

Sales Forecasting of Plug-in Board Products Using Linear Regression: A Case Study of Yaqi Factory

Quanrui Li^{1*}, and Bavornwit Rojsuwan²

Master of Science in Management of Logistics, Rangsit University, Thailand

**Corresponding author. E-mail: joser.quanrui@gmail.com^{1*}*

Received May 8, 2024; **Revised** June 18, 2025; **Accepted** July 5, 2025

Abstract

This study aims to develop a sales forecasting model for Yaqi Factory's plug-in board products, thereby optimizing inventory management and enhancing operational efficiency, by collecting sales data for 31 months, selecting five best-selling products, and using the linear regression method to analyze the relationship between sales volume and key factors such as price, inventory, and production volume. The results showed that the model constructed based on six-month data had high predictive ability (R value of 0.816), explained 81.6% of sales variation, and had significantly lower prediction error than the model based on three-month data (absolute error of 607 vs. 1528). In addition, simple moving average graphical analysis shows that the model still has room for optimization when facing products with significant sales fluctuations, such as B5330. This study provides empirical evidence and methodological references for establishing effective sales forecasting and inventory management systems for small and medium-sized manufacturing enterprises.

Keywords: inventory management; plug-in boards sales prediction; linear regression

Introduction

Inventory management is a crucial aspect of a company's operations, encompassing the ordering, storage, utilization, and sale of raw materials, components, and finished products (Hayes, 2022). Effective inventory management is crucial for a company's operational efficiency and profitability, as it directly affects production and the completion of customer orders. Different inventory management methods, such as first-in, first-out (FIFO), last-in, first-out (LIFO), and the weighted average cost method, each have their own advantages and disadvantages and are

suitable for different company needs. In any factory, inventory management is considered a key aspect in reducing costs.

As a plug-in board manufacturer, the Yaqi factory, located in Dongguan, China, faces inventory management challenges, particularly in predicting sales volumes and maintaining optimal inventory levels. For example, overstocking of models like B5330 due to inaccurate forecasts has resulted in higher storage costs, while understocking of models such as UU4320 has led to order delays. These issues are mainly due to the lack of data-driven forecasting tools. Incorrect inventory management may lead to cost waste and delivery delays.

This study aims to apply a linear regression model to predict plug-in sales, thereby optimizing inventory management strategies and reducing costs associated with inventory surplus and shortage. The purpose of this study is to apply a linear regression model to forecast future sales of plug-in goods and optimize inventory management tactics using this data. Linear regression, a traditional statistical analysis tool, can create a linear relationship between product sales volume and time or other variables by evaluating historical sales data and related factors that influence sales (Teja Reddy & Malathi, 2022). On this basis, businesses can better understand sales trends, develop reasonable production and inventory plans, and cut expenses associated with inventory surplus and shortage.

Although many studies have explored inventory control methods, fewer have focused on applying linear regression to medium-sized manufacturing companies, such as Yaqi. This study addresses that gap and provides a practical approach for improving inventory accuracy and operational performance.

Research Objectives

1. Determine a model of the sales forecasting plugin-board product for Yaqi Factory.
2. Find out a correlation between the effects of seven factors of plug-in board product attributes on sales.
3. To create and execute optimal inventory management strategies, reduce inventory costs, and enhance operational efficiency.

Scope of Research

Content of the Study: The study's main objective is to predict Yaqi Factory's plug-in board product sales using linear regression techniques. It examines how product parameters, including

pricing, inventory levels, and manufacturing volume, influence sales volume. The study assesses the model's performance and reliability in real-world forecasting by comparing prediction accuracy across various periods (3 months vs. 6 months).

Population and Sample: The population consists of plug-in board items manufactured by Yaqi Factory. The five best-selling models—B5220, B5330, B5320, UU4222, and UU4320—are included in the sample. Data on 31 months of monthly sales, production, and inventory were gathered.

Area: Yaqi Factory in China provides the data used in this study. Nonetheless, the approach and model can be applied more widely in the electronics and manufacturing sectors, which face similar challenges with sales and inventory forecasting.

Time: Sales data from January 2022 to July 2024 (31 months) were used in this study. Forecasting models were trained on 3-month and 6-month data windows and validated against the actual historical data from the same period.

Literature Review

Introduction Inventory Management

Effective planning, control, and supervision of enterprise resources to meet production or sales needs while lowering costs and hazards constitute inventory management, a crucial component of supply chain management (Waters, 2020). Finding the ideal supply-demand balance is the primary goal of inventory management, which enables businesses to continue operating effectively even in the face of shifting market conditions, supply chain disruptions, and fluctuating demand. As a result, inventory management has a direct impact on business profitability in addition to product availability and customer service standards.

Businesses often utilize models such as safety stock and economic order quantity (EOQ) in traditional inventory management systems to establish optimal ordering cycles and inventory levels. While safety stock is intended to mitigate supply or demand unpredictability and prevent inventory shortages, the economic order quantity model helps businesses determine the optimal quantity for each order by balancing ordering and holding costs (Wild, 2017). In settings with solid supply chains and high levels of predictability, these techniques can be somewhat successful. However, traditional inventory management techniques have steadily shown their limitations and are unable to handle

the complex and constantly shifting market environment resulting from globalization and ongoing changes in market demand (Silver et al., 1998).

Additionally, using big data and machine learning technologies in inventory management enables businesses to identify trends in vast amounts of historical data and produce more precise demand projections. (Zhang et al., 2021). It is noted that, compared to traditional models, machine learning-based prediction models are more accurate in forecasting demand because they can effectively manage more complex variables, such as market trends, seasonal variations, and economic conditions. Businesses can enhance their ability to adapt to market changes and dynamically adjust their inventory management methods by integrating these predictive models. Businesses can optimize inventory levels, reduce the risk of inventory shortages and surpluses, and enhance the overall effectiveness of their supply chain by integrating forecasting with inventory control.

Linear Regression Model

The linear regression model has several essential statistics and components that elucidate the model's fitting efficacy and the impact of independent variables on the dependent variable. The regression coefficient indicates the magnitude and direction of each independent variable's influence on the dependent variable, with positive values denoting a positive relationship and negative values indicating a negative relationship (Montgomery et al., 2021). Standardized regression coefficients facilitate the comparison of the relative influence of independent variables (Frost, 2020). The intercept represents the anticipated value of the dependent variable when the independent variable equals zero, typically requiring no more elucidation in practical contexts (James et al., 2021). The goodness-of-fit index R^2 indicates the extent to which the model accounts for the variability of the dependent variable. A value near 1 signifies robust explanatory power of the model, yet it does not entirely represent predictive performance (James et al., 2021); the adjusted R^2 accounts for the effect of the number of independent variables, rendering it especially appropriate for multiple regression (Seber & Lee, 2018). However, a higher R^2 does not necessarily mean that the model is the optimal model. Researchers need to consider factors such as research objectives, model assumptions, and residual distribution comprehensively (Hayes, 2022). In recent years, with the increasing dimensionality of data and the growing importance of feature engineering, researchers have not only focused on model prediction accuracy but also on model interpretability and simplicity (Adeyemo & Olatunji, 2021). During this process, adjusted R^2 often interacts with AIC, BIC, cross-

validation error, and other indicators, which are used in combination to assist in selecting the optimal feature combination. When certain variables contribute minimally to R^2 and do not significantly improve the adjusted R^2 , they can often be removed to obtain a more concise and robust regression model.

The p-value assesses the importance of regression coefficients. If it is below the significance threshold (often 0.05), the variable is seen to exert a significant influence on the dependent variable; a lower significance correlates with a higher confidence level (Field, 2018). The residual represents the discrepancy between the anticipated and actual values. In contrast, residual analysis can evaluate model assumptions, such as linear relationships, independence, and homoscedasticity, and help identify systematic errors or outliers (Pallant, 2020).

Research Methodology

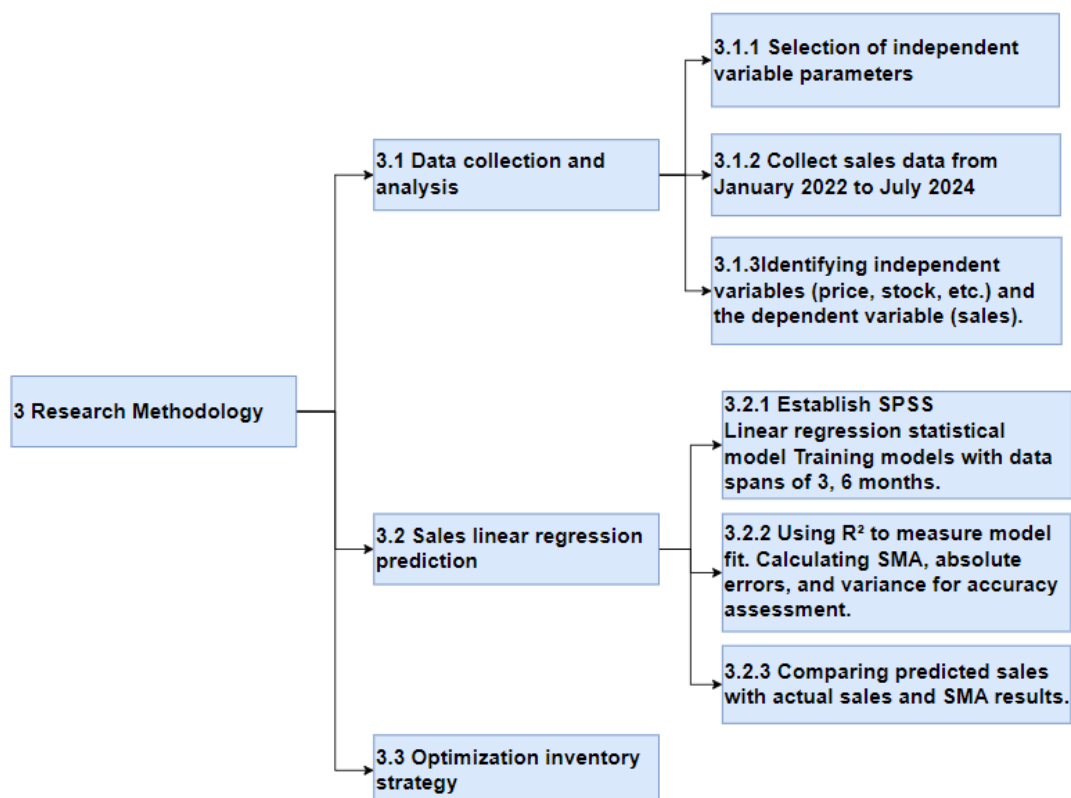


Figure 1: Conceptual framework

Factors Affecting the Sales of Plug-in Boards

The YaQi plug-in board factory's information system indicates that there are numerous packages of data monitoring on book sales. In this study, ten factors, including sales volume, product code, pricing, length of plug-in extension cable, safety certification, number of socket positions, maximum power, number of promotions, initial stock, and final stock, that are closely associated with significant sales impacts were chosen based on the product parameters of the product itself, the years of sales experience of YaQi Plug-in Factory, and other pertinent literature. Ultimately, these variables were chosen as the pertinent variables for the YaQi plug-in manufacturing product sales linear regression prediction.

These variables were chosen based on a combination of the product's physical and functional parameters, YaQi Factory's years of practical sales experience, and findings from relevant literature. Among these factors, pricing, cable length, and the number of socket positions have a direct influence on customer purchase preferences. At the same time, safety certification and maximum power are closely tied to regulatory compliance and product safety—key considerations in customer decision-making. Inventory-related variables such as initial and final stock, as well as production and promotional quantities, influence supply capability and marketing intensity.

Create sample data to forecast the sales condition of YaQi plug-in board factories after filtering the plug-in board data, including these characteristics: plug-in board models and sales revenue. The standard distribution analysis of the significance of the pertinent factors can be obtained after using IBM SPSS Statistics 26 software (henceforth referred to as SPSS software) for linear regression analysis and prediction. This can then be adjusted to determine whether to optimize and eliminate some factors that have little impact. Lastly, this model may be applied in real-world scenarios to forecast the sales of all previous data plug-in boards of YaQi plug-in board manufacturing products. The whole operation and management are more positively impacted by thorough statistical analysis.

Table 1 Linear Regression Dependent Variables

NO.	Dependent variables
1	Sales Forecasting

Table 2 Linear Regression Independent Variables

NO.	Independent variables
1	Product Model
2	Month
3	Price
4	Final Stock
5	Production

Research Results

This section presents the research results structured according to the three research objectives. The R^2 values from linear regression models are analyzed at different periods (3 months and 6 months), focusing on the top five products with the highest sales: B5220, B5330, B5320, UU4422, and UU4320. The dataset spans from January 2022 to July 2024 (31 months).

Objective 1: To analyze the explanatory power of linear regression at different periods

Table 3 Regression Summary Table for Different Periods

Period (months)	R	R^2	Adjusted R^2
3	0.801	0.642	0.499
6	0.903	0.816	0.786

Table 3 shows that the R^2 value gradually increases with the increase of period, from 0.642 at 3 months to 0.816 at 6 months. This suggests that as the period increases, linear regression models can better explain the changing trends in sales data and possess higher explanatory power. The 6-month time span has the highest R^2 value (0.816) and adjusted R^2 value (0.786), indicating that the model performs very stably over long-term data and effectively captures the relationship between sales performance and product attributes.

Objective 2: To evaluate the predictive equation derived from the 3-month linear regression model

Using a Three-month Data Linear Regression Model, Objective 2: To evaluate the predictive equation derived from the 3-month linear regression model

Comparison between the three-month SMA sales value and the three-month linear regression prediction sales.

The following equation was obtained from applying linear regression to predict sales:

Predicted Sales Value = $3578.579 + (\text{Price} * 18.925) + (\text{Initial Stock} * 0.409) + (\text{Final Stock} * -0.613) + (\text{Production} * 0.414)$

Objective 3: To compare predicted sales with actual and SMA sales

Figure 2 shows a comparison chart of sales predicted by a linear regression model using 3-month data and SMA sales. The horizontal axis represents the sales distribution of the B5220 product in 28 months, while the vertical axis represents the sales value. The sales line predicted by the model aligns closely with the actual sales line in some months. The predicted sales revenue of the B5220 product in the 22nd month is significantly lower than the actual sales revenue, showing a significant deviation. In most cases, the sales line of SMA falls between actual sales and forecasted sales. In some months with significant sales fluctuations, the gap between SMA sales and actual sales remains substantial.

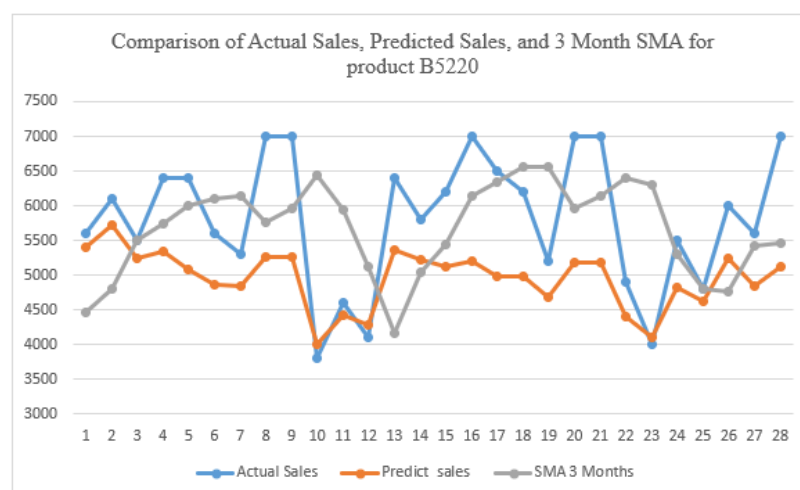


Figure 2: Comparison of Actual sales, Predicted sales, and 3-Month SMA for B5220

Figure 3 shows the results of predicting the sales of the B5330 product using a linear regression model based on a 3-month data range and compares them with actual sales and SMA (Simple Moving Average) sales. The horizontal axis of the chart represents the sales distribution of B5330 products in 28 months, while the vertical axis represents the sales value. According to the graph, the sales predicted by the model are relatively close to the actual sales in certain months. There is a significant deviation between the predicted value and the actual sales revenue, particularly in the 22nd month, indicating a substantial discrepancy. SMA sales are located between actual sales and predicted sales, reflecting the model's ability to balance prediction errors. In months with significant sales fluctuations, the gap between SMA sales and actual sales remains large.

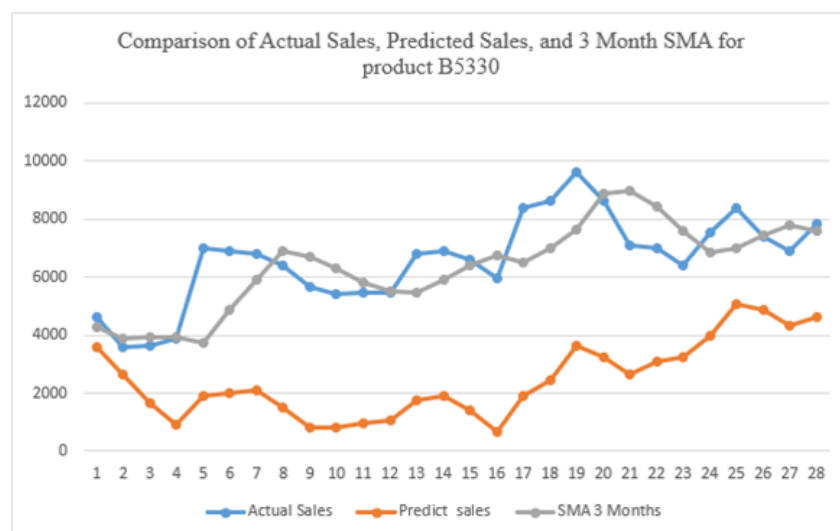


Figure 3: Comparison of Actual sales, Predicted sales, and 3-Month SMA for B5330

Figure 4 shows a comparison of sales predicted using a linear regression model with 3-month data and SMA sales. The horizontal axis represents the sales distribution of the B5220 product in 28 months, and the vertical axis represents the sales value. The sales line predicted by the model approaches the actual sales line in certain months. The expected sales revenue of the B5220 product in the 22nd month is significantly lower than the actual sales revenue, showing a significant deviation. In most cases, the sales line of SMA falls between actual sales and forecasted sales. In some months with significant sales fluctuations, the gap between SMA sales and actual sales remains substantial.

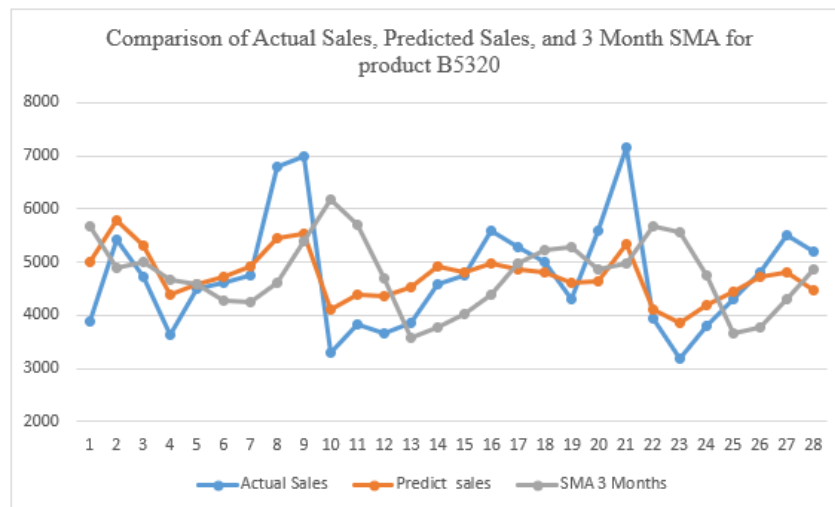


Figure 4: Comparison of Actual sales, Predicted sales, and 3–Month SMA for B5320

Figure 5 shows a comparison of sales predicted using a linear regression model with 3–month data and SMA sales. The horizontal axis represents the sales distribution of the UU4222 product in 28 months, and the vertical axis represents the sales value. The sales line predicted by the model approaches the actual sales line in certain months. The expected sales revenue of the UU4222 product in the 28th month is significantly lower than the actual sales revenue, showing a significant deviation. In most cases, the sales line of SMA falls between actual sales and forecasted sales. In some months with significant sales fluctuations, the gap between SMA sales and actual sales remains substantial.

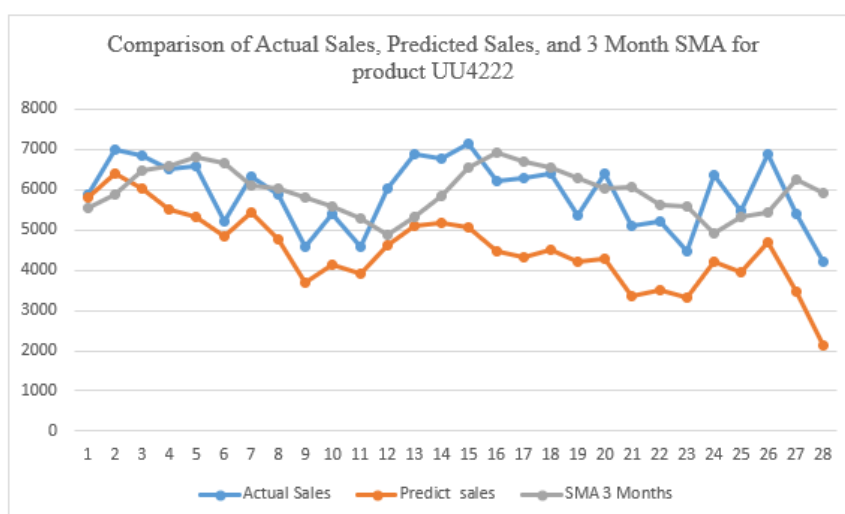


Figure 5: Comparison of Actual sales, Predicted sales, and 3–Month SMA for UU4222

Figure 6 shows a comparison of predicted sales using a linear regression model with 3-month data and SMA sales. The horizontal axis represents the sales distribution of the UU4320 product in 28 months, and the vertical axis represents the sales value. The sales line predicted by the model approaches the actual sales line in certain months. The expected sales revenue of the UU4320 product in the 22nd month is significantly lower than the actual sales revenue, showing a significant deviation. In most cases, the sales line of SMA falls between actual sales and forecasted sales. In some months with significant sales fluctuations, the gap between SMA sales and actual sales remains substantial.

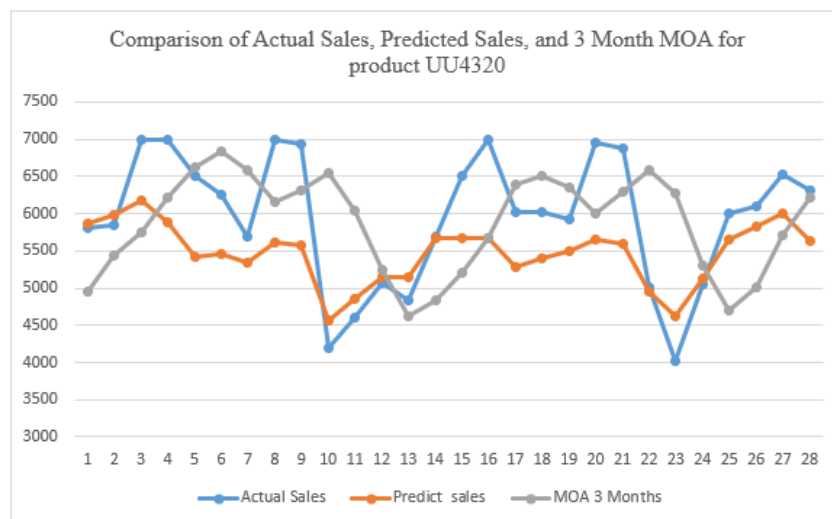


Figure 6: Comparison of Actual sales, Predicted sales, and 3-Month SMA for UU4320

Comparison between sales forecast based on 3-month data and actual sales in 2023

Table 4 presents the error analysis results of sales forecasting for five products (B5220, B5330, B5320, UU4222, and UU4320) using a linear regression model. The model is trained based on sales data from January to March 2022 and compared with actual sales data from 28 months (2022–2024) using predicted values to evaluate the predictive performance of the model on different products.

The overall absolute error is 1528, indicating the average deviation level of the model across all products. The prediction error of B5330 is the largest, with an absolute error of 4,156 and a standard deviation of 1,453, indicating that the sales data for this product fluctuates significantly and the model struggles to capture its trend accurately.

The prediction performance of B5320 is the best, with an absolute error of only 581 and a standard deviation of 456, indicating that the model's sales trend prediction for this product is relatively stable and accurate.

The errors and absolute errors of B5320 and UU4320 are relatively small, indicating that the model's predictions on these products have a certain level of reliability.

Table 4: Linear regression prediction error analysis based on sales data of five products in three months

Product Model	Error	Absolute Error	Standard Deviation
B5220	-845	881	614
B5330	-4156	4156	1453
B5320	-14	581	456
UU4222	-1403	1403	581
UU4320	-481	616	458
All Models	-1380	1528	1564

Using a Six-month data linear regression model

Comparison between the six-month SMA sales value and the three-month linear regression prediction sales.

The following equation was obtained from applying linear regression to predict sales:

Predicted Sales $1537.726 + (\text{Price} * 16.199) + (\text{InitialStock} * 0.626) + (\text{FinalStock} * -0.7) + (\text{Production} * 0.706)$

Figure 7 illustrates the results of applying linear regression to predict sales revenue, actual sales revenue, and SMA using 6-month sales data. The obtained R-squared value is 0.816, indicating that the model explains 81.6% of the variation in the data. The horizontal axis represents the sales distribution of the B5220 product within 25 months, and the vertical axis represents the sales volume.

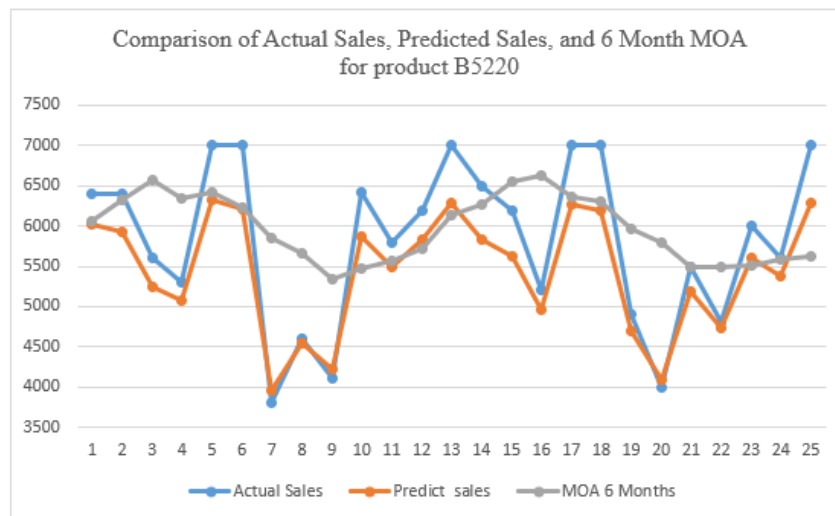


Figure 7: Comparison of Actual sales, Predicted sales, and 6–Month SMA for B5220

Figure 8 illustrates the results of applying linear regression to predict sales revenue, actual sales revenue, and SMA using 6–month sales data. The obtained R-squared value is 0.816, indicating that the model explains 81.6% of the variation in the data. The horizontal axis represents the sales distribution of the B5330 product within 25 months, and the vertical axis represents the sales volume.

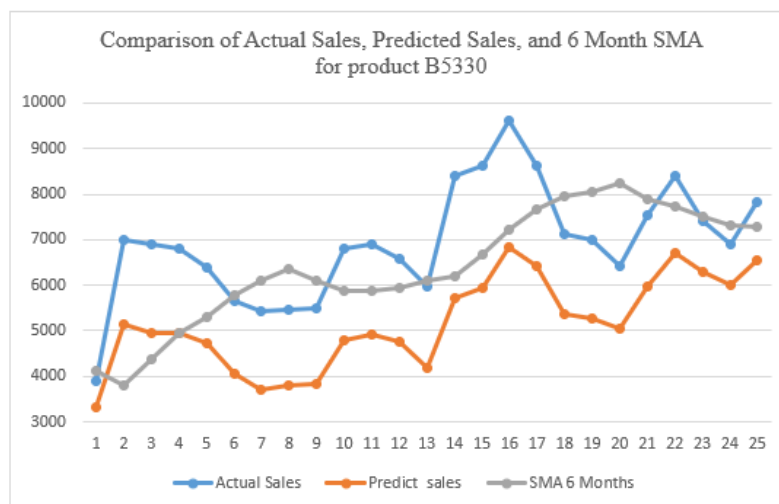


Figure 8: Comparison of Actual sales, Predicted sales, and 6–Month SMA for B5330

Figure 9 illustrates the results of applying linear regression to predict sales revenue, actual sales revenue, and SMA using 6–month sales data. The obtained R-squared value is 0.816, indicating that the model explains 81.6% of the variation in the data. The horizontal axis represents

the sales distribution of the B5320 product within 25 months, and the vertical axis represents the sales volume.

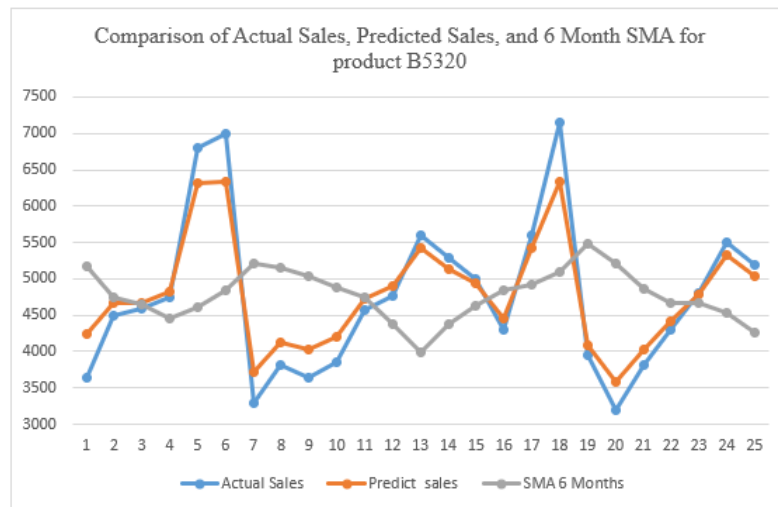


Figure 9: Comparison of Actual sales, Predicted sales, and 6-Month SMA for B5320

Figure 10 illustrates the results of applying linear regression to predict sales revenue, actual sales revenue, and SMA using 6-month sales data. The obtained R-squared value is 0.816, indicating that the model explains 81.6% of the variation in the data. The horizontal axis represents the sales distribution of the UU4222 product within 25 months, and the vertical axis represents the sales volume.

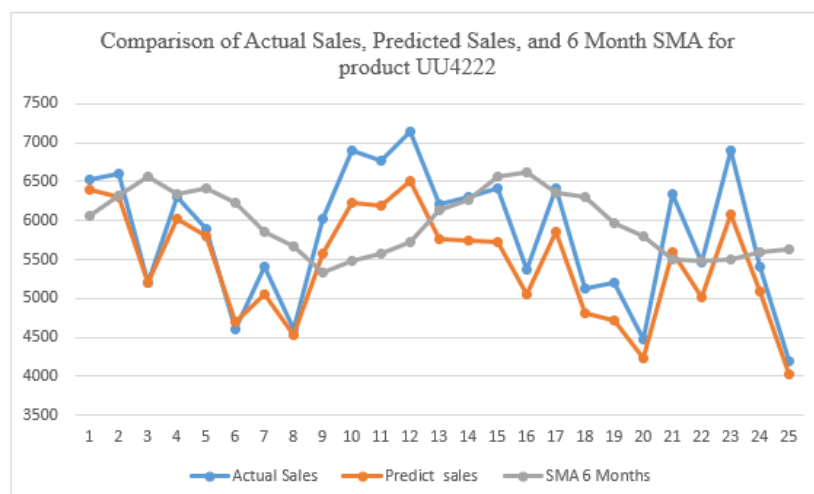


Figure 10: Comparison of Actual sales, Predicted sales, and 6-Month SMA for UU4422

Figure 11 illustrates the results of applying linear regression to predict sales revenue, actual sales revenue, and SMA using 6-month sales data. The obtained R-squared value is 0.816, indicating that the model explains 81.6% of the variation in the data. The horizontal axis represents the sales distribution of the UU4222 product within 25 months, and the vertical axis represents the sales volume.

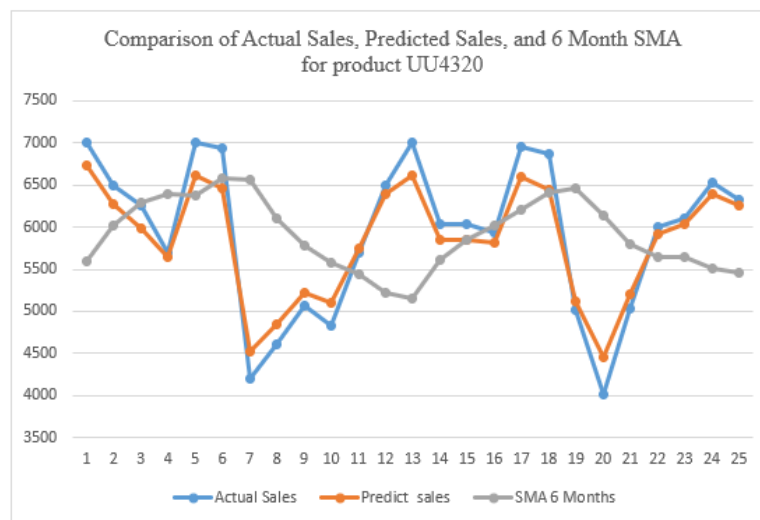


Figure 11: Comparison of Actual sales, Predicted sales, and 6-Month SMA for UU4320

Comparison between sales forecast based on 6-month data and actual sales in 2023

Table 5 shows the error analysis results of using a linear regression model for sales forecasting of five products (B5220, B5330, B5320, UU4222, and UU4320). The model is trained based on sales data from January to June 2022 and compared with actual sales data from 28 months (2022-4 to 2024-7). The overall absolute error of the model is 607, indicating the average deviation level of the model on all products. The prediction error of B5330 is the largest, with an absolute error of 1,751 and a standard deviation of 506, indicating that the sales data for this product fluctuates significantly and the model struggles to capture its trend accurately.

Compared to others, UU4320 has the best predictive performance with an absolute error of only 220 and a standard deviation of 132, indicating that the model's prediction of the product's sales trend is relatively stable and accurate. The errors and absolute errors of B5320 and UU4222

are relatively small, although the model's predictions for some products (such as B5330) still need improvement.

Table 5: Linear regression prediction error analysis based on sales data of five products in six months

Product Model	Error	Absolute Error	Standard Deviation
B5220	-378	407	243
B5330	-1751	1751	506
B5320	32	262	204
UU4222	-384	393	228
UU4320	-81	220	132
All Models	-513	607	646

Summary of Results

The R^2 value significantly increases with the increase of period: from 0.642 at 3 months to 0.816 at 6 months. This indicates that as the training data span extends, the model's explanatory power for sales changes significantly. The overall absolute error decreased from 1528 at 3 months to 607 at 6 months, indicating that the model became more accurate as the amount of data increased. For products with good stability, such as UU4320 and B5220, the prediction error decreases significantly with an increase in the time span. For example, the absolute error of UU4320 decreased from 458 in 3 months to 132 in 6 months. The absolute error of UU4222 decreased from 1403 at 3 months to 393 at 6 months.

For products with high volatility (such as B5330), although the error is still significant, it has decreased from 4156 at 3 months to 1751 at 6 months, indicating that the model has improved in capturing complex trends. The overall standard deviation has also decreased from 1453 at 3 months to 506 at 6 months, indicating a gradual reduction in the volatility of the predicted results.

As time increases, the predictive performance and stability of linear regression models improve significantly. The model based on 6 months of data has the highest R^2 value and the lowest error, making it suitable for long-term sales forecasting. For products with high volatility (such as B5330), although the prediction error has been reduced, some deviation remains, and further model optimization or the introduction of additional variables is needed to improve prediction accuracy.

Discussions

The results indicate that the linear regression model based on six months of data exhibits high predictive performance, with an R-squared value of 0.816, explaining approximately 81.6% of the variation in sales. The study also identified seven key product attributes—including price, promotional quantity, and product model—that show significant correlations with sales volume. This supports the selection of these variables in the linear regression model for accurate sales prediction. According to Long and Liu's (2021) findings, the explanatory power and prediction accuracy of the model significantly improve as the training data period increases. Specifically, the R^2 value increased from 0.642 at 3 months to 0.816 at 6 months, indicating a significant improvement in the model's explanatory power for sales changes. According to Xiong et al. (2024), the study's findings suggest that linear regression is a reliable and comprehensible technique for sales forecasting. Platform profitability can be increased by utilizing sales forecasting models to predict product demand, enhance inventory control, and efficiently schedule procurement. According to Gopalakrishnan et al. (2018), the sales predicted by linear regression had an accuracy of 84% when compared to actual sales in 2014. Zmart can enhance its inventory strategy and boost sales by forecasting sales.

Meanwhile, the overall absolute error decreased from 1528 at 3 months to 607 at 6 months, further demonstrating the model's accuracy improvement with increasing data volume.

Conclusion

This study established a sales forecasting model for patch panel products applicable to the YaQi factory and revealed the impact mechanism of product attributes on sales through data analysis. The research results have important practical application value in improving inventory management efficiency, reducing inventory costs, and optimizing enterprise operations. In the future, the model can be further optimized.

By using a linear regression model and combining sales data for power strip products, the impact of various periods (3 months and 6 months) on the prediction results was studied. The model based on 6 months of data performs the best, with an R^2 value of 0.816, which explains approximately 81.6% of the sales changes. This indicates that with the extension of the data period, the predictive accuracy and stability of the model significantly improve, providing a reliable tool for formulating long-term sales strategies and inventory planning.

Using SPSS software, a correlation analysis was conducted between seven key attributes of power strip products (such as price, initial inventory, final inventory, production volume, etc.) and sales data. The research results indicate that these variables have a statistically significant impact on sales, providing data support for optimizing prediction models and inventory management strategies in the future. By comparing sales forecasts with actual sales data, the model's prediction errors (such as absolute error and standard deviation) were quantified, and the performance of models with different periods was evaluated.

Research has shown that models based on 6 months of data can effectively reduce prediction errors (with an overall absolute error of 607), improve inventory management accuracy, and avoid inventory surplus or shortage. At the same time, a visual evaluation of the forecast results' accuracy was conducted in conjunction with SMA sales, which helps develop more targeted inventory replenishment and production strategies.

Furthermore, this study contributes new knowledge by identifying that a six-month data window significantly enhances the accuracy and stability of sales forecasting models in industrial settings, offering guidance on optimal data selection for regression analysis. The practical implications extend beyond YaQi Factory, providing a replicable approach for other manufacturing environments to tailor inventory and production planning based on data-driven sales forecasts. Overall, the findings offer valuable insights for both academia and industry, supporting improved operational efficiency and cost reduction through precise inventory management. Future research could explore adapting this model to different product categories and integrating more advanced predictive techniques to enhance performance further.

Suggestions

Based on the findings of this study, it is recommended that Yaqi Factory utilize the sales forecasting model to develop detailed operational plans for production scheduling, inventory checks, and storage management. By aligning manufacturing and inventory strategies with accurate sales predictions, the factory can reduce costs, avoid stockouts or overstocking, and improve overall operational efficiency.

Limitations

Although this study has achieved specific results in sales forecasting and inventory management optimization, it still has some limitations. The data only covers the sales records of the

five best-selling plug-in board products from the YaQi factory, excluding the sales data of other product lines and potential new products, which may limit the universality of the model. Although the model exhibits high accuracy (R^2 value of 0.816), there is still some deviation in predicting products with high volatility (such as B5330), indicating that the model has limited adaptability to high-volatility sales data. This study primarily relied on SPSS software for regression analysis. It did not compare it with more complex machine learning methods, such as support vector machines and neural networks, which may not have fully explored the data's potential.

Future Outlook

In future research, incorporating more product line data and cross-industry sales records is recommended to further enhance the universality and applicability of the model.

Extend the time frame for data collection and analyze the impact of special market conditions (such as major holiday promotions or force majeure events) on sales to improve the predictive ability of the model. Based on existing linear regression models, nonlinear prediction models (such as support vector regression, random forest, and deep learning) can be introduced to better capture complex sales trends and fluctuations. By collaborating with the supply chain management system, we can enhance the responsiveness and efficiency of the entire supply chain.

New Knowledge

Firstly, in terms of prediction methods, research has shown that linear regression models based on more extended periods (6 months) have significant advantages in prediction accuracy and model stability, reducing the absolute error from 1528 for 3-month models to 607. Secondly, in terms of identifying key factors, research has confirmed significant correlations between variables such as product price, initial inventory, final inventory, production volume, and sales volume. These factors should be given priority consideration in sales forecasting and inventory planning. In terms of inventory optimization applications, research results support the use of accurate sales forecasts to guide inventory management, thereby reducing the risks of excess or shortage and improving the operational efficiency of enterprises. Finally, the study also revealed the limitations of linear regression models in dealing with high-volatility products (such as B5330) and suggested the introduction of nonlinear prediction models (such as random forests or neural networks) in the future to enhance predictive capabilities.

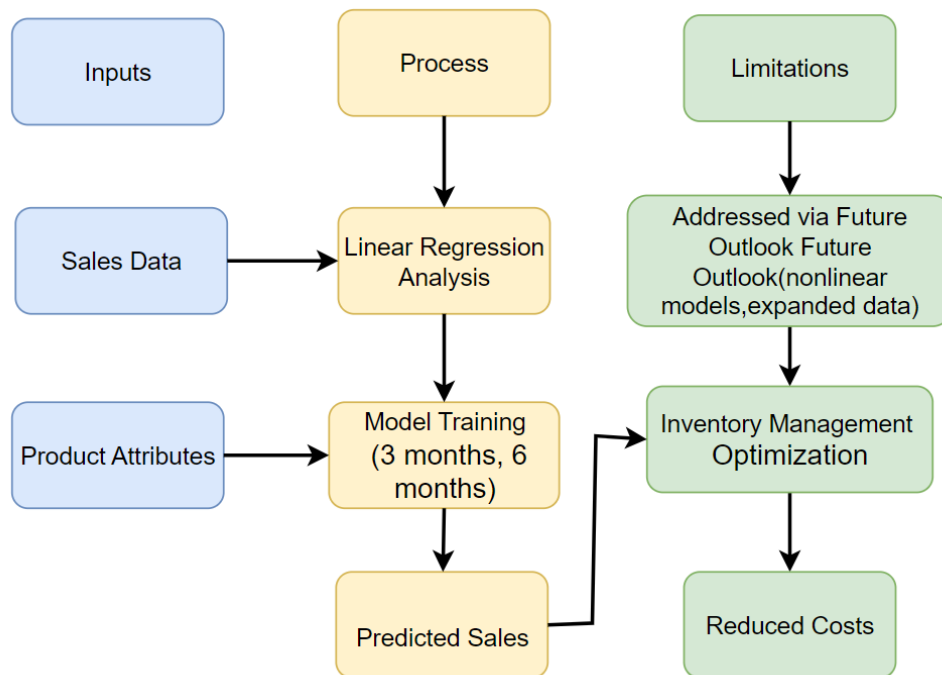


Figure 12: Knowledge Structure of Sales Forecasting and Inventory Optimization Using Linear Regression.

References

- Adeyemo, B. T., & Olatunji, K. B. (2021). Short-Term Load Forecasting Using Weighted Moving Average Method. *IEEE Access*, 9, 43453–43460.
<https://doi.org/10.1109/ACCESS.2021.3065597>
- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). Sage.
- Frost, J. (2020). *Regression analysis: An intuitive guide for using and interpreting linear models*. Statistics By Jim Publishing.
- Gopalakrishnan, T., Choudhary, R., & Prasad, S. (2018). Prediction of sales value in online shopping using linear regression. In *Conference: 2018 4th International Conference on Computing Communication and Automation (ICCCA)*. DOI:10.1109/CCAA.2018.8777620
- Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach* (3rd ed.). The Guilford Press.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021) *An Introduction to Statistical Learning: with applications in R* (2nd ed.). Springer. <https://doi.org/10.1007/978-1-0716-1418-1>

- Long, S., & Liu, Q. (2021). Research on new energy vehicle sales forecast and product optimization based on data mining. In *2021 2nd International Conference on Electronics, Communications and Information Technology (CECIT)* (pp. 1019–1024).
<https://doi.org/10.1109/CECIT53797.2021.00181>
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis* (5th ed.). John Wiley & Sons.
- Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (7th ed.). Open University Press.
- Teja Reddy, S. R., & Malathi, P. (2022). Linear regression and artificial neural networks based efficient sales forecasting model with increased prediction accuracy. In *2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, (pp. 1100–1105). <https://doi.org/10.1109/ICIRCA54612.2022.9985619>
- Seber, G. A. F., & Lee, A. J. (2018). *Linear regression analysis* (2nd ed.). Wiley.
- Silver, E. A., Pyke, D. F., & Peterson, R. (1998). *Inventory management and production planning and scheduling* (3rd ed.). John Wiley.
- Waters, D. (2020). *Inventory control and management* (3rd ed.). Wiley.
- Wild, T. (2017). *Best practice in inventory management* (3rd ed.). Routledge.
- Xiong, R., Yu, L., & Li, Y. (2024). Research on sales forecasting model based on linear regression algorithm. In *Conference: 2024 International Conference on Data Science and Network Security (ICDSNS)*. DOI:10.1109/ICDSNS62112.2024.10690972
- Zhang, Y., Chai, Y., & Ma, L. (2021). Research on multi-echelon inventory optimization for fresh products in supply chains. *Sustainability*, 13(11), 6309. <https://doi.org/10.3390/su13116309>