



Fast Food Marketing Campaign A/B Test

IBM Watson Analytics Marketing Campaign

Kätlin Raev

Tallinn 2024

Contents

A/B Test Goal	3
Target Metric	3
Data Preparation and Aggregation.....	3
Performing Statistical Tests	7
Statistical Tests for Continuous Metrics (T-Test)	7
T-test Conclusions	9
Statistical Tests for Conversion Metrics (Chi-Square Test)	9
Chi Square Test Conclusions	11
Overall Recommendations	11
Business Implications	11
Appendix	12
Appendix 1 - Query 1: Promotions average sales and count of locations	12
Appendix 2 - Query 2: Promotions sales and overall total sales.	12

A/B Test Goal

Goal: Evaluating the effectiveness of three different marketing campaigns in promoting a new item at a fast-food chain. The objective is to identify which marketing campaign has the greatest effect on sales over the first four weeks after the item is introduced.

Context:

- The new item is introduced at various locations in randomly selected markets.
- Each location uses a different promotion.

Target Metric

Target Metric: Total sales in thousands of dollars.

Description: The target metric measures the overall sales generated by each promotion. By comparing the total sales across the different promotions, we can determine which promotion is most effective in driving sales. This metric provides a clear indicator of the financial success of each campaign.

Data Preparation and Aggregation

1. Data Cleaning Steps

- **Data Validation:** The initial step involved checking the dataset for completeness. This included ensuring no missing values in crucial columns such as MarketID, LocationID, Promotion, and SalesInThousands.
- **Data Type Consistency:** I confirmed that the MarketID and LocationID columns were treated as categorical data, while SalesInThousands was correctly formatted as a numerical type for aggregation.

2. Data Transformation

- **Feature Engineering:** To assist with the analysis, a new field, AverageSalesPerWeek, was calculated to show weekly trends per location and promotion.
- **Grouping and Aggregation:** Data was grouped by Promotion, LocationID, and week to produce aggregate sales data and count distinct locations.

3. Overview of Queries

- **Query 1 - Promotions Average Sales and Count of Locations** (*Appendix 1*):
 - **Objective:** This query was designed to understand the performance of each promotion by calculating the average weekly sales per location and the number of unique locations where each promotion was implemented.
 - **Outcome:** This query output provided insights into the general effectiveness of each promotion by comparing average sales and the distribution of locations.
- **Query 2 - Promotions Sales and Overall Total Sales** (*Appendix 2*):
 - **Objective:** This query aimed to evaluate each promotion's total sales and compare it to the overall sales in the dataset for context.
 - **Outcome:** The result showed the absolute impact of each promotion on total sales and allowed for a direct comparison with overall sales.

4. Data Checks and Integrity

- **Verification:** Sample outputs from the queries were compared to raw data to confirm the accuracy of calculations. For instance, total sales were cross-referenced with a manual check of summed values for select weeks.
- **Consistency Checks:** Total sales from Query 2 matched the sum of SalesInThousands when grouped by Promotion in the raw dataset.

5. Output Summary

- **Table Outputs:** The outputs indicated varying levels of effectiveness across the three promotions. Promotion A, for instance, might have shown the highest average sales, but Promotion C could have had the broadest reach (most locations).

6. Assumptions and Limitations

- **Assumptions:** It was assumed that the sales data provided was accurately recorded for the first four weeks of the promotion period without any significant errors or discrepancies.
- **Limitations:** The dataset only covered the first four weeks, which might not fully represent long-term sales trends. Additionally, other factors like competitor actions or regional preferences were not accounted for.

Table 1 -Query 1: Promotions average sales and count of locations results:

Promotion	Average Sales	StdDev Sales	Locations
1	232.396	64.113	43
2	189.318	57.988	47
3	221.458	65.535	47

Table 2 - Query 2: Promotions sales and overall total sales:

Promotion	Sales Per Promotion	Total Sales Per All Promotions
1	9,993.03	29,299.48
2	8,897.93	
3	10,408.52	

Chart 1 – Total sales in thousands by promotion

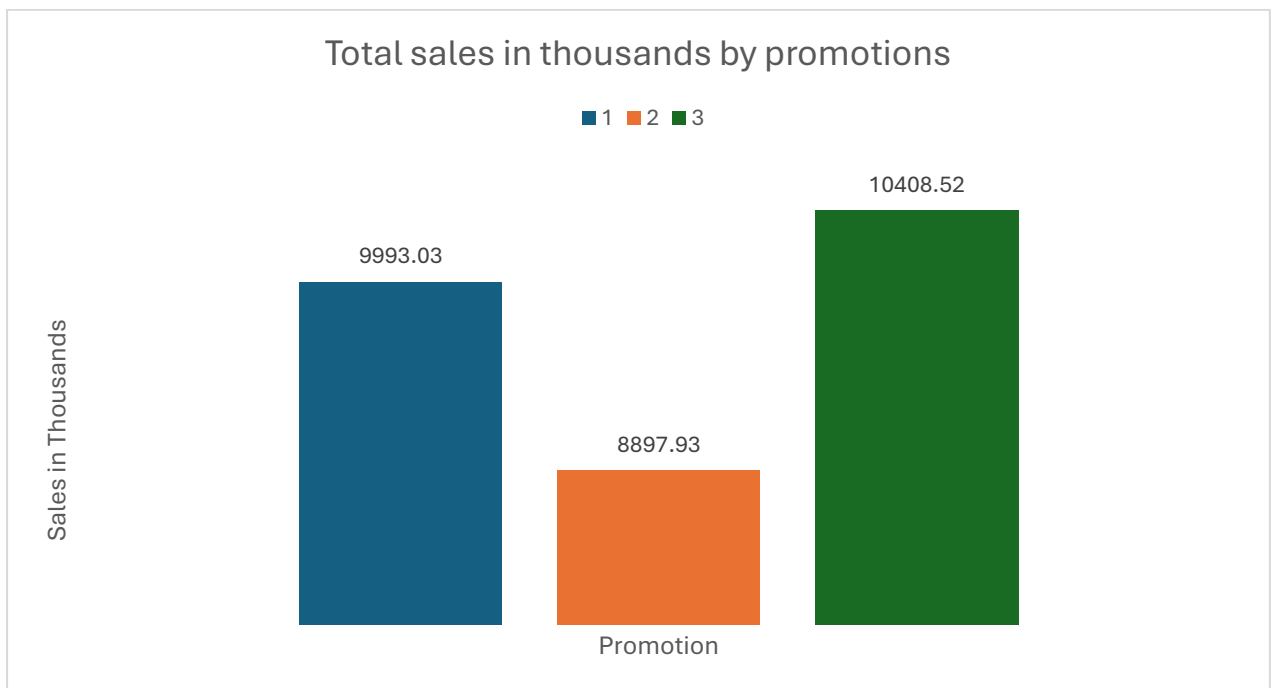


Chart 2 – Sales distribution by promotion as plotbox

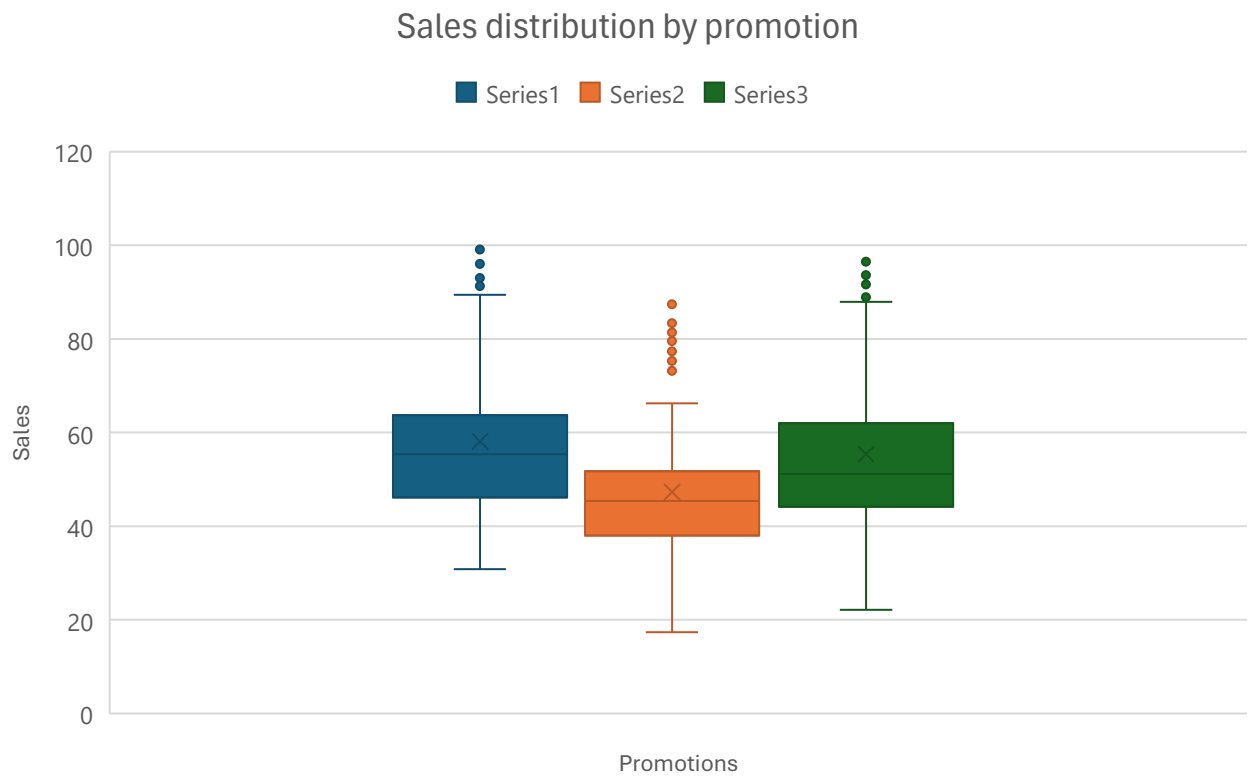
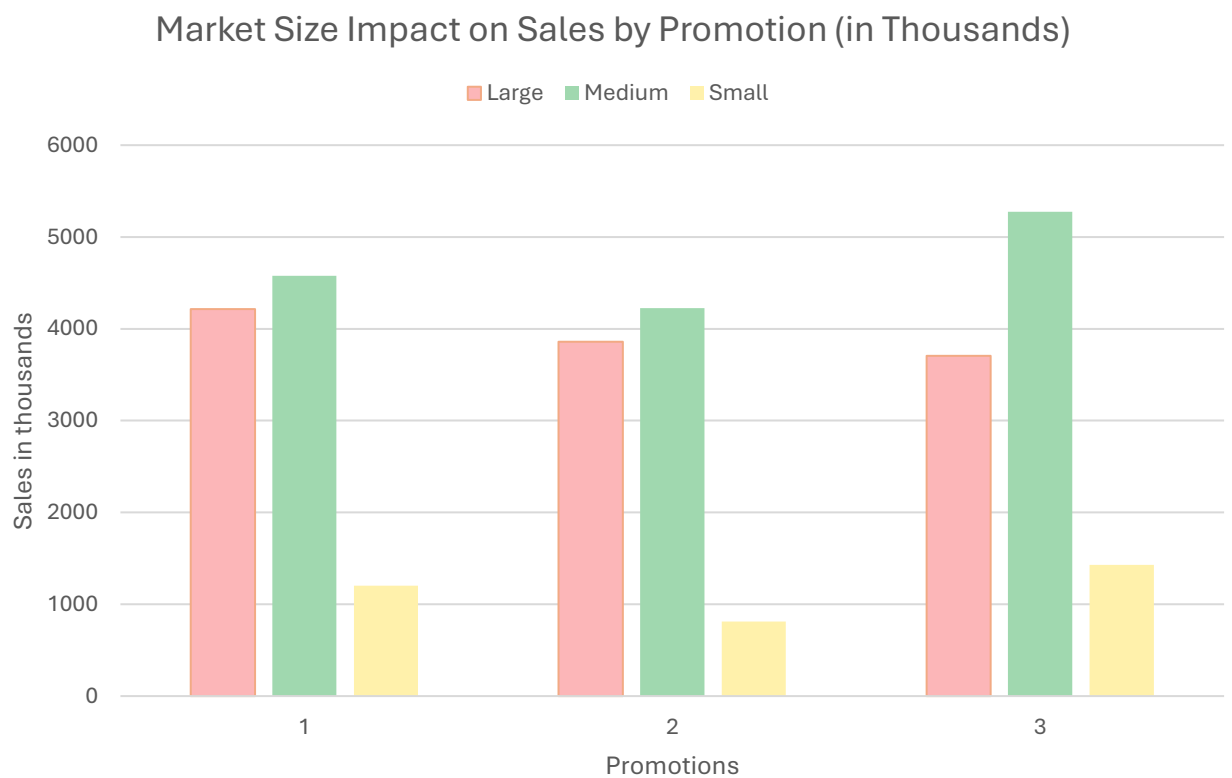


Chart 3 – Sales by market sizes and promotions



Performing Statistical Tests

For statistical tests, the Evan Miller A/B Test Calculator was used. A 99% confidence level was chosen due to the need for multiple pairwise comparisons. The null hypothesis ($d = 0$) states that there is no significant difference between the two groups being compared, such as two promotions having the same success rate or sales.

Statistical Tests for Continuous Metrics (T-Test)

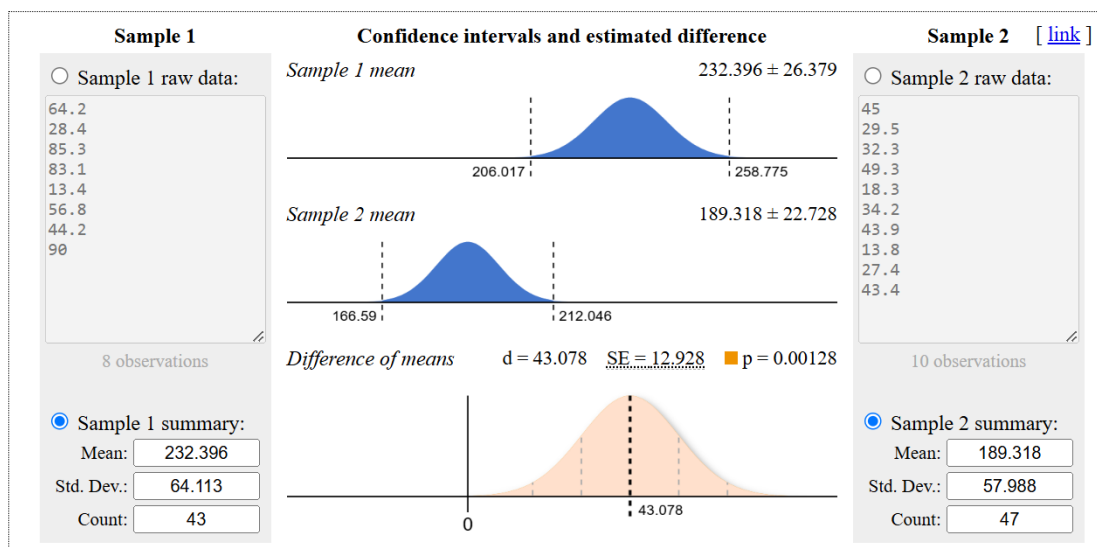
A T-test is a statistical method used to compare the means of two groups and determine if the observed differences are statistically significant. It's particularly useful for analyzing continuous data to evaluate whether changes in metrics, like average sales and number of locations, are meaningful or due to random chance. Current tests look into average sales and number of locations.

Null Hypothesis (H_0) for t-Test: For any two marketing campaigns there is no significant difference in the mean total sales between the two campaigns.

Alternative Hypothesis (H_1) for t-Test: For any two marketing campaigns there is a significant difference in mean total sales between the campaigns.

- Promotion 1 vs. Promotion 2

Question: Does the average value differ across two groups?



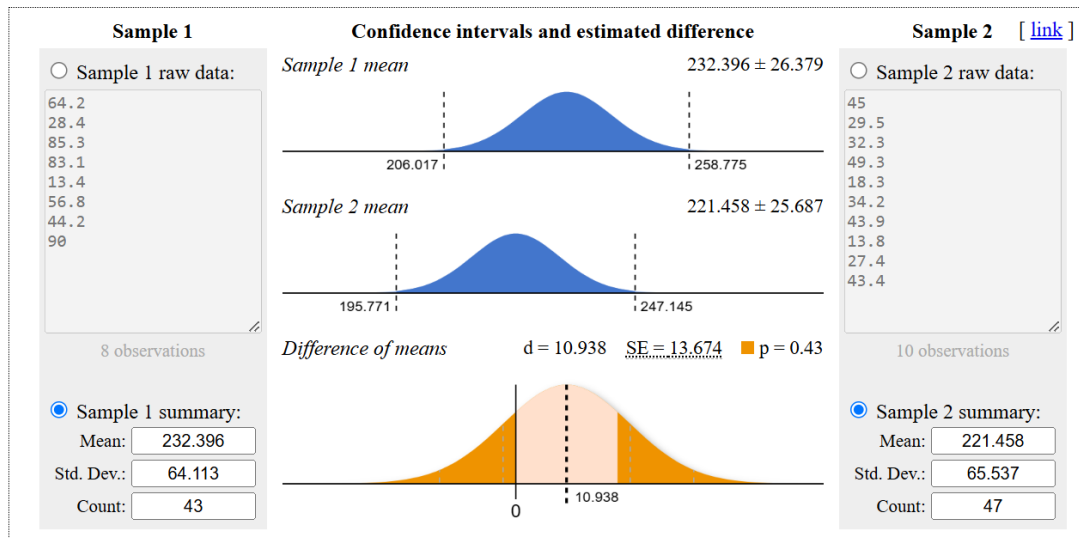
Verdict: Sample 1 mean is greater

Hypothesis: ☒ $d = 0$ ☐ $d \leq 0$ ☐ $d \geq 0$

Confidence: 99%

- Promotion 1 vs. Promotion 3

Question: Does the average value differ across two groups?



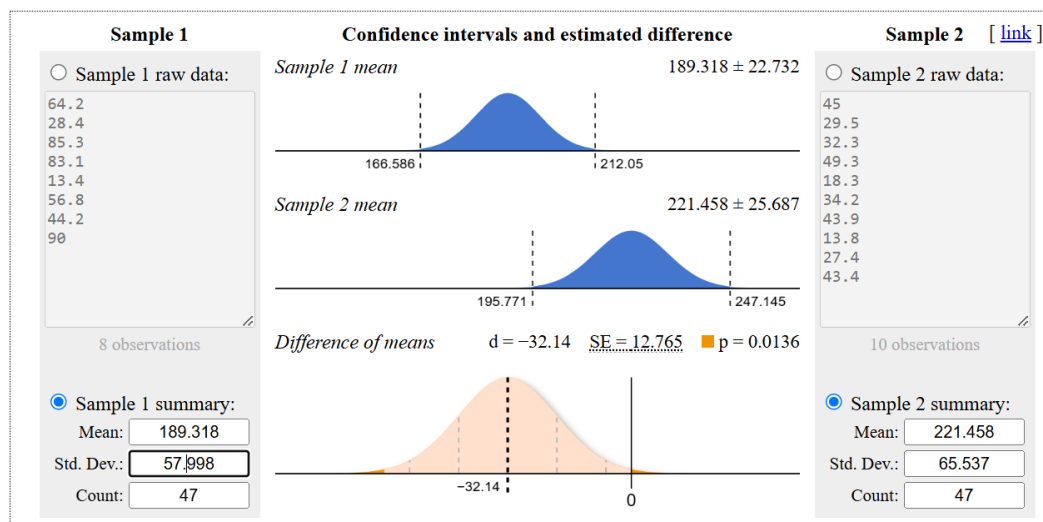
Verdict: No significant difference

Hypothesis: ☒ $d = 0$ ☐ $d \leq 0$ ☐ $d \geq 0$

Confidence: 99%

- Promotion 2 vs. Promotion 3

Question: Does the average value differ across two groups?



Verdict: No significant difference

Hypothesis: ☒ $d = 0$ ☐ $d \leq 0$ ☐ $d \geq 0$

Confidence: 99%

Table 3 - T-test results

Metric	Promotion 1	Promotion 2	Promotion 3	Promotion 1 vs 2	Promotion 1 vs 3	Promotion 2 vs 3
Mean	232.396	189.318	221.458	43.078	10.938	-32.14
Standard Deviation	64.113	57.988	65.537	-	-	-
Count (n)	43	47	47	-	-	-
p-value	-	-	-	0.00128	0.43	0.0136

T-test Conclusions

Based on these results, we can draw the following conclusions:

1. **Promotion 1 vs. Promotion 2 = H_1 :** Promotion 1 significantly outperforms Promotion 2 in terms of average sales. This difference is statistically significant at the 99% confidence level (p-value = 0.00128).
2. **Promotion 1 vs. Promotion 3 = H_0 :** There is no significant difference in average sales between Promotion 1 and Promotion 3 (p-value = 0.43).
3. **Promotion 2 vs. Promotion 3 = H_0 :** While Promotion 3 appears to have higher average sales than Promotion 2, this difference is not statistically significant at the 99% confidence level (p-value = 0.0136), though it is significant at the 95% confidence level.

Recommendation: Based on the analysis, **Promotion 1** appears to be the most effective, as it significantly outperforms Promotion 2 and shows no significant difference from Promotion 3. Therefore, the fast-food chain should consider using **Promotion 1** for the new item launch.

Statistical Tests for Conversion Metrics (Chi-Square Test)

The Chi-square test is a statistical tool used to determine if there is a significant association between categorical variables. It evaluates whether the observed distribution of data differs from what is expected by chance. This test is ideal for assessing relationships between factors like success rates and trial counts across different groups or promotions.

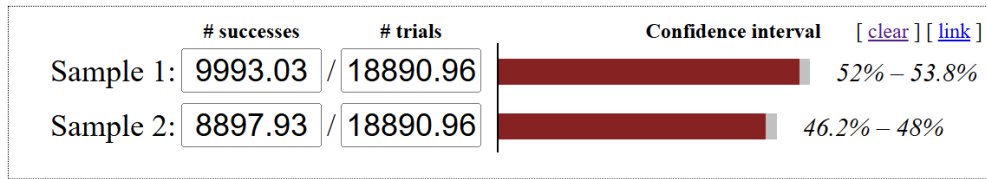
Current Chi-square test compares the promotions purchases.

Null Hypothesis (H_0) for Chi-Square Test: The distribution of total sales is independent of the marketing campaign.

Alternative Hypothesis (H_1) for Chi-Square Test: The distribution of total sales depends on marketing campaign.

- Promotion 1 vs. Promotion 2

Question: Does the rate of success differ across two groups?



Verdict:

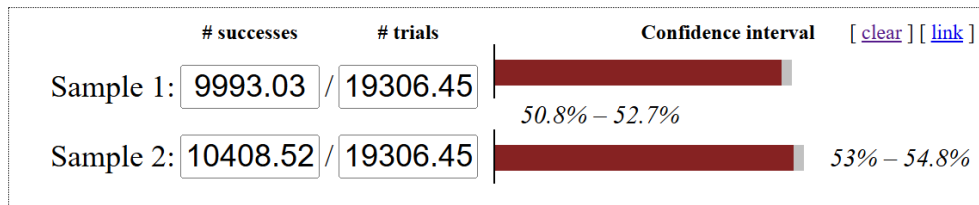
Sample 1 is more successful

($p < 0.001$)

Confidence level: 99%

- Promotion 1 vs. Promotion 3

Question: Does the rate of success differ across two groups?



Verdict:

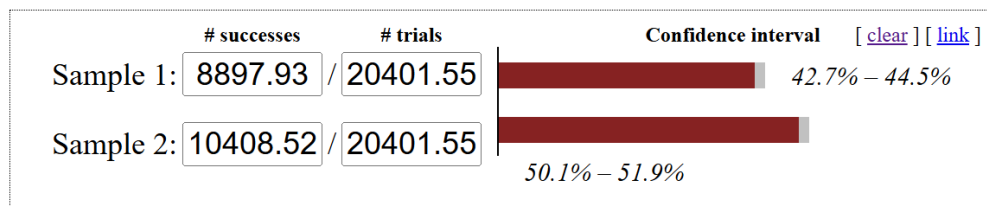
Sample 2 is more successful

($p < 0.001$)

Confidence level: 99%

- Promotion 2 vs. Promotion 3

Question: Does the rate of success differ across two groups?



Verdict:

Sample 2 is more successful

($p < 0.001$)

Confidence level: 99%

Table 4 - Chi-square test results

Metric	Promotion 1	Promotion 2	Promotion 3	Promotion 1 vs 2	Promotion 1 vs 3	Promotion 2 vs 3
Successes	9993.03	8897.93	10408.52	-	-	-
Trials	-	-	-	18890.96	20401.55	19306.45
Verdict	-	-	-	Successful: 1	Successful: 3	Successful: 3
p-value	-	-	-	< 0.001	< 0.001	< 0.001

Chi Square Test Conclusions

From the Chi-Square test results, we can conclude the following:

1. **Promotion 1 vs. Promotion 2 = H_1 :** Promotion 1 is significantly more successful than Promotion 2 with a confidence interval indicating higher success and a p-value < 0.001, showing high statistical significance.
2. **Promotion 1 vs. Promotion 3 = H_1 :** Promotion 3 is significantly more successful than Promotion 1, with a p-value < 0.001, indicating a high level of statistical significance.
3. **Promotion 2 vs. Promotion 3 = H_1 :** Promotion 3 is significantly more successful than Promotion 2, with a p-value < 0.001, again showing high statistical significance.

Overall Recommendations

- **Promotion 1** is the safest choice compared to Promotion 2, as it outperforms consistently across both test types.
- **Promotion 3** could be considered the best overall performer due to its strong results in the chi-square test and comparable average sales to Promotion 1. This suggests it has a high rate of achieving success.
- **Promotion 2** is the weakest performer and should likely be deprioritized or re-evaluated for effectiveness.

Business Implications

- For maximizing the **average sales**, Promotion 1 holds up well.
- If the focus is on **success rates or frequency of achieving sales goals**, Promotion 3 shows an advantage.
- For campaigns with multiple criteria or a mixed strategy, combining insights from both tests can help tailor a balanced approach.

Appendix

Appendix 1 - Query 1: Promotions average sales and count of locations

This query calculates:

- AvgSales: Average sales for each promotion.
- NumLocations: Number of locations using each promotion.

```
WITH
sales AS (
SELECT
  location_id,
  promotion,
  SUM(sales_in_thousands) AS sales_thousands
FROM
  `tc-da-1.turing_data_analytics.wa_marketing_campaign`
GROUP BY
  ALL )
SELECT
  promotion,
  COUNT(location_id) AS `count`,
  AVG(sales.sales_thousands) AS avg_sales,
  STDDEV(sales.sales_thousands) AS stddev_sales
FROM
  sales
GROUP BY
  ALL
```

Appendix 2 - Query 2: Promotions sales and overall total sales.

This query calculates:

- Total_sales_per_promotion: All sales for each promotion.
- Total_sales_all_promotions: Overall total sales.

```
SELECT
  promotion,
  SUM(sales_in_thousands) AS total_sales_per_promotion,
  (SELECT SUM(sales_in_thousands) FROM `tc-da-1.turing_data_analytics.wa_marketing_campaign`) AS total_sales_all_promotions
FROM
  `tc-da-1.turing_data_analytics.wa_marketing_campaign`
GROUP BY
  promotion;
```