Goal of the Test

The main goal of this A/B test was to test wether any of the three Promotions for the new menu would improve sales of our fast-food chains. For analysis of the conducted A/B tests we will use a confidence level of 99%.

Target Metric

The dataset provides two metrics:

age_of_store - the age of the store in which the promotions occured.

sales_in_thousands - the amount of sales done in each specific store.

Since the goal of the A/B test was to compare the best promotional campaings for our fast-food chains. I've decided to use the **sales_in_thousands** metric to find out which of the Promotional campaing garnered the most sales.

I will not use the age_of_store metric as it has little to no effect on the efficacy of Promotional campaings. Also to calculate our samples I will use the **location_id** to count each distinct fast-food chain location that participated in the Promotional campaigns.

Therefore, the target metric of the A/B test is the median sales of each promotion in thousands, coded as sales_in_thousands in the dataset.

Hypothesis and prior assumptions.

As we have no prior assumptions to our data. The directionality of the difference is not assumed, allowing us to test for any difference in either direction. Which means that in our calculations later on we will be using a **Two-tailed test**. As we should aggregate by **LocationID** and **PromotionID** before conducting the statistical tests. I've calculated the samples for each promotion. **Promotion 1(n) = 43, Promotion 2 and Promotion 3 (n) = 47**. As we will calculate the **means (μ)** of each Promotion and our sample sizes **(n) >= 30** we will conduct a **t-test** instead of a z-test.

For my first experiment I'v chosen to measure Promotion 1 and Promotion 2.

The H0 (null) hypothesis is that there is no significant difference between the two promotions (H0: Mean Sales Promotion 1 = Mean Sales Promotion 2)

The H1 (alternative) hypothesis is that there is a significant difference in median sales between Promotion 1 and Promotion 2. (H1: Mean Sales Promotion 1 ≠ Mean Sales Promotion 2

For my second experiment I'v chosen to measure Promotion 1 and Promotion 3.

The H0 (null) hypothesis is that there is no significant difference between the two promotions (H0: Mean Sales Promotion 1 = Mean Sales Promotion 2)

The H1 (alternative) hypothesis is that there is a significant difference in median sales between Promotion 1 and Promotion 3. (H1: Mean Sales Promotion 1 ≠ Mean Sales Promotion 3)

For my third experiment I'v chosen to measure Promotion 2 and Promotion 3.

The H0 (null) hypothesis is that there is no significant difference between the two promotions (H0: Mean Sales Promotion 2 = Mean Sales Promotion 3)

The H1 (alternative) hypothesis is that there is a significant difference in median sales between Promotion 2 and Promotion 3. (H1: Mean Sales Promotion 2 ≠ Mean Sales Promotion 3)

Limitations

The main limitations of our experiments are coinciding with the samples which are a bit different. For example the sample for Promotion 1 is smaller by 4 locations (8,5%) compared to Promotion 2 and 3. Also there are differences in the sample sets that might make small mismatches to our data for example the market sizes of locations are not equally dispersed between each promotion. Most places that have a Large market size are in Promotion 2 (16), compared to (14, 14,3 % less) in Promotion 1 and (12, 25 % less) in Promotion 3. Similar differences occur all throught the provided samples.

Row	location_count	promotion	market_size
1	14	1	Large
2	16	2	Large
3	12	3	Large
4	5	1	Small
5	4	2	Small
6	6	3	Small
7	24	1	Medium
8	27	2	Medium
9	29	3	Medium

Promotion	avg_sales	stddev_sales	num_stores	t_statistic	Degrees of Freedom	p-value	avg_age	std_dev_pooled	Cohens d	Sales Variance	Pooled Variance	Standard Error	t-critical	Margin of Error	Difference in Means	Confidence Interval
1	232.396	64.113	43	3.347	88	0.00120	8.279	60.988	0.706	4110.463	3719.562	12.87016226	2.369472275	30.49549264	43.078	12.58 to 73.57
2	189.318	57.988	47				7.979			3362.653						

df	F-statistic	p-value
42	1.222386742	0.253
46		
P-Value > 0.05: Suggests that the variances are equal (we do not reject the null hypothesis of equal variances).		

Promotion	avg_sales	stddev_sales	num_stores	t_statistic	Degrees of Freedom	p-value	avg_age	std_dev_pooled	Cohens d	Sales Variance	Pooled Variance	Standard Error	t-critical	Margin of Error	Difference in Means	Confidence Interval
1	232.396	64.113	43	0.799	88	0.43	8.279	64.860	0.169	4110.463	4206.872	13.687298	2.369472275	32.43167312	10.938	-21.49 to 43.37
3	221.458	65.535	47				9.234			4294.897						

df	F-statistic	p-value
42	1.044869411	0.441
46		
P-Value > 0.05: Suggests that the variances are equal (we do not reject the null hypothesis of equal variances).		

Promotion	avg_sales	stddev_sales	num_stores	t_statistic	Degrees of Freedom	p-value	avg_age	std_dev_pooled	Cohens d	Sales Variance	Pooled Variance	Standard Error	t-critical	Margin of Error	Difference in Means	Confidence Interval
2	189.318	57.988	47	-2.518	92	0.0135	8.279	61.877	-0.519	3362.653	3828.775	12.76427043	2.36756577	30.22024975	-32.140	-62.36 to -1.92
3	221.458	65.535	47				9.234			4294.897						

df	F-statistic	p-value
46	1.277234515	0.205
46		
P-Value > 0.05: Suggests that the variances are equal (we do not reject the null hypothesis of equal variances).		

A/B test no.1 (Promotion 1 and Promotion 2)

Sample 1

Sample 1 raw data:

64.2
28.4
85.3
83.1
13.4
56.8
44.2
90

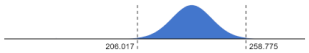
8 observations

Sample 1 summary:

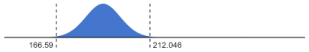
Mean:	232.396
Std. Dev.:	64.113
Count:	43

Confidence intervals and estimated difference

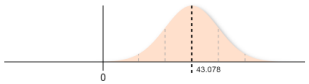
Sample 1 mean 232.396 ± 26.379



Sample 2 mean 189.318 ± 22.728



Difference of means d = 43.078 SE = 12.928 p = 0.00128



Verdict: Sample 1 mean is greater

Hypothesis: d = 0 d ≤ 0 d ≥ 0

Confidence: 99%

Sample 2 [link]

Sample 2 raw data:

45
29.5
32.3
49.3
18.3
34.2
43.9
13.8
27.4
43.4

10 observations

Sample 2 summary:

Mean:	189.318
Std. Dev.:	57.988
Count:	47

A/B test no.1 (Promotion 1 and Promotion 3)

Sample 1

Sample 1 raw data:

64.2
28.4
85.3
83.1
13.4
56.8
44.2
90

8 observations

Sample 1 summary:

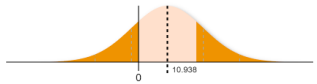
Mean:	232.396
Std. Dev.:	64.113
Count:	43

Confidence intervals and estimated difference

Sample 1 mean 232.396 ± 26.379

Sample 2 mean 221.458 ± 25.686

Difference of means d = 10.938 SE = 13.674 p = 0.43



Verdict: No significant difference

Hypothesis: d = 0 d ≤ 0 d ≥ 0

Confidence: 99%

Sample 2 [link]

Sample 2 raw data:

45
29.5
32.3
49.3
18.3
34.2
43.9
13.8
27.4
43.4

10 observations

Sample 2 summary:

Mean:	221.458
Std. Dev.:	65.535
Count:	47

A/B test no.1 (Promotion 2 and Promotion 3)

Sample 1

Sample 1 raw data:

64.2
28.4
85.3
83.1
13.4
56.8
44.2
90

8 observations

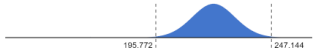
Sample 1 summary:

Mean:	189.318
Std. Dev.:	57.988
Count:	47

Confidence intervals and estimated difference

Sample 1 mean 189.318 ± 22.728

Sample 2 mean 221.458 ± 25.686



Difference of means d = −32.14 SE = 12.764 p = 0.0136

Verdict: No significant difference

Hypothesis: d = 0 d ≤ 0 d ≥ 0

Confidence: 99%

Sample 2 [link]

Sample 2 raw data:

45
29.5
32.3
49.3
18.3
34.2
43.9
13.8
27.4
43.4

10 observations

Sample 2 summary:

Mean:	221.458
Std. Dev.:	65.535
Count:	47

Looking through all of the data and the 3 A/B tests done we can conclude that the least plausible option would be to choose the 1st promotion plan, as it provided with the best results overall. With a Cohens d of 0.706, which argues of the medium to large size of effect of Promotion 1 as well as a very low p-value of 0.0012. Although when tested between Promotion 1 and Promotion 3 we find that there is no significant difference between the two. I would still argue that because of the Cohens d and the very low p-value that Promotion 1 gives it would be a more viable option to choose it as the most effective way to promote our new menu to the customers. Another addition that should leverage the power of Promotion 1 is positive on both sides. i.e. we have a higher likelyhood of a stronger effect on our promotions efficacy.

Similar query was done for each A/B pair

```
WITH sales AS (
SELECT
COUNT(DISTINCT location_id) AS location_count,
promotion,
SUM(sales_in_thousands) AS sum_sales,
AVG(sales_in_thousands) AS avg_sales,
STDDEV(sales_in_thousands) AS stddev_sales,
VARIANCE(sales_in_thousands) as variance_sales,
AVG(age_of_store) AS avg_age
FROM
(
SELECT
location_id,
promotion,
SUM(sales_in_thousands) AS sales_in_thousands,
AVG(age_of_store) AS age_of_store
FROM
`tc-da-1.turing_data_analytics.wa_marketing_campaign`
GROUP BY
location_id,
promotion
) AS check_for_quadruplicates
GROUP BY
promotion
ORDER BY
promotion
),
comparison AS (
SELECT
a.avg_sales AS avg_sales_1,
a.stddev_sales AS stddev_sales_1,
a.location_count AS num_stores_1,
a.variance_sales as variance_sales_1,
```

```
b.avg_sales AS avg_sales_2,
b.stddev_sales AS stddev_sales_2,
b.location_count AS num_stores_2,
b.variance_sales as variance_sales_2,
FROM
sales a
JOIN
sales b
ON
a.promotion = 1 AND b.promotion = 2
),
comparison_w_pooled_variance as (
SELECT *,
((num_stores_1 - 1 ) * variance_sales_1 + (num_stores_2 - 1 ) * variance_sales_2) / (num_stores_1 + num_stores_2 - 2) as pooled_variance,
num_stores_1 + num_stores_2 - 2 as df
FROM comparison )
SELECT
avg_sales_1,
stddev_sales_1,
num_stores_1,
variance_sales_1,
variance_sales_2,
avg_sales_2,
stddev_sales_2,
num_stores_2,
pooled_variance,
(avg_sales_1 - avg_sales_2) / SQRT(pooled_variance * (1 / num_stores_1 + 1 / num_stores_2)) AS t_statistic
FROM
comparison_w_pooled_variance;
```