1. **Goal**

The main goal of this A/B test is to determine which promotion among the three has the best positive effect on sales for the new menu item across all locations.

**2.0 Target Metric**

Based on the goal, the main metric to focus on is the Average Weekly Sales per Promotion. This metric directly compares the average sales performance of each promotion in all locations. By doing this, the company would be able to determine which promotion consistently drives the highest sales.

**3.0 Calculations**

Table 3.1 contains the numbers necessary to analyse the A/B test to decide which promotion is best. Calculations for average weekly sales, standard deviation and sample size was done in SQL (see appendix for query).

*Table 3.1: Summary of results for each promotion*



*Source: Author’s construct from wa marketing campaign data base, October 2024.*

Although, the average weekly sales for Promotion 1 gives an indication that it is the best campaign to use among the three, it does not determine whether the observed differences in performance are statistically significant. Therefore, it is imperative to conduct a statistical test.

*3.1 Description of statistical test*

This analysis is to test 3 promotions therefore, a pairwise comparison between the three promotions (Promotion 1 and 2, Promotion 1 and 3 and Promotion 2 and 3) using the T-test is done.

For each comparison the following hypothesis is defined.

* **Null hypothesis (H₀):** There is no difference in the average weekly sales between the two promotions.
* **Alternative hypothesis (H₁):** There is a difference in the average weekly sales between the two promotions.

The T-test (Evan Miller A/B Test Calculator) is used because it is suitable to compare the means of two groups, commonly used when dealing with sample sizes and handles unequal sample sizes which is the case in this analysis (see table 3.1).

However, when comparing more than 2 groups, the greater the risk of a Type I error. To mitigate this, a more stringent confidence level of 99%and a significance level of 0.01 are set. To further reduce the chance of a false positive, the Bonferroni correction is applied. The significance level of 0.01 is divided by 3 since the comparison is between 3 promotions. Hence 0.01/3 =0.0033. This means that each pairwise comparison will be tested against a significance level of 0.0033. Hence, reducing the chance of false positives.

* 1. **Decision**

Based on the pairwise comparisons made for the three promotions, these were the outcome.

* 1. *Promotion 1 vs Promotion 2*

Difference in means (d): 43.08

Standard Error (SE): 12.928

p-value: 0.00128 (significant at 0.0033 after Bonferroni correction)

**Interpretation**: Promotion 1 significantly outperforms Promotion 2, with a large difference in average weekly sales of $43.08, and this difference is statistically significant (p < 0.0033). Based on this a confident conclusion can be made that Promotion 1 is better than Promotion 2.

* 1. *Promotion 1 vs. Promotion 3:*

Difference in means (d): 10.94

Standard Error (SE): 13.674

p-value: 0.43 (not significant)

**Interpretation**: There is a small difference between Promotion 1 and Promotion 3, however this difference is not statistically significant (p > 0.0033). This suggests that Promotion 1 and Promotion 3 have similar effects on sales. The observed difference may be due to random chance.

* 1. *Promotion 2 vs. Promotion 3:*

Difference in means (d): 32.14

Standard Error (SE): 12.765

p-value: 0.0136 (not significant at 0.0033 after Bonferroni correction)

**Interpretation**: While Promotion 2 shows a sales difference of $32.14 compared to Promotion 3, this result is not statistically significant after applying the Bonferroni-corrected significance level (p = 0.0136 > 0.0033). Therefore, it cannot be confidently claimed that Promotion 2 outperforms Promotion 3.

Based on these findings, **Promotion 1** is the best choice because it significantly outperforms Promotion 2 and shows no statistically significant difference with Promotion 3.

In conclusion, if other factors such as cost are similar, choosing **Promotion 1** would likely result in the highest sales performance for the Food Chain’s new menu item across all locations without the risk of losing out on significant gains from the other promotions.

**APPENDIX**

A). SQL Query

WITH TotalSales AS (

    SELECT

        location\_id,

        Promotion,

        ROUND(SUM(sales\_in\_thousands), 2) AS Total\_Sales

    FROM

        `tc-da-1.turing\_data\_analytics.wa\_marketing\_campaign`

    GROUP BY

        location\_id,

        Promotion

)

SELECT

    Promotion,

    ROUND(AVG(Total\_Sales), 2) AS Avg\_Weekly\_Sales,

    ROUND(STDDEV(Total\_Sales),2) AS Standard\_Deviation,

    COUNT(\*) AS Sample\_size

FROM

    TotalSales;

B). Link to Evan Miller A/B Test calculator

*Promotion 1 and 2*

[*https://www.evanmiller.org/ab-testing/t-test.html#!232.4/64.11/43;189.32/57.99/47@99*](https://www.evanmiller.org/ab-testing/t-test.html#!232.4/64.11/43;189.32/57.99/47@99)

*Promotion 1 and 3*

[*https://www.evanmiller.org/ab-testing/t-test.html#!232.4/64.11/43;221.46/65.54/47@99*](https://www.evanmiller.org/ab-testing/t-test.html#!232.4/64.11/43;221.46/65.54/47@99)

*Promotion 2 and 3*

<https://www.evanmiller.org/ab-testing/t-test.html#!189.32/57.99/47;221.46/65.54/47@99>