In learning the application of decision tree models and integration methods (Random Forest, Gradient Boosting) in SAS Enterprise Miner, I mastered how to model and predict data. By adjusting the model parameters and tree depth, I learned how to balance the performance of the model to avoid overfitting or underfitting problems. In addition, I gained insight into the importance of data partitioning and learned how to partition data into training and validation sets and use different evaluation metrics on the training and validation sets to assess the accuracy and generalization ability of the model.
In terms of the analysis of variable importance, I learned how this analysis can be used to identify the features that are most critical to the model's predictions and thus better explain the model's decision-making process. Finally, I gained a deeper understanding of the benefits of integration methods, particularly methods such as Random Forest and Gradient Boosting, and how they can yield stronger predictive performance by integrating multiple models. These learnings enabled me to apply decision trees and integration methods more effectively for data modeling and analysis, enhancing my data science and predictive modeling skills.

Challenges:
Addressing potential overfitting by adjusting the tree depth and parameters of the integration method.

Insights：
A comprehensive look at the importance of variables, decision tree analysis, and integration methods tells us that the number of items purchased is the most critical influence in predicting customer behavior, followed by average rating, days since last purchase, and satisfaction level. Decision tree analysis shows the importance of the depth of the model, too shallow may lead to underfitting while too deep may trigger overfitting, thus better generalization ability is achieved at medium depth. Of the integrated methods, Random Forest demonstrated strong performance by aggregating predictions from multiple decision trees, while Gradient Boosting excelled in determining feature importance and was effective in capturing complex relationships.

Based on these insights, the following business strategies are developed:

- In response to the importance of the number of items purchased, it is possible to increase sales by focusing on marketing strategies that will motivate customers to increase their purchases of goods.
- The impact of average rating, days since last purchase and satisfaction level shows how customers feel about and are loyal to the product or service, therefore it is recommended to focus on the quality of the product or service and timely interaction with customers to increase the satisfaction level.
- Optimize forecasting using integrated methods for more accurate customer engagement and personalized marketing campaigns. This can be done by accurately predicting customer needs based on information such as purchase history and reviews, providing personalized services and recommendations, enhancing the customer experience, and prompting repeat purchases.