Jurnal Akuntansi, Ekonomi dan Manajemen Bisnis
Vol. 13 No. 1, Juli 2025, 63-68
E-ISSN: 2548-9836

Article History
Received May, 2025
Accepted July, 2025

Evaluation of Fast Food Promotion Effectiveness Using a Hybrid Approach of Robust Non-Parametric Methods and Monte Carlo Simulation Based on Market Size Segmentation

I Gusti Ngurah Sentana Putra

IPB University
Statistics and Data Science
Jl. Raya Darmaga Kampus IPB, Babakan, Kec. Dramaga, Kabupaten Bogor, Jawa Barat 16680
E-mail: ngurahsentana@apps.ipb.ac.id

Abstract

This study evaluates the effectiveness of fast food promotions using a hybrid approach combining robust non-parametric methods, Random Forest, and Monte Carlo simulation. The analysis focuses on segmenting the market, by Market Size to identify the most impactful promotional strategies for each segment. Non-parametric A/B testing using the Kruskal-Wallis test revealed significant differences in sales across promotions, with Promotion 1 emerging as the most effective overall. The Random Forest model highlighted LocationID as the most critical factor influencing sales, particularly in large markets. Monte Carlo simulation further demonstrated that Promotion 1 yields the highest Expected Monetary Value, making it the optimal choice for long-term sales growth. The findings emphasize the importance of tailoring promotional strategies to specific market segments, considering factors such as location, timing, and store history. This study provides actionable insights for businesses to optimize promotional campaigns, enhance sales performance, and achieve sustainable growth in this industry.

Keywords: Fast Food, Promotion Effectiveness, Market Size Segmentation, Non-Parametric Methods, Monte Carlo Simulation

1. Introduction

The fast food industry is arguably one among the most dynamic and competitive fields, promotions being a key factor affecting the sales and retaining the customer loyalties. However, gauging how effective promotions are is not an easy calculation. Sales data tend to be characterized by high variability, outliers, and non-normal distributions, making traditional statistical techniques such as ANOVA or t-tests likely to yield unreliable results. This is why strong non-parametric techniques like the Kruskal-Wallis test are favored — as they make no assumptions of normal distributions and are more resistant to outliers. In addition, the effect of promotions could differ by market attributes (Market Size (large, medium, small)). Market Segmentation by Market Size companies need to help identify which target market will respond to which type of promotional methods.

However, relying solely on non-parametric statistical tests may not provide a comprehensive understanding. So, such a hybrid work, which includes strong non-parametric models along with Monte Carlo simulation could offer more insights. With Monte Carlo simulation, organizations can incorporate uncertainty in the data and compute Expected Monetary Value (EMV), a measure of the anticipated monetary value of promotional activities.

In this study, use a hybrid framework. First, start with Non-Parametric A/B testing approach to check the significance of differences between distinct promotional strategies and also show the results per Market Size segment. Second, we obtain feature importance obtained with Random Forest to quantify which variables are the most relevant in predicting sales for each Market Size segment. Finally, a Monte Carlo simulation is used to estimate the long-term effect of each promotional strategy and to compute its EMV. This blend of methodologies allows companies to make more informed and data-led decisions.

One perfect case study for this type of lens is the Fast Food Marketing Campaign. The fast food sector is highly dynamic, competitive, and driven by changing consumer demand. Companies utilize sales data of various Market Size segments to compare and contrast

which promotional strategies work the best and provide the maximum value. By setting the foundation for future business growth, the aim is to replicate our study findings, which focus on strategic recommendations that can boost sales in the short-term.

2. Theoretical Review

2.1. Promotion

Promotion plays a crucial role in the fast food industry, aiming to enhance consumer awareness, attract new customers, and retain existing ones. According to Kotler and Keller (2016), effective promotional strategies, such as advertising, discounts, and loyalty programs, can significantly boost sales and strengthen brand positioning in a competitive market. Research by Aaker (2018) highlights that well-executed promotions can lead to increased customer engagement and brand loyalty.

2.2. Kruskal-Wallis Test

The Kruskal-Wallis test is a non-parametric statistical method used to compare three or more groups. It is particularly useful when the assumption of normality is not met. This test assesses whether there are significant differences between the medians of multiple groups. The formula for calculating the Kruskal-Wallis statistic (H) is:

$$H = (\frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}) - 3(N+1) \qquad (1)$$

where N is the total number of observations, k is the number of groups, R_i is the number of ranks in the ith group, and n_i is the number of observations in the ith group.

2.3. Feature Importance Random Forest

Random Forest is a machine learning algorithm used for classification and regression tasks. One of its strengths is the ability to compute feature importance, which indicates the contribution of each feature to the model's predictions. Breiman (2001) explains that feature importance can be assessed by measuring the decrease in model accuracy when a specific feature is permuted. This provides insights into which features are most influential in making predictions.

2.4. Expected Monetary Value

Expected Monetary Value (EMV) is a decision-making technique used to evaluate risks and uncertainties. EMV is calculated by multiplying the value of each outcome by the probability of that outcome occurring and summing all these values.

2.5. Monte Carlo

The Monte Carlo method is a statistical technique used to estimate the outcomes of a process involving uncertainty. This method employs random sampling to generate a distribution of results. Rubinstein and Kroese (2016) note that Monte Carlo can be applied across various fields, including finance, engineering, and computer science, for risk analysis and decision-making.

3. Methods

This research was conducted systematically using data obtained from Kaggle, a platform that provides datasets for data analysis and machine learning. The dataset used in this study contains information about fast food marketing campaign information, variables: Promotion, Market Size, Sales in thousands, Location ID, Age of Store & Week; Python on Google Colab — where one can code, process, and visualize in a cloud-based environment. I used python libraries Pandas and NumPy to import the dataset into Google colab.

To check the structure and distribution of the data. To assess how effective the different promotions were, A/B testing was carried out with Kruskal-Wallis test. The Wilcoxon rank-sum test was selected as it does not assume a normal distribution and is robust to outliers. This was conducted to compare sales across promotion types (Promotion 1,2,3) overall and within each Market Size segment (Large, Medium, Small). Where large differences were identified, a post-hoc Mann-Whitney U test with Bonferroni correction was performed to establish which of the promotions were different to each other.

Random Forest Regressor was used to determine the important factors determining sales in each of the Market Size segment. The independent variables were Promotion, Location ID, Age of Store and Week and the dependent variable was Sales in Thousands. We divided the Market Size value into subsets and trained a different Random Forest model on each of the subsets. Feature importance values from the models were evaluated to see which variable contributed most to sales. This step allowed us to derive actionable insights regarding the factors driving sales in various market segments.

Monte Carlo simulation was performed to estimate the long-term impact of promotional strategies and the expected monetary value. - For each promotion type, the probability of its implementation being a success (or not) was used to generate a failure/success matrix that would be multiplied with the reward matrix (which contains the revenue values) to generate the EMVs. To ensure robust results, the same simulation was run for multiple iterations 100,000. As such, we were able to perform a probabilistic evaluation of the financial implications of various promotional approaches, informing our subsequent decision-making.

4. Result and Discussion

4.1. Statistic Descriptive

The distribution of promotions (Promotion 1, 2, and 3) across different Market Sizes:

TABLE I

OVERVIEW PROMOTION BASED ON MARKET SIZE

MarketSize	Promotion 1	Promotion 2	Promotion 3
Large	56	64	48
Medium	96	108	116
Small	20	16	24

From the table above, it is evident that the distribution of promotions varies across Market Sizes. Promotion 3 is the most utilized promotion in the medium markets (116), while Promotion 2 is more predominant (64) in the Large markets. This variation indicates that the power of promotions may vary by market size and deserves a deeper dive.

4.2. A/B Testing with Kruskal-Wallis test

The overall non-parametric A/B testing using the Kruskal-Wallis test revealed a statistically significant difference in sales across the three promotion types (Promotion 1, 2, and 3). The test statistic was 53.29, with a p-value of 2.67e-12, which is much lower than the significance level of 0.05. This indicates that there is strong evidence to reject the null hypothesis, confirming that at least one promotion type significantly differs from the others in terms of sales performance. This finding underscores the importance of evaluating promotional strategies to identify the most effective approach.

When the analysis was segmented by Market Size, the results further highlighted significant differences in sales across promotions within each segment. For Medium markets, the Kruskal-Wallis test yielded a test statistic of 45.67 and a p-value of 1.21e-10, indicating significant differences. Similarly, in small markets, the test statistic was 21.05 with a p-value of 2.69e-05, and in Large markets, the test statistic was 41.03 with a p-value of 1.23e-09. These results demonstrate that the effectiveness of promotions varies significantly across different market sizes, emphasizing the need for tailored promotional strategies based on market characteristics.

4.3. Mann-Whitney U with Bonferroni Correction

The Mann-Whitney U test is a non-parametric statistical test used to compare two independent groups when the assumptions of normality are not met (McKnight & Najab, 2010). It is particularly useful for analyzing ordinal data or data with outliers, as it relies on rank-based methods rather than raw values (Fagerland & Sandvik, 2009). When multiple comparisons are conducted, the Bonferroni correction is applied to adjust the p-values, reducing the risk of Type I errors (false positives) by dividing the significance level (e.g., 0.05) by the number of comparisons (Bland & Altman, 1995). In this study, the Mann-Whitney U test with Bonferroni correction was conducted to identify which specific promotion pairs showed significant differences in sales after the Kruskal-Wallis test indicated overall significance. Below is a summary of the results:

TABLE II

MANN-WHITNEY U WITH BONFERRONI CORRECTION RESULTS

Promo 1	Promo 2	Statistic U	P-value	P-value (Corrected)
3	2	23251	1.20E-07	3.59E-07
3	1	14089	3.51E-02	1.05E-01
2	1	9378.5	5.85E-12	1.75E-11
Medium Market				
Promo1	Promo2	Statistic U	P-value	P-value (Corrected)
3	2	8829	1.21E-07	3.64E-07
3	1	4848	1.06E-01	3.17E-01
Small Market				
Promo1	Promo2	Statistic U	P-value	P-value (Corrected)
1	3	251.5	0.795418	2.386254
1	2	288	0.000049	0.000148
3	2	338	0.000059	0.000177
Large Market				
Promo1	Promo2	Statistic U	P-value	P-value (Corrected)
3	1	1339	9.77E-01	2.93E+00
3	2	2499.5	1.50E-08	4.49E-08
1	2	2788.5	1.61E-07	4.84E-07

Source : Processed Researcher 2025

The results of the post-hoc Mann-Whitney U test with Bonferroni correction reveal significant differences in the effectiveness of promotional strategies across various Market Sizes. Promotion 2 consistently demonstrates a highly significant advantage over Promotion 1 in all market segments, with corrected p-values ranging from 1.75e-11 (overall) to 4.84e-07 (Large markets). Similarly, Promotion 3 also shows a significant edge over Promotion 2 in most cases, with corrected p-values such as 3.59e-07 (overall) and

4.49e-08 (Large markets). However, the comparison between Promotion 3 and Promotion 1 does not show significant differences, with corrected p-values like 0.105 (overall) and 2.93 (Large markets), indicating that while Promotions 2 and 3 are more effective than Promotion 1, there is no clear winner between Promotions 3 and 1.

When examining the results by Market Size, the trends remain consistent but with some nuances. In Medium markets, Promotion 2 significantly outperforms Promotion 1 and Promotion 3, while Promotion 3 and Promotion 1 show no significant difference. In Small markets, both Promotion 1 and Promotion 3 significantly outperform Promotion 2, but there is no significant difference between them. In Large markets, Promotions 1 and 3 both significantly outperform Promotion 2, but again, there is no significant difference between Promotions 1 and 3. These findings highlight the importance of tailoring promotional strategies to specific market segments, as the effectiveness of promotions varies significantly depending on the Market Size.

4.4. Feature Importance with Random Forest

Feature importance is a metric used in machine learning models, such as Random Forest, to quantify the contribution of each input feature (variable) to the model's predictions. In Random Forest, feature importance is often measured using Gini Importance or Mean Decrease in Impurity (MDI), which evaluates the total reduction in impurity achieved by splitting the data on a particular feature across all trees in the forest (Louppe, 2014). The importance values are normalized so that the sum of all feature importance values equals 1, allowing for easy comparison of the relative importance of features within the model. This method is widely used in predictive analytics to identify key drivers of outcomes, such as sales performance in marketing campaigns (Hastie et al., 2009).

TABLE III

FEATURE IMPORTANCE RESULTS

Feature	Medium	Small	Large
LocationID	0.685144	0.400895	0.870036
week	0.127567	0.237191	0.031653
Promotion	0.101824	0.129929	0.068841
AgeOfStore	0.085465	0.231985	0.029469

Source : Processed Researcher 2025

The feature importance analysis revealed that LocationID is the most critical factor influencing sales across all Market Sizes, with the highest importance values in large markets (0.870), followed by Medium (0.685) and small markets (0.401). This indicates that the location of the store has a dominant influence on sales, likely due to factors such as foot traffic, local demographics, and competition. In contrast, the Promotion variable, while still significant, has a relatively smaller impact, with importance values ranging from 0.069 (Large) to 0.130 (Small). This suggests that while promotions do affect sales, their effectiveness is secondary to the store's location. Additionally, the week variable shows moderate importance, particularly in small markets (0.237), indicating that the timing of promotions (e.g., specific weeks) can influence sales, especially in smaller markets where seasonal or weekly trends may have a stronger impact.

The results highlight the importance of tailoring strategies based on Market Size. In Small markets, factors such as week (0.237) and AgeOfStore (0.232) play a relatively larger role, suggesting that timing and store history are more influential in these segments. In contrast, in large markets, LocationID (0.870) dominates, with other factors having minimal impact. This implies that companies should prioritize location selection in large markets, while in smaller markets, they should also consider the timing of promotions and the historical performance of stores. Overall, the findings emphasize that while promotions are effective, their impact is context-dependent, and companies should focus on optimizing location and timing to maximize sales performance. These insights can guide decision-making, helping businesses allocate resources more effectively and design targeted strategies for each market segment.

4.5. Long-term Impact Using Monte Carlo Simulation

The Monte Carlo simulation was conducted to predict the long-term impact of promotional strategies by modeling uncertainty in sales data. The process began by grouping sales data based on the type of promotion (Promotion 1, 2, and 3) and calculating the median sales for each promotion. Next, 100,000 iterations of the simulation were performed, where each iteration randomly selected a promotion based on predefined probabilities. For each selected promotion, a sales value was randomly generated using a normal distribution with the median sales as the mean and the standard deviation of the sales data as the measure of variability. The simulation results were then collected and analyzed to calculate the Expected Monetary Value (EMV) for each promotion.

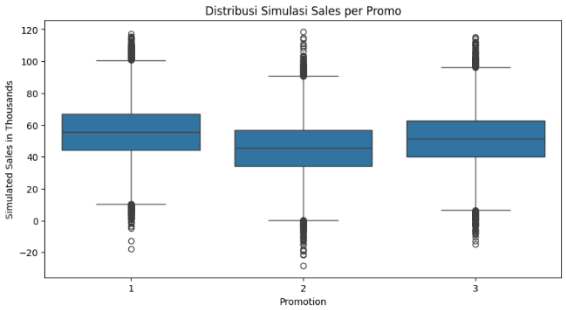

Figure 1: Boxplot Monte Carlo Simulation

Source: Processed Researcher 2025

The Monte Carlo simulation results revealed that Promotion 1 had the highest Expected Monetary Value (EMV) of 55.49, followed by Promotion 3 with an EMV of 51.43, and Promotion 2 with an EMV of 45.42. These findings suggest that Promotion 1 has the greatest potential in terms of expected revenue generation. Furthermore, the visualization of the simulated sales distributions indicated that Promotion 1 not only had a higher average sales performance but also demonstrated lower variability, reflecting greater consistency and reliability. Based on these results, Promotion 1 appears to be the most promising strategy for enhancing long-term sales.

However, it is important to note that the EMV calculations presented do not account for promotion costs. Therefore, while Promotion 1 may yield the highest revenue potential, its financial efficiency remains uncertain without considering the associated expenses. The company should also take into account additional factors such as promotional cost structures, market segmentation, and target audience responsiveness before making a final decision. A comprehensive cost-benefit analysis is recommended to ensure that the selected promotion strategy is not only effective but also economically viable.

4.6. Bussiness Recommendations

Based on the findings of this study, it is evident that Promotion 1 is the most effective strategy for increasing sales across all market segments, particularly in the long term, as demonstrated by the Monte Carlo simulation results. Companies should prioritize Promotion 1 in their marketing campaigns to maximize sales performance. However, it is also crucial to consider the location of stores, as LocationID emerged as the most influential factor across all Market Sizes. In Large markets, where location dominates, businesses should focus on selecting high-traffic areas to establish their stores. In Small markets, where timing (week) and store history (AgeOfStore) play a significant role, companies should align their promotional activities with seasonal trends and leverage the historical performance of their stores to optimize sales.

Additionally, the results highlight the importance of tailoring promotional strategies to specific market segments. For instance, in medium markets, Promotion 2 showed significant effectiveness, while in Small markets, both Promotion 1 and Promotion 3 outperformed Promotion 2. This suggests that a one-size-fits-all approach may not be optimal. Companies should adopt a segmented strategy, customizing promotions based on the unique characteristics of each market. By combining the strengths of Promotion 1 with location optimization and market-specific adjustments, businesses can enhance their overall sales performance and build a strong foundation for long-term growth.

5. Conclusion and Suggestion

In conclusion, this study demonstrates the effectiveness of using a hybrid approach combining robust non-parametric methods, Random Forest, and Monte Carlo simulation to evaluate fast food promotional strategies. The results reveal that Promotion 1 is the most effective in driving sales, while LocationID is the most critical factor influencing sales across all Market Sizes. The findings also emphasize the importance of considering market-specific factors, such as timing and store history, particularly in smaller markets. By leveraging these insights, companies can make data-driven decisions to optimize their promotional strategies and improve sales performance.

For future research, it is recommended to incorporate additional variables such as customer demographics, competitor activities, and economic conditions to further refine the analysis. Additionally, exploring the cost-effectiveness of different promotional strategies could provide a more comprehensive understanding of their impact on profitability. Companies should also consider conducting real-world experiments to validate the findings of this study and ensure the practical applicability of the recommended strategies. By continuously refining their approach based on data-driven insights, businesses can stay competitive in the dynamic fast food industry.

References

Aaker, D. A., “Building Strong Brands”, Simon and Schuster, 2018.

Bland, J. M., & Altman, D. G., “Multiple significance tests: The Bonferroni method”, BMJ, 310(6973), 170,1995. https://doi.org/10.1136/bmj.310.6973.170

Breiman, L., “Random Forests. Machine Learning”, 45(1), 5-32, 2001.

https://doi.org/10.1023/A:1010933404324

Conover, W. J., “Practical Nonparametric Statistics”,

3rd ed. Wiley, 1999.

Fagerland, M. W., & Sandvik, L., “Performance of five two-sample location tests for skewed distributions with unequal variances. Contemporary Clinical Trials”, 30(5), 490-496, 2009. https://doi.org/10.1016/j.cct.2009.06.007

Hastie, T., Tibshirani, R., & Friedman, J., “The Elements of Statistical Learning: Data Mining, Inference, and Prediction”, Springer, 2009.

Kotler, P., & Keller, K. L., “Marketing Management (15th ed.), Pearson, 2016.

Louppe, G., “Understanding Random Forests: From Theory to Practice”, arXiv preprint arXiv:1407.7502, 2014.

McKnight, P. E., & Najab, J., “Mann-Whitney U Test. In The Corsini Encyclopedia of Psychology”, John Wiley & Sons, Inc, 2010. https://doi.org/10.1002/9780470479216.corpsy0524

Rubinstein, R. Y., & Kroese, D. P., “Simulation and the Monte Carlo Method (3rd ed.), Wiley, 2016.