



Evaluation of Promotion Strategies: An Analytical Report

Nima Azar

Contents

1	Executive Summary	3
2	Introduction	4
3	Data Overview	4
4	Recommendations for Data Collection	5
5	Methodology	5
5.1	Exploratory Data Analysis	5
5.1.1	Normality Test of Sales Data	5
5.1.2	Sales Distribution by Promotion Type	6
5.1.3	Sales by Market Size	8
5.1.4	Sales by Age of Store	8
5.2	A/B Testing	9
5.2.1	Kruskal-Wallis Test	9
5.2.2	Post-Hoc Analysis	10
5.2.3	Bias Assessment	11
5.3	Regression Analysis with Location as a Fixed Effect	13
5.3.1	Isolating the Effect of Promotion Type	13
5.3.2	Key Outputs	13
5.3.3	Interpretation	14
5.3.4	Percentage Difference in Sales Between Promotion 1 and Other Promotions	14
5.4	Key Performance Indicators (KPIs)	15
6	Conclusion	17

1 EXECUTIVE SUMMARY

This report evaluates the effectiveness of three promotion strategies used by a fast-food chain to increase the sales of a newly promoted item. The main goals were to evaluate the A/B testing results to determine which marketing strategy worked best and to identify the key performance indicators (KPIs) for analysis.

KEY FINDINGS

- **Promotion Effectiveness:** Promotion 1 was the most effective in generating higher sales of the newly promoted item compared to Promotions 2 and 3.
- **Percentage Difference:** Based on regression analysis, Promotion 1 outperformed Promotion 2 by approximately 14.3% and Promotion 3 by approximately 36.1% in terms of sales.
- **Impact of Location:** Sales varied significantly across different locations, which influenced the overall results. This factor was controlled for in the analysis.
- **Statistical Insights:** Significant differences were identified between Promotions 1, 2, and 3, with the true difference between Promotions 1 and 3 becoming clear only after controlling for location.
- **Other Factors:** While market size and store age also impacted sales, their distribution across promotions did not affect the validity of our A/B test.

FRAME 1: CONCLUSIONS

Promotion 1 was the most effective strategy for increasing sales. After accounting for location, Promotion 1 outperformed Promotion 2 by 14.3% and Promotion 3 by 36.1%.

RECOMMENDATIONS

- **Prioritize Promotion 1:** Implement Promotion 1 in future campaigns due to its proven effectiveness.
- **Account for Location:** Future strategies should take location-specific effects into account to optimize results.
- **Utilize Pre-Test Sales Data:** Gathering sales data before future A/B tests could provide valuable insights for trend analysis.

2 INTRODUCTION

In the fast-food industry, effective promotions are key to boosting sales. To find out which promotion strategy works best, our fast-food chain tested three different strategies for a new menu item across various locations.

This report analyzes the results of these tests, aiming to identify the promotion that drives the highest sales and to determine the key performance indicators (KPIs) that will guide future marketing efforts. The analysis uses a complete dataset that includes Market Size, Location, Age of Store, Promotion Type, and Sales in Thousands. While the main focus is on comparing the promotion strategies, the analysis also considers how factors like location, market size, and store age might affect the results.

By identifying the most effective promotion, the company can enhance its marketing strategy and improve sales of the new item.

3 DATA OVERVIEW

The dataset used for this analysis includes 548 records, each showing weekly sales data from different store locations where three promotion strategies were tested for four weeks. The data is complete, with no missing values. Key details about the data are:

- **MarketID:** ID for the market where the promotion ran.
- **MarketSize:** Size of the market (Small, Medium, Large).
- **LocationID:** ID for the store location.
- **AgeOfStore:** Age of the store in years.
- **Promotion:** The promotion type (1, 2, or 3).
- **Week:** The week number of the experiment (1 to 4).
- **SalesInThousands:** Sales of the new item in thousands of dollars for that week.

The data includes both numbers and categories, and with no missing values, it provides a solid basis for analyzing the promotion strategies.

4 RECOMMENDATIONS FOR DATA COLLECTION

Based on the analysis conducted, several recommendations can be made to improve future data collection and analysis efforts:

- **Collect Location Data in Coordinates:** In future analyses, consider collecting location data in the form of geographic coordinates (latitude and longitude). This would allow for more advanced visualizations and a better spatial understanding of the impact of different promotion strategies. Given the small size of our dataset, we could benefit from more detailed types of information.
- **Include Pre-Test Sales Data:** Having access to sales data from before the start of the A/B test would provide valuable context and help in understanding trends over time. This historical data could improve the accuracy of the analysis and provide a clearer baseline for comparison.
- **Assess Adequacy of Sample Size:** It is currently unclear whether the number of observations is sufficient for the analysis, as we lack information on the total number of restaurants involved. Future studies should ensure that the sample size is appropriate to confidently draw conclusions about the effectiveness of different promotion strategies.

By addressing these points, future analyses can be more robust, providing deeper insights and more reliable results.

5 METHODOLOGY

This analysis was conducted in several stages, including exploratory data analysis, statistical and A/B testing, and also regression analysis.

5.1 EXPLORATORY DATA ANALYSIS

The Exploratory Data Analysis (EDA) was conducted to gain initial insights into the dataset and to guide the selection of appropriate statistical methods for further analysis. The EDA focused on the distribution of sales, differences across promotion types, and the impact of market size, store age, and location on sales.

5.1.1 NORMALITY TEST OF SALES DATA

The Shapiro-Wilk test was performed to assess the normality of the sales data. The results showed a test statistic of 0.935 and a p-value less than 0.001, indicating that the data does not follow a normal distribution. As a result, non-parametric tests were deemed appropriate for further analysis.

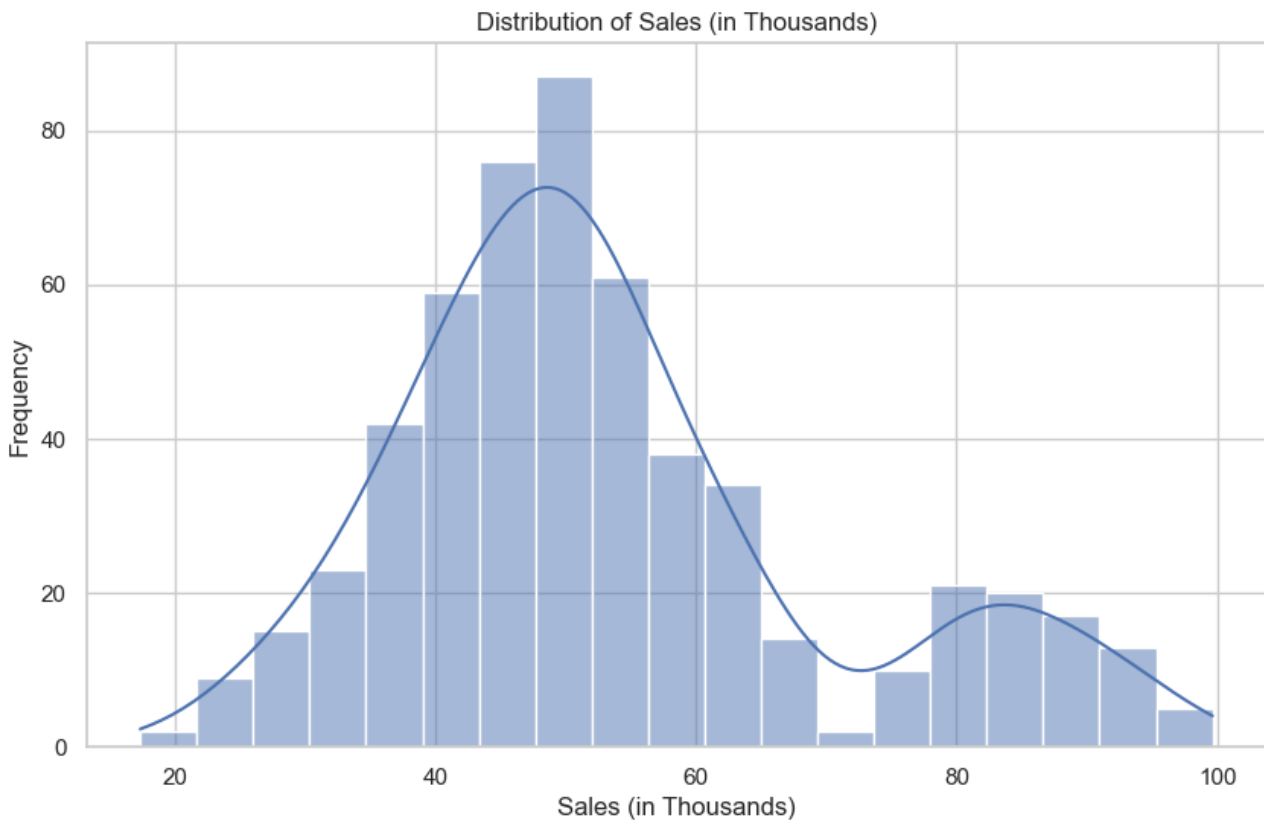


FIGURE 1: DISTRIBUTION OF SALES (IN THOUSANDS)

Shapiro-Wilk Test: Statistic=0.935, p-value=9.38e-15

Conclusion: The sales data does not follow a normal distribution. Non-parametric tests were recommended for subsequent analysis.

5.1.2 SALES DISTRIBUTION BY PROMOTION TYPE

The average sales over the 4-week period were analyzed for each promotion type. The following plots show the average sales and the distribution of sales by promotion type, revealing differences in how each promotion strategy performed.

Key Observations:

- Promotion 1 appears to generate higher average sales compared to the other promotions.
- The distribution of sales varies across promotion types, indicating potential differences in effectiveness.

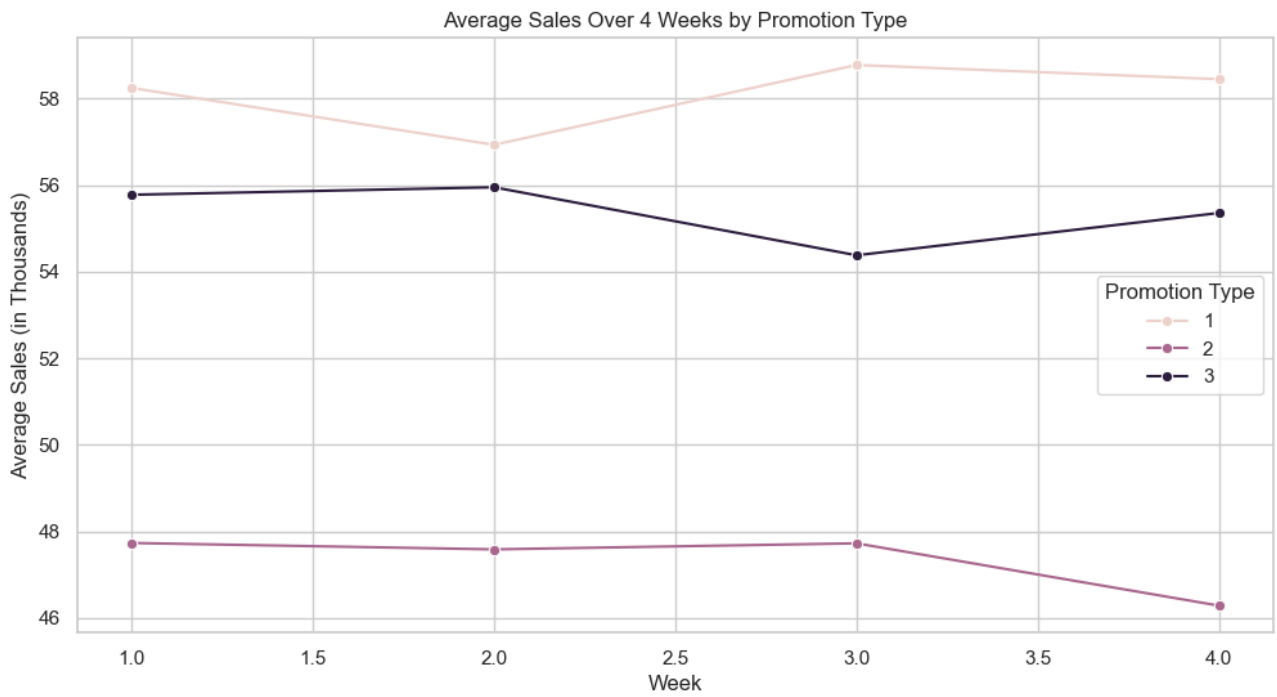


FIGURE 2: AVERAGE SALES OVER 4 WEEKS BY PROMOTION TYPE

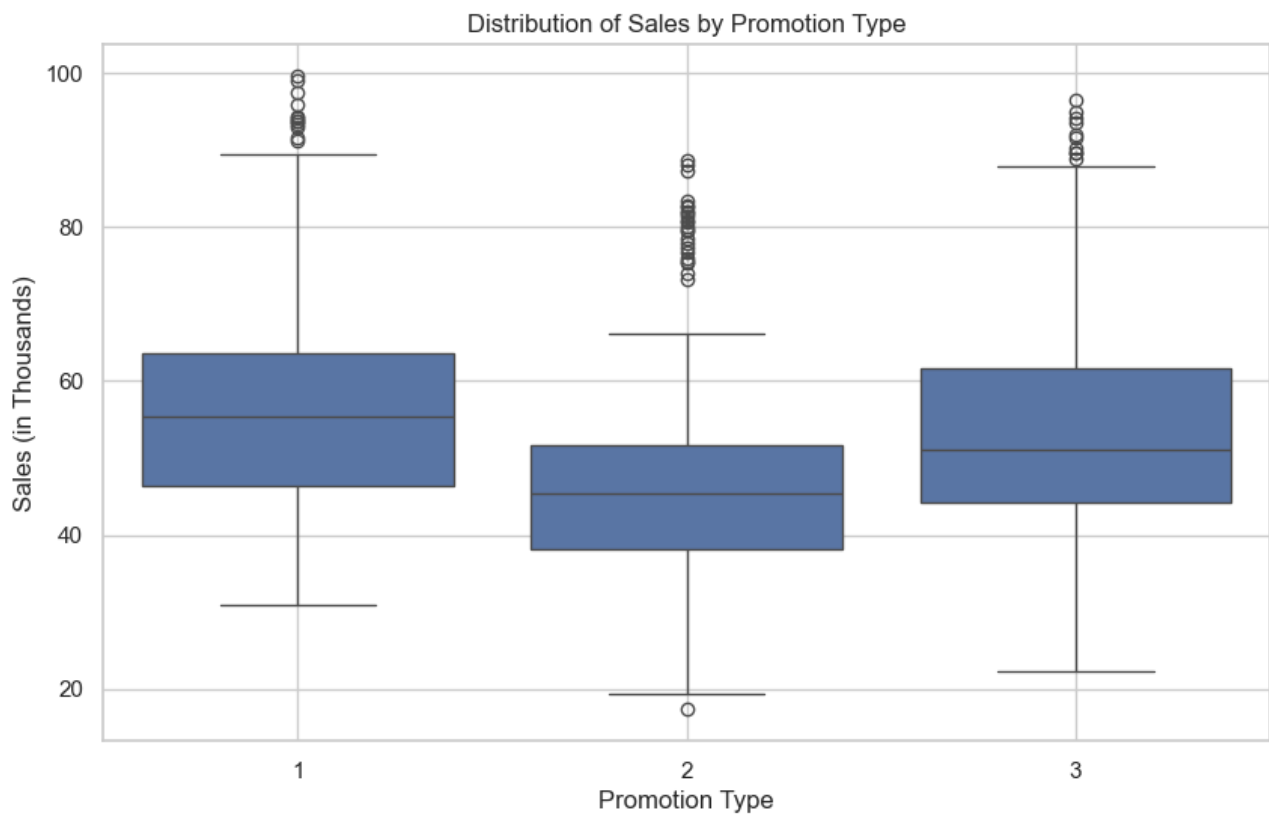


FIGURE 3: DISTRIBUTION OF SALES BY PROMOTION TYPE

5.1.3 SALES BY MARKET SIZE

Market size was analyzed to understand its impact on sales performance. Descriptive statistics and a boxplot were generated to compare sales across different market sizes.

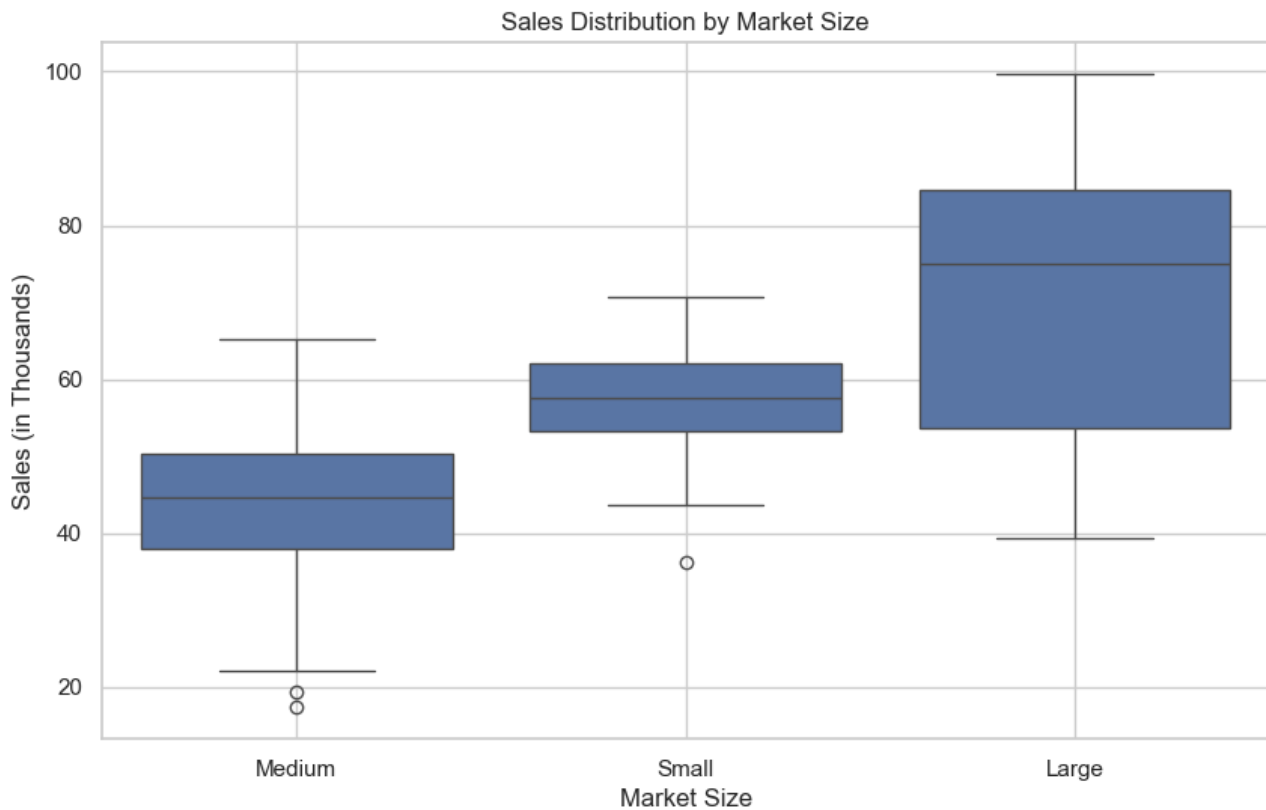


FIGURE 4: BOXPLOT OF SALES BY MARKET SIZE

Descriptive Statistics:

- **Large Markets:** Higher average sales (mean $\approx 70.12K$) with a wider spread (std $\approx 17.05K$).
- **Medium Markets:** Lower average sales (mean $\approx 43.99K$) with less variation (std $\approx 9.05K$).
- **Small Markets:** Average sales in between (mean $\approx 57.41K$) with less variation (std $\approx 6.63K$).

Boxplot Insights:

- Large Markets have the highest range of sales, suggesting market size is a significant factor.
- Medium Markets show more consistent sales, while Small Markets also display consistency but at a lower range.

5.1.4 SALES BY AGE OF STORE

The age of the store was explored to determine how it influences sales. Descriptive statistics and a boxplot were generated to compare sales across different store ages.

Descriptive Statistics:

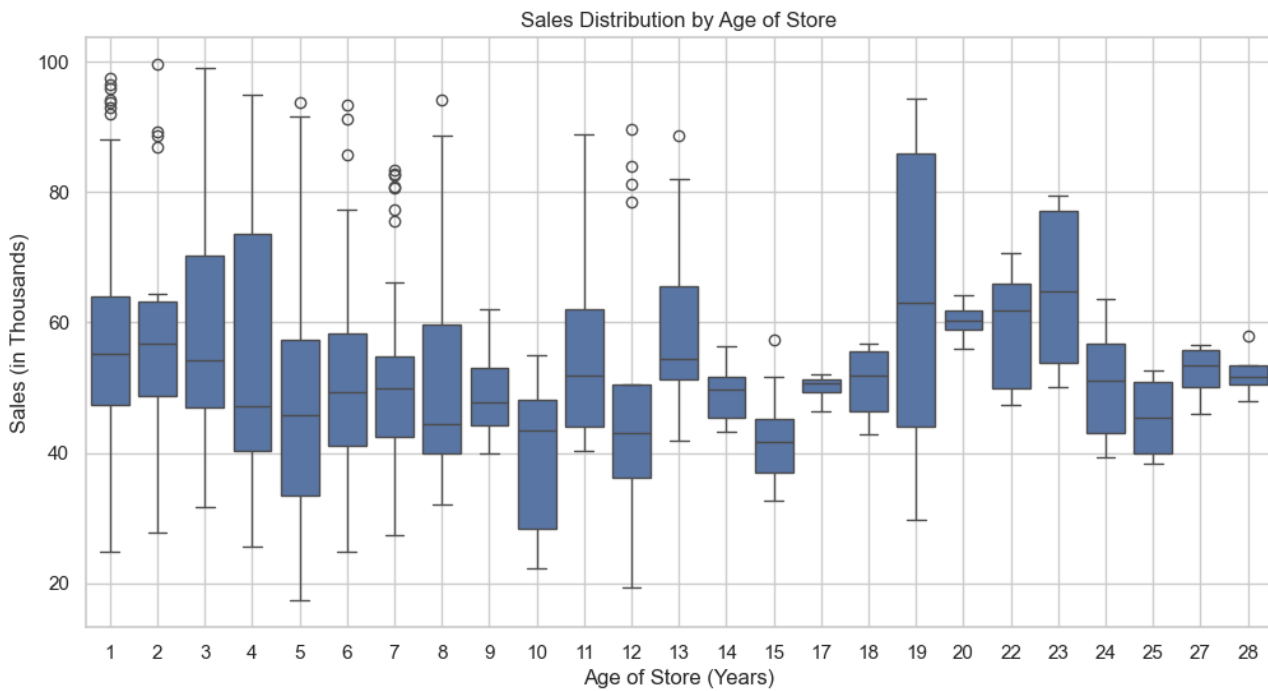


FIGURE 5: BOXPLOT OF SALES BY AGE OF STORE

- Younger stores (1-3 years old) show higher average sales.
- Stores around 19 years old also have high average sales but with significant variability.
- Older stores (10 years and beyond) generally show lower average sales.

Boxplot Insights:

- Sales vary significantly by store age, with younger stores often performing better.
- Some older stores, particularly those around 19 years old, exhibit high variability in sales, indicating differences in performance across locations.

5.2 A/B TESTING

The A/B testing was conducted to evaluate the effectiveness of three different promotion strategies on the sales of a newly promoted item. This section outlines the statistical analysis performed to determine if there were significant differences between the promotions and whether any potential biases might have influenced the results.

5.2.1 KRUSKAL-WALLIS TEST

To formally assess whether there was a significant difference in sales performance across the different promotions, the Kruskal-Wallis test was performed. This non-parametric test was chosen due to the non-normal distribution of the sales data.

FRAME 2: HYPOTHESES

- **Null Hypothesis (H0):** There is no difference in the effectiveness of the different promotions.
- **Alternative Hypothesis (H1):** At least one promotion is significantly more effective than the others.

Test Results:

- Test Statistic: 53.29
- p-value: 2.67×10^{-12}

Interpretation: The extremely low p-value indicates a statistically significant difference in sales performance across the different promotions. This suggests that at least one of the promotions significantly outperforms the others.

5.2.2 POST-HOC ANALYSIS

Following the Kruskal-Wallis test, a post-hoc Dunn's test was conducted to identify which specific pairs of promotions were significantly different.

Dunn's Post-Hoc Test Results:

- The matrix below shows the p-values for pairwise comparisons between promotion groups, adjusted using the Bonferroni correction.
- Significant differences were found between Promotion 1 and Promotion 2, as well as between Promotion 2 and Promotion 3.
- No significant difference was observed between Promotion 1 and Promotion 3.

	Promotion 1	Promotion 2	Promotion 3
Promotion 1	1.000000e+00	6.461842e-12	1.458660e-01
Promotion 2	6.461842e-12	1.000000e+00	7.083069e-07
Promotion 3	1.458660e-01	7.083069e-07	1.000000e+00

TABLE 1: DUNN'S POST-HOC TEST P-VALUES FOR PAIRWISE COMPARISONS

Summary of Findings:

- **Significant Differences:** There is a significant difference between Promotion 1 and Promotion 2, and between Promotion 2 and Promotion 3.
- **No Significant Difference:** There is no significant difference between Promotion 1 and Promotion 3.

5.2.3 BIAS ASSESSMENT

An analysis was conducted to determine whether the distributions of Market Size, Age of Store, and Location were balanced across the different promotion groups, as any imbalance could introduce bias into the A/B testing results.

Statistical Test Results:

- **Market Size Distribution:** The Chi-Square test yielded a p-value of 0.313, indicating no significant difference in Market Size distribution across the promotions.

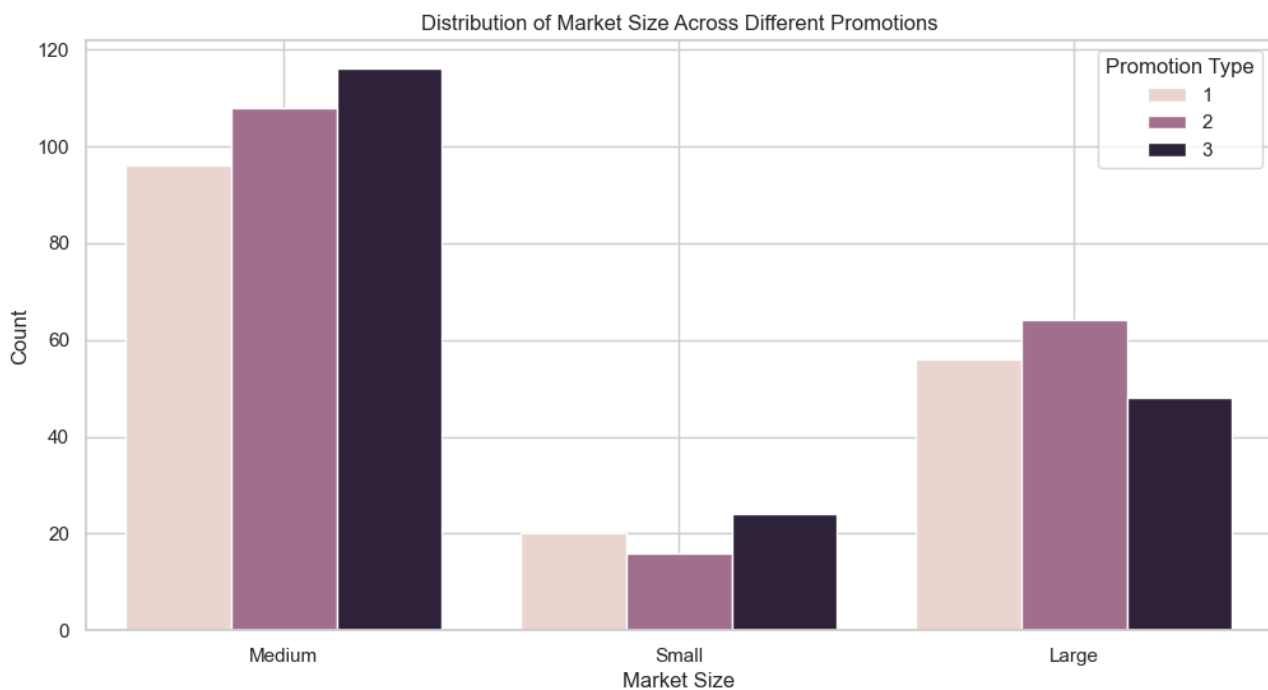


FIGURE 6: DISTRIBUTION OF MARKET SIZE ACROSS DIFFERENT PROMOTIONS

- **Age of Store Distribution:** The Kruskal-Wallis test resulted in a p-value of 0.099, suggesting no significant difference in the distribution of store ages across the promotions.
- **Location Distribution:** The Chi-Square test for Location distribution showed a p-value of 2.72×10^{-99} , indicating a significant difference in location distribution across the promotions.

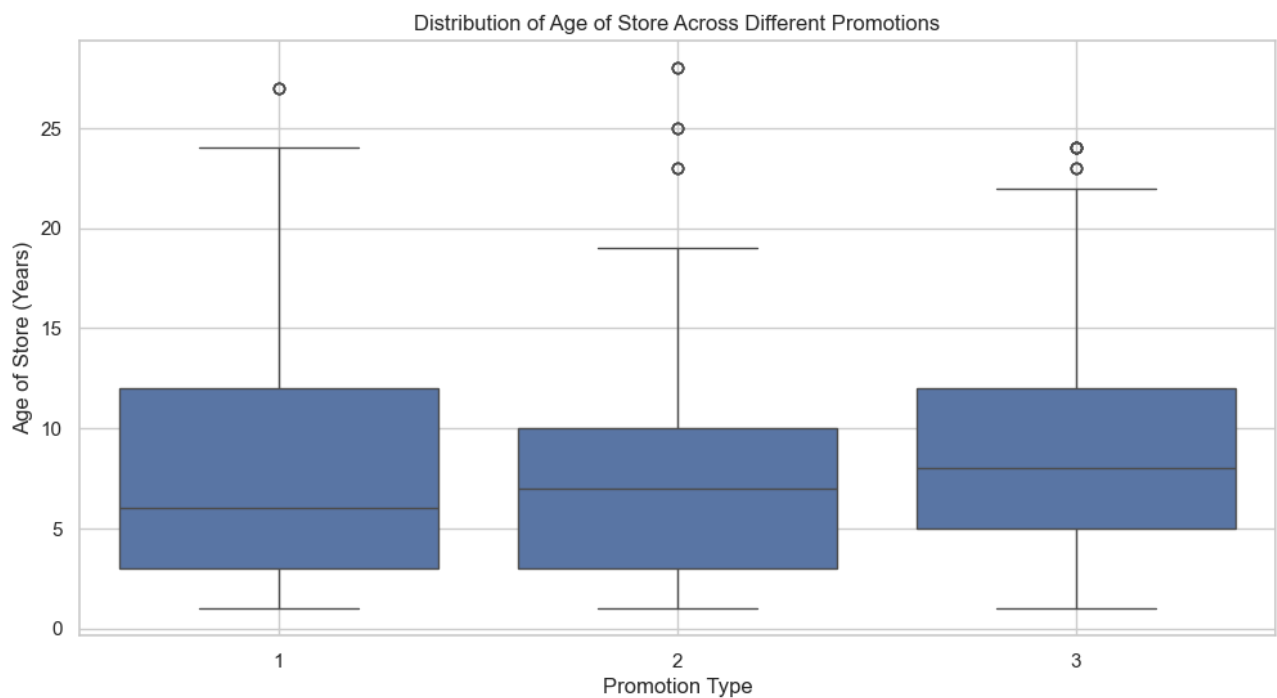


FIGURE 7: DISTRIBUTION OF AGE OF STORE ACROSS DIFFERENT PROMOTIONS

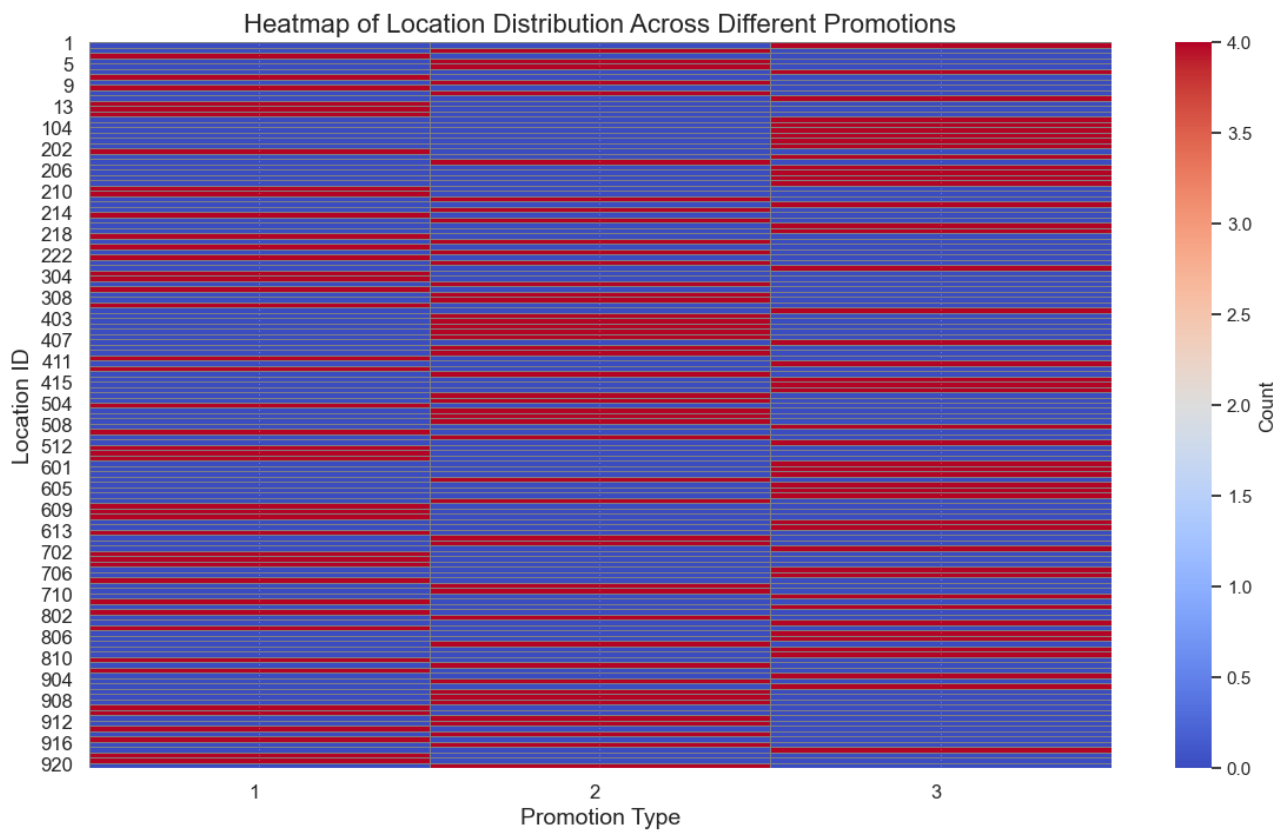


FIGURE 8: HEATMAP OF LOCATION DISTRIBUTION ACROSS DIFFERENT PROMOTIONS

Interpretation:

- Market Size and Age of Store are well balanced across the promotion groups, suggesting they do not introduce bias into the analysis.
- However, certain locations are over-represented in one or more promotion groups, which could introduce bias. If each restaurant (location) is assigned only one promotion type, then it is expected that the location distribution across promotions would be different. This setup inherently means that the locations for different promotions are completely distinct, which complicates the analysis because location-specific factors might confound the observed effects of the promotions.

5.3 REGRESSION ANALYSIS WITH LOCATION AS A FIXED EFFECT

5.3.1 ISOLATING THE EFFECT OF PROMOTION TYPE

To isolate the impact of the promotion type on sales, a linear regression model was employed. In this model, the promotion type was the primary independent variable, while location was included as a fixed effect (a categorical variable). This approach allowed us to control for location-specific effects, ensuring that the analysis accurately reflects the influence of the promotions.

Method: A linear regression model was used with promotion type as the independent variable and location as a fixed effect.

Advantage: This method controls for location-specific factors, isolating the true impact of each promotion type on sales.

5.3.2 KEY OUTPUTS

- **R-squared:** 0.929
The model explains 92.9% of the variance in sales, indicating a very strong fit.
- **Promotion Effects:**
 - **Promotion 2:** The coefficient is -8.5791 ($p < 0.001$), suggesting that, on average, sales under Promotion 2 are significantly lower than under Promotion 1 (the reference category).
 - **Promotion 3:** The coefficient is -21.6711 ($p < 0.001$), indicating that, on average, sales under Promotion 3 are significantly lower than under Promotion 1.
- **Location Effects:**
 - The coefficients for different locations vary significantly, reflecting the expected location-specific factors that affect sales independently of the promotion type.

5.3.3 INTERPRETATION

- **Promotion Effectiveness:** After controlling for the effect of location, both Promotion 2 and Promotion 3 show significantly lower sales compared to Promotion 1. This suggests that Promotion 1 might be the most effective in driving sales.
- **Location Impact:** The significant variation in coefficients for different locations confirms that location-specific factors strongly influence sales. By including these in the model, we control for their effect, allowing a clearer view of the impact of the promotion type.

5.3.4 PERCENTAGE DIFFERENCE IN SALES BETWEEN PROMOTION 1 AND OTHER PROMOTIONS

Using the coefficients from the regression analysis, we calculate the percentage difference in sales between Promotion 1 and the other two promotions.

- **Promotion 2:** Sales are lower by approximately \$8.5791K compared to Promotion 1.
- **Promotion 3:** Sales are lower by approximately \$21.6711K compared to Promotion 1.

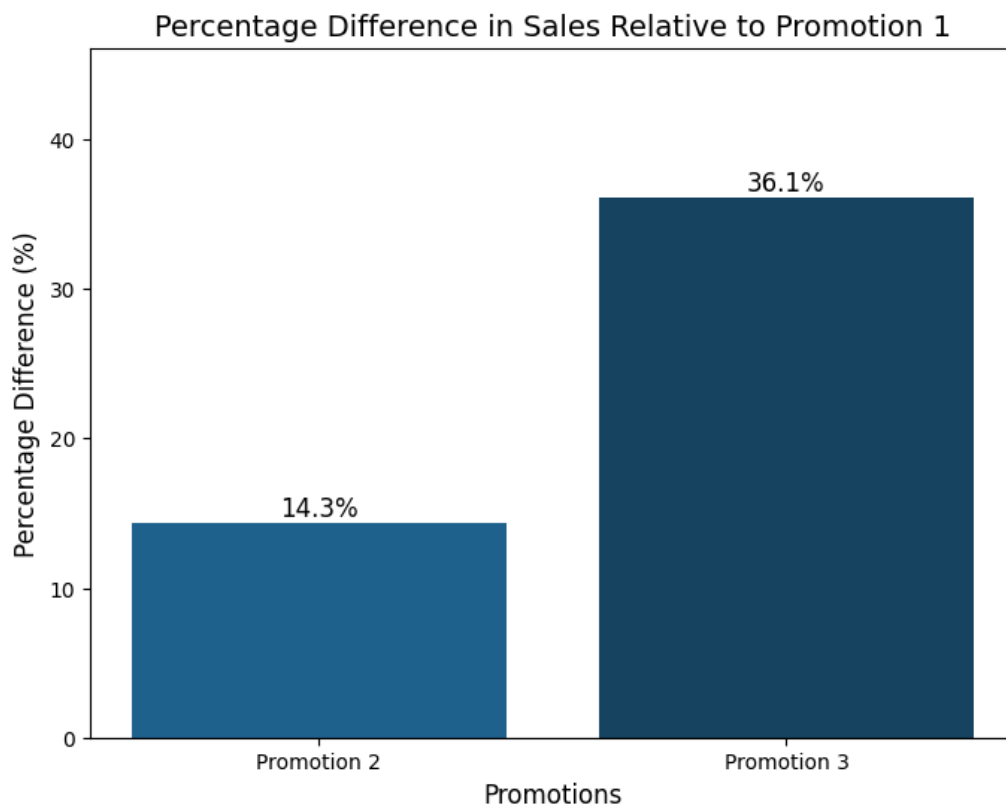


FIGURE 9: PERCENTAGE DIFFERENCE IN SALES RELATIVE TO PROMOTION 1

To express the differences in percentage terms, we use the following formula:

$$PercentageDifference = \left(\frac{DifferenceinSales}{SalesofPromotion1} \right) \times 100$$

- **Promotion 1 vs. Promotion 2:** Promotion 1 led to approximately 14.3% higher sales compared to Promotion 2.
- **Promotion 1 vs. Promotion 3:** Promotion 1 led to approximately 36.1% higher sales compared to Promotion 3.

These percentages highlight how much more effective Promotion 1 was in driving sales compared to the other two strategies.

5.4 KEY PERFORMANCE INDICATORS (KPIs)

When selecting KPIs for this analysis, the goal was to choose metrics that effectively measure the success and impact of the promotions on sales. Below are the KPIs used in this analysis, along with their justification and how they were applied:

1. AVERAGE SALES (IN THOUSANDS OF DOLLARS)

Definition: The average revenue generated from sales, expressed in thousands of dollars.

Purpose: This KPI directly measures the financial outcome of the promotions and serves as a straightforward indicator of how well each promotion drives revenue.

Application: Average sales were calculated for each promotion group to compare the overall effectiveness of the different promotions. By evaluating which promotion led to the highest average sales, we could identify the most successful strategy.

2. SALES DISTRIBUTION

Definition: The spread of sales data, including measures like median, quartiles, and range.

Purpose: This KPI helps to understand the variability in sales, highlighting whether a promotion consistently performs well or if it leads to unpredictable results.

Application: Boxplots and histograms were used to visualize the distribution of sales across promotion groups. This helped identify potential outliers and the consistency of each promotion's performance, guiding decisions on risk and reliability.

3. R-SQUARED (COEFFICIENT OF DETERMINATION)

Definition: A statistical measure indicating the proportion of variance in the dependent variable (sales) that can be explained by the independent variables (promotion type, location, etc.).

Purpose: This KPI assesses the overall fit of the regression model, indicating how well the

model explains the observed sales data.

Application: The R-squared value was used in the regression analysis to determine how much of the variation in sales could be attributed to the promotion types and location effects. A higher R-squared value indicated a better model fit.

4. PERCENTAGE DIFFERENCE

Definition: The relative difference in sales between Promotion 1 and the other promotion types, expressed as a percentage, based on the regression analysis using fixed effects for location.

Purpose: This KPI provides a clear, quantitative measure of how much better or worse each promotion performed compared to Promotion 1. By using the regression coefficients, it accounts for location-specific factors, offering a more accurate and fair comparison of the effectiveness of the promotions.

Application: The percentage difference was calculated from the regression coefficients, which reflect the difference in sales between Promotion 1 (the reference group) and Promotions 2 and 3. By taking the coefficients and expressing them as a percentage relative to Promotion 1, this KPI allows us to evaluate how much Promotion 1 outperformed or underperformed in relation to the other promotions, while controlling for location-specific effects. This approach ensures that the comparison is not biased by external factors, such as the store's location.

5. P-VALUES OF REGRESSION COEFFICIENTS

Definition: The probability that the observed relationship between the independent and dependent variables occurred by chance.

Purpose: This KPI is crucial for determining the statistical significance of the relationship between promotion types and sales.

Application: In the regression analysis, p-values were used to test the null hypothesis that the promotion type had no effect on sales. By focusing on p-values below 0.05, we identified which promotions significantly influenced sales outcomes.

6. POST-HOC TEST RESULTS

Definition: Statistical comparisons made after an initial test, such as ANOVA or Kruskal-Wallis, to determine which specific groups differ.

Purpose: Post-hoc tests help pinpoint where significant differences lie between promotion groups after a global test indicates some difference exists.

Application: The Dunn's post-hoc test with Bonferroni correction was applied following the Kruskal-Wallis test. It provided detailed insights into which specific pairs of promotions were

significantly different from each other, guiding the conclusions on the relative effectiveness of each promotion.

7. ADJUSTED SALES

Definition: Sales data adjusted to control for confounding factors, such as location or market size.

Purpose: This KPI provides a clearer view of the promotion's impact by isolating the effect of the promotion from other influencing factors.

Application: In the regression analysis, sales were adjusted for location effects to ensure that the observed differences between promotions were not confounded by location-specific factors.

JUSTIFICATION FOR KPI SELECTION

The KPIs were chosen to provide a comprehensive understanding of the effectiveness of the promotions from multiple angles; financial performance (Average Sales), consistency and reliability (Sales Distribution), model fit (R-squared), statistical significance (P-values), specific group differences (Post-Hoc Test Results), and control for confounding factors (Adjusted Sales).

6 CONCLUSION

This report analyzed the effectiveness of three different promotion strategies implemented by a fast-food chain to increase the sales of a newly promoted item. The analysis was conducted using a comprehensive dataset and a combination of statistical tests, regression analysis, and key performance indicators (KPIs).

KEY FINDINGS

The analysis revealed several important insights:

- **Promotion Effectiveness:** Promotion 1 was found to be the most effective strategy, consistently driving higher sales compared to Promotions 2 and 3. This conclusion was supported by the A/B testing results followed by the regression analysis with location as a fixed effect.
- **Location Impact:** Location-specific factors significantly influenced sales outcomes. By including location as a fixed effect in the regression model, we were able to isolate the true impact of each promotion type.

- **Statistical Significance:** The Kruskal-Wallis test and subsequent post-hoc analysis identified significant differences between the promotions, particularly between Promotion 1 and Promotion 2, and between Promotion 2 and Promotion 3. However, the true difference between Promotions 1 and 3 became clear only after controlling for location..
- **Bias Considerations:** The analysis confirmed that market size and store age were well balanced across the promotion groups, minimizing potential biases. However, the significant variation in location distribution highlighted the importance of controlling for location effects in future analyses.

RECOMMENDATIONS

Based on these findings, the following recommendations are proposed:

- **Prioritize Promotion 1:** Given its effectiveness, Promotion 1 should be prioritized in future marketing campaigns to maximize sales.
- **Review and Improve Promotions 2 and 3:** Promotions 2 and 3 should be reassessed and refined to enhance their impact on sales.
- **Incorporate Location Data in Future Analyses:** Future analyses should continue to control for location-specific factors to ensure accurate and unbiased results.
- **Collect Pre-Test Sales Data:** Collecting sales data before the start of any future promotions would provide a useful baseline for comparison and improve the robustness of the analysis.

FINAL WORDS

The comprehensive approach taken in this analysis, including the use of KPIs, statistical testing, and regression analysis, has provided clear and actionable insights into the effectiveness of the promotion strategies. By implementing the recommendations outlined in this report, the fast-food chain can optimize its promotional efforts and drive higher sales for new menu items in the future.