

SALES FORECASTING FOR FURNITURE STORE

A Machine Learning Approach
to Predict Future Sales

Presented by Vismaya R



TABLE OF CONTENTS

01. Introduction

02. Data Preprocessing

03. Model Selection

04. Model Performance

05. Feature Importance

06. Residual Analysis

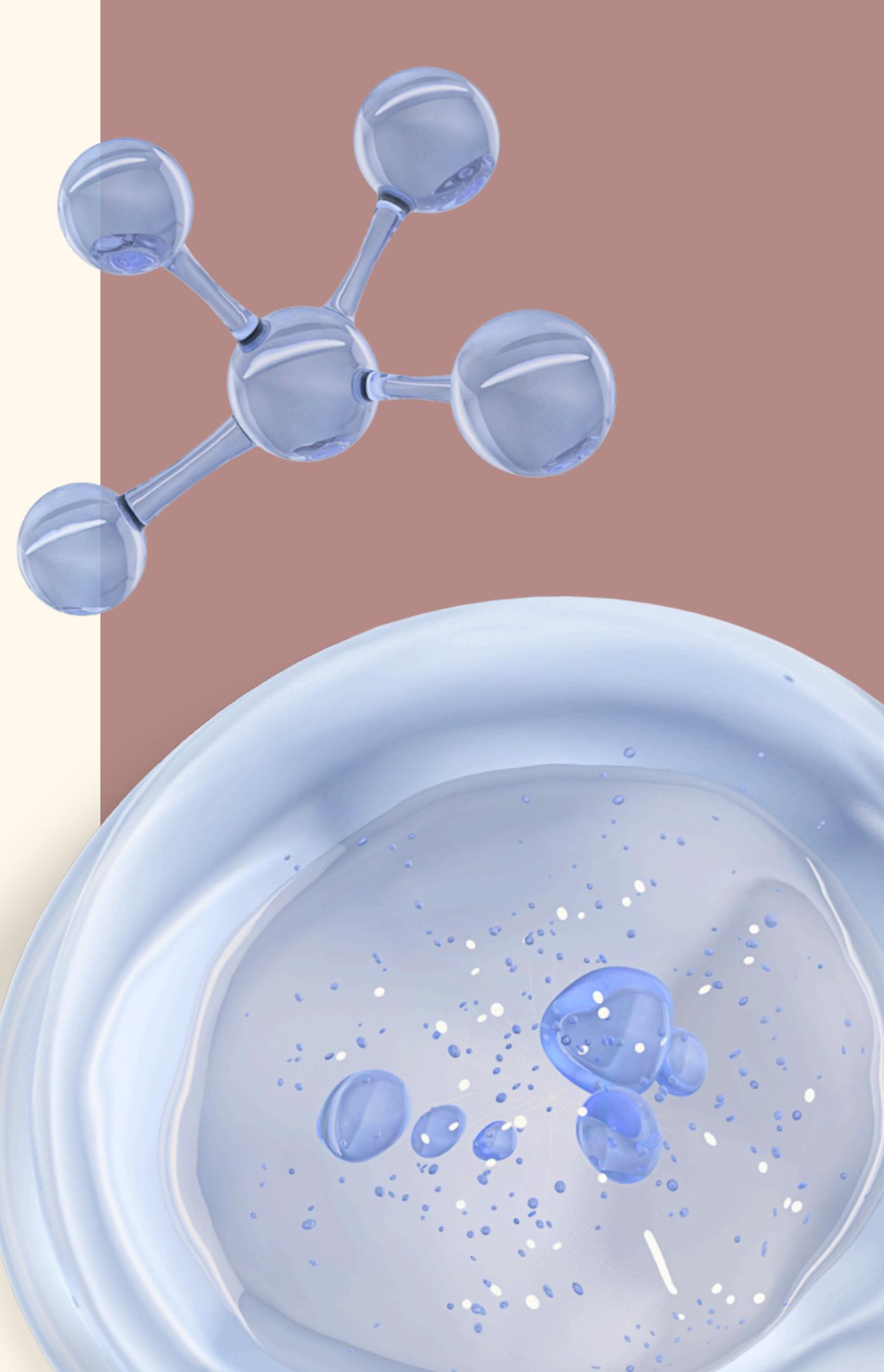
07. Actual vs Predicted Sales

08. Sales Trend Analysis

09. Conclusion

INTRODUCTION

The primary goal of this project is to develop a machine learning model that accurately forecasts future sales for a furniture store. Sales forecasting is crucial for optimizing inventory management, improving marketing strategies, and enabling more effective financial planning.



DATA PREPROCESSING

DATA CLEANING

Missing values in the dataset were handled through imputation, and unnecessary columns such as 'Row ID', 'Customer Name', and 'Order ID' were removed to streamline the analysis.

DATE CONVERSION

The 'Order Date' and 'Ship Date' columns were converted to the datetime format. Features such as year, month, and day of the week were extracted to capture time-based patterns.

FEATURE ENGINEERING

New features such as `Discount_Quantity_Interaction` were created to capture relationships between discounts and sales volume. Categorical variables like 'Region' and 'Category' were label encoded for use in machine learning models.



MODEL SELECTION

- After evaluating various machine learning algorithms, Gradient Boosting Regressor was selected for its ability to handle non-linear relationships and prevent overfitting.
- GridSearchCV was employed to optimize the model's hyperparameters. The following parameters were tuned: learning_rate, max_depth, n_estimators, and subsample.
- Gradient Boosting was chosen because of its ensemble nature, allowing it to combine the predictions of multiple models to improve accuracy. It is particularly effective for structured/tabular data.



MODEL PERFORMANCE

- **EVALUATION METRICS**

Mean Absolute Error (MAE): 126.21

Mean Squared Error (MSE): 87,659.63

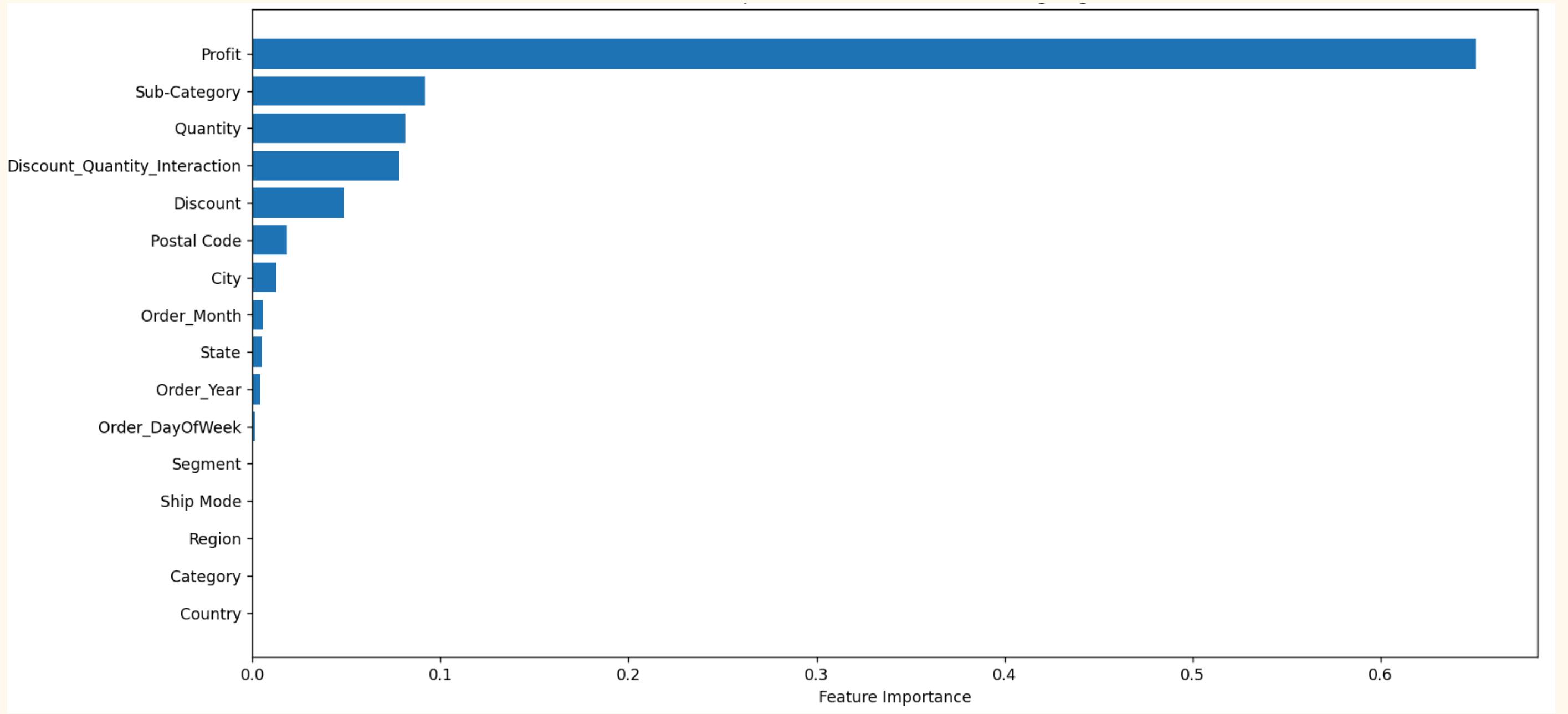
Root Mean Squared Error (RMSE): 296.07

R² Score: 0.71

- **INTERPRETATION**

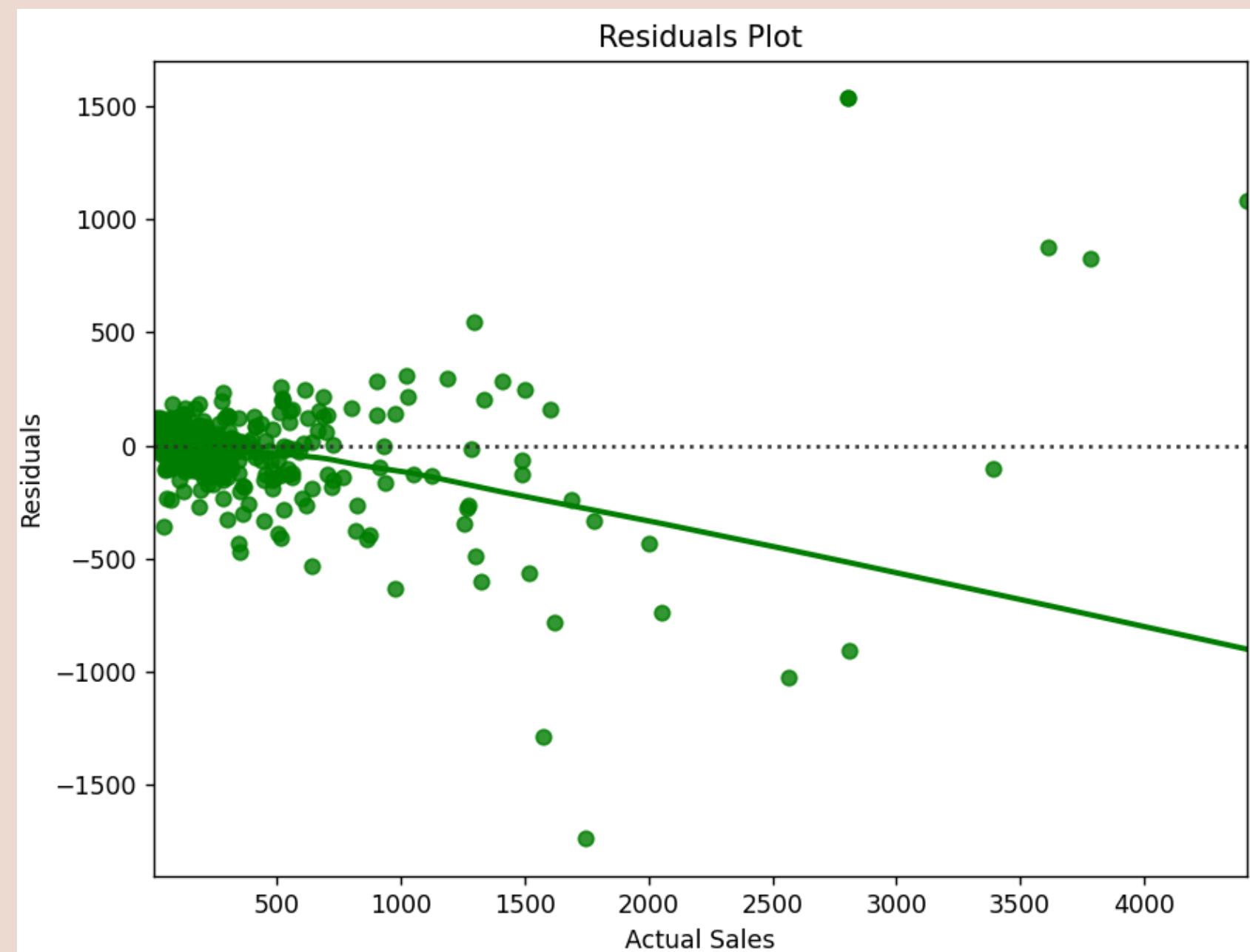
The model achieved a reasonably good R² score of 0.71, indicating that 71% of the variance in sales can be explained by the model. The relatively low RMSE shows that the predictions are close to the actual sales values, though there is still room for improvement.

FEATURE IMPORTANCE



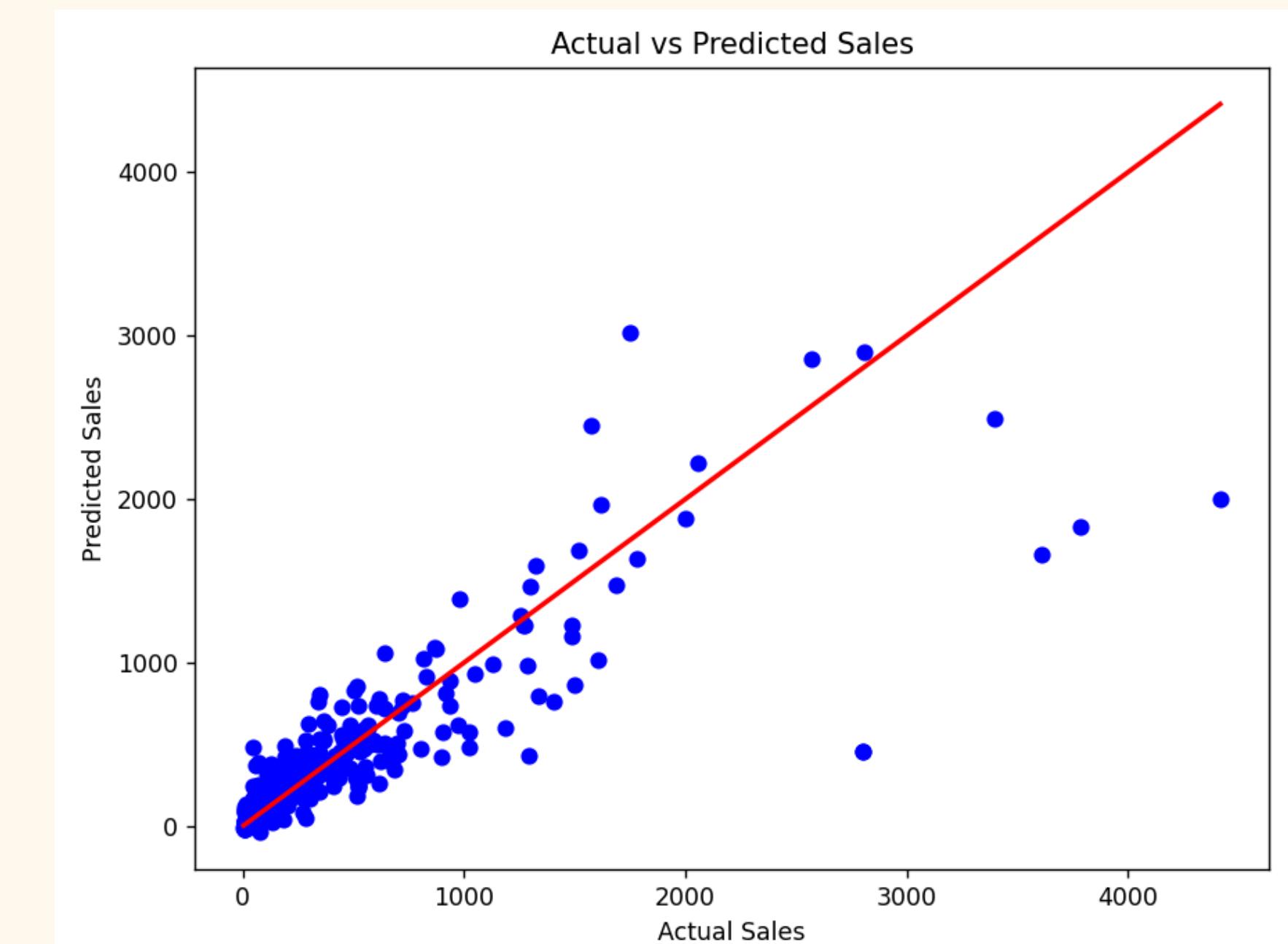
RESIDUAL ANALYSIS

- The residual plot shows the difference between actual and predicted sales. Ideally, residuals should be randomly scattered around zero
- The model residuals appear fairly distributed, indicating that the model captures the general trend well but may have some scope for improvement in specific areas



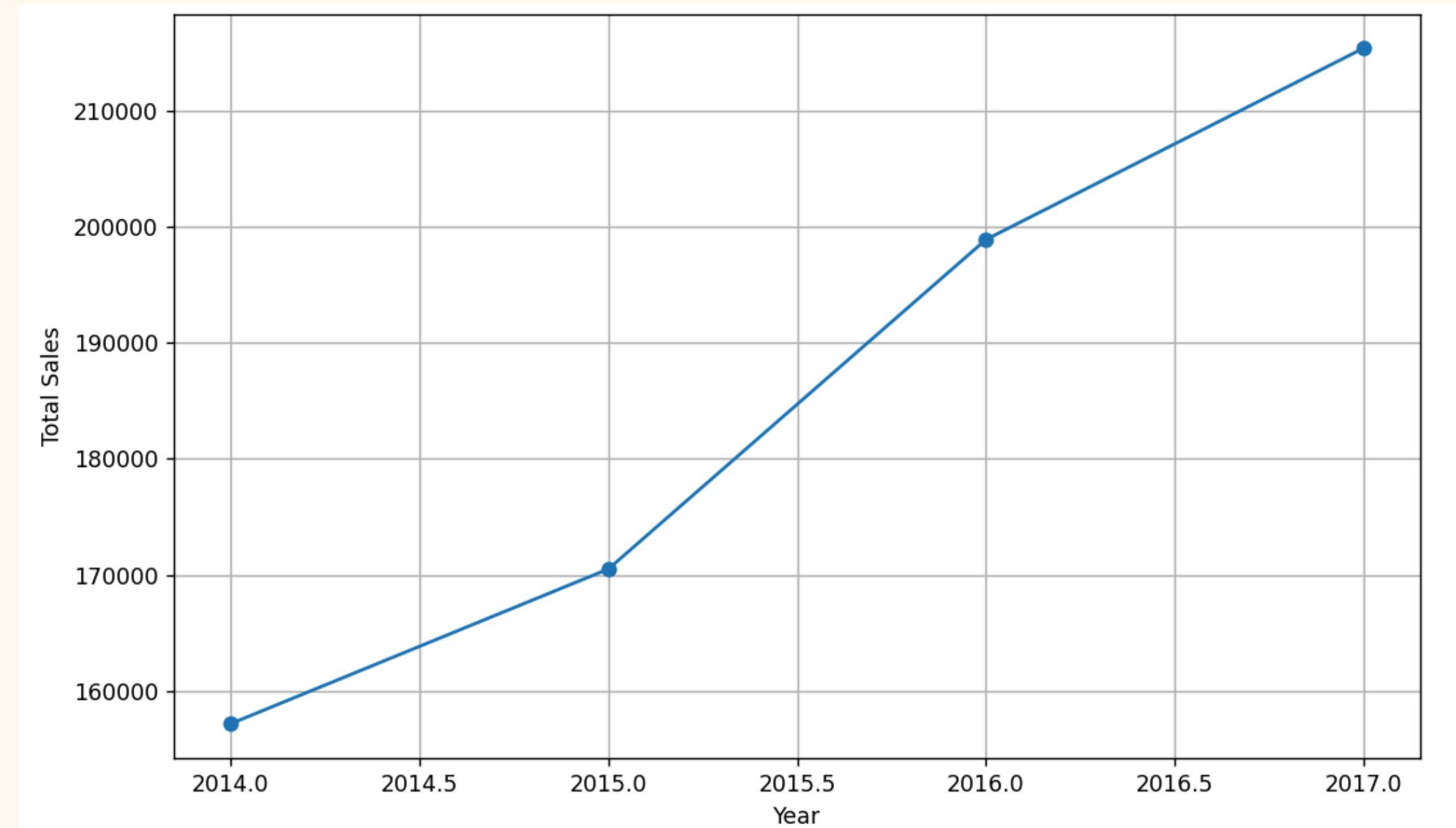
ACTUAL VS PREDICTED SALES

- Each point represents an order, with the x-axis showing the actual sales value and the y-axis showing the predicted sales value
- The red line represents the ideal scenario where actual sales equal predicted sales. The model predictions are reasonably close to actual sales, reflecting its effectiveness in forecasting



SALES TREND ANALYSIS

- The trend analysis shows how sales have fluctuated over the years
- Sales peak during certain periods, reflecting seasonal demand or promotional activities



CONCLUSION

- The Gradient Boosting Regressor performed well in forecasting furniture store sales, with an R^2 score of 0.71. Key factors like Profit significantly influence the sales prediction
- Future work could include improving the model's performance by incorporating additional external data such as marketing campaigns, economic indicators, or customer reviews to further refine predictions. Additionally, deploying the model into a real-time forecasting system for continuous monitoring can enhance decision-making

THANK
YOU!