

Fresh E-Commerce Platform Consumer Status and Market Opportunity Mining Survey-Taking Chongqing as an Example

Lu Jiang, Ying Guo*, Wanting Huang, Yong Dai

Chongqing Institute of Engineering, Chongqing, China

**Corresponding author*

Keywords: Fresh e-commerce, consumer satisfaction, correlation analysis, k-means clustering analysis

Abstract: With the economic development and the rapid improvement of national living standards in the Internet society, the consumption level and concept of consumers have also changed. Affected by the traditional consumption habits of residents, the fresh market is mainly based on traditional offline sales. Major e-commerce companies and traditional fresh channels have laid out fresh e-commerce markets. Since the so-called "first year of fresh e-commerce" in 2012, and a large amount of capital has leveraged the market to grow rapidly. However, while the scale is increasing year by year, a series of development bottlenecks inevitably appear in the industry. Therefore, in order to analyze the consumption status and market opportunities behind the fresh food market, based on the results of the questionnaire survey, this paper analyzes the respondents of different genders, ages and occupations to understand their understanding and use of the fresh food e-commerce platform.

1. Introduction

In recent years, China's fresh retail market has maintained steady growth. In 2020, the scale of China's fresh retail market exceeded CNY 5 trillion. As one of China's basic consumer goods, fresh products are expected to continue to grow in the future with the increase of per capita disposable income and consumption expenditure. By 2025, the scale of China's fresh retail market will reach CNY 6.8 trillion.

Affected by the new coronavirus epidemic and the 'house economy', China's fresh e-commerce market developed rapidly in 2020, and the user penetration rate soared sharply as shown in Figure 1. In 2021, the penetration rate of fresh e-commerce reached 15.8%. With the development of fresh e-commerce and the maturity of the model, the cultivation of users' online shopping habits, the increasingly wide coverage of fresh e-commerce users and the increasingly mature technology.

Fresh food has the characteristics of high frequency rigid demand and is a non-standard product with instant demand. Offline channel is the main retail channel of fresh food. With the increase of online retail year by year, the continuous penetration of the Internet, and the gradual development of consumers online consumption habits, online consumer groups are younger, and consumption pays

more attention to convenience and distribution speed. At present, fresh e-commerce has entered the stage of both economic model optimization and scale growth. In the future, fresh e-commerce should explore a sustainable economic model. It is expected that fresh e-commerce will continue to maintain rapid growth for a period of time in the future. By 2023, the scale of fresh e-commerce industry will exceed trillions, and the multi-mode fresh e-commerce situation will continue.

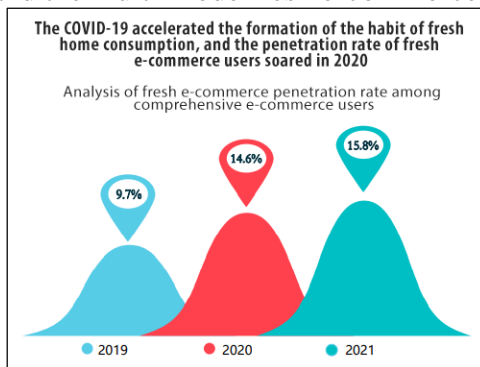


Figure 1: Analysis of the penetration rate of fresh e-commerce among comprehensive e-commerce users.

2. Survey General Basic Situation

The basic situation of the sample is: the number of respondents in the survey sample is 500, after screening out the invalid questionnaire, the remaining 478 questionnaires, and then screening out the invalid questionnaire, the remaining 453 questionnaires, the effective utilization rate of the overall questionnaire is 90.6 %.

According to the fifth question of the questionnaire ' Have you used the fresh e-commerce platform to buy fresh products ', the survey results obtained are shown in Figure 2. There are 274 people (60.4 %) who have used the platform and 179 people (39.6 %) who have not used it.

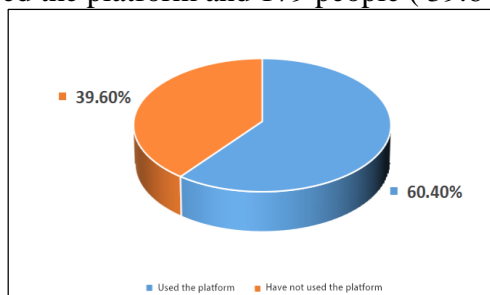


Figure 2: Proportion of use.

2.1. Ways to Understand the Fresh Platform

The results of the questionnaire survey show that as shown in Figure 3, 31.54 % of the population are fresh e-commerce platforms learned through online advertising, followed by friends or their own initiative to try, through the rest of the way. The proportion is relatively small.

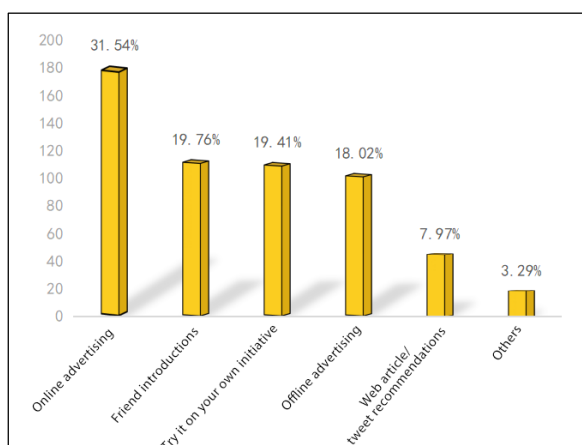


Figure 3: Understanding the proportion of ways.

2.2. Reasons for Using Fresh E-Commerce Platform

As shown in Figure 4, the results show that the reasons why most people use the platform are 'affordable', 'convenient and fast', followed by 'delivery speed'.

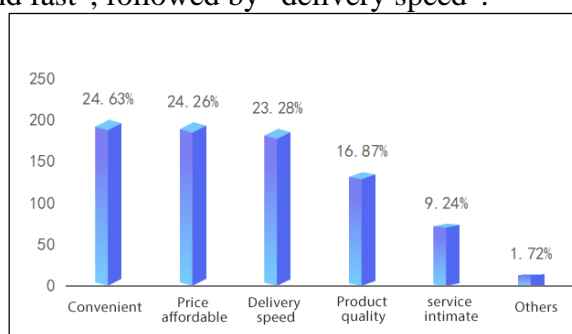


Figure 4: Proportion of reasons for use.

It is worth noting that the choice of product quality is relatively low, and the service is only less than 1%. This shows that the product quality and service of the e-commerce platform are not generally recognized by consumers.

2.3. Number of Purchases and Average Cost

We investigate the number and average cost of fresh food purchased by users per week, as shown in Figure 5 and Figure 6. The results show that most people have 1-2 purchase records per week, followed by 3-4 times. More people spend less than 50 yuan on average, and only 3% spend more than 200 yuan.

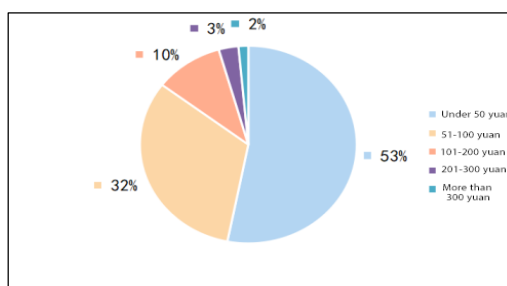


Figure 5: Weekly purchases.

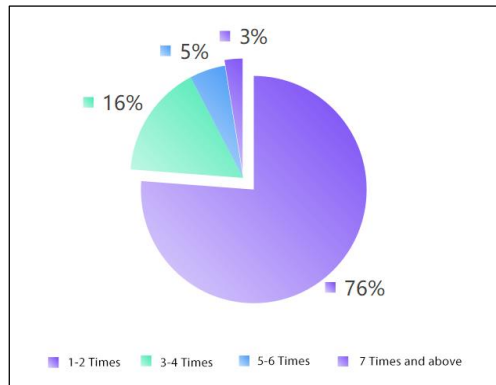


Figure 6: Average weekly expenditure.

2.4. Purchase of Fresh Types

We set the topic ' What products do you usually buy on the fresh e-commerce platform? According to the survey results, as shown in Figure 7, the most purchased was fruit, accounting for 30.38 %, followed by vegetables, accounting for 23.8 %.

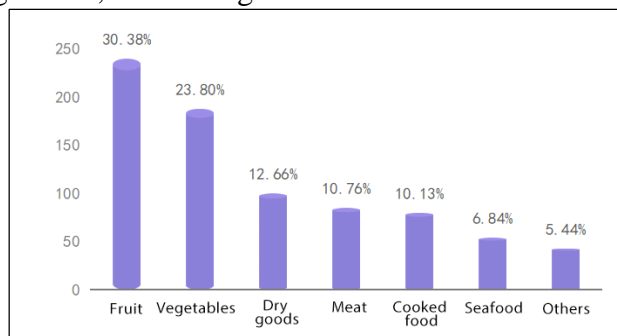


Figure 7: Proportion of fresh species purchased.

2.5. Advertising Factors that Influence Consumers ' purchasing Decisions

As shown in Figure 8, it is noted that most consumers buy goods because of store discounts, followed by short video campaigns, while only a few decide to buy through television and newspapers.

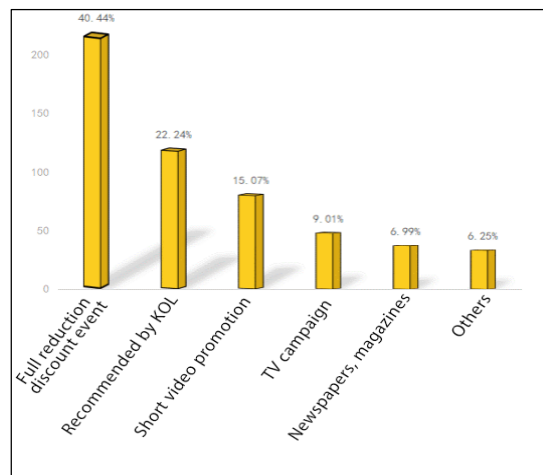


Figure 8: Proportion of advertising types affecting decision making.

3. Investigation Results and Cause Analysis

3.1. Logistic Regression Analysis of Consumer Satisfaction

Define a mathematical statistic odds, called odds ratio, and its relationship with probability is:

$$\text{odds} = \frac{p}{1-p} = e^{f(x)} \quad (1)$$

It can be found that odds is not a linear model and cannot be analyzed by linear regression. Considering the nature that natural logarithm can reduce the order of the index, we take the natural logarithm on both sides of the above formula, and have:

$$\ln \frac{p}{1-p} = f(x) = \beta_0 + \sum_{i=1}^k \beta_i X_i \quad (2)$$

There are six aspects of the use satisfaction scale, of which five aspects are related to consumers' satisfaction with the fresh e-commerce platform, namely: platform use characteristics, expected satisfaction, product satisfaction, after-sales satisfaction and logistics satisfaction. Each aspect corresponds to several items. We take a project as an independent variable and consumer satisfaction as a dependent variable. Through the Logistic regression model, we construct a multiple linear function relationship between multiple independent variables and a dependent variable, so as to predict the overall satisfaction of consumers.

We divided the five subscales of satisfaction into 1, 2 and 3 as 'unsatisfactory', 4 and 5 as 'satisfactory'. Encode the project as shown in Table 1.

Satisfaction. In the model, the dependent variable is consumer satisfaction, the item code corresponding to 'satisfaction' is '1', and the item code corresponding to 'dissatisfaction' is '0'. The item code ABCDE in the above table is a binary variable, non-0 is 1, which is used to explore the impact of project factors on consumer satisfaction^[1].

We import the survey data into SPSS for analysis^[2]. In the results table, we focus on the two tables of 'Omnibus Tests of Model Coefficients' and 'Hosmer and Lemeshow Test', and extract the key parameter test values of these two test results. The results are shown in Table 2.

Table 1: Consumer satisfaction items.

aspect	project	project code	aspect	project	project code
platform usage feature	save time, convenient and quick	A ₁	expectation satisfaction degree	can buy the required products	B ₁
	more discounts, cheaper prices	A ₂		the product quality achieves the expectation.	B ₂
	more safe	A ₃		the platform service meets expectations	B ₃
product satisfaction degree	product freshness is good	C ₁	after-sales satisfaction degree	smooth and diverse communication channels	D ₁
	rich variety of products	C ₂		handle after-sales problems timely	D ₂
	the product appearance is beautiful	C ₃		after-sales solution is reasonable	D ₃
physical distribution satisfaction degree	product delivery is intact.	E ₁			
	accurate delivery arrival time	E ₂			
	distribution personnel service is good	E ₃			

Table 2: Parameters for the fitness test of the overall model (platform usage characteristics)

test method	monitor value	P
Omnibus	17.334	0.001
Hosmer Lemeshow	2.257	0.632

According to the analysis results, the chi-square value of the Omnibus test was 17.334, P = 0.0010.05, indicating that the information in the current data was fully extracted and the model fit was high. The results show that the independent variable can effectively explain the dependent variable.

Therefore, we analyze the expected satisfaction, product satisfaction, after-sales satisfaction and logistics satisfaction according to the same method. The probability that consumers are satisfied with the fresh e-commerce platform only because ' product distribution is intact ' is 0.604, the probability that consumers are satisfied with the fresh e-commerce platform only because ' delivery arrival time is accurate ' is 0.441, and the probability that consumers are satisfied with the fresh e-commerce platform only because ' distribution personnel serve well ' is 0.547. Among them, the probability of making consumers satisfied with the fresh e-commerce platform by the two factors of ' product distribution intact ' and ' good service of distribution personnel ' is the highest, 0.648 ; at the same time, the probability determined by the three factors is 0.593, and the significance is also large.

3.2. Evaluation Model of Commodity Correlation Degree Based on Bayesian Theorem

We will form a combination of seven kinds of goods in pairs (including the combination formed between the goods themselves, the combination of goods (A, B) and (B, A) to calculate the two combinations, the purpose is to facilitate the preparation of the solution program, eliminating unnecessary data processing) Substituted into the correlation model equation, such a combination of a total of 49 kinds^[3]. We take the combination of vegetables and fruits as an example to illustrate the process of model solving.

Vegetables in the first row of the purchase matrix, N₁ is 188; fruit in the second row of the purchase matrix, N₂ is 240 ; through data screening by MATLAB, it is concluded that the number of times that vegetables and fruits are 1 at the same time is 170^[4]. We put these three parameters into the relational degree model equation, there are:

$$K_{12} = \frac{N_{12}^2}{N_1 N_2} = \frac{170^2}{188 \times 240} = 0.64 \quad (3)$$

Therefore, the correlation between vegetables and fruits is 0.64. In this way, we calculate the correlation degree between any two kinds of fresh products, and put the obtained correlation value K_{ij} in a matrix as follows. We call this matrix the correlation matrix as shown in Table 3:

Table 3: Correlation matrix (two-dimensional).

$\begin{matrix} j \\ i \end{matrix}$	1	2	3	4	5	6	7
1	1.00	0.64	0.34	0.16	0.17	0.22	0.05
2	0.64	1.00	0.25	0.16	0.26	0.37	0.12
3	0.34	0.25	1.00	0.35	0.15	0.10	0.01
4	0.16	0.16	0.35	1.00	0.21	0.07	0.02
5	0.17	0.26	0.15	0.21	1.00	0.24	0.06
6	0.22	0.37	0.10	0.07	0.24	1.00	0.09
7	0.05	0.12	0.01	0.02	0.06	0.09	1.00

It can be seen that the value of the main diagonal of the correlation matrix is 1, because the commodity corresponding to the main diagonal is the correlation analysis between itself and itself,

so the correlation value is 1. In addition to the main diagonal, it can be found that the maximum correlation value is 0.64, corresponding to $i = 1, j = 2$ or $i = 2, j = 1$, that is, the value of K_{12} or K_{21} is 0.64, which corresponds to vegetables and fruits in the purchase matrix^[5]. That is to say, vegetables and fruits have the highest correlation.

3.3. K-means Clustering Analysis Model of Potential Consumers

We use Python to cluster potential consumers into four categories^[6]. The results are shown in Table 4.

Table 4: Cluster center table

Potential consumer categories	1	2	3	4
age	age between 18 to 25	age between 18 to 25	age between 26 to 35	age between 26 to 35
sexuality	female	males	female	female
income	3000 yuan and below	3000 yuan and below	5001 - 8000 Yuan	8000 Yuan and above
occupation	pupil	pupil	Enterprise workers	civil servant
purchase intention	Will use	Will use	considered to use	considered to use

According to the results of the above table, we make a feature analysis based on the characteristics of the four groups, as shown in Figure 9:

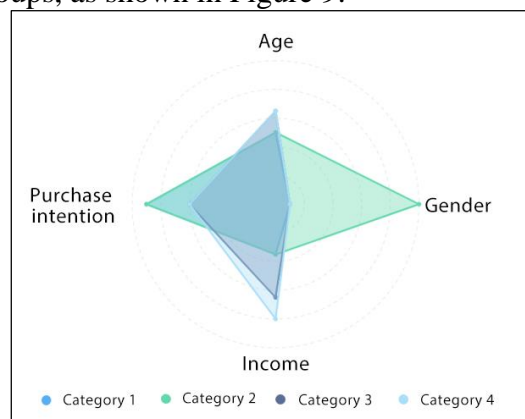


Figure 9: Characteristic analysis

According to the cluster center table and the feature analysis diagram, we define the four categories of consumers as: important potential consumers, more important potential consumers, secondary potential consumers and general potential consumers. Next describe their basic characteristics^[6].

3.3.1. Important Potential Consumers and More Important Potential Consumers

This type of consumer is a category 1 consumer, characterized by: 18-25 years old, student occupation, income below 3000, and said to use the fresh shopping platform. Consumers in this category have strong self-learning ability, are easy to accept new things, and are willing to try new things. The clustering results show that these users are the most potential consumers.

3.3.2. More Important Potential Consumers

This type of consumer is a category 2 consumer, characterized by: 18-25 years old, male, student, income below 3000, and said to use the fresh shopping platform. Consumers in this category have strong self-learning ability. The only difference between them and important potential consumers is that the gender is male. The clustering results show that such users are more important potential

consumers.

3.3.3. Secondary Potential Consumers

This type of consumer is a category 3 consumer, characterized by: 26-35 years old, female, professional enterprise workers, income of 5001-8000 yuan, and said it would consider using the fresh e-commerce platform. This category of consumers is in the golden period of work and the rising period of career, and has a certain economic foundation. However, considering their busy work, they may not spend a lot of time to understand the use of fresh e-commerce platforms. Compared with the previous two categories of consumers, the use probability of such customers may be lower, so the clustering results show that such consumers are secondary potential consumers.

3.3.4. General Potential Consumers

This type of consumer is a category 4 consumer, characterized by: 26-35 years old, female, civil servant, income of more than 8000 yuan, and said it would consider using the fresh e-commerce platform. This category of users generally work in public institutions and national government agencies, and have a certain economic foundation, but they do not use the fresh e-commerce platform much. This is because of the nature of their work. There may be untrustworthy factors for the e-commerce platform, which leads to the general potential of such people to use the fresh e-commerce platform in the future. Therefore, the clustering results show that such users are general potential consumers.

Of course, for people outside the above four categories, their potential to use fresh e-commerce platforms in the future will be relatively lower.

4. Conclusion and Suggestion

4.1. Conclusion

In recent years, with the development of science and technology and the change of consumer demand, many types of e-commerce platforms have emerged. The emergence of the epidemic situation and the 'house economy' has given fresh e-commerce platform development opportunities. The influx of a large amount of capital into the fresh e-commerce layout market has a certain impact on people's lives and the development of the national economy. Therefore, we investigate the understanding and consumption of fresh e-commerce in different genders, ages, occupations, and draw conclusions and relevant suggestions.

4.1.1. The Overall Use Degree Needs to be Improved, and the Reliability Propaganda is Worth Digging

It can be found from the respondents that nearly 40 % of the respondents have not used the fresh e-commerce platform. Compared with online, they prefer offline fresh physical stores. When investigating the reasons for not using the fresh e-commerce platform, it was found that most of them were lack of understanding of the fresh e-commerce industry and whether there was a market around to provide related services. From the analysis of influencing consumers' purchasing decisions, it can be seen that most of the used fresh e-commerce surveys are more affected by acquaintances' recommendation, online advertising, physical store activities and their own initiative, while offline advertising and online article recommendation. The number of people in other ways is small, indicating that the influence of offline advertising is low for consumers of fresh e-commerce

in Chongqing, and the reliability propaganda of fresh e-commerce is not enough.

4.1.2. Fresh E-Commerce Platform Tangible Services are Satisfactory

From the surveyed members, it was found that most of the investigators liked the price and time spent on fresh e-commerce and the speed of delivery. With the development of the economy, people's pace of life is getting faster and faster. Consumers choose fresh e-commerce because they are busy. When they do not have time, they will buy in fresh e-commerce, and the platform price is lower than the offline fresh entity price. At present, China's logistics coverage is high and the distribution speed is fast. With the help of the existing logistics foundation, consumers can meet the distribution speed.

4.1.3. The Empathy Service of Fresh E-Commerce Platform still Needs to be Improved

In the investigation and analysis of consumer satisfaction, we can find products that can meet the needs of consumers and the quality of products can meet expectations is crucial to the satisfaction of the entire platform, and the impact of single product freshness on consumer satisfaction Fresh e-commerce platform is more significant. In the analysis of after-sales evaluation of purchased products and the analysis of investigators who did not use fresh e-commerce platform, we can also find that product freshness and after-sales service have a significant impact on fresh platform.

For fresh products, most of the after-sales disputes are difficult to solve the problem of deterioration of products received by consumers. Fresh goods have a short shelf life and products are prone to loss. Fresh goods are generally produced in rural areas, and their consumers are mainly concentrated in cities. There is a spatial distance from the upstream base of the supply chain to the end users. How to maintain the freshness of products and reduce the after-sales disputes caused by deterioration is the key to the development of cold chain logistics technology of fresh products, which is a difficult problem that China needs to think about and overcome at this stage.

4.1.4. Purchase Times, Purchase Amount is Low and the Product Correlation is Strong

From the number of purchases and usual expenses, it can be seen that most people buy less frequently; more people's average expenses are in a low consumption state and less high consumption. From the analysis of the types of fresh purchases, it can be concluded that the most purchased are fruits, followed by vegetables.

The price of meat, seafood, cooked food and other products is relatively small compared with the higher proportion of product purchases. From the analysis of the product correlation analysis of consumer purchases, the correlation between vegetables and fruits can be very strong, followed by dry fruit, meat seafood, vegetable meat and fruit cooked food. From the above analysis, it can be concluded that the number of purchases and the intensity of purchases are not high, and the products purchased are highly correlated.

4.2. Recommendations

4.2.1. Grasp the Target Population, Increase Publicity, Online and Offline Precision Marketing

For the overall use of fresh e-commerce is not high, fresh e-commerce platform can consider, product segmentation, precision marketing. For groups such as important potential consumers and more important potential consumers, their income is relatively low. They can carry out holiday promotion discount activities, regularly issue red envelopes to consumers, reduce consumer

shopping costs, and increase their attention to fresh e-commerce platforms. For secondary potential consumers and general potential consumers, income has a certain pursuit of quality of life compared with higher income. For such groups, they will cooperate with suppliers to promote the construction of traceability system, ensure product quality from the source, and emphasize that every step is a visible concept.

For the offline promotion of fresh e-commerce, the fresh e-commerce platform can establish an offline experience center offline, issue promotional coupons that can only be used offline to attract people who are not familiar with fresh e-commerce to experience and provide explanation and experience services to improve the popularity of the platform.

4.2.2. To Ensure the Quality of Service, Enhance the Consumer Purchase Experience

In terms of customer satisfaction, fresh e-commerce platform is an important channel for consumers to understand and purchase goods. Consumers' feelings when using the platform directly affect their trust in the platform and their willingness to repurchase. The platform should firmly grasp the after-sales service, carry out professional skills and literacy training for the customer service personnel in the platform, improve the ability of the after-sales personnel to solve the user's complaints quickly and effectively, eliminate the customer's dissatisfaction in time, improve customer satisfaction, and solve the worries of consumers.

4.2.3. Improve Cold Chain Logistics Technology and Service Level

For the empathy service of fresh e-commerce, the government guides and supports the construction of cold chain logistics infrastructure, encourages fresh e-commerce enterprises to invest in the construction of cold storage and other refrigeration equipment according to their shares, and promotes the sharing of logistics information and logistics resources among enterprises. Strengthen the supply chain management of the whole process, create competitive advantages, strengthen the construction of cold chain logistics, and promote the progress of the whole fresh e-commerce industry^[7].

4.2.4. Grasp Product Quality, Enrich Product Variety, Related Product Bundling

Focus on product quality, improve product cost performance.

In terms of product quality, fresh e-commerce platforms should strictly control the quality of agricultural products and screen suppliers that meet the quality access level. While improving the overall quality of agricultural products, it can form a stable cooperative relationship with the local production base to purchase goods at a lower price. It can also use the Internet of Things technology to shorten the intermediate links and reduce the price difference between middlemen, so that consumers can buy high-quality, cheap, healthy and fresh agricultural products.

Launch a diversified combination of products to increase the order rate.

According to the correlation degree of commodities, the platform can adopt bundling sales, such as vegetables + fruits, fruits + dry goods, etc., increase the variety of business products, and carry out close cooperation with fruit and vegetables, livestock and poultry meat, and aquatic product production bases. Enrich the supply types of products in different seasons and regions of the platform, and provide consumers with more purchase options such as clean vegetables and semi-clean vegetables. It can also introduce a variety of combinations of products, consumers can spend the same funds to buy a variety of fresh products. Based on this, the fresh e-commerce platform can form its own competitive advantage and promote consumers to form brand loyalty to the platform.

Acknowledgements

This work was supported by a project grant from Innovation and Entrepreneurship Training Plan for College Students: Investigation on the Current Situation of Fresh E-commerce Consumers and Market Opportunity Mining in Chongqing (Grant No.CXCY2022166).

References

- [1] Wang Jichuan, Guo Zhigang. *Logistic Regression Model: Methods and Applications*. Higher Education Press, 2001.
- [2] Yu Jianying. *Statistical analysis and SPSS application*. People's Posts and Telecommunications Publishing House, 2003.
- [3] Yao Xiteng. *Analysis of Bayesian theorem and its application*. *Education modernization*, 2017 (06): 260-262.
- [4] Yao Hesheng. *Apply MATLAB programming to realize the operation of a large number of matrices*. *Journal of Anqing Normal University: Natural Science Edition*, 2020, 13 (1): 3.
- [5] Zhang Xianda. *Matrix analysis and application*. Tsinghua University Press, 2013.
- [6] Li Fei, Xue Bin, Huang Yalou. *K-Means clustering algorithm with initial center optimization*. *Computer Science*, 2021, 29 (7): 94-96.
- [7] Li Jie. *Construction and Research on Competitiveness Evaluation System of Fresh E-commerce Cold Chain Logistics*. *Journal of Shanxi Finance and Taxation College*, 2019, 18 (4): 67-69.