase record, getting the score, accumulating based on time series, we get the relationship between reputation and time.

In the third part, based on the previously calculated star ratings and evaluation score, using multiple linear regression model, we find optimal combinations of star ratings and reviews to predict sales. In order to research the connection between star ratings and ratings: 1) Assuming that the numbers of accumulative data form a time series, analysis whether a certain number of special star ratings will cause special reviews. 2) We count the number of star ratings (5,4,3,2,1) and reviews (good, average, bad) combinations that appear in a purchase record over a while, building a 3*5 matrix, draw 3D images, observing whether text-based reviews strongly associated with rating levels.

We further discuss how to design the best strategy, analyzing the conclusions obtained above and do further researches on the data provided, extracting keywords from each review to identify potentially important design features that would enhance product desirability.

Finally, we conduct sensitivity analysis, dissect the pros and cons of our model and write a suggestion letter for Sunshine Company.

Keywords: NLP; TOPSIS;Entropy Weight Method;Grey Relational Analysis;Multiple Linear Regression

Contents

1 Introduction

1.1 Promblem Statement

With the rise of e-commerce and the popularity of online shopping, Amazon provides customers with an opportunity to rate and review purchases. Buyers can use a scale of 1 to 5 to indicate satisfaction with the product, and they can also submit reviews to express their opinions on the product. Other customers can decide whether to buy the product based on the star rating and evaluation.

Sunshine Company plans to launch three new products in the online market: pacifiers, microwave ovens, hair dryers. They have hired our team as consultants to use time-based data to determine whether the data is helpful for the company to sell products, analyzing the relationship between the data, and the impact of the data, and finally inform the company's marketing strategy.

To meet the needs of the Sunshine Company, our team built TOPSIS Model based on Entropy Weight Method that can handle the relationships between the variables and evaluate them. According to the results of model evaluation, we conduct further analysis and finally propose strategies.

1.2 Our Work

We build a work-flow diagram based on our model.

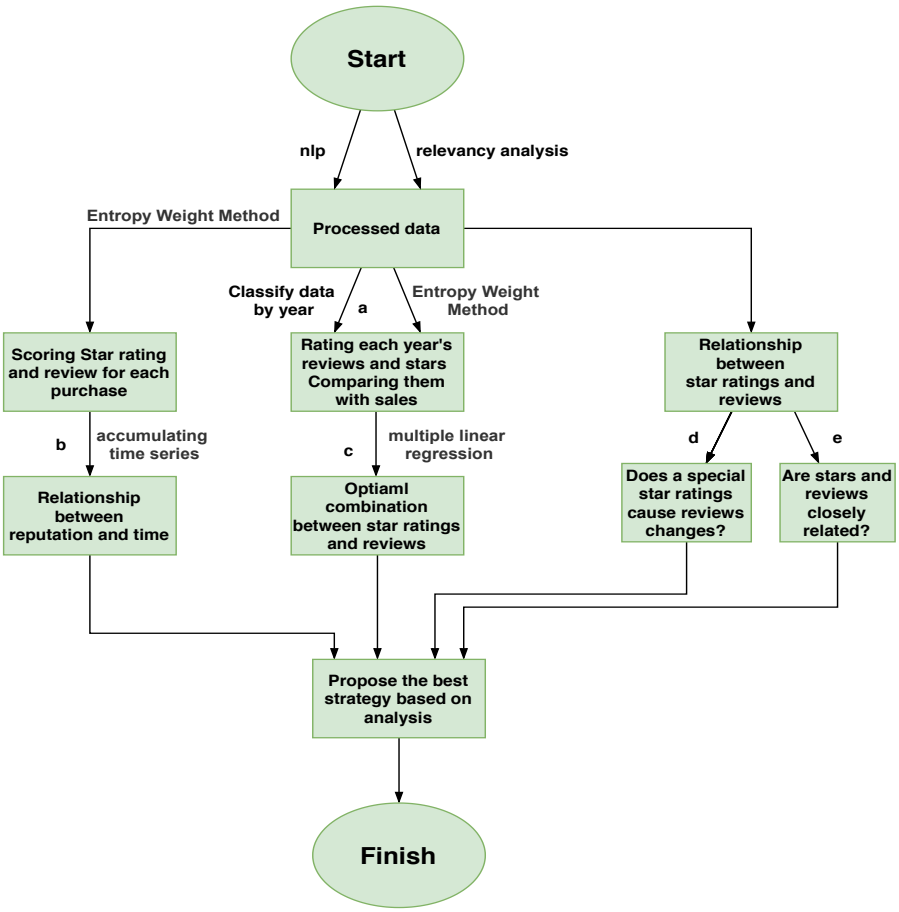

Figure 1: work-flow diagram

2 Definitions and Assumptions

2.1 Definitions

ă Denotation	Definitions
P_{ij}ă	ăProportion of the i-th sample value in the index under the j-th indexă
ă E_jă	ăEntropy of the j-th indexă
ă W_jă	ăWeight of the j-th indexă
ă F_iă	ăFinal score for the i-th sampleă
ă TOPSISă	ăTechnique for Order Preference by Similarity to an Ideal Solutionă
ă NLPă	ăNatural Language Processingă
ă	

2.2 Assumptions

- Datas on the same day are ordered.
- 3 star is equivalent to neutral comment.
- The relationship among stars, reviews and sales is linear.
- The content of data is all purchase records.
- The number of accumulative data is equivalent to time series.

3 Data Processing

3.1 NLP

Using natural language processing (NLP) and result quantization processing, and in order to save training time, after comparing many APIs, Baidu Cloud Packaged NLP engine was selected. Manually process very small amounts of data and bad values that some code programs cannot handle.

Perform NLP processing on the 'review_body' column, judging the emotional polarity category (positive, negative, neutral) of the 'review_body' text of each purchase information, directly corresponding to enthusiastic, disappointed, middle, and give the corresponding confidence. In the actual running process, it is found that the recognition effect of English comments will have a certain deviation, so a Python script is used to translate the comments into Chinese, then NLP processing, and the final result is output to an excel form.

At the same time, because the amount of data is large, in order to prevent errors in individual data, the code has undergone error correction processing. A very small number of unrecognizable comments will be marked as None, processing with manual judgment. This part of the data is small.

In the post-quantification process, the mood is enthusiastic, disappointed, and middle are quantified corresponding to 2,0,1. Stars are quantified as 5, 4, 3, 2, 1.

3.2 Error Message Handling

By performing NLP analysis on the product_title column in the pacifier.tsv file and extracting keywords, it was found that some purchase records were not related to pacifier, and some star data was obviously bad value, so these data were deleted.

4 Model Interpretation

The TOPSIS model based on the entropy weight method uses the combination of the entropy weight method and the TOPSIS model to obtain the objective weight of the evaluation index. The entropy method is a measure of the amount of information and can also determine the amount of valuable information of known data. The principle of the weight of each indicator obtained by this method is to estimate the value coefficient of the indicator information. If a certain indicator has a large value coefficient, the indicator will have a great contribution to the evaluation result, so the weight will also increase. Big. Therefore, it is possible to obtain m indicators for n evaluation objects.

4.1 Data normalization

The unity of the same evaluation index, the higher the value of the high-quality index, the better. Usually, the inverse method (or formula in the code) can be used for the reverse monotonous index.

Because the range of values of each indicator is different, and the dimension and meaning are different, in order to eliminate these effects, the data needs to be normalized. Suppose there are n samples in total, each sample has m indicators, and the collected data is x_{ij} (i = 1,2,, n; j = 1,2,, m). The larger each value is, the more favorable it is. The normalized processing method uses the following formula:

$$y_{ij} = \frac{x_{ij} - x_{jm}^*}{x_{jM}^* - x_{jm}^*} \tag{1}$$

Among them:$x_{jM}^* = max_{ij} \quad x_{jm}^* = min_{ij}$

After the above transformation, all data is changed to [0,1], which is convenient for subsequent processing for unified processing.

4.2 Weight Calculating

The processed data is y_{ij} (i = 1,2,, n; j = 1,2,, m). Among them, the calculation formula of the information entropy of the j-th index is:

$$E_j = -\frac{\sum_{i=1}^{n} p_{ij} \ln p_{ij}}{\ln n} \qquad j = 1, 2, ..., m \tag{2}$$

$0 \leq E_j \leq 1$, among them$p_{ij} = \frac{y_{ij}}{\sum_{i=1}^{n} y_{ij}}$, if$p_{ij} = 0$, define $p_{ij} \ln p_{ij} = 0.00001$.

E_i, the smaller,the greater the difference between the data, so the greater the information provided, the greater the weight of the index.

E_j, the larger, the closer the data are to each other, so the less information is provided, the smaller the weight of the index is.

Objective weight formula:

$$W_j = \frac{1 - E_j}{m - \sum_{j=1}^{m} E_j} \qquad j = 1, 2, ..., m \tag{3}$$

Weighting average of all indicators for each sample can get the comprehensive score. The calculation formula is:

$$f_i = \sum_{j=1}^{m} w_j \cdot y_{ij} \qquad (i = 1, 2, ..., n) \tag{4}$$

5 Model Construction

5.1 Problem One

5.1.1 Invalid Data Processing

For each column of 'Data Set Definitions' in the data , we analyzed and processed redundant information. The 'marketplace' is all US, so it is streamlined; because every purchase record is taken as the analysis object, so the purchaser is not considered, so 'customer_id', 'review_id' is streamlined; 'product_id', 'product_parent' is found to be the same as the same product after comparison, so only one is kept; the 'product_category' in three data sets are exactly the same, so they are simplified.

5.1.2 Important Data Processing

Use excel to perform a fitting test of whether the 'vine' is 'Y' and the total 'nonhelpful' votes obtained by ('total_votes'-'helpful_votes'). It was found that whether the 'vine' is 'Y' has a small weight impact, and there were almost no vine users before 2008, so The column is discarded.

'verified_purchase' is 'Y' to prove that the buyer did use the product, the review has high credibility, and after fitting with 'helpful_votes', it proves that the two are positively correlated and have a higher weight, so the 'verified_purchase' is retained.

5.1.3 Relation Analysis of Remaining Data

The relationship among stars, comments, and helpful votes is shown in the figure:

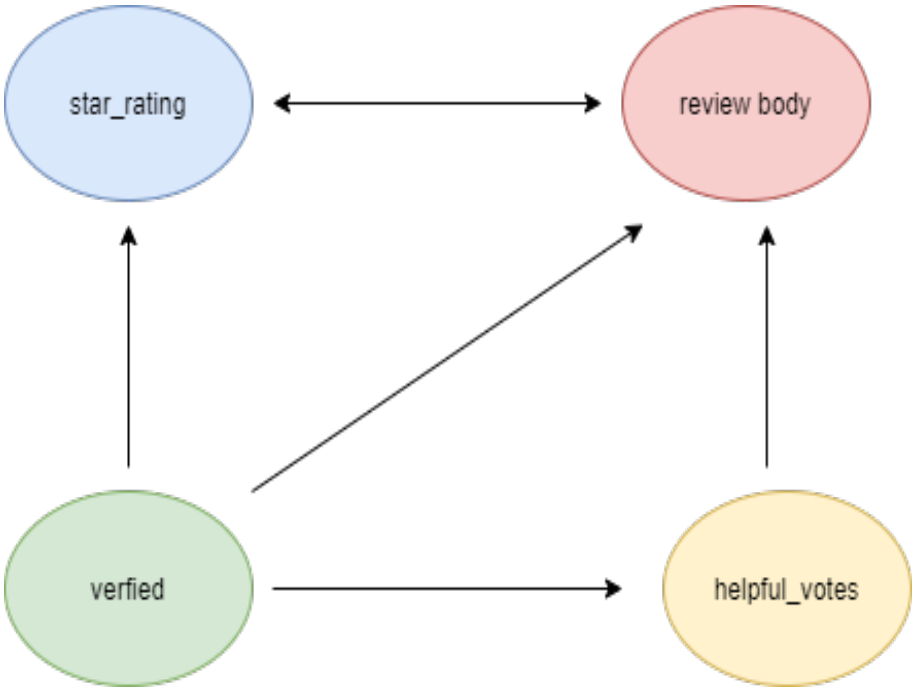

Figure 2: relationship diagram

5.2 Problem Two

5.2.1 a

When consumers face reviews with the same star rating or the same emotional color (enthusiasm, disappointment), they do not exactly generate the same desire to buy.

The trust in treating the same emotional color comments is related to other parameters, such as the number of useful comments for the comment and whether the comment has 'verified-purchase'.

Based on the data processing of task one, we can conclude:

- Product star rating is related to 'verpurchase' parameter.
- Product reviews are related to useful votes, useless votes, 'verified_purchase'.

In order to effectively evaluate product star ratings and review indicators over a period of time, we introduce "star ratings" and "evaluation scores". For each of the three data sets, the number of star ratings and the number of positive reviews, as well as the number of parameters related to star ratings and reviews, are counted in years. The TOPSIS model is established for each year's "star rating" and "evaluation score". After normalizing the data, the entropy weight method is used to assign weights to relevant parameters of the "star score" and "evaluation score", and finally the specific values of "star score" and "review score" are obtained year by year.

The following are the proportions of the 'star rating' and 'review score' parameters of three categories of products:

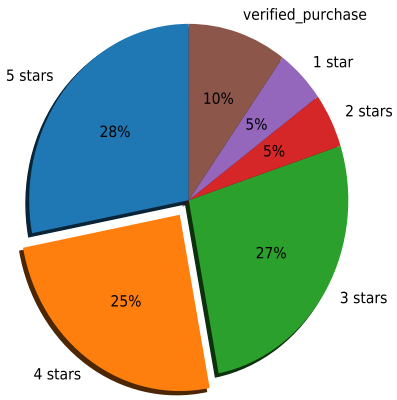

Figure 3: pacifier 'star score'

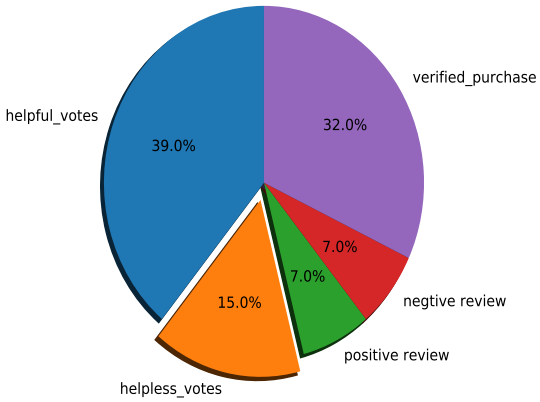

Figure 4: pacifier 'review score'

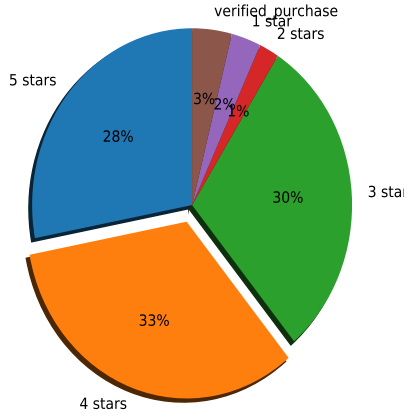

Figure 5: microwave 'star score'

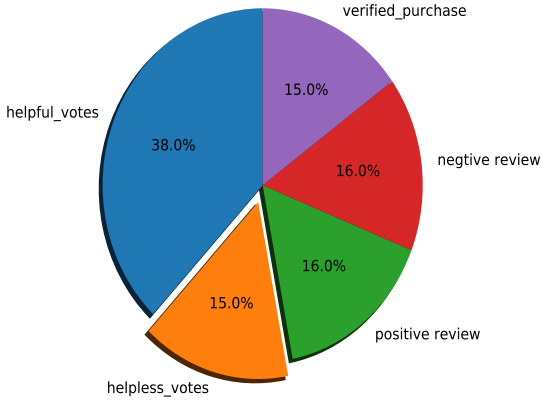

Figure 6: microwave 'review score'

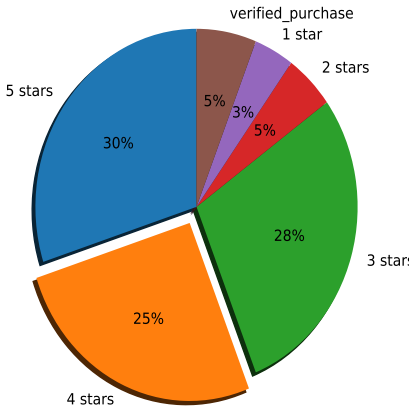

Figure 7: hair_dryer 'star score'

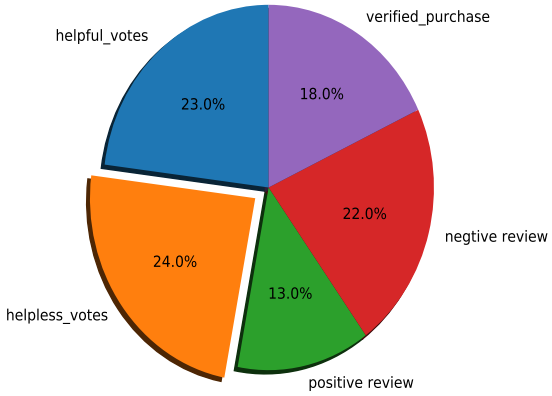

Figure 8: hair_dryer 'review score'

5.2.2 b

Simply fitting the trend of the number of positive reviews over time cannot explain the reputation of the product.

Star rating+reviews establishes a new indicator-"product score" for each purchase record, which is used to indicate product reputation. Product reputation is a combination of star ratings and positive reviews. The entropy weight method is used to derive the "star rating" and "review" ratio parameters. A "product score" is calculated for a single purchase record. Under the "time-based measurement mode and requirements", the purchase quantity is accumulated on the time series, and the average value of the "product score" before a certain point in time is used as its reputation, and the "time-reputation" curve is finally obtained.(We make blue lines)

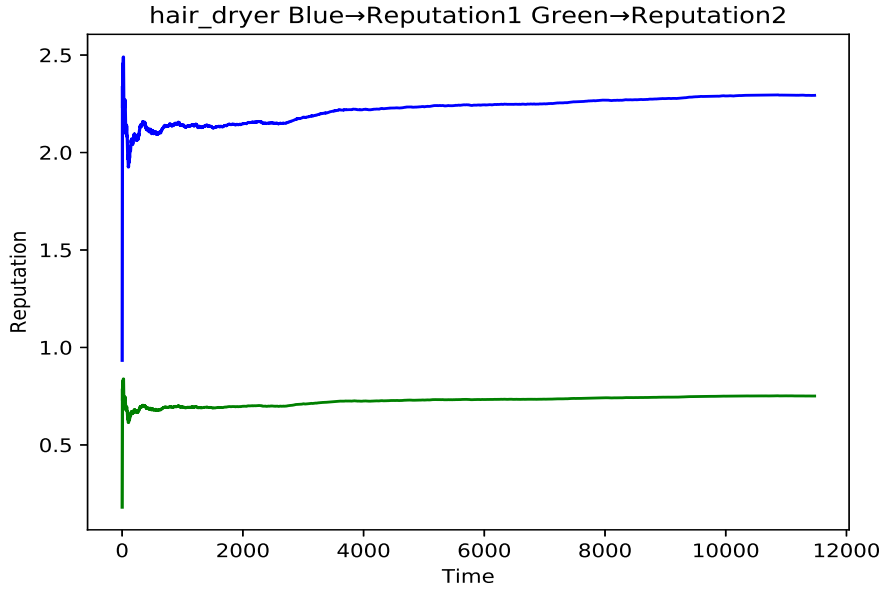

Figure 9: hair_dryer time-reputation

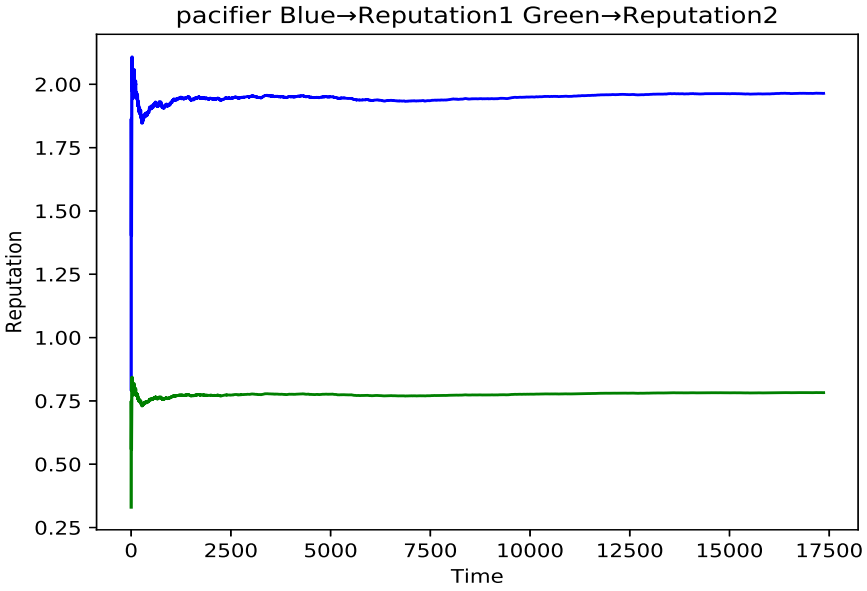

Figure 10: pacifier time-reputation

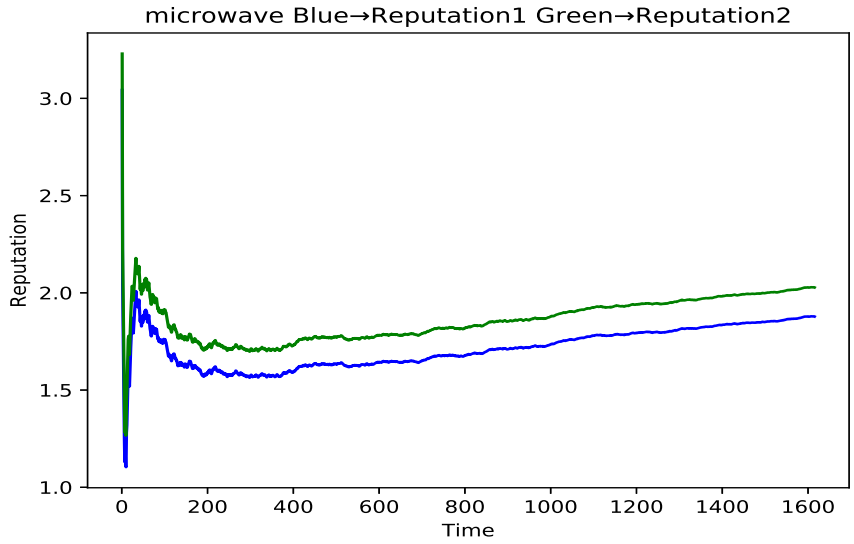

Figure 11: microwave time-reputation

5.2.3 c

In the previous question, the annual star ratings and evaluation scores of the three products have been obtained using the entropy weight method. These two ratings are directly related to product sales. We use a binary linear regression model and use Python to analyze the data. In training, the independent variable coefficients are obtained according to the given data, and the regression equation is fitted and predicted. Finally, visual processing is carried out.

$$y = ax_1 + bx_2 + c \tag{5}$$

y:sales x_1:the score of star rating x_2:the score of review

Data example(hair_dryer):

Table 1: year-review score-star score- sales table

	star_score	review_score	sales
2002	0.1019	0.502	8
2003	0.0868	0.5989	10
2004	0.0927	0.3222	8
2005	0.1086	0.4158	20
2006	0.1146	0.3362	71
2007	0.156	0.5945	182
2008	0.1651	0.4974	233
2009	0.1897	0.524	301
2010	0.2523	0.5081	580
2011	0.2849	0.5655	634
2012	0.349	0.5428	918
2013	0.6871	0.563	2149
2014	0.9091	0.5873	3206
2015	0.9153	0.5506	3143

After fitting the three products, regression equations and visualization graphics are obtained:

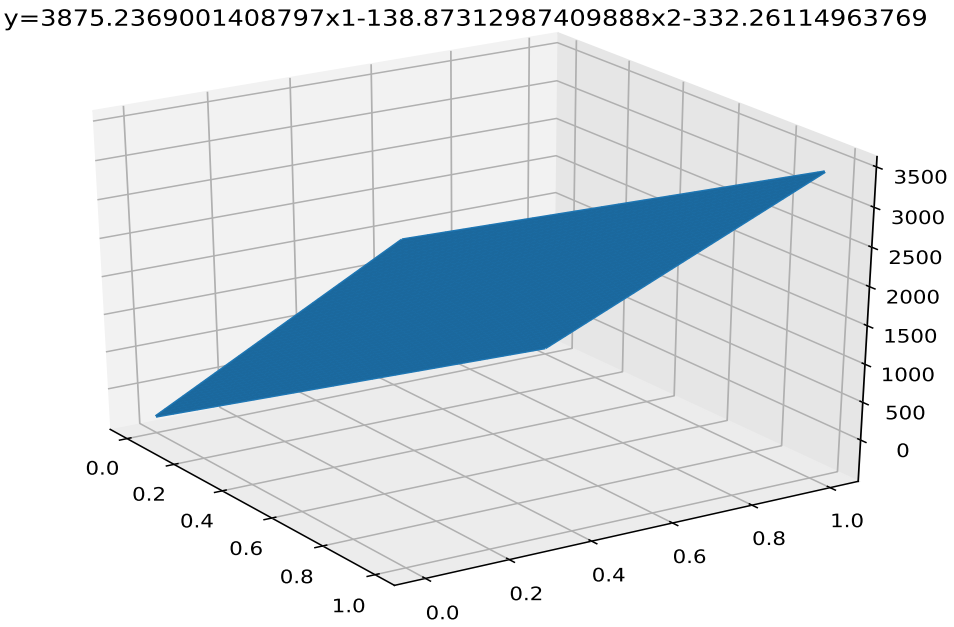

Figure 12: hair_dryer fitted graph

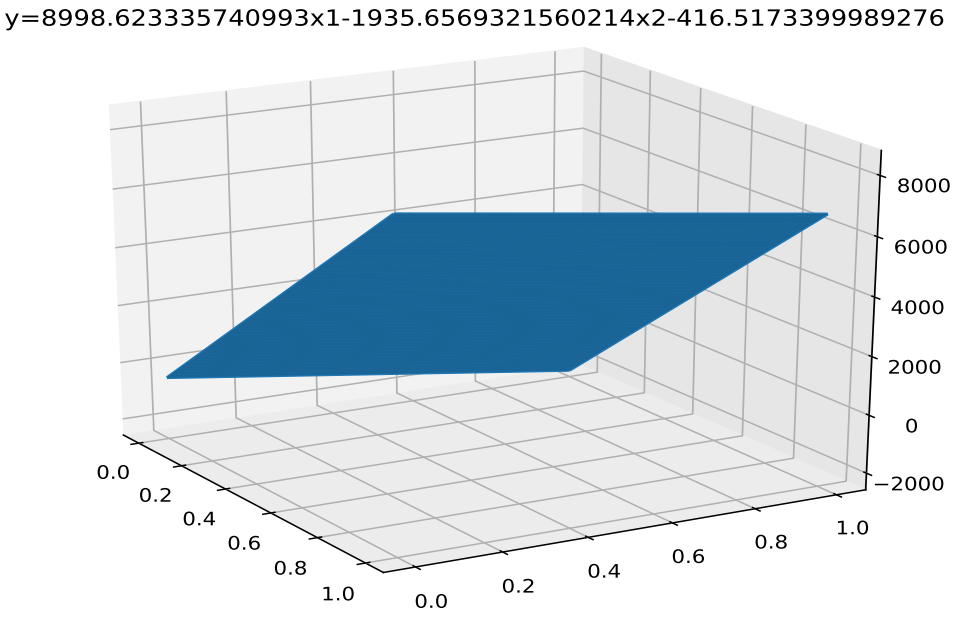

Figure 13: pacifier fitted graph

Hair_dryer and pacifier products have negative x_2 coefficients, that is, sales are negatively correlated with review scores. The reason is that the source of the data is different from the weight obtained. From the data point of view, the star rating is significantly positively related to the year, but the relationship between the review score and the year has been fluctuating, and the useful votes in the calculation of the review score have a greater weight on it, so it is easily affected by the subjective view of people in a certain period. People usually pay more attention to the negative reviews so that they pay more attention to the shortcomings of the product. So

the useful votes of many negative reviews will be higher than the simple praise. It can also be seen that people pay more attention to the substance of reviews, and less attention to star ratings. At the same time, the negative coefficient of x_2 is much smaller than the positive coefficient of x_1, so the degree of negative correlation is also small.

This reflects the mathematics and practicality of the model data, conforms to objective laws and has the function of reference prediction.

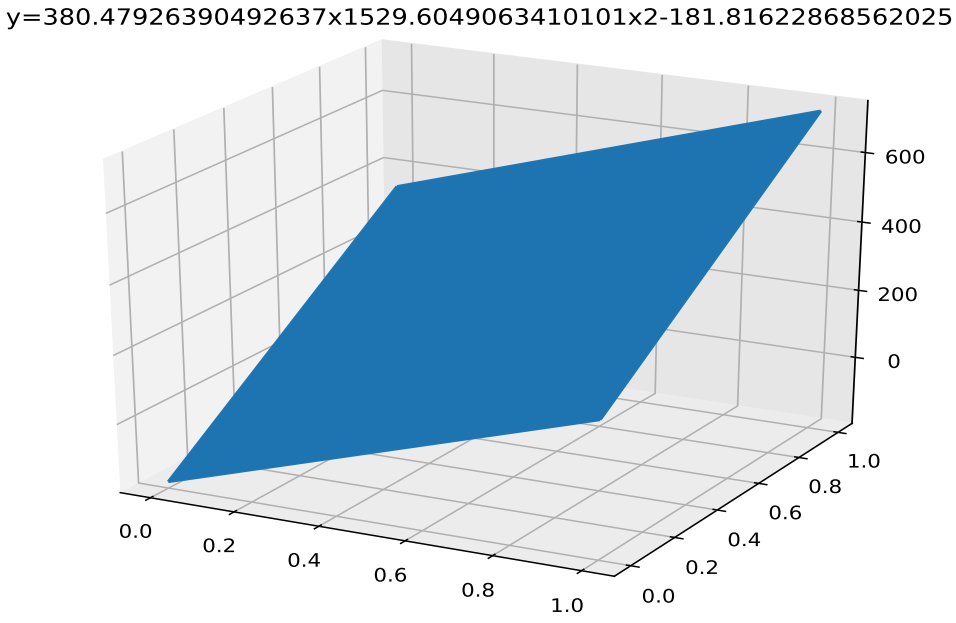

Figure 14: microwave fitted graph

5.2.4 d

In order to investigate whether people's comments are affected by the star ratings given by others before them, to reflect whether the buyer has a certain herd mentality.The star ratings and reviews have been quantified before and based on this we have drawn scatter plots for the three products.

In order to ensure that the graph clearly reflects the rules, the data volume is reduced, and a continuous time interval is randomly selected from each product to draw the "star rating-time diagram" and "review-time diagram".At the same time, for the convenience of observation, the pictures of each product will be combined into one picture and the time in the "Comment-Time Picture" will be delayed by three units to better observe the stars in the previous time period. So you can better observe the impact of star ratings in the previous period on review in the later period.

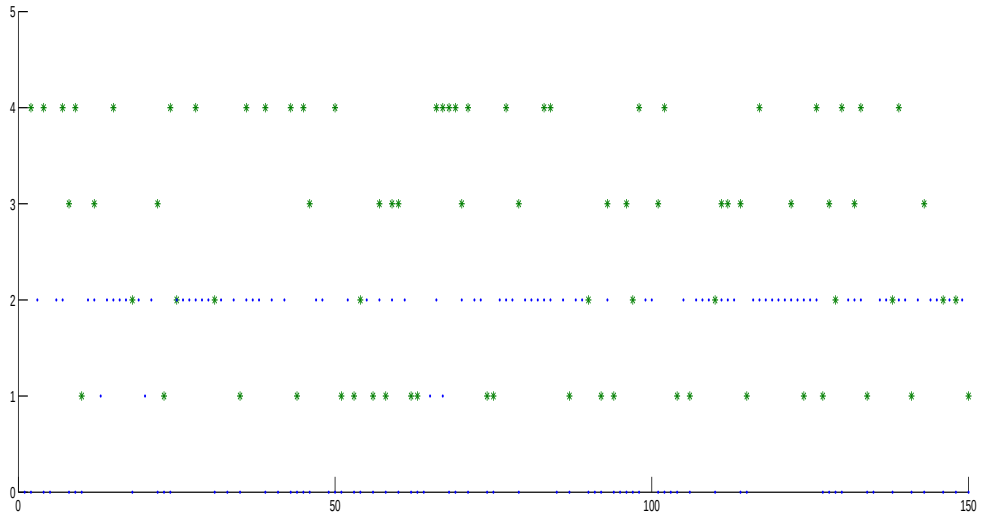

Figure 15: hair_dryer 'star rating"review'distribution

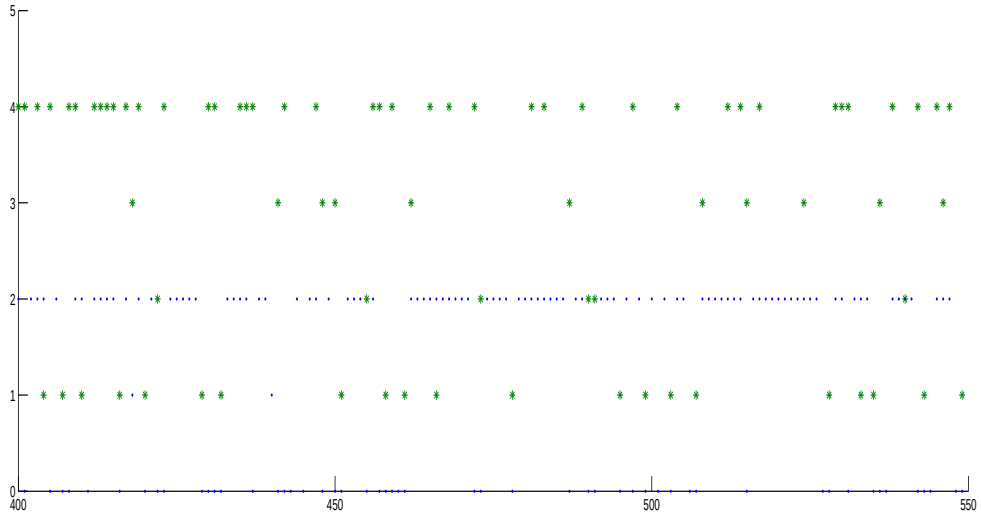

Figure 16: pacififer 'star rating"review'distribution

From the scatter plot, it can be seen that the continuous high star ratings in the previous time period will be accompanied by the continuous positive reviews in the later time period, while most of the low star ratings that appeared continuously in the previous period will be accompanied by the negative reviews that continuously appear in the later period. It shows that people's reviews will be affected by the star rating at the previous time and they have a certain herd mentality which also reflects people's high attention to reviews and star ratings. The statistical results are in line with reality and objective.

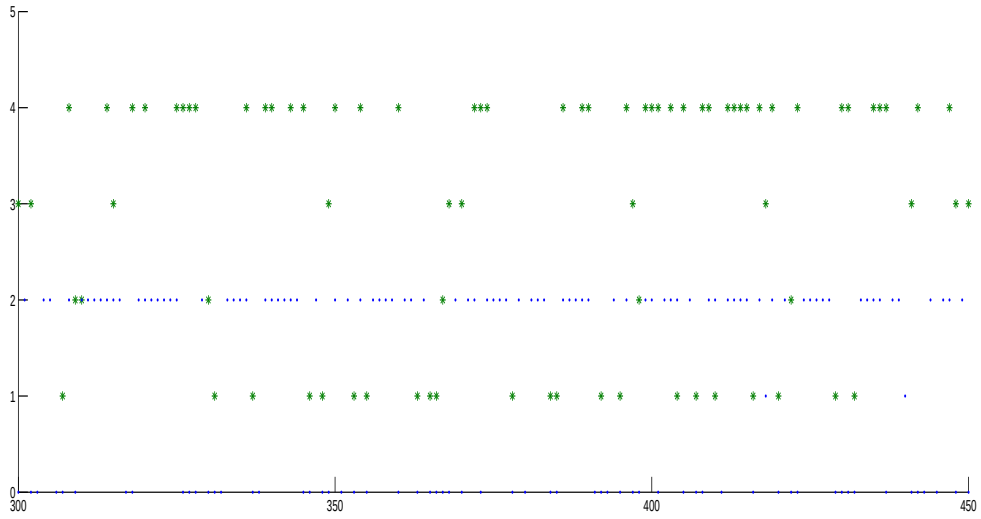

Figure 17: microwave 'star rating"review'distribution

5.2.5 e

To explore whether there is a certain relationship between review mood and star ratings, under normal circumstances we subjectively think that high star ratings are accompanied by positive reviews and low star ratings are accompanied by poor reviews.As mentioned earlier, we performed NLP on the comments to extract three moods: enthusiastic, middle disappointed and quantified them as 2,1,0. At the same time, the star rating is quantified to 5,4,3,2,1.Therefore, there are 15 types of correspondence between stars and reviews. We use 3D waveforms and scatter plots to explore the laws. Draw a three-dimensional image with reviews, stars as the x and y axes, and the corresponding sales volume as the z axis.

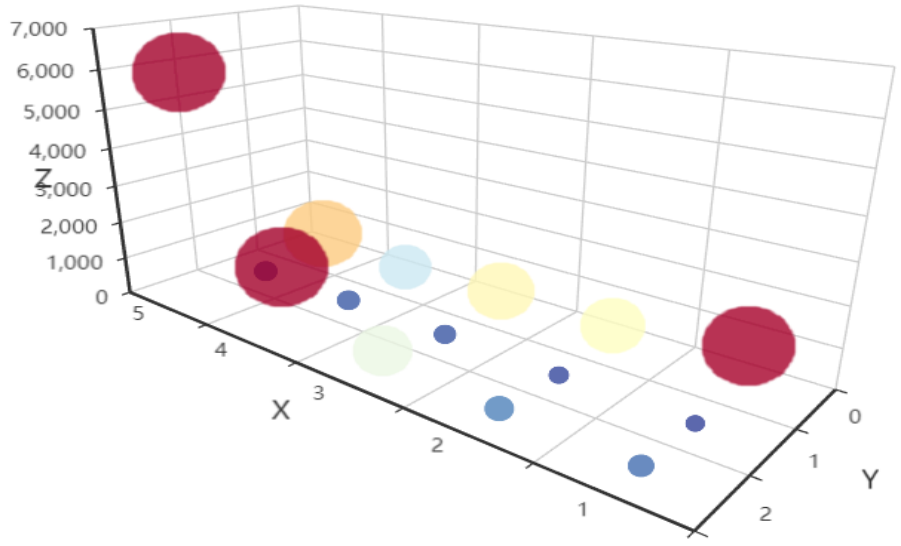

Figure 18: hair_dryer 3D Scatter plot

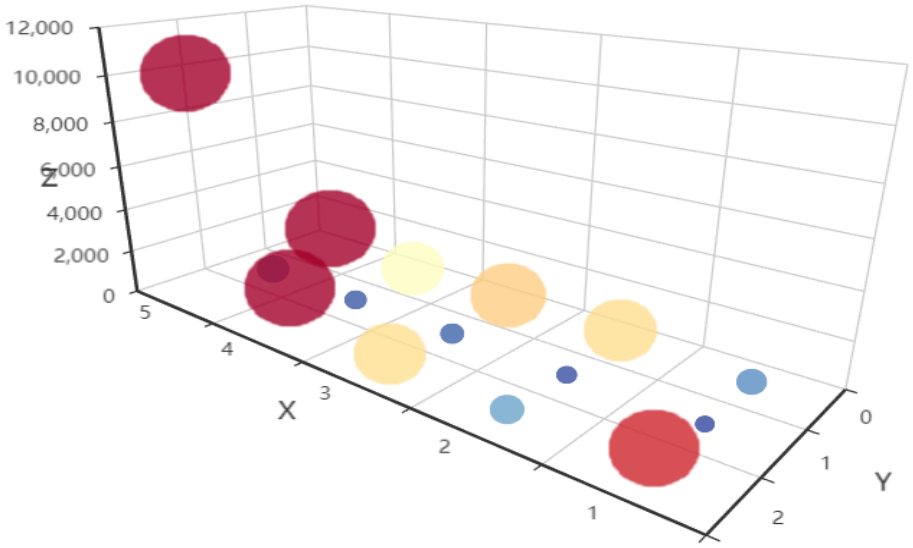

Figure 19: pacififer 3D Scatter plot

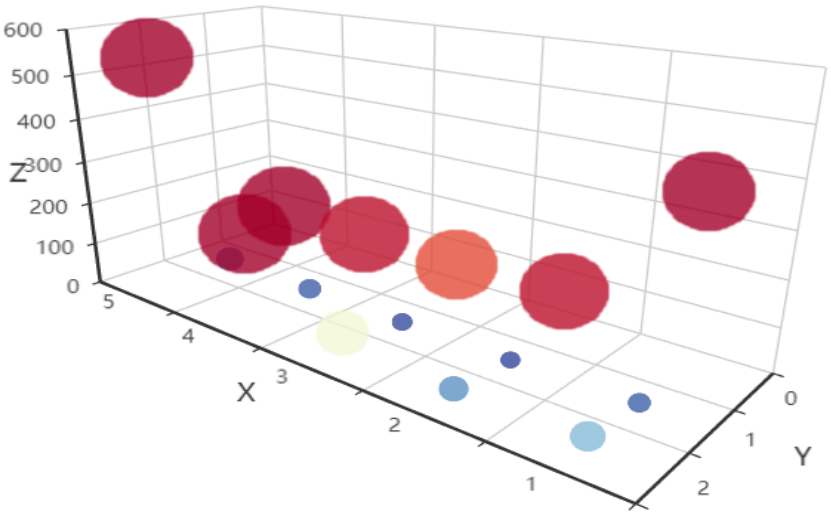

Figure 20: microwave 3D Scatter plot

It can be seen from the figure that the "five star-positive review" and "one star-negtive review" peaks in the waveform chart and is also highly in the scatter chart. So it proves that most of the data meets subjective cognition, that is, high star rating will be accompanied by positive reviews, low star ratings will be accompanied by negtive reviews. Star ratings are positively related to review mood.However, there are still a certain number of high star ratings and low star ratings. For example, the "five star-negtive rating" in hair_dryer also has a small peak. We analyze this part of the data may be that the buyer has not intentionally filled in a review or the appearance of robots and navies, leading to the phenomenon of brushing high, low star and positive, negtive reviews.

6 Proposal of Strategy

6.1 Decision on Priority Listing

In order to determine the best sales strategy, we categorize them by product_title and rank them. The top five product names with the highest sales volume and their sales data are as follows. The products in the table are far more competitive than similar products, so we recommend Sunshine to prioritize the products in Table 2.

Table 2: Sales ranking

	product_id	totle sales
hair_dryer	B003V264WW	555
	B0009XH6TG	535
	B003V264WW	424
	B000FS1W4U	407
	B00O0WRO08	363
microwave	B0052G14E8	394
	B005GSZB3M	79
	B005GSZB9Q	78
	B005GSZB7I	76
	B001QFYDSI	74
pacifier	b0045i6iay	833
	B003CK3LDI	521
	B00793CZAE	493
	B0028IDXDS	475
	B001FGL9X0	259

7 Sensitivity Analysis

Among the data given by Sunshine, there are many types of data and the relationship between the data is complicated which makes it difficult to process. Task b proposes time-based metrics and patterns to analyze product reputation. Among most methods of reputation analysis: Analytic Hierarchy Process, Fuzzy Cluster Analysis and other methods require that the processing staff have high experience in commodity analysis, so there is a defect that subjective factors are too strong. Objective weight analysis methods such as: Entropy Weight Method, Standard Deviation Method, CRITIC Method, etc. are also applicable to multi-object, multi-index weight analysis. In question b, we used the Entropy Weight Method to analyze the weights of reputation-related parameters "star rating" and "review". In order to perform the sensitivity analysis of the model, we changed the analysis method of reputation parameter weights and analyzed the weights using the Standard Deviation Method. Refit the product's 'reputation-time'plot. (9)(10)(11)(We make green lines)

The blue curve represents the "reputation-time diagram" obtained from the entropy weight analysis. The green curve represents the "reputation-time diagram" obtained from the entropy

weight analysis. It can be seen from the figure that the two methods are basically the same for reputation evaluation under the same time pattern.

8 Strength and Weakness

8.1 Strength

8.1.1 Regression Model

- Show a significant relationship between independent and dependent variables.
- Shows the influence intensity of multiple independent variables on a dependent variable.
- Can be used to predict subsequent data, and flexibly adjust according to the relationship between the variables and the actual situation.
- Modeling fast, no complicated calculations are needed, running fast even with large amounts of data.

8.1.2 Entropy Weight Method Model

- The Entropy Weight Method is an objective weighting method, it only depends on the discreteness of the data itself.
- It can avoid subjectivity when assigning weights, simple calculation formula and easy data processing.
- No conflict when used with other methods.

8.2 Weakness

8.2.1 Regression Model

- The regression model is relatively simple, and the algorithm is relatively low-level.
- The regression equation is only a speculation, which affects the diversity of the independent variables and the untestability of some independent variables, making the regression analysis limited in some cases.
- Fit non-linear data well. But we need to determine whether the variables are linear.

8.2.2 Entropy Weight Method Model

- If the change in the index value is small or suddenly becomes large or small, the entropy weight method has limitations.
- Entropy Weight accuracy decreases when time series data for unit indicators is insufficient.
- The use is limited and can only be used for weight distribution calculation.

9 Referance

References

[1] BaiducloudNLPAPI:https://ai.baidu.com/ai-doc/NLP/tk6z52b9z

[2] BaidutranslatorAPI:https://api.fanyi.baidu.com/doc/21

[3] EntropyWeightMethodtodeterminetheweightfromzhihu:https://zhuanlan.zhihu.com/p/28067337

[4] Multiplelinearregressionalgorithmandimplementation:https://www.cnblogs.com/python-frog/p/8934599.html

[5] Neuralnetworksandtheirtraining:https://www.cnblogs.com/royhoo/p/9149172.html

[6] ResearchonLinearRegressionModel:https://max.book118.com/html/2018/0628/7052032114001135.shtm

[7] Entropyweightsdetermineweights:https://zhuanlan.zhihu.com/p/28067337

10 A Suggestion Letter to Sunshine Company

Dear SunShine Company

Our team after a series of data processing. According to product_title for three products sales statistics, get the most competitive several products priority recommended sunshine company listed. In order to facilitate the evaluation of the market reputation of sunshine company after the product is put into the market, we use the Entropy Weight Method to assign weight to the parameters related to the reputation (stars, comments) for easy evaluation. Under the measurement mode based on time, the "time-reputation" curve is made to reasonably evaluate the reputation, so as to facilitate the market feedback of the listed products of sunshine company.After the product is launched, TOPSIS model is introduced to evaluate the success or failure of the product. We conduct "star rating" and "review rating" on the products on an annual basis, and then use the sales volume of linear regression fitting to evaluate the success or failure of the products, and obtain relatively ideal results.

After analyzing the correlation degree between star rating and comment, we find that there is indeed a "copycat effect". Because people pay more attention to the substance of the reviews when they buy products, they pay less attention to the stars, and a series of low stars can easily lead to bad reviews. Therefore, sunshine company should pay more attention to the product bad review content to improve the product after listing the product, so as to avoid too much bad review leading to the decline of public praise. At the same time, in most cases, high star with good, low star with bad, star and comment mood has a positive correlation. However, there is still a situation where the star rating is inconsistent with the comment, which may be the robot or the water army malicious comment. Therefore, it is recommended that sunshine regularly delete such comments without reference value.

Our team built a TOPSIS model to evaluate specific indicators of the product so that we could get market feedback based on a small amount of data after sunshine launched the product.

Moreover, the Entropy Weight Method used to deal with the weight of related parameters is based on the product data itself, which is very objective, making the analysis of our team very convincing.

Sincerely yours,

Your consultants