*Fast Food A/B Testing Analysis and Insights*

***Contents***

[I. Project overview 1](#_Toc184650972)

[II. Goal 3](#_Toc184650973)

[III. Target metric 3](#_Toc184650974)

[IV. Calculations 4](#_Toc184650975)

[V. Decision 7](#_Toc184650976)

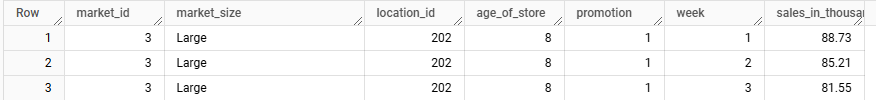
[VI. Appendix 7](#_Toc184650977)

# Project overview

|  |  |
| --- | --- |
|  | **Scenario:**  A fast-food chain plans to add a new item to its menu. However, they are still undecided between three possible marketing campaigns for promoting the new product. In order to determine which promotion has the greatest effect on sales, the new item is introduced at locations in several randomly selected markets. A different promotion is used at each location, and the weekly sales of the new item are recorded for the first four weeks. |

**Database**: `turing\_data\_analytics.wa\_marketing\_campaign`

**Snippet of database:**



**Shape**: The dataset contains 548 rows and 7 columns.

**Null Values**: There are no missing values in the dataset.

**Columns:**

* *MarketID*: unique identifier for market
* *MarketSize*: size of market area by sales
* *LocationID*: unique identifier for store location
* *AgeOfStore*: age of store in years
* *Promotion*: one of three promotions that were tested
* *week*: one of four weeks when the promotions were run
* *SalesInThousands*: sales amount for a specific *LocationID*, *Promotion*, and *week*

# Goal

The primary goal of the A/B test is to evaluate the **effectiveness of three marketing campaigns** (marked by p*romotion – 1, 2, 3*) on customer spending, aggregated by *LocationID*. The objective is to determine which campaign yields the highest average revenue or spending per location while minimizing the chances of false positives due to the **multiple testing problem**. After done that, it will be possible to determine which marketing strategy works best with statistically significant differences among the test groups.

To address the multiple testing problem, for analysis of A/B test results we will use a confidence level of 99%. This increases the limit for statistical significance, reducing the chance of false positives.

# Target metric

Identify Relevant Metrics:

* ***Total Sales***: The sum of sales\_in\_thousands for each location\_id and promotion.
* ***Average Weekly Sales***: The mean of sales\_in\_thousands per week for a given promotion at each location.
* ***Number of Weeks***: The total number of weeks observed for each location\_id and promotion.

***Total Sales*** is chosen as the target metric because it directly reflects the revenue generated by each marketing campaign. This makes it the most relevant measure of the campaign's financial success. By analyzing total sales, we can evaluate the overall effectiveness of each promotion in driving revenue across all locations, which aligns with the primary goal of identifying the best-performing campaign.

**Null Hypothesis (H0)**:

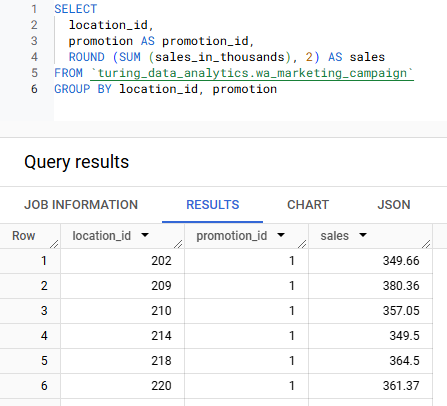
There is **no significant difference** in the total sales generated by the three marketing campaigns. If the null hypothesis is rejected, it means that at least one campaign differs significantly from the others in total sales performance.

**Alternative Hypothesis (HA):**

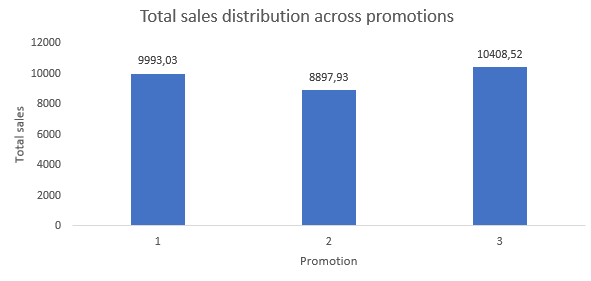
There **is a significant difference** in the total sales generated by at least one pair of marketing campaigns.

# Calculations

First step is to aggregate data by location\_id and promotion to prepare for statistical analysis. We calculate metrics based on the sales\_in\_thousands column using SQL and BigQuery environment. Query is provided in Appendix. Snippet below shows part of aggregated data.

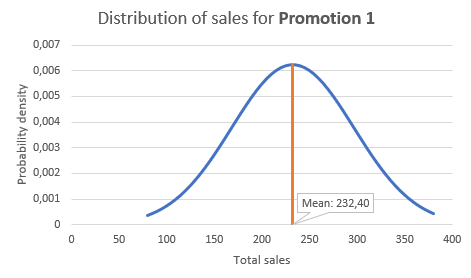
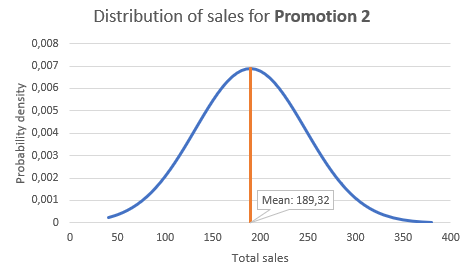
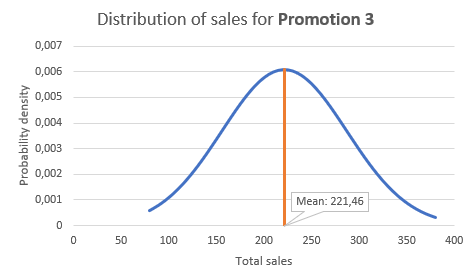


The graph below displays the total sales distribution across different promotions. According to the data, Promotion 3 generated the highest sales at 10,408.52k USD during the measured period. While Promotion 3 appears to generate the highest revenue based on this initial metric, we need to conduct additional statistical tests to confirm its statistical significance and overall effectiveness.



Next, let's visualize the distribution of sales for each promotion. The bell curves below represent normal distributions, showing how sales are spread out around the mean for each promotion.

A smaller standard deviation indicates a narrower, more concentrated distribution, while a larger standard deviation indicates a wider, more dispersed distribution.

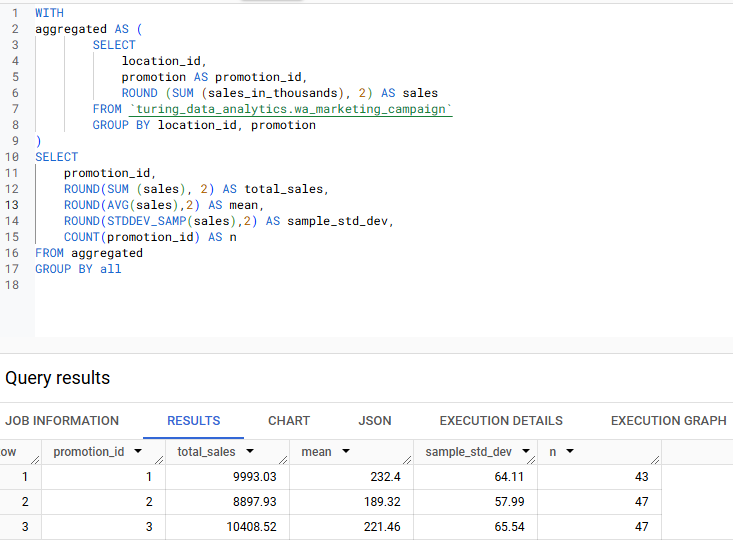
  

Comparing the three promotions, Promotion 2 has the lowest standard deviation, indicating that its performance is the most consistent across locations. Promotion 3 has the highest standard deviation, implying that it experienced the greatest variability in sales performance.

Now, it’s time for statistical tests. For each pair of campaigns:

* Calculate the **mean** and **standard deviation** of total\_sales for each group.
* Perform a **two-sample t-test** assuming unequal variances.

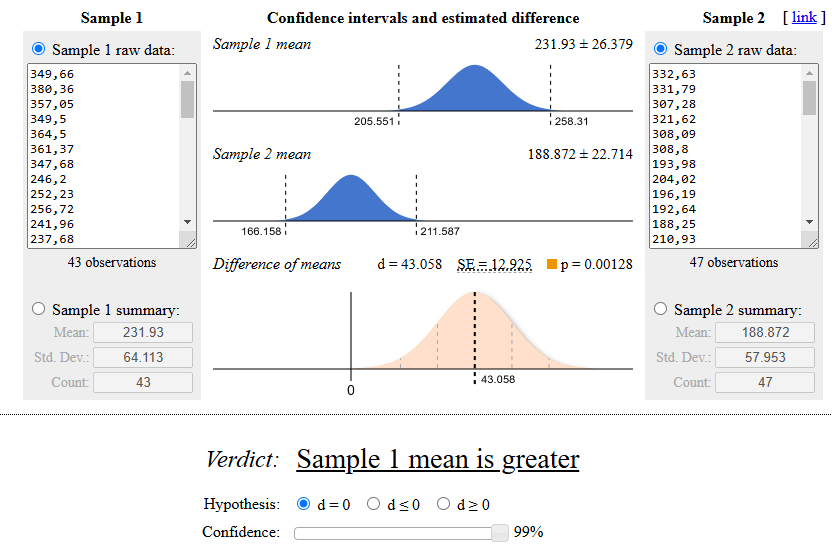
All statistical calculations are done using SQL and BigQuery environment. Query is provided in Appendix. Below is provided results for each promotion campaign.



Now that the data is aggregated and the target metric is total\_sales, the next step is to conduct pairwise comparisons between the three campaigns (1, 2, 3) to determine which one performs the best.

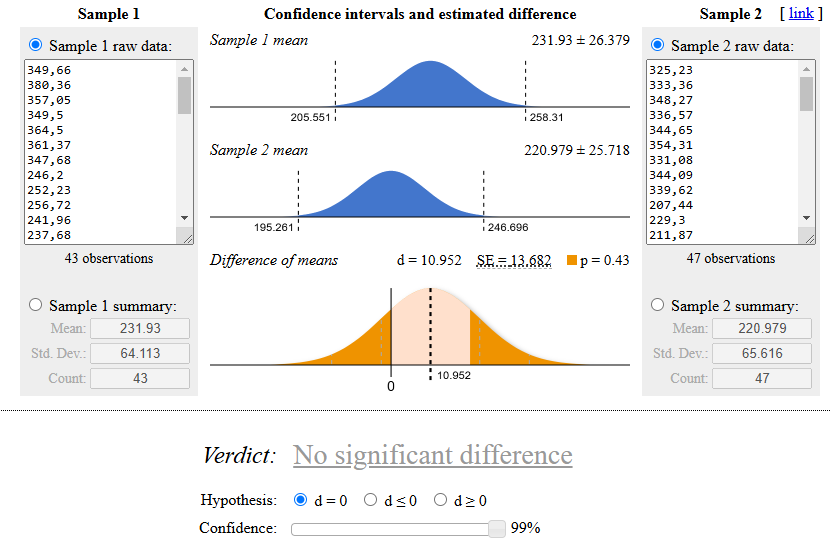
Since total\_sales is a continuous variable, a **t-test** is appropriate for comparing the means between two groups. For multiple tests, we address the multiple testing problem by using a 99% confidence level. For statistical tests we use the Evan Miller A/B Test Calculator (<https://www.evanmiller.org/ab-testing/> ).

***Comparison Promotion 1 vs. Promotion 2***



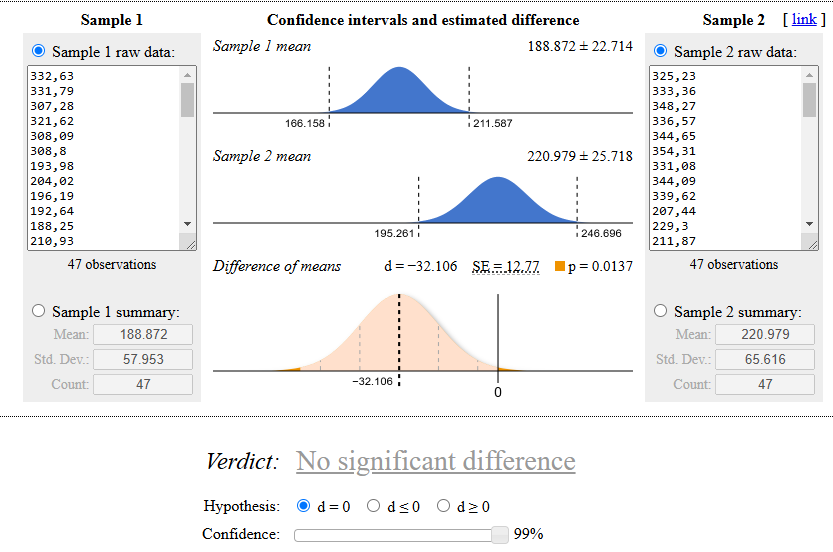
The difference between campaigns 1 and 2 ***is statistically significant***. Campaign 1 performs better than campaign 2 with a difference of means of 43 thousand dollars. In other words, over the period of the test, campaign 1 generated an estimated 43 thousand dollars more than campaign 2.

***Comparison Promotion 1 vs. Promotion 3***



We fail to reject the null hypothesis for the comparison between Promotion 1 and Promotion 3. There is no statistically significant difference between the two promotions' total sales (p > 0.01).

***Comparison Promotion 2 vs. Promotion 3***

******

Comparing Promotion 2 and Promotion 3 has no significant difference at the 99% confidence level, but marginally significant at the 95% level (p < 0.05)

# Decision

After evaluating all tests and calculations, I have decided to choose ***Promotion 1*** for promoting the new product, as it is the most effective in maximizing sales.

Recommendations for possible actions in future:

* Start rolling out Promotion 1 in targeted locations for a period of time (e.g., 2 months) to validate its effectiveness. Monitor total sales performance during this period and assess whether the promotion continues to generate positive results.
* Segment sales data by market size (e.g., small, medium, large) and reanalyze the promotion’s impact on total sales in each segment. Conduct A/B tests on each segment to see if Promotion 1 performs better in one market size over another.
* For more reliable results, it's crucial to increase the sample size. The current sample sizes of 43 and 47 are relatively small and may lead to less reliable conclusions. A larger sample size will reduce the chance of Type I and Type II errors, narrow confidence intervals, and increase the power of statistical tests, allowing for more precise estimates of the impact of different promotions.

# Appendix

1. Query for data aggregation:

SELECT

  location\_id,

  promotion AS promotion\_id,

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

FROM `turing\_data\_analytics.wa\_marketing\_campaign`

GROUP BY location\_id, promotion

1. Query for statistical calculations:

WITH

aggregated AS ( -- aggregate data before conducting statistical test and calculations

        SELECT

            location\_id,

            promotion AS promotion\_id,

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

        FROM `turing\_data\_analytics.wa\_marketing\_campaign`

        GROUP BY BY location\_id, promotion

)

SELECT

    promotion\_id,

    ROUND (SUM (sales), 2) AS total\_sales,

    ROUND (AVG(sales),2) AS mean,

    ROUND (STDDEV\_SAMP(sales),2) AS sample\_std\_dev,

    COUNT(promotion\_id) AS n

FROM aggregated

GROUP BY all