Research on Key Factors for Chinese Consumers to Purchase Sleep Aid Supplements based on Association Rules

Xiaoxin Zhuang, Shanwen Wang, Jialing Shen, Yi Chen, Wandian Hu

School of finance and economics, Jiangsu University, Zhenjiang 212013, China

Abstract

To explore the key factors that influence Chinese consumers to purchase sleep aid supplements, this paper uses the Apriori algorithm in association rules to analyze the correlation between the factors that consumers pay attention to when purchasing sleep supplements, and the correlation between consumer characteristics and key factors of concern. The study showed that the product's safety, ingredient, price and reputation were the most important factors influencing consumers to select sleep aid supplements, and female undergraduates aged from 18 to 25 with monthly living expenses from 1001 to 3000 CNY were more concerned with the safety factor.

Keywords

Sleep Aid Supplements; Association Rules; Apriori Algorithm.

1. Introduction

With the rapid pace of people's lives and the increasing work pressure, sleep problems are becoming more and more serious in China, and sleep health has become a public concern. According to *the 2022 China National Healthy Sleep White Paper*, more than 300 million people across the country have sleep disorders, and the COVID-19 pandemic has exacerbated sleep problems among young people in particular; 94% of Chinese people's sleep quality does not meet relevant health standards. The huge population with sleep disorders has given rise to a prosperous sleep economy market. According to the relevant data, the size of the sleep economy market in China has grown from CNY 261.63 billion to CNY 377.86 billion from 2016 to 2020, which has increased by 44.42%, and the total amount has been expected to exceed trillions of yuan, with steady growth. At the same time, in its market segments, the sleep aid supplement market has an average annual growth rate of 15%, and the penetration rate has been expected to further increase. As a result, the sleep aid supplement market is a blue ocean market with more certain growth prospects.

In view of the current market development of sleep aid supplements in China, the author collected 729 valid questionnaires as data support and used correlation rule analysis to mine the key factors for Chinese consumers to choose and purchase sleep aid supplements, in order to explore the potential of the sleep health products market and promote the quality development of the sleep aid supplement industry.

2. Literature Review and Theoretical Basis

2.1. Association Rules

Association rules are used to find frequent patterns, associations, correlations or causal structures between sets of items or objects in transactional data, relational data or other information vehicles [1]. It is an implication shaped as X→Y. X and Y are relatively called the antecedent and consequent of association rules. The association rule was first proposed in 1993 by scholars Agrawal, Imieliński and Swami, in order to solve the problem of market basket analysis [2], i.e. to analyze the shopping habits of customers from the association between

different products placed in their shopping baskets, so as to rationalize the placement of goods and facilitate customer selection[3]. The discovery of this rule helps businesses to identify which categories of goods are mostly purchased by consumers at which times of the day, so that they can target their sales patterns and plans to maximize profits with better sales strategies.

Association rule analysis uncovers frequent dependencies and correlations from large amounts of data, which is of great importance to real-world business decisions. To develop more appropriate marketing strategies, increase product sales, improve the consumer shopping experience and find potential consumers, merchants often use association rule analysis in shelf merchandising, shop cross-category recommendations, joint promotional activities and so on.

2.2. Apriori Algorithm

The Apriori algorithm is used to mine frequent itemsets for Boolean association rules, whose core is based on a recursive algorithm of the idea of two-stage frequency sets; it is a single-dimensional, single-level and Boolean association rule in terms of classification. The set of all items with support greater than the minimum support is called the frequent itemset, or frequency set for short.

The basic idea of the algorithm is listed as follows:

(1) Find all the frequency sets in the data transaction set by recursive methods[4]. (2) Generate strong association rules from the frequency sets that satisfy the minimum support and minimum confidence levels[5].

(3) Use the aforementioned frequency set to generate expectation rules, and generate all rules that contain only the items of the set, where each rule has only one item in the right-hand part[6]. If the above rules are generated, only those rules that are greater than the minimum confidence given by the user need to be retained.

The Apriori algorithm is used in a wide variety of fields to analyze and mine data for correlations, which can facilitate making strategic decisions effectively.

3. Association Analysis of Key Factors in Consumers' Purchase of Sleep Aid Supplements

3.1. Data Preparation

Table 1. Fact sheet on concerned factors for consumers purchasing sleep aid supplements

Variables	1	2	3	...	297	...	544	...	711	...
Package	0	0	0		1		0		1	
Price	1	1	1		1		1		1	
Ingredient	1	1	0		1		1		1	
Taste	0	1	1		1		1		0	
Safety	1	1	1		1		1		0	
Effective Speed	0	0	0		1		1		1	
Edible Way	0	1	0		0		0		1	
Advertisement	0	0	0		0		1		1	
Brand Awareness	0	0	0		1		1		1	
Spokesman	0	0	0		0		1		0	
Reputation	1	0	0		0		0		0	
Purchasing Convenience	0	0	0		0		0		0	
Aftersales Service	0	0	0		0		0		0	
Others	0	0	0		0		0		1	

A total of 729 valid questionnaires were collected for this study. Question 19 was used as the primary data, and the 14 factors in the questionnaire were drawn from relevant research at home and abroad. The data was presented in a fact table with the variable name being the concerned factors for consumers purchasing sleep aid supplements, with the variable values taken as 0 and 1, where 0 means that consumers will not pay attention to such factors when purchasing sleep supplements and 1 means that consumers will pay attention to such factors when purchasing sleep supplements. The 14 factors in the questionnaire are presented in Table 1 below.

3.2. Data Pre-processing

The above data was stored in an SPSS file named "Consumer Data. sav" and the SPSS Modeler software was opened to collect, process and model the data in the flow editing area. Remove unwanted fields from the "Consumers Data. sav" through the filter, read the data types and preview the original data records in the table in preparation for building the web and model.

3.3. Web Drawing

Before building the association rule model node, the "web" node is added to connect the web node to the filter node. The web allows us to visualize the extent to which the factors that consumers pay attention, giving us a direct sense of the relationship. As shown in Figure 1, the thickness and depth of connecting lines represent the strength of the association, from which it can be seen that factors such as safety and ingredient, price and safety, and ingredient and price are strongly associated.
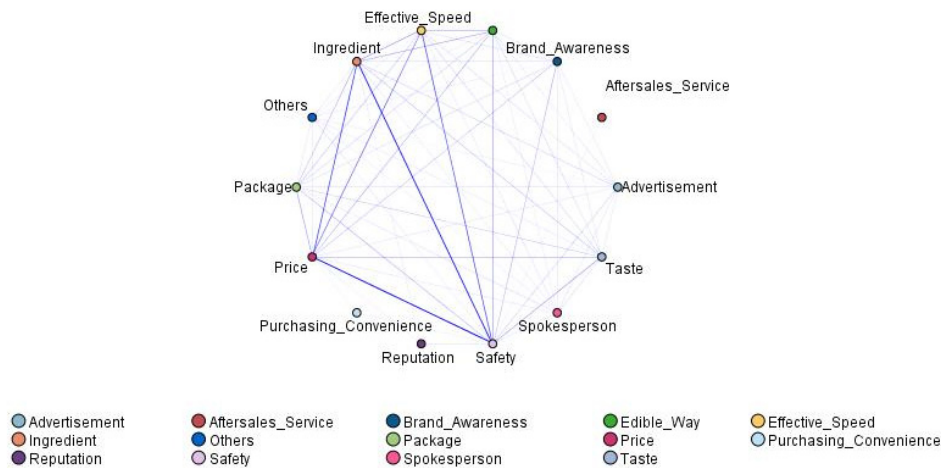

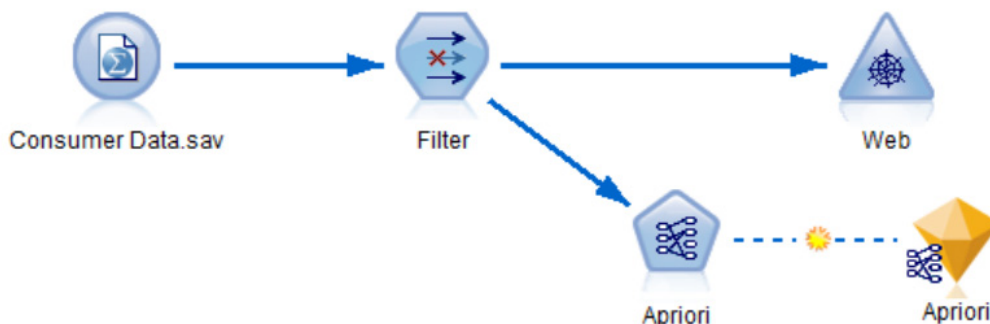**Figure 1.** Web

3.4. Model Building

Figure 2. Association rule flow diagram

Based on the data pre-processing, the association rule model was constructed through the Apriori algorithm in SPSS. The Apriori nodes were dragged to the data flow editing area, connections between the filter and Apriori node were established, and the minimum support, minimum confidence and maximum antecedent were set. The association rule flow diagram is shown in <u>figure 2</u> below.

Based on the results shown in <u>Table 2</u>, 27 valid association rules were found when the minimum support threshold was 3%, the minimum confidence threshold was 80%, and the maximum antecedent was 5. When the minimum support threshold became 4% and the rest of the conditions remain unchanged, 7 valid association rules could be found.

Table 2. Output of Apriori algorithm analysis (minimum support threshold of 4%)

Back items	Previous items	Percentage of Support	Percentage of Confidence
Safety	Reputation Price	5.075	91.892
Safety	Reputation Ingredient	4.938	91.667
Safety	Brand awareness Ingredient Price	4.39	81.25
Price	Effective Speed Ingredient Safety	7.819	80.702
Safety	Brand awareness Effective Speed Price	4.115	80.0
Effective Speed	Edible Way Packaging Price	4.115	80.0
Safety	Edible Way Effective Speed Ingredients Price	4.115	80.0

3.5. Analysis Results

When the minimum support threshold was set at 4%, the minimum confidence threshold at 80%, and the maximum antecedent at 5, the 7 valid association rules were mined. For example, rule 1: reputation, price → safety, a support level of 5.075% means that 5.075% of consumers will pay attention to reputation and price factors when buying sleep supplements; a confidence level of 91.892% means that when consumers pay attention to reputation and price selecting sleep supplements, they will also pay attention to safety in 91.892% of cases. For another example, rule 4: effective speed, ingredient, safety → price, a support level of 7.819% indicates that 7.819% of consumers will pay attention to effective speed, ingredient and safety when buying sleep supplements: a confidence level of 80.702% indicates that when consumers pay attention to effective speed, ingredient and safety selecting sleep supplements, they will also pay attention to price in 80.702% of cases. Similarly, the other rules also reveal the correlation between the factors that consumers pay attention to when purchasing sleep supplements, which have a strong practical value. From the output, we can assume that safety, ingredient, price and reputation are closely related.

4. Association Analysis between Consumer Characteristics and Key Factors in Consumer Purchase of Sleep Aid Supplements

Based on the above results, we further analyzed the links between consumer characteristics and the factors of safety, ingredient, price and reputation.

4.1. Data Preparation

The consumer characteristics in this study consisted of gender, age, academic qualification, occupation and monthly income. The influencing factor data on sleep supplements consumers remained as shown in Table 1 above and the consumer characteristics were shown in Table 3 below.

Table 3. Fact sheet on the consumer characteristics of sleep aid supplements

	Gender	Age	Academic qualification	Occupation	Monthly income
1	Female	18-25	Undergraduate	Student	CNY 0 - 1000
2	Male	Over 50	Lower Secondary	Worker	CNY 3001-5000
3	Female	18-25	Undergraduate	Student	CNY 1001-3000
…					
297	Female	Over 50	Undergraduate	Individual businessman	CNY 1001-3000
…					
544	Male	31-40	Undergraduate	Individual businessman	CNY 3001-5000
…					
711	Male	26-30	Tertiary	Others	CNY 3001-5000
…					

4.2. Correlation Analysis

According to the output results, when the minimum support threshold was set at 10%, the minimum confidence threshold at 80%, and the maximum antecedent at 5, two valid association rules could be mined, as shown in Table 4.

Table 4. Apriori algorithm analysis output results

Back items	Previous items	Percentage of support	Percentage of Confidence
Safety	Occupation = 1 Monthly income = 2 Academic qualification = 5 Gender = 2	10.974	81.25
Safety	Occupation = 1 Monthly income = 2 Academic qualification = 5 Gender = 2 Age = 2	10.974	81.25

Note: The numbers in the previous items above are SPSS data codes and can be referred to the previous data descriptions, e.g. gender = 1 for male and gender = 2 for female.

4.3. Analysis of Results

With the two association rules obtained above, we can see that consumers whose occupation is student, monthly income is CNY 1001-3000, academic qualification is bachelor's degree, gender is female and age is 18-25 years old will pay attention to the safety factor when buying sleep aid supplements. We can assume that since the consumer group of health care products in China tends to be younger, the post-90s and post-00s gradually become the main consumers. Meanwhile, women pay more attention to health and product safety than men, so female undergraduates with monthly living expenses from 1001 to 3000 CNY will pay more attention to the safety factor.

5. Conclusions and Recommendations

The association rules show that consumers are more concerned about safety, ingredient, price and reputation when choosing to buy sleep aid supplements. At the same time, female undergraduates aging 18-25 years old and living on a monthly budget of CNY 1001-3000 pay more attention to the safety factor.

Based on these findings, we make the following recommendations for sleep aid supplement companies.

Firstly, enrich the product range of sleep supplements. Based on the results of the questionnaire analysis, we found that Chinese consumers have a certain distrust of the science, usefulness and safety of sleep supplements in the present stage, but "safety" is still one of the most important factors in the purchase process. Therefore, sleep aid supplement companies should invest more in product development, promote the diversification and naturalization of sleeping ingredients, and improve the credibility and trust of sleep aid supplements.

Secondly, promote the rationalization of product price. According to the output results, price is one of the key factors for consumers to buy sleep supplements. Therefore, in terms of price strategy, sleep aid supplement companies should use differential pricing methods for different categories of sleep health products. In addition, for the consumers' various needs and acceptable degrees of differential pricing, companies ought to improve the cost performance of sleep supplements and the overall satisfaction of consumers.

Third, open up the market for the young group. According to the analysis results, young groups aged from 18 to 25 are increasingly focused on health and wellness, so for sleep aid supplement companies, the target customer groups are not only the middle-aged and elderly, but also need to fully consider the new generation of young groups, especially women's groups, to fully cater to the consumer mentality. Moreover, companies should develop personalized and novel products and improve the sense of consumer experience for young groups.

References

[1] M. Niu. A Basic Study of Data Mining Methods and Functions[J]. Computer Knowledge and Technology, 2018,14(14):6-7.

[2] R. Agrawal, T. Imieliński, A. Swami. Mining Association Rules between Sets of Items in Large Databases[J]. Acm Sigmod Record,1993,22(2):207-216.

[3] J.W. Han. Data Mining Concept and Technology[M]. Beijing: China Machine Press, 2012.

[4] R. Agrawal, R. Srikant. Fast Algorithms for Mining Association Rules[C]. Proceedings of the 20th International Conference on Very Large Data Bases. IEEE, 1994.

[5] W. Xue, H. Chen. Data Mining based on SPSS Modeler[M]. People's University of China Press,2014.10.

[6] C.Y. Wang. Research and Application of Association Rule Apriori Algorithm[D]. North China Electric Power University (Beijing),2018.