COMPARING MAXIMUM AND AVERAGE BIDDING: AN A/B TEST IN FACEBOOK ADS

Created by: Ricko Januar H.

https://github.com/Lt-Dan-Taylor/

PROJECT STRUCTURE

01 Introduction

02 Data Pre-Processing

03 Exploratory Data Analysis

04 A/B Testing

05 Conclusion

INTRODUCTION

The core of this project centers around the A/B test, a statistical technique enabling a direct comparison between two groups. In this context, our focus is on examining the impact of two different bidding strategies in Facebook advertising, Maximum bidding and Average bidding.

- Maximum bidding, a manual approach where advertisers establish a fixed maximum bid.
- Average bidding, an automated method in which Facebook's algorithm adaptively adjusts bids in response to real-time auction dynamic.

Main Question

1. Do different bidding strategies yield varying results in terms of impressions, ad-clicks, purchases and earnings?
2. Which bidding strategy produces the most favorable outcomes?
3. Are the result statistically significant for decision-making?

DATA PRE-PROCESSING

Before the dataframe is ready for analysis, we need to ensure that we have all the necessary data. Below is the raw data and the descriptions of each.

Raw Data

	dtype	Nan	num_unq	example
Impression	int64	0	39	[82529, 98050, 82696, 109914]
Click	int64	0	40	[6090, 3383, 4168, 4911]
Purchase	int64	0	37	[665, 315, 458, 487]
Earning	int64	0	39	[2311, 1743, 1798, 1696]

- Impression: the number of times an ad is shown to users.
- Click: the number of times users interact with an ad by clicking on it.
- Purchase: the completion of a transaction where a user buys a product or service.
- Earning: the total income generated over a specified period.

The dataset consists of two dataframes: control (maximum bidding data) and test (average bidding data), both sharing identical columns with different values. Below are the steps to prepare the analysis-ready dataframe:

1. Create a 'Date' column by generating random dates.

2. Add a 'Group' column for each dataframe.

3. Merge the two dataframes into one.

The sample of the final dataframe result:

Main Dataframe

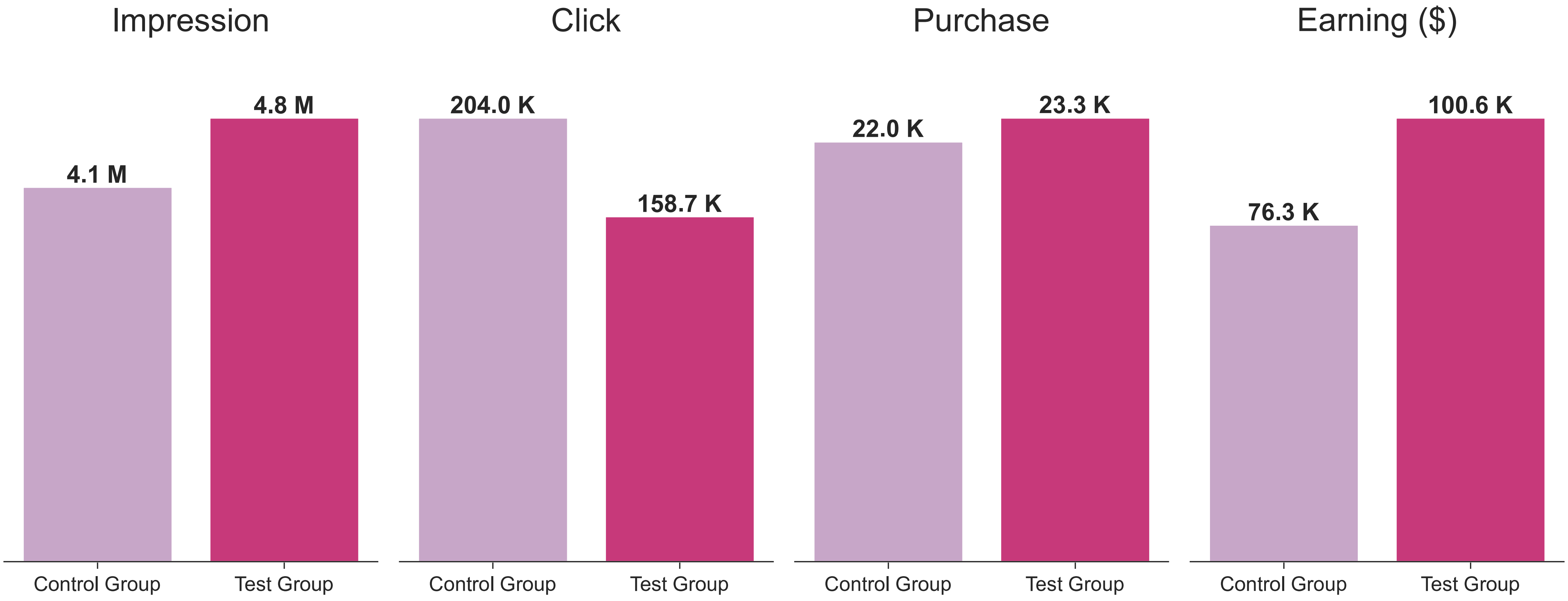
	Group	Date	Impression	Click	Purchase	Earning
0	Control Group	2022-01-01	82696	4168	458	1798
1	Test Group	2022-01-04	137222	4042	677	2260
2	Control Group	2022-01-18	147539	3857	329	2144
3	Test Group	2022-01-18	119878	3623	689	2812
4	Control Group	2022-01-22	92045	4667	729	2497
5	Control Group	2022-01-26	83677	4273	386	2174
6	Test Group	2022-01-29	130702	3626	450	2531
7	Control Group	2022-02-07	121086	4285	590	1289
8	Test Group	2022-02-19	141368	3925	501	2786
9	Test Group	2022-02-21	129801	4244	629	2756
...	...	...	...	...	...	...

EXPLORATORY DATA ANALYSIS

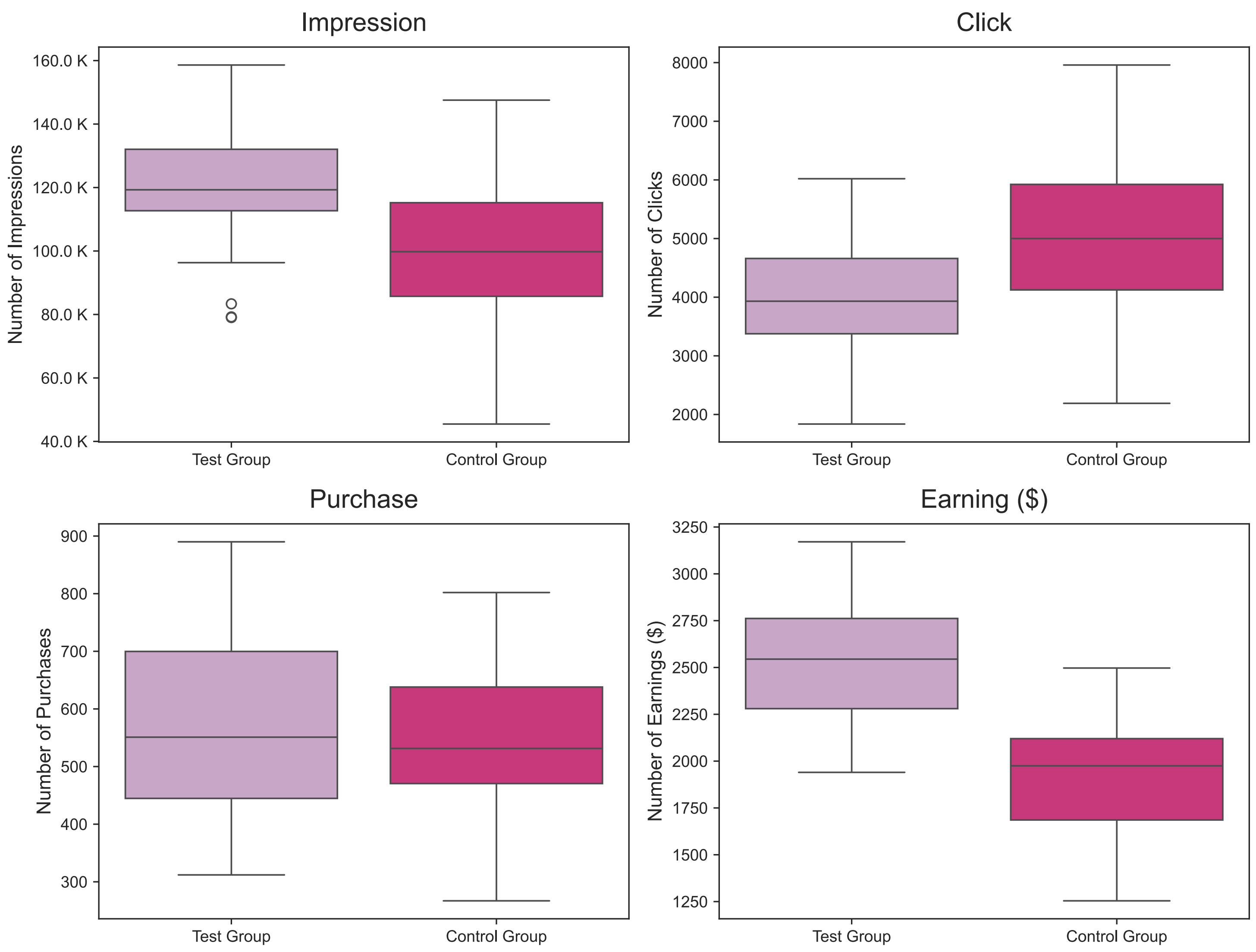
Performing exploratory data analysis and visualization is a crucial step to learn the distribution of the data and it underlying patterns. We began by asking question that can extract insight from the data.

1. What are the total impressions, clicks, purchases, and earnings from each group?

- ❑ Impression: the Test group outperforms the Control group slightly, showing up to 7k more impressions or a 17% increase.
- ❑ Click: the Control group significantly outperforms the Test group, registering up to 45.3k more clicks or a 28.5% increase.
- ❑ Purchase: the Test group demonstrates a slight improvement over the Control group, recording up to 1.3k more purchases or a 5% increase.
- ❑ Earning: the Test group significantly outperforms the Control group, yielding $24.3k more or a 31.8% increase in earnings.



Impression
4.1 M
4.8 M
Click
204.0 K
158.7 K
Purchase
22.0 K
23.3 K
Earning ($)
76.3 K
100.6 K
Control Group
Test Group



2. What are the descriptive statistics for the data in each group?

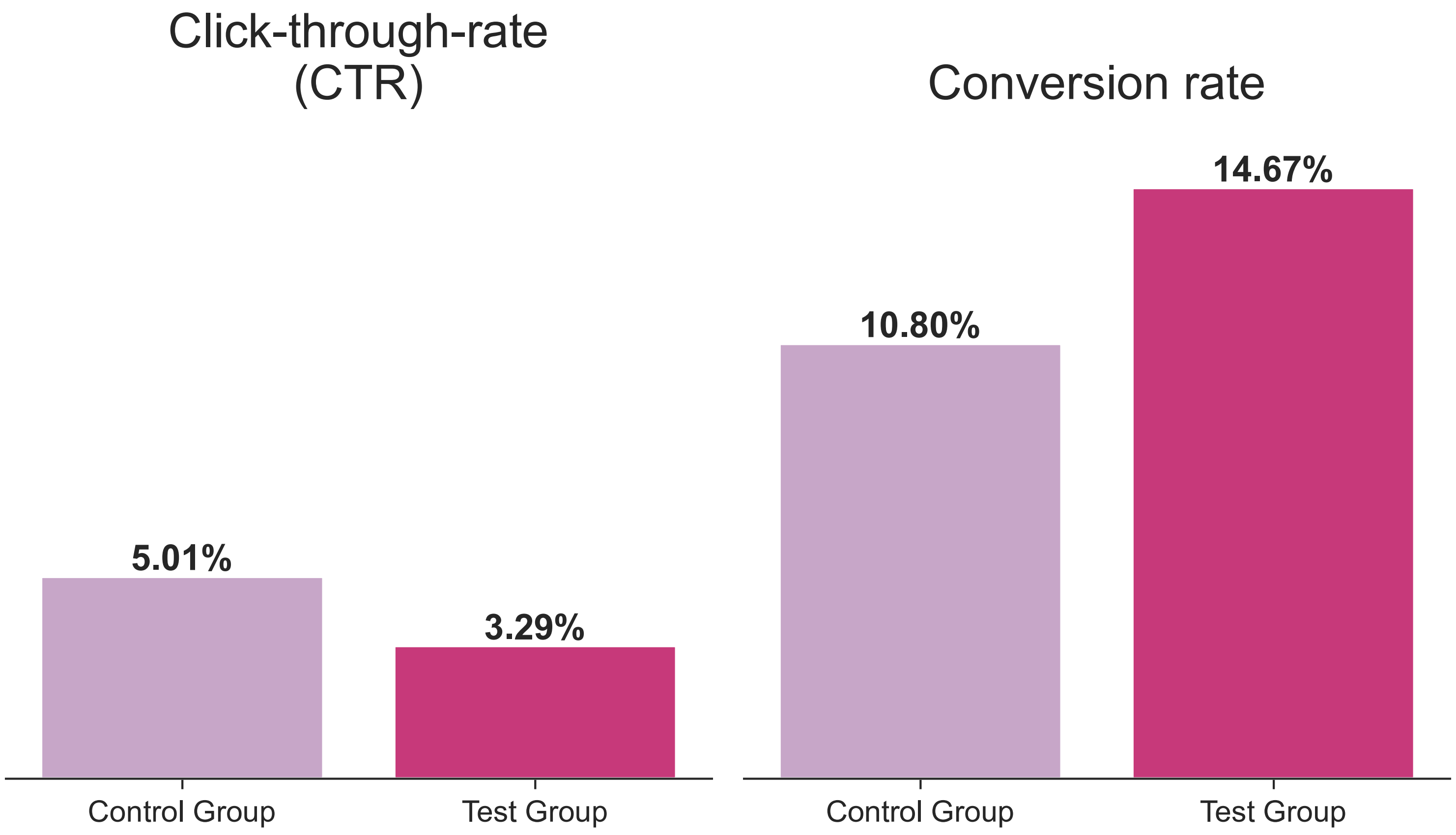
- In terms of impressions, purchases, and earnings, the Test group has a higher median, a higher maximum value, and a lower minimum value than the Control group.

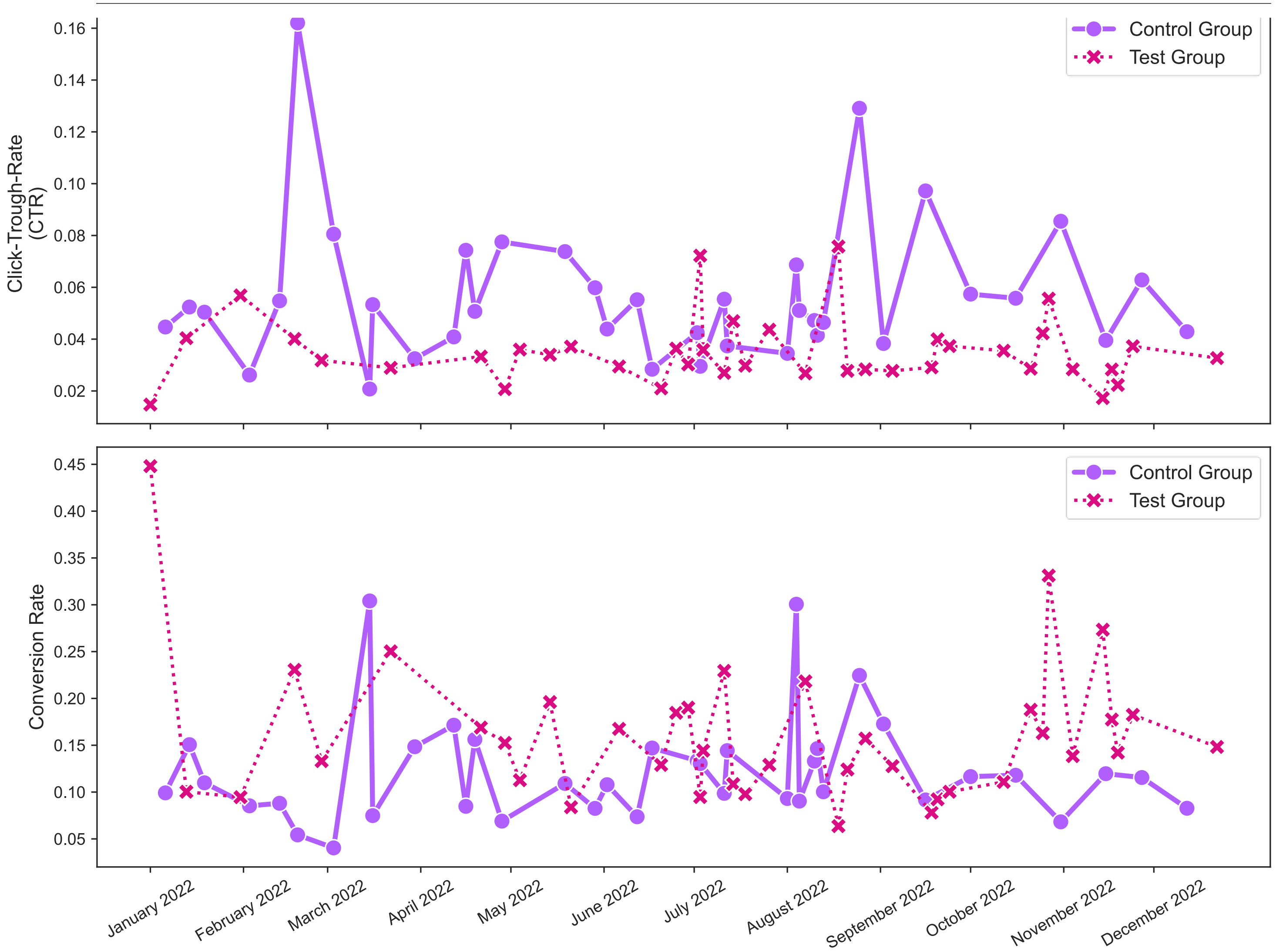
- In terms of clicks, the Control group has a higher median and a higher maximum value than the Test group.

3. What are the Click-Through-Rate (CTR) and Conversion rates for each group?

- The Control group performs slightly better than the Test group in terms of Click-through-rate, which is measured by the number of ad-clicks per impression.

- The Test group performs slightly better than the Control group in terms of Conversion rate, which is measured by the number of purchases per ad clicked.





4. How does the trend in CTR and Conversion rate vary for each group?

- The Control group consistently dominates the trend in CTR most of the time, with its highest rate reaching up to 16%
- In terms of Conversion rate, the Test group emerged as the winner, achieving its highest rate of a staggering 45% in the starting month.

A/B TESTING

In this section, our objectives are as follows:

Generate bootstrap samples with sample sizes identical to the original dataset.

Calculate the difference in 'converted' probabilities between these samples.

Create a sampling distribution of this difference through 10,000 iterations.

Perform a hypothesis test to assess the significance of the difference in 'converted' probabilities.

Repeat the test for click-through-rate, conversion rate and earning.

1. Click-through-rate (CTR)

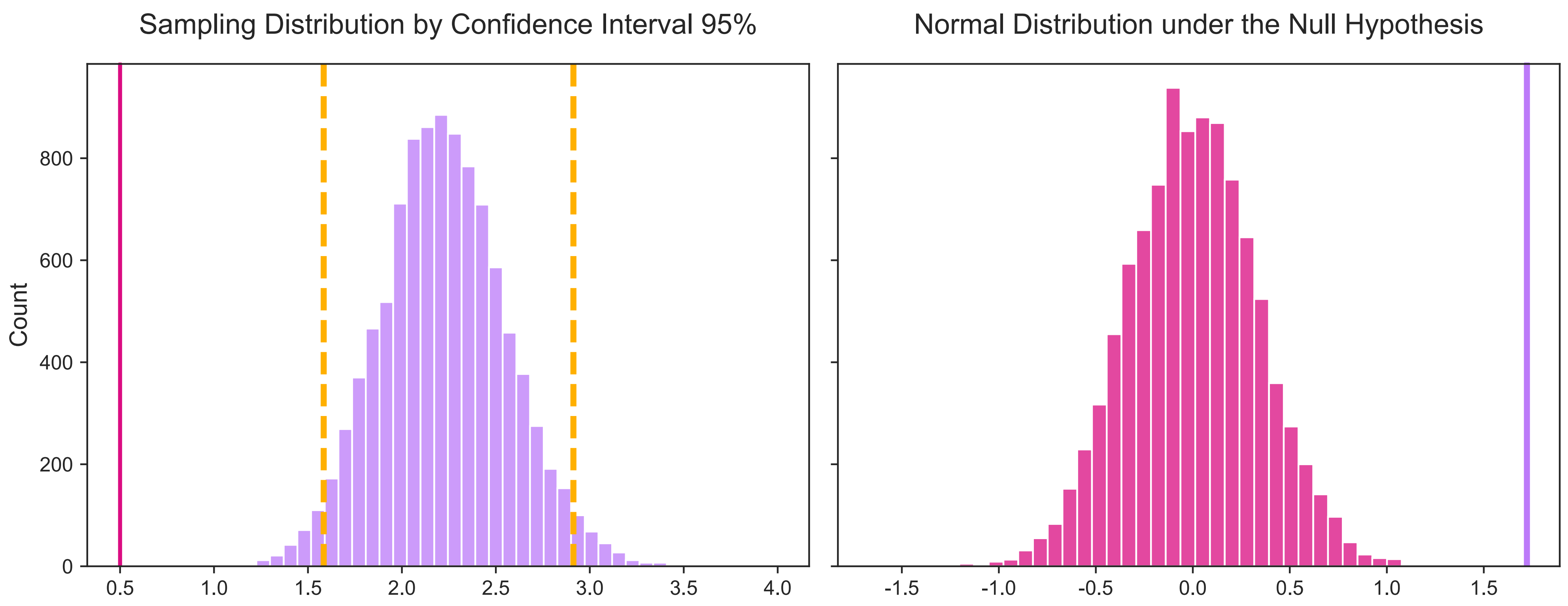
As we observed in the EDA section, the Click-through-rate (CTR) for the Control group is higher than that of the Test group. We will now determine if this difference is statistically significant.

Null Hypothesis (H0):

$$H_0 : P(control) - P(test) \leq 0$$

Alternative Hypothesis (HA):

$$H_A : P(control) - P(test) > 0$$



- Result: Reject the Null Hypothesis (H0)
- The Control group's CTR performance is significantly superior to the Test group's CTR performance.

2. Conversion rate

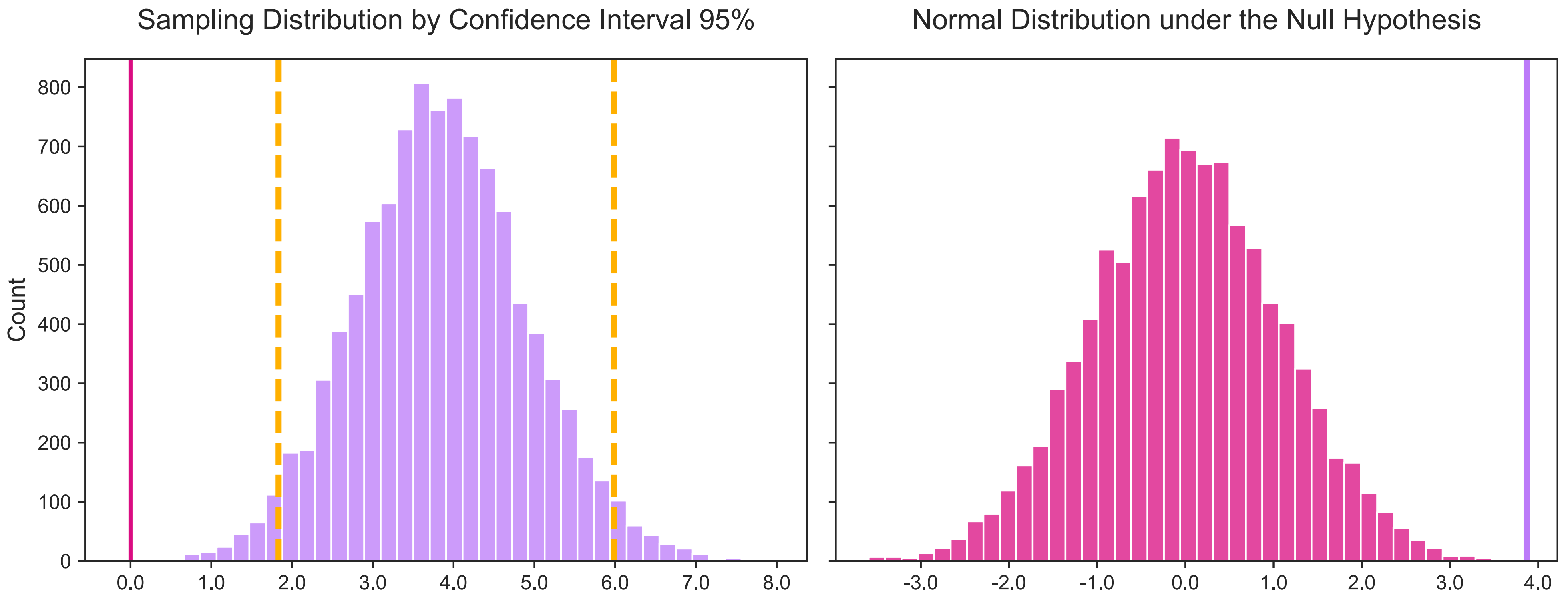
Conversion rate for Test Group is much higher than Control Group, so we will now assess whether this difference is statistically significant.

Null Hypothesis (H0):

$$H_0 : P(test) - P(control) \leq 0$$

Alternative Hypothesis (HA):

$$H_A : P(test) - P(control) > 0$$



- ❑ Result: Reject the Null Hypothesis (H0)
- ❑ The Test group's Conversion rate is significantly superior to the Control group's Conversion rate.

3. Earnings

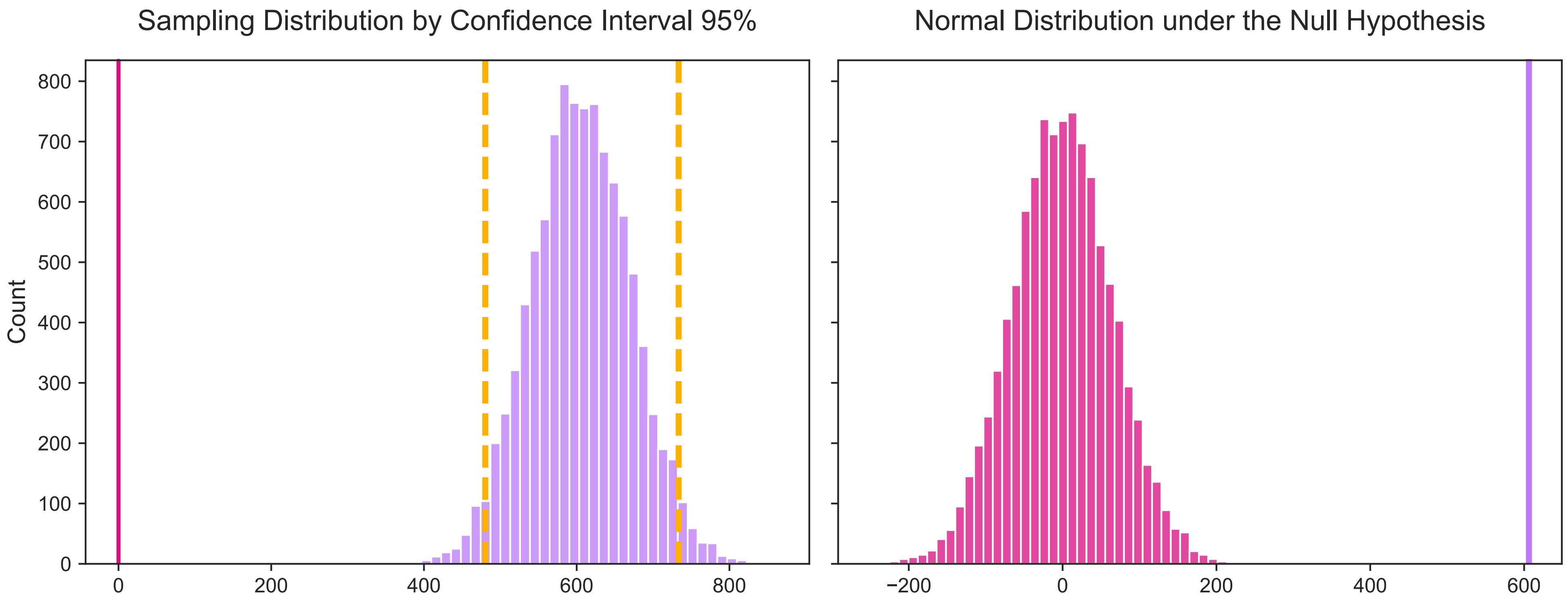
Earning for Test Group is higher beyond the Control Group, so we will now assess whether this difference is statistically significant.

Null Hypothesis (H0):

$$H_0 : P(test) - P(control) \leq 0$$

Alternative Hypothesis (HA):

$$H_A : P(test) - P(control) > 0$$



- ❑ Result: Reject the Null Hypothesis (H0)
- ❑ The Test group’s Earning is significantly superior to the Control group’s Earning.

CONCLUSION

Here are some of the key findings and insights from this project:

- Average Bidding appears to be more effective in terms of impressions, purchases,and earnings, while Maximum Bidding performs better in generating ad clicks.
- Maximum Bidding demonstrates a significant superior Click-through-rate (CTR) compared to Average Bidding.
- On the other hand, Average Bidding significantly outperforms Maximum Bidding in conversion rate and earnings.

These insights underline the flexibility of different bidding strategies, by carefully selecting the bidding strategy, advertisers can optimize their campaigns and maximize their ROI.

References:

Check out the following resources to learn more about the datasets and tools/libraries used in this project:

- Optimizing Ad Bidding | Facebook's A/B Test Story: https://kaggle.com/datasets/furth3r/facebook-ab-test-of-bidding-feature
- Matplotlib user guide: https://matplotlib.org/3.3.1/users/index.html
- Seaborn user guide: https://seaborn.pydata.org/tutorial.html
- opendataset Python library: https://github.com/JovianML/opendatasets
- SciPy user guide : https://docs.scipy.org/doc/scipy/tutorial/index.html

Source code: https://github.com/Lt-Dan-Taylor/facebook-ads-AB-testing