

Research on Seasonal Orderliness and Linear Correlation of Vegetable Sales Data

Zilu Wen¹, Xiangyue Zhang², Haoxu Miao¹

¹ Department of electrical information, Shandong University of Science and Technology, Jinan, Shandong, 250031, China

² Swinburne College, Shandong University of Science and Technology, Jinan, Shandong, 250031, China

Abstract: The key to the study of seasonal orderliness and linear correlation of vegetable sales data is to understand how sales patterns are affected by seasonal factors and the linear relationship between these patterns and other variables. Seasonal ordering reflects cyclical changes in sales data over time that are influenced by factors such as climate, traditional festivals, and consumption habits. For example, certain vegetables may sell better in the winter while others in the summer. By analysing historical sales data, time series models such as Autoregressive Moving Average (ARMA) can be used to forecast future sales trends, which in turn support inventory management and pricing strategies. In addition, linear correlation analysis helps to identify the relationship between sales volume and factors such as price, cost and market demand, as well as the interactions between different vegetables, including complementary and substitution effects. This information is crucial for supply chain optimisation, promotional campaign planning and risk management. With the development of big data and analytics, researchers can more accurately model and predict market behaviours to support strategic decision-making for vegetable producers and retailers. Research that integrates seasonal ordering and linear correlation will provide insights into the complex dynamics of the vegetable market and provide data to support market participants in developing strategies to adapt to the changing environment.

Keywords: Vegetable Marketing; Seasonality; Orderliness; Linear Correlation; Data Analysis.

1. Introduction

Vegetables, as key agricultural products, play an important role in people's daily diet. Therefore, in-depth analyses of their sales data are essential for understanding consumption patterns, predicting market trends, and formulating effective supply chain strategies. The seasonality of vegetable sales data shows clear regularity, which is closely related to natural growth cycles and consumption habits. For example, certain vegetables may be in high demand during the winter months and relatively less so during the summer months; meanwhile, festivals and local customs may also have an impact on vegetable sales [1]. Understanding and exploiting this seasonality can help farmers and retailers to better plan their planting and marketing strategies to maximise returns.

Orderliness is another important aspect of vegetable sales data that relates to stability and predictability of sales volumes. By evaluating ordered patterns in historical data, supply chain efficiency can be improved and waste can be reduced [2]. For example, by analysing sales data, it is possible to determine which vegetables have stable sales volumes and which fluctuate over a given period of time, thereby optimising inventory management.

Linear correlation analysis, on the other hand, can reveal the relationship between vegetable sales volumes and other variables, such as price, climate change, and population growth. This type of analysis helps to identify which factors have a direct impact on changes in sales volume, providing data to support the development of pricing strategies and sales forecasts [3]. For example, if a strong linear relationship is found between the sales volume of a particular type of vegetable and its price, then price adjustment can be used as an effective sales incentive.

In summary, the study of seasonality, orderliness and linear correlation of vegetable sales data is of great significance in optimising the vegetable market. This can not only help

agricultural market participants to grasp the market dynamics and provide a scientific basis for strategic adjustments, but also provide consumers with a more stable, high-quality supply of vegetables. Through data analysis, a more intelligent and precise agricultural product sales system can be constructed to promote the sustainable development of the entire agricultural industry chain.

2. Related Work

The study of seasonal orderliness and linear correlation of vegetable sales data is concerned with the regularity of changes in the volume of vegetable sales across seasons and the linear relationship with other factors. Such research helps to understand sales patterns, predict market trends, optimise supply chain management, and guide the cultivation and deployment of produce.

The research usually begins with the collection and collation of historical sales data, including sales volumes, prices, places and times of sale of different types of vegetables. With these data, researchers can construct time-series models to analyse the seasonal patterns of vegetable sales data, i.e. which vegetables see a rise or fall in sales during different seasons of the year, and whether such changes are regular and predictable [4]. For example, summer may see an increase in watermelon and cucumber sales, while winter may be the peak season for root vegetables.

Next, the researcher will use statistical methods to assess the ordinarieness of the sales data, i.e., whether the changes in sales volume show a certain pattern, such as a similar trend of increase or decrease in the sales volume of vegetables in each season for several consecutive years [5]. Orderliness analysis helps to confirm the stability of the sales pattern and provide decision support for vegetable production and supply.

On the other hand, researchers explore the linear correlation between sales data and other factors, such as climatic conditions, festive events, and changes in consumer

preferences. This involves the use of statistical tools such as correlation coefficients and regression analysis to analyse the relationship between different variables. For example, climatic conditions have a direct impact on the growth and availability of vegetables, which may be reflected in sales data [6]. By analysing this data, it is possible to reveal how specific factors affect the sales trends of vegetables.

Finally, the researcher will synthesise the results of the seasonal ordering and linear correlation analyses to come up with explanations for vegetable sales patterns and make recommendations for vegetable planting plans, inventory management, and marketing strategies. For example, if the sales volume of a certain vegetable in a particular season is found to have a strong correlation with climate change, the planting and supply strategy can be adjusted on the basis of climate prediction to adapt to the market demand and minimise losses [7].

The whole study should combine data analysis and industry knowledge to continuously optimise the model and improve the accuracy of the prediction, so as to provide a scientific basis for the sustainable development of the vegetable industry. Through these studies, relevant enterprises and farmers can better understand market dynamics and make effective business decisions.

3. Establishment and Solution of the Model

In order to explore the statistical patterns of the sales volume of vegetable categories and the sales volume of individual products, the sales volume of 251 individual products was first analysed by descriptive statistics. These products were classified into six different categories and those with sales volume below 500 kg were categorised as 'Others'. This categorisation allowed for the study of the distribution pattern of sales of each product within its category. It is noteworthy that six products, namely Wuhu green pepper (1), broccoli, Chinese cabbage, Yunnan lettuce and enoki mushrooms (boxed), had sales volume of more than 15 000 kg. More than half of the individual products sold less than 1 000 kg. The aubergine category had the smallest total sales volume of 22 113 kg, while the leafy vegetables category had the largest sales volume of 203 212 kg. The leading selling products in their respective categories were turnip peppers (1), broccoli, Chinese cabbage, enoki mushrooms (boxed), purple aubergines (2) and peeled lotus roots (1). The leafy vegetables category had a wide range of products, while the cauliflower category was known for its simple product range.

Subsequently, correlation analyses were conducted to assess the association between vegetable categories and individual product sales. As the sales volume-frequency curves plotted at the preprocessing stage did not conform to a normal distribution, the Kendall rank correlation coefficient - a non-parametric test - was used in order to prevent side-by-side rankings between product data [8]. In addition, given the large number of individual products, the analyses were conducted at the category level rather than the individual product level.

Analysing vegetable marketing data in this way provides a nuanced understanding of market dynamics, which is essential for stakeholders in the agricultural sector to make informed decisions. The combination of descriptive statistics and correlation analysis provides a powerful framework for predicting trends and adjusting marketing strategies

accordingly. The findings show that some products consistently dominate sales within their categories, suggesting the potential for targeted cultivation and supply chain management [9]. By identifying the complex relationship between product sales and categorisation, this study contributes to strategic planning and sustainable development of the vegetable industry.

In this study, Kendall's rank correlation coefficients were analysed to investigate the consistency of sales trends between two vegetable categories in order to assess the statistical dependence of their sales classes. To conduct Kendall's correlation analysis, certain prerequisites must be fulfilled: the sales volume of the two categories should be ordered or continuous, and the trends of the two categories should have a common direction. The sales data for each category were real and seasonally varying, thus ensuring orderliness; furthermore, the sales volumes of the different vegetable categories showed a linear correlation in general [10]. Notably, there were significant positive correlations between the sales volumes of the chilli and aubergine categories, cauliflower and aubergine categories, aquatic roots and aubergine categories, and aquatic roots and chilli categories. This analytical improvement underscores the importance of a methodical study of inter-category relationships and contributes to a deeper understanding of sales patterns in the vegetable market, which is important for supply chain optimisation and market forecasting [11].

In column a , x_a , and column b , x_b , of the matrix, define the Kendall τ coefficients, noting that the category homogeneous pairs are (i, j) , and measuring whether the categories are homogeneous or not in terms of $x_{a,i} - x_{a,j}$ and $x_{b,i} - x_{b,j}$, and noting that the adjusted term of the knot value in the normalisation constant is $Tau-b$:

$$\tau = \frac{2K}{m(m-1)} \quad (1)$$

$$K = \sum_{i=1}^{m-1} \sum_{j=i+1}^m \xi \quad (2)$$

$$\xi = \begin{cases} 1 & \text{if } (x_{a,i} - x_{a,j})(x_{b,i} - x_{b,j}) > 0 \\ 0 & \text{if } (x_{a,i} - x_{a,j})(x_{b,i} - x_{b,j}) = 0 \\ -1 & \text{if } (x_{a,i} - x_{a,j})(x_{b,i} - x_{b,j}) < 0 \end{cases} \quad (3)$$

The range of correlation coefficient is indicated by ξ . When the correlation coefficient tends to 0, it means that there is no relationship between the columns; when the correlation coefficient tends to 1, it means that the two column orders are the same; when the correlation coefficient tends to -1, it means that the two column orders are inverse of each other [12].

The following quantitative descriptive statistics were analysed based on the mean, median and standard deviation of each category of vegetables and the results are plotted in the table:

Table 1. Descriptive statistics of vegetable categories

Vegetables	Average	Median	Standard deviation
Cauliflower	8357.82	5825.59	11340.94
Foliage	2032.12	142.03	3630.88
Chilli	2037.03	283.00	4791.78
Eggplant	2211.33	672.42	4154.66
Mushrooms	1057.39	138.50	2427.13
Aquatic Roots	1357.32	50.82	3278.08

Table 1 presents the descriptive statistical analysis of the

vegetable category revealing significant trends in sales performance. The Cauliflower category had the highest mean sales volume of 8357.82 units with a median of 5825.59 units and a high standard deviation of 11340.94 indicating high variability in sales volume [13]. On the contrary, the sales volume of edible mushrooms remained stable but low with a mean of 1057.39 units, a median of 138.50 units and a relatively small standard deviation of 2427.13. Sales figures for leafy vegetables and aquatic roots and tubers were skewed, with occasional exceptionally high sales figures. This skewness, as well as the high volatility of sales in the cauliflower and chilli categories, suggests the presence of outliers or sporadic peak sales events, which may be due to seasonal demand or market promotions [14]. These insights will not only help in developing targeted marketing strategies but also help in refining inventory management practices in the vegetable industry.

4. Conclusion

In this study, vegetable sales data were analyzed in depth to explore the seasonal orderliness and linear correlation, and the results revealed significant seasonal patterns and sales trends. The study shows that there is a significant correlation between vegetable sales and seasonal variations, which is closely linked to the growth cycle of vegetables and the market supply and demand patterns [15]. Specific seasons, such as late spring, summer and early autumn, saw an increase in the supply of vegetables due to the facilitating effect of climatic conditions on crop growth, which correspondingly boosted sales. In addition, festivals and traditional customs also have an impact on vegetable sales, such as the significant increase in vegetable sales around the Chinese New Year.

In terms of seasonal orderliness, by analysing the impact of seasonal changes on vegetable types, it was found that the peak sales of different vegetables occurred in different seasons. For example, vegetables with higher sales in summer were usually sown in spring, while vegetables with hot sales in autumn and winter were implanted in summer. The study also found that climate anomalies had a perturbing effect on vegetable sales, demonstrating the impact of climate change on agricultural markets [16].

Linear correlation analyses showed that vegetable sales data exhibited a linear relationship with climatic factors such as temperature and precipitation. Higher temperatures usually lead to increased sales of certain vegetables, while changes in precipitation may affect the supply chain and price volatility of vegetables, which in turn affects sales data. In addition, economic factors such as consumer purchasing power, price fluctuations, and market share of imported vegetables also show some linear correlation with vegetable sales volume.

Overall, the study of seasonal orderliness and linear correlation of vegetable sales data helps to understand the dynamics of the vegetable market and provides a scientific basis for vegetable production arrangements, supply chain management, and market forecasting. The findings emphasise the importance of paying attention to meteorological information and market demand forecasting, which are of great practical significance for improving the efficiency of vegetable sales and ensuring the stability of market supply. Future work could further explore the marketing patterns of different regions and vegetable varieties, as well as the possible long-term impacts on vegetable marketing in the context of global changes.

5. Discussion

In this study, we delved into the seasonal orderliness and linear correlation of vegetable sales data with a view to unravelling sales patterns and providing guidance for market strategies. The analyses revealed that vegetable sales volumes are strongly correlated with seasonal factors, as evidenced by the fact that climate change directly affects the growth cycle and availability of vegetables. For example, warm spring and favourable autumn weather conditions promote vegetable growth and bring about sales peaks. In addition, festivals and cultural practices play an equally important role in driving vegetable sales at certain times of the year.

We also note that seasonal variations have varying impacts on the sales of different vegetable types, with certain vegetables selling well in the summer months while others are more popular in the autumn and winter months. This variability in seasonal sales peaks points to the need for adjustments in growing and marketing strategies [17]. The negative impact of climatic anomalies such as droughts or floods on vegetable sales also shows the sensitivity of fragile supply chains to environmental factors, which needs to be taken into account in future strategies.

Linear correlation analyses further demonstrated the association of vegetable sales data with climatic factors such as temperature and precipitation. This relationship suggests that climate forecasting tools can be an effective aid in predicting vegetable market trends. Economic factors, such as consumer purchasing power and price fluctuations, also have a significant impact on vegetable sales volume, suggesting that we must pay attention to the possible impact of the macroeconomic environment on the vegetable market.

In summary, the seasonal orderliness and linear correlation of vegetable sales data provide insights that help us understand market dynamics and guide production and marketing strategies. The study highlights the importance of weather information and provides a scientific basis for market demand forecasting. Future research could focus on the long-term effects of regional differences, variety characteristics, and global climate change on vegetable marketing, with a view to providing more comprehensive guidance on agricultural production and market supply.

Future research will focus on the seasonal orderliness and linear correlation of vegetable sales data, aiming to improve the accuracy of sales forecasts and the effectiveness of market strategies through in-depth analyses. Seasonal characteristics are determined by a combination of factors such as climatic conditions, festive customs and consumer preferences, and are manifested as an increase in the sales volume of certain vegetables within a given season. For example, more root vegetables may be sold in winter, while summer may be dominated by melons and fruits. This seasonal ordering recurs every year and provides a reliable basis for sales forecasting. In the future, we plan to use time-series analysis methods, such as the ARMA model, to forecast future vegetable sales trends based on historical data and to conduct long-term trend analyses taking into account the impact of climate change on crop growth cycles.

The linear correlation study will explore the sales relationship between different vegetables and the interaction between vegetable prices, sales volume and seasonal factors. By calculating correlation coefficients, we can identify which vegetables have complementary or competitive relationships with each other. Linear regression models help to further

analyse the dynamic effects of these variables on each other and provide a theoretical basis for market adjustment.

The study of integrated seasonal ordering and linear correlation can not only optimise inventory management and supply chain scheduling, but also support the development of marketing strategies. In addition, technological advances and the development of big data analysis methods will greatly enhance the efficiency and precision of these studies, enabling us to understand the complexity of vegetable markets with a more nuanced perspective. Through a deeper understanding of seasonal patterns and linear correlations, we will be able to better predict market dynamics, leading to more efficient market operations and more precise consumer satisfaction. Future research will aim to build a comprehensive model that not only includes seasonality and linear correlation, but also incorporates additional factors that affect sales, such as weather changes, agricultural policies, and global economic trends, in order to remain competitive in a changing market.

References

- [1] Pathan, S., Ndunguru, G., Islam, M. R., Jhumur, S. T., & Ayele, A. G. (2023). Production of Quinoa Leafy Greens in High Tunnel for Season Extension in Missouri. *Horticulturae*, 9(2), 209.
- [2] Borrero, J. D., & Borrero-Domínguez, J. D. (2023). Enhancing Short-Term Berry Yield Prediction for Small Growers Using a Novel Hybrid Machine Learning Model. *Horticulturae*, 9(5), 549.
- [3] Spiker, M. L., Welling, J., Hertenstein, D., Mishra, S., Mishra, K., Hurley, K. M., ... & Lee, B. Y. (2023). When increasing vegetable production may worsen food availability gaps: A simulation model in India. *Food Policy*, 116, 102416.
- [4] Chang, H. H., & Meyerhoefer, C. D. (2021). COVID-19 and the demand for online food shopping services: Empirical Evidence from Taiwan. *American Journal of Agricultural Economics*, 103(2), 448-465.
- [5] Miliou, I., Xiong, X., Rinzivillo, S., Zhang, Q., Rossetti, G., Giannotti, F., ... & Vespignani, A. (2021). Predicting seasonal influenza using supermarket retail records. *PLOS Computational Biology*, 17(7), e1009087.
- [6] Trasberg, T., Soundararaj, B., & Cheshire, J. (2021). Using Wi-Fi probe requests from mobile phones to quantify the impact of pedestrian flows on retail turnover. *Computers, Environment and Urban Systems*, 87, 101601.
- [7] Falatouri, T., Darbanian, F., Brandtner, P., & Udokwu, C. (2022). Predictive analytics for demand forecasting—a comparison of SARIMA and LSTM models in retail SCM. *Procedia Computer Science*, 200, 993-1003.
- [8] Sirisha, U. M., Belavagi, M. C., & Attigeri, G. (2022). Profit prediction using Arima, Sarima and LSTM models in time series forecasting: A Comparison. *IEEE Access*, 10, 124715-124727.
- [9] Tadayonrad, Y., & Ndiaye, A. B. (2023). A new key performance indicator model for demand forecasting in inventory management considering supply chain reliability and seasonality. *Supply Chain Analytics*, 3, 100026.
- [10] Torres-Sánchez, R., Martínez-Zafra, M. T., Castillejo, N., Guillamón-Frutos, A., & Artés-Hernández, F. (2020). Real-time monitoring system for shelf life estimation of fruit and vegetables. *Sensors*, 20(7), 1860.
- [11] Dolega, L., Rowe, F., & Branagan, E. (2021). Going digital? The impact of social media marketing on retail website traffic, orders and sales. *Journal of Retailing and Consumer Services*, 60, 102501.
- [12] Vogel, C., Crozier, S., Penn-Newman, D., Ball, K., Moon, G., Lord, J., ... & Baird, J. (2021). Altering product placement to create a healthier layout in supermarkets: Outcomes on store sales, customer purchasing, and diet in a prospective matched controlled cluster study. *PLoS Medicine*, 18(9), e1003729.
- [13] Banerjee, T., Sinha, S., & Choudhury, P. (2022). Long term and short term forecasting of horticultural produce based on the LSTM network model. *Applied Intelligence*, 1-31.
- [14] Wudad, A., Naser, S., & Lameso, L. (2021). The impact of improved road networks on marketing of vegetables and households' income in Dedo district, Oromia regional state, Ethiopia. *Heliyon*, 7(10).
- [15] Adeagbo, O. A., & Adejumo, O. O. (2020). Economic analysis of dry season vegetable production in Ogun State, Nigeria. *African Journal of Economic and Management Studies*, 11(3), 427-441.
- [16] Li, Y., Zhou, H., Lin, Z., Wang, Y., Chen, S., Liu, C., ... & Xia, J. (2020). Investigation in the influences of public opinion indicators on vegetable prices by corpora construction and WeChat article analysis. *Future Generation Computer Systems*, 102, 876-888.
- [17] Stone, T. F., Thompson, J. R., Rosentrater, K. A., & Nair, A. (2021). A life cycle assessment approach for vegetables in large-, mid-, and small-scale food systems in the midwest US. *Sustainability*, 13(20), 11368.