Frontiers in Business, Economics and Management
ISSN: 2766-824X | Vol. 14, No. 3, 2024

Study of Commodity Relationships Based on Integrated Modelling

Zijun Wang, Xuyi Hui, Fanxiang Yan

Department of Gonggongweisheng, Hangzhouyixueyuan University, Hangzhou, 311399, China

Abstract: With the times, sound business decisions are crucial for superstores. Multiple metrics such as replenishment and pricing decisions, future sales volume expectations, etc. tend to have an important place in operational decisions. Firstly, data visualisation is performed to study the distribution pattern of category sales volume and find outliers. Second, the relationship between categories and categories, and between individual products and individual products is analysed by Spearman's correlation coefficient. Then the connection between category and single product is described by grey correlation analysis, and the sample single product is classified by clustering, so as to describe the distribution law of single product. At the same time, the Fourier transform was used to find the existence of cyclical changes in each category of vegetables, which further describes the sales distribution of the categories.

Keywords: Grey correlation, Spearman's correlation coefficient, merchandise replenishment, Fourier Transform.

1. Introduction

In reality, market demand is usually a random variable, so many scholars have researched on random inventory management strategies[1-3] . Vegetables, which are often purchased in the market in daily life, have very high water content, so they are prone to rot and deterioration after picking. The freshness period of vegetable goods is usually short, and the quality of vegetables will deteriorate with the increase of time. Therefore, if most vegetables are not sold on the same day, there is no way to sell them again on the next day. Therefore, supermarkets will replenish their stocks according to the daily sales and demand of each commodity during the daily stock replenishment process. Due to the wide variety of vegetable origins and types, reliable market demand analyses are conducted before making replenishment decisions and pricing decisions. There is usually some correlation between sales volume and time of day for vegetable items based on market demand. Supermarkets have limited selling space, so they need to choose a reasonable mix of stocking volume and balance the relationship between vegetables[4-5] .

2. Materials and Methods

This study collects data on the purchase and sales of vegetable category in shopping malls from 1 January 2022 to 31 December 2022 in superstores in Hangzhou, Zhejiang Province. Observe the collection of data in the superstore, there are no gaps or missing in the data. The type of sales is divided into two types: sales and returns, because returns belong to the personal behaviour of customers and it is impossible to judge whether the returned goods can be sold for a second time, so the data of returns are not considered and only the sales part is considered.

The basic idea of grey correlation analysis is to determine whether the curves are closely related by analysing the degree of similarity of their geometrical shapes. The closer the curves are, the more connected and similar the sequences are, and conversely, the less connected they are. Through grey correlation analysis, sales weekly sales volume was determined as an indicator to find the relationship between categories and individual products. A combination of multiple models is used to comprehensively describe the sales distribution of categories and individual products in three different time scales: monthly, weekly and daily, and to establish the progressive linkage models of individual products and individual products, individual products and categories, and categories and categories, so as to make the model complete.

3. Modelling and Solving

Translate this problem into four research directions. Direction 1: To solve the distribution pattern of vegetables in each category; Direction 2: To find the interrelationship between categories; Direction 3: To find the interrelationship between categories and individual products; Direction 4: To find the interrelationship between different individual products.

3.1. Modelling

(1) Direction 1: Fourier transform to find the distribution of sales volume of vegetables by category

According to the data preprocessing, the sales volume of various types of vegetables may show certain cyclicality, and the supply varieties show certain seasonal changes.

The formula for expressing the Fourier transform of a continuous time signal x(t) is:

$$x(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t}\, dt \tag{1}$$

Let x(n) be a periodic sequence of period N. The discrete Fourier transform is given by:

$$x(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk} \tag{2}$$

The complex form of Fourier transform can be obtained after Fourier transform, and after extracting the real and imaginary parts of the complex, the frequency intensity of each time domain sampling value can be calculated according to the following formula.

$$I(f_i) = \frac{N}{2}(a_i^2 + b_i^2) i=1,2,\dots .k \quad (3)$$

$$T_i = \frac{1}{f_i} \quad (4)$$

Based on the above equation, plot the frequency intensity graph with frequency as the horizontal coordinate and frequency intensity as the vertical coordinate, analyse the graph and draw a conclusion.

(2) Direction 2: Spearman's correlation yields interrelationships between categories

There is some correlation between the categories. After data visualisation it was found that the sales volume of the categories did not conform to normal distribution. Therefore, after deriving the periodicity of the sales volume of the categories, Spearman spearman correlation coefficient was used to derive the correlation between the sales volume of the categories and the categories.

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \quad (5)$$

(3) Direction 3: Establishment of grey correlation models

Firstly, the relationship between categories and individual products is intuitively analysed through data visualisation.

Step 2: Determine the analytical series. Determine the sales volume of a category as the parent, or reference, series, denoted C0 . The sales volume of a single vegetable item is the subsequence, or comparison, series, denoted C1 , C2 , . . . , Cm , and m is the number of subsequences.

In the third step, the variables were preprocessed to narrow them down and simplify the calculations. Firstly, the mean value of each indicator was obtained, and then each element was divided by the mean value.

In the fourth step, the correlation coefficient of each indicator in the subsequence with the parent sequence is calculated. The sub-series and the parent series are shown below.

$$\begin{cases} C_0 = (C_0(1), C_0(2), \dots C_0(n))^T \\ C_1 = (C_1(1), C_1(2), \dots C_1(n))^T \\ \cdots \\ C_m = (C_m(1), C_m(2), \dots C_m(n))^T \end{cases} \quad (6)$$

Definition: bipolar minimum difference g =min min$|C_0(k) - C_i(k)|$, bipolar maximum difference$\hbar$ = max max$|C_0(k) - C_i(k)|$

$$\gamma(C_0(k), C_i(k)) = \frac{g+\hbar\rho}{|C_0(k)-C_i(k)|+\hbar\rho} \quad (7)$$

whereρ is the resolution coefficient (generally taken as 0.5), i = 1,2,...,m, k=1, 2,..., m Finally, the following equation is obtained as the grey correlation between C0 and Ci.

$$\gamma(C_0, C_i) = \frac{1}{n}\sum_{k=1}^{n} \gamma(C_0(k), C_i(k)) \quad (8)$$

Conclusions can be drawn by comparing the correlation of several subsequences with the parent sequence.

(4) Direction 4: Interrelationships among individual vegetables based on Spearman correlation and clustering

The requirement was to study the interrelationships of different individual vegetables based on the sales volume of individual items. Firstly, a mathematical model of cluster analysis is considered and the sample to sample distance is used to classify the vegetables by using squared Euclidean distance. And write the distance matrix between samples.

$$D_0 = \begin{bmatrix} 0 & & & & \\ d\left(\vec{x}_2, \vec{x}_1\right) & 0 & & & \\ d\left(\vec{x}_3, \vec{x}_1\right) & d\left(\vec{x}_3, \vec{x}_2\right) & & & \\ \cdots & \cdots & \cdots & \ddots & \\ d\left(\vec{x}_i, \vec{x}_1\right) & d\left(\vec{x}_i, \vec{x}_2\right) & \cdots & d\left(\vec{x}_i, \vec{x}_j\right) & 0 \end{bmatrix} \begin{matrix} \vec{x}_1 \\ \vec{x}_2 \\ \vec{x}_3 \\ \vdots \\ \vec{x}_i \end{matrix} \quad (10)$$

where$d\left(\vec{x}_i, \vec{x}_j\right)$ is the Euclidean distance between the ith sample and the jth sample, considering each sample as a class. The cluster analysis dendrogram was obtained, but according to the final out results it was found that the clustering results were not good enough, so the same direction two same Spearman correlation was considered to study the interrelationships of different individual vegetables.

3.2. Solving the model:

(1) Direction one model solution

We came up with frequency intensity plots for the six types of vegetables, as shown in Figure I. Looking at the frequency intensity plot, take care to look for the frequency components that have the highest intensity, called peak frequencies. These peak frequencies can indicate the major frequency components in the signal. Peak frequencies are 78 weeks for foliar and cauliflower, 52 weeks for aquatic roots, eggplant and edible mushrooms, and 156 weeks for chilli. This results in a 78-week cycle for foliar and cauliflower species, which is about one and a half years, and is considered to be related to cultivation and growing habits. Aquatic roots and tubers, eggplants and edible mushrooms have a cycle of 52 weeks, which is exactly one year, so it is considered that these three types of vegetables have a cycle of one year. Peppers have a cycle of 156 weeks, which is nearly three years, so it is considered that peppers do not have a cycle.

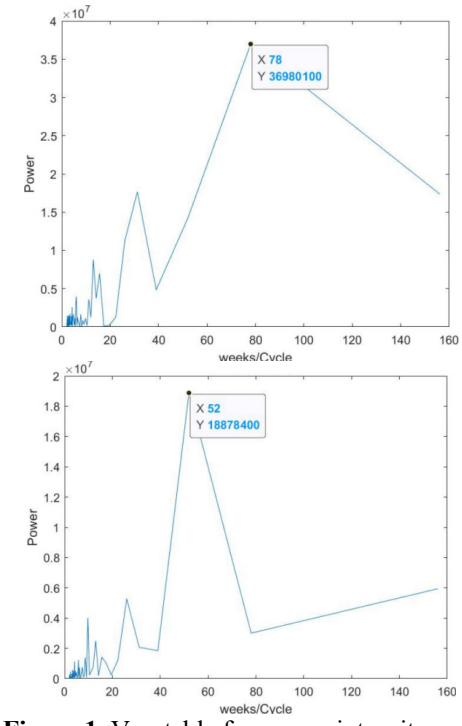

Figure 1. Vegetable frequency intensity map

(2) Solution of the Direction 2 model

Orientation two modelling yields Spearman coefficient values as shown in Table 1.

Based on the table below, it was concluded that the correlation between cauliflower and foliar species was the highest at 0.6889, showing a positive correlation. This is followed by edible mushrooms with aquatic rhizomes. This is followed by chilli with aquatic rhizomes, chilli with edible mushrooms, cauliflower with chilli, edible mushrooms with cauliflower, aquatic root meridian with cauliflower, aquatic root meridian with cauliflower, chilli with cauliflower, and edible mushrooms with cauliflower.

Table 1. Sales Relationship Table

correlation coefficient	philodendron	cauliflower	Aquatic rhizomes	eggplant	chilli	edible mushroom
philodendron	1.0000	0.6889	0.5522	0.1951	0.5854	0.5712
cauliflower	0.6889	1.0000	0.5595	0.1819	0.5218	0.5049
Aquatic rhizomes	0.5522	0.5595	1.0000	-0.1552	0.6077	0.6231
eggplant	0.1951	0.1819	-0.1552	1.0000	0.1004	-0.1077
chilli	0.5854	0.5218	0.6077	0.1004	1.0000	0.5927
edible mushroom	0.5712	0.5049	0.6231	-0.1077	0.5927	1.0000

Looking at the table below, the test hypothesis, only the eggplant and aquatic roots have a p-value of 0.0523 which is greater than 0.05 and the original hypothesis cannot be rejected. All others are less than 0.05 and the original hypothesis is rejected. Combined with Table 1, it was concluded that there was no significant relationship between eggplant and other categories, while any two other categories showed a strong positive correlation. Therefore, there is no relationship between eggplant and other categories, but the other five categories have a positive relationship with each other, i.e., the sales volume of one category increases.

Table 2. Sales Relationship Table

p-value	philodendron	cauliflower	rootstock	eggplant	chilli	edible mushroom
philodendron	1.0000	0.0000	0.0000	0.0143	0.0000	0.0000
cauliflower	0.0000	1.0000	0.0000	0.0226	0.0000	0.0000
rootstock	0.0000	0.0000	1.0000	0.0523	0.0000	0.0000
eggplant	0.0143	0.0226	0.0523	1.0000	0.2110	0.1792
chilli	0.0000	0.0000	0.0000	0.2110	1.0000	0.0000
edible mushroom	0.0000	0.0000	0.0000	0.1792	0.0000	1.0000

(3) Direction three model solution

In the third question, a grey correlation model is developed to solve the relationship between each vegetable item and the vegetable category. The correlation coefficients of each type of vegetables are plotted by grey correlation analysis. Figure 2 shows the correlation coefficients between leafy vegetables and edible mushrooms.

According to the above figure, it can be seen that the relationship between the sales volume of various vegetable items and various vegetables. Among the sales volume of leafy vegetables, the 17th leafy vegetable, that is, Chinese cabbage, has the largest correlation coefficient with leafy vegetables, reaching 0.69273, which is the closest relationship, followed by Yunnan lettuce and Yunnan lettuce (share). The rest of the various vegetables also showed correlations of different sizes. As for the sales volume of edible mushroom vegetables, the 63rd type of vegetables, i.e., enoki mushrooms (box), has the largest correlation coefficient with edible mushroom vegetables, reaching 0.738128, which is the closest relationship. It is followed by Xixia Shiitake Mushroom (1). The rest of the individual items also show different sizes of correlation, but the correlation is not as strong as that between the Golden Needle Mushroom (box) and the Xixia Shiitake Mushroom (1).

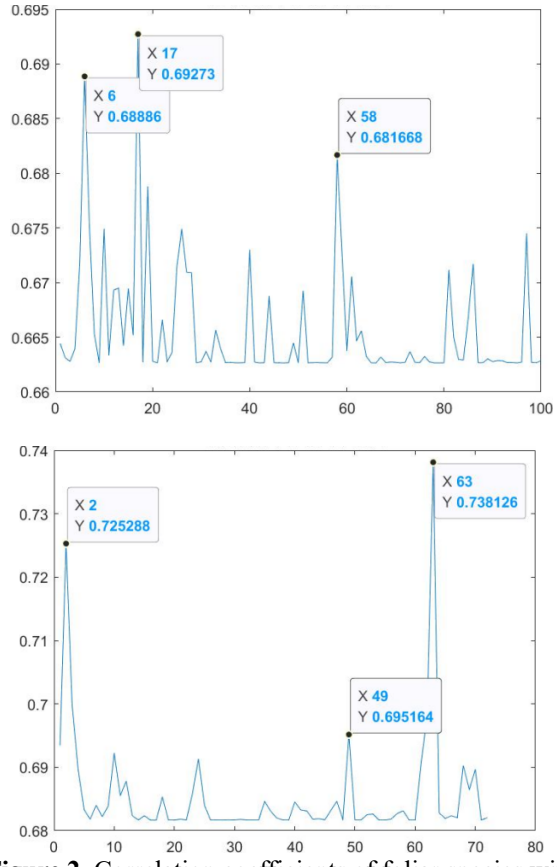

Figure 2. Correlation coefficients of foliar species with edible mushroom individual items

(4) Direction four model solution

Direction four looks for the relationship between the individual items and the sales volume of the individual items. Firstly, clustering method was used to categorise the vegetables into 8 categories based on a tree diagram and it was found that most of the vegetable items were infrequently bought or the sales volume was often very low. These vegetables may be seasonally available or have low sales volume. These 243 vegetables were categorised into the first category. Broccoli and net root (1) were the fifth category with a high degree of similarity, considering that consumers are used to buying broccoli and net root (1) at the same time and have a similar customer base. The remaining categories are, in order, the second category of Yunnan lettuce, the third category of Chinese cabbage, the fourth category of Yunnan lettuce (copies), the sixth category of bubble peppers, the seventh category of Wuhu green peppers , and the eighth category of enoki mushrooms (boxes). Except for the first category, all the other categories have high sales volume and are available on a daily basis. A heat map of correlation coefficients of sales volume of each individual item among the categories of vegetables was plotted. As can be seen from the graph, most of the individual vegetable items are not correlated with each other, but a few have some correlation. For example, the correlation coefficient between category 7 and category 8 of the aubergine category, i.e. green aubergine (2) and purple aubergine (1), reaches 0.9798, which is a strong correlation.

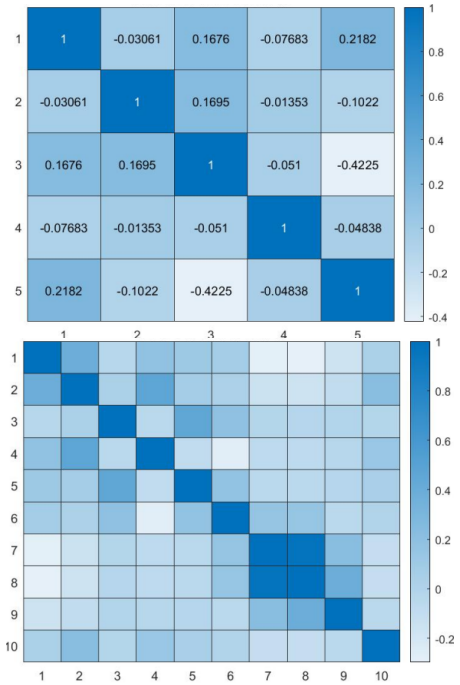

Figure 5. Heat map of correlation coefficients of sales volume of each individual item among categories of vegetables

4. Discussion:

For the planning of supermarket purchasing goods, this study adopts a combination of multiple models to comprehensively describe the sales distribution of categories and individual products on three different time scales: monthly, weekly, and daily, and establishes the progressive linkage model of individual products and individual products, individual products and categories, and categories and categories, which elaborates on the relationship between the three pairs in a detailed manner, but the traditional grey model has a subjective factor. Further research can select multiple evaluation models for comprehensive analysis.

5. Conclusion

This study portrays the relationship between vegetable commodities purchased by superstores through four aspects of research, constructs a comprehensive analysis model containing Fourier transform, Spearman correlation coefficient analysis, grey correlation analysis, and derives the three influencing relationships between vegetable category and category, single product and single product, single product and category, compared to the single relationship analysis and single model analysis, the analytical method of this study is more comprehensive and objective, and it provides a valuable reference for future superstore This study is more comprehensive and objective, and provides a valuable reference for future merchandising decisions. However, if we want to further provide merchants with stocking decisions, we should further analyse the sales volume and time, and use time series characteristics to predict the sales volume changes, so as to provide merchants with more accurate stocking strategies.

References

[1] Chopra S, Rein hardt G, Dada M. The effect of lead time uncert aint y on safety st ocks[J] . Decision Science, 2004, 35 (1) : 635- 642.

[2] Ouyang L Y, Wu K S . Mixture inventory model involving variable lead time w ith a service level constraint [J] . Comp uters a nd Operations Research , 1997, 24 (9) : 875 - 882.

[3] Gupt a A, Maranas C D. Managing demand uncertainty in supply chain planning [J] . Compu ters and Chemical E ngineering , 2003, 27: 1219- 1227.

[4] Hakeem-Ur R. Demand Forecasting, Lot-sizing and Scheduling in the Multistage Production System with Dynamic Information Updates[D]. Shanghai Jiaotong University, 2019.

[5] Jayant A, Chaudhary N. Decision Support Framework for Smart Implementation of Green Supply Chain Management Practices[M]. Springer Nature Switzerland AG 2020.