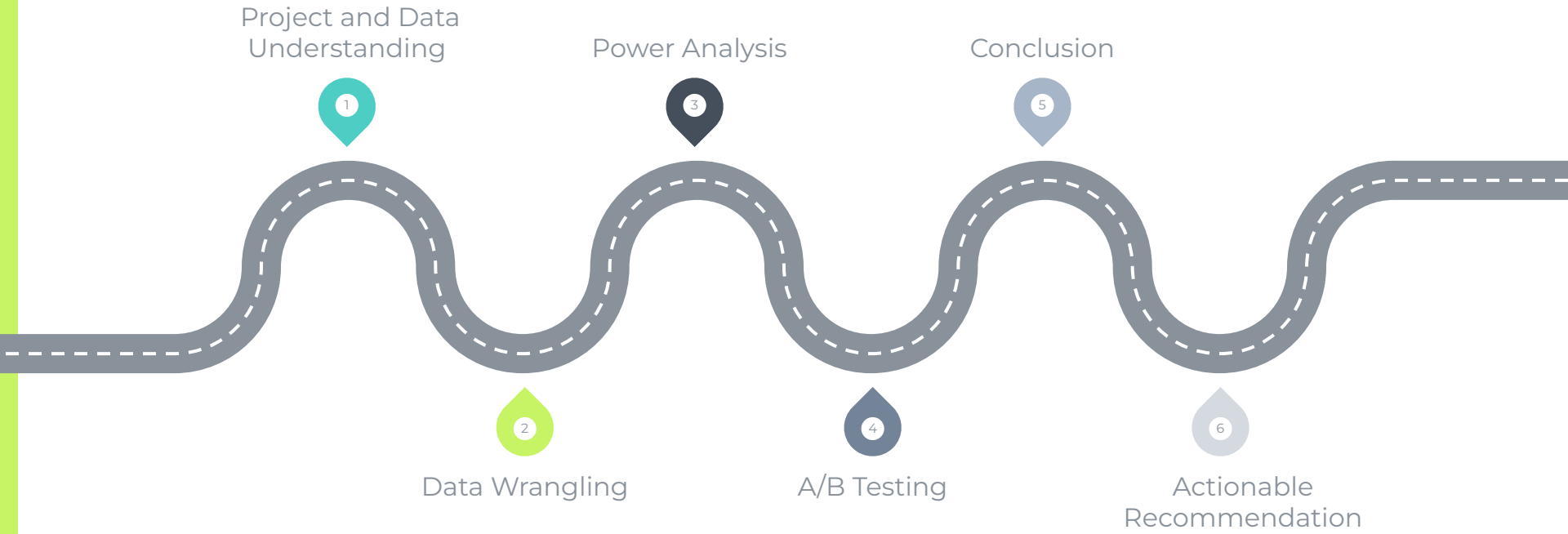
Evaluate the Effectiveness of New Campaign With A/B Testing

By Riza Arfiqih

Overview


Project and Data Understanding
1
Power Analysis
3
Conclusion
5
2
Data Wrangling
4
A/B Testing
6
Actionable Recommendation

Background

A company has developed a new web page design and new advertising campaign with the aim of increasing the number of users who convert (i.e., pay for the company's product). This is a crucial business objective as conversion rates directly impact revenue and overall business performance.

Project Outline

Objective

This project aims to determine if a new ad campaign significantly increases revenue compared to PSAs and if a new web page design significantly improves conversion rates compared to the existing design. This will be achieved through A/B testing and statistical analysis.

Outcomes

- Identify whether the new web page design has a statistically significant impact on conversion rates.
- Quantifying the potential revenue gain from the new ad campaign.
- Determining the statistical significance of the difference in conversion rates between the new ad campaign and PSAs.
- Providing actionable recommendations to the company based on the findings.

Methodology

- Data Cleaning and Preprocessing
- Exploratory Data Analysis (EDA)
- Statistical Analysis (Hypothesis Testing)
- Actionable Recommendations

Data Overview

new_design_page_test.csv

The dataset was obtained from Kaggle.com. E-commerce A/B testing | [https://www.kaggle.com/datasets/ahmedmohameddawoud/ecommerce-ab-testing/data

#	Column	Non-Null Count	Dtype
0	user_id	294478 non-null	object
1	time	294478 non-null	object
2	group	294478 non-null	object
3	landing_page	294478 non-null	object
4	converted	294478 non-null	int64

This dataset describes of an A/B test results run by an e-commerce website. The company has developed a new web page in order to try and increase the number of users who convert, meaning the number of users who decide to pay for the company's product.

It consists of 5 features and more than 290.000 rows:

- **user_id**: User ID (unique)
- **time**: Time of a session
- **group**: Contains 2 variables, control and treatment
- **landing_page**: What version of a site a user saw, old_page or new_page
- **converted**: If a user made a purchase (1) or not (0)

marketing_AB.csv

The dataset used in this project describes the results of an A/B test conducted by the company. It is titled "Marketing A/B Testing" and can be found on Kaggle. [https://www.kaggle.com/datasets/faviovaz/marketing-ab-testing/data].

#	Column	Non-Null Count	Dtype
0	user_id	588101 non-null	int64
1	test_group	588101 non-null	object
2	converted	588101 non-null	int64
3	total_ads	588101 non-null	int64
4	most_ads_day	588101 non-null	object
5	most_ads_hour	588101 non-null	int64

This dataset describes an A/B test results run by a company that wanna find if the ads were successful and how much the company can make from the ads.

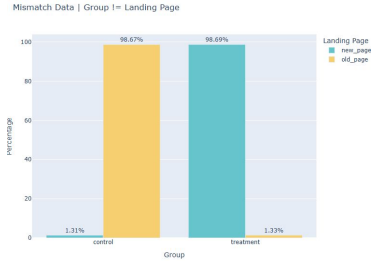
It consists of 6 features and more than 588.101 rows:

- **user_id**: User ID (unique)
- **test_group**: If "ad" the person saw the advertisement, if "psa" they only saw the public service announcement
- **converted**: If a person bought the product then 1, if not 0
- **total_ads**: Amount of ads seen by person
- **most_ads_day**: Day that the person saw the biggest amount of ads
- **most_ads_hour**: Hour of day that the person saw the biggest amount of ads

Data Inspection

New Design Page

- **Missing Values**
  There are no missing values!
- **Duplicate Data**
  Drop a duplicate data in 'user_id' == '773192'
- **Mismatch Data**
  The correct data are on the feature group, control should have old_page then treatment should have new_page.



Drop mismatch data that consist small portion of 3893 rows out of 294478 (only 1% - 2%).

CODE

New Ads Campaign

- **Missing Values**
  There are no missing values!
- **Duplicate Data**
  There is no duplicated data

The dataset is clean!

CODE

Power Analysis

New Design Page

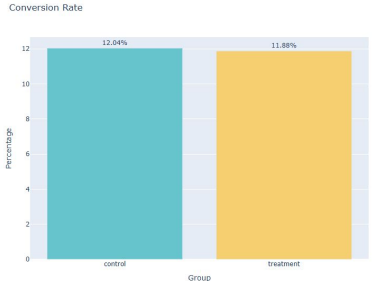
Control Group Conversion Rate: 12.04 %

The conversion rate (number of purchases/number of visits) in the company before was 12.04 % and the increase expected will be set to 14 % to ensure that a new design page of the site will certainly increase the conversion.

For this experiment, we need at least 4433 data for each group

Test Group	Count
Treatment	145310
Control	145274

The data fulfilled the requirement for doing the experiment since both groups had more than 145.000 data.



It shows that the new design page performs poor than the old page. The change in conversion rate does not suggest that there is an improvement when we use the new page (11.9 %), it is slightly lower than the old page (12 %).

CODE

New Ads Campaign

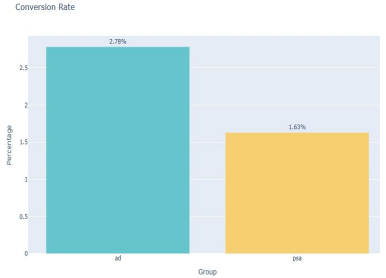
Ads Conversion Rate: 1.79 %

The conversion rate in the company when only using public service announcements was 1.79 % and the expected increase will be set to 3 % to ensure that the new ads campaign will certainly increase the conversion.

For this experiment, we need at least 2514 data for each group

Test Group	Count
ad	564577
psa	23524

The data fulfilled the requirement for doing the experiment. In this case, I use 2514 sample data on each test group.



It shows that the new ads campaign performs slightly well than the psa. The change in conversion rate does suggest that there is an improvement of 1.15 % when we use the new ads campaign (2.78%), it is slightly higher than the psa (1.63 %).

CODE

Let's validate whether this difference is statistically significant!

I'll perform a statistical test to compare these two groups with using z-test to calculate the p-value considering the large sample size.

New Design Page

Does this new design page improve the conversion rates?

- **H0**: The new design page does not significantly improve the conversion rates.
- **H1**: The new design page does significantly improve the conversion rates.

Test Group	Conversion Rate	Metrics Different	P-Value	Z Statistic	Confidence Interval
Control	12.04 %	⬇ -0,16 %	0.190	1.31	[0.119, 0.122]
Treatment	11.88 %				[0.117, 0.120]

P-Value > 0.05

"Fail to reject the null hypothesis"

It statistically validates that the new design page **does not significantly improve** the conversion rates.

CODE

New Ads Campaign

Does this new ads campaign improve the conversion rates?

- **H0**: The new ads campaign does not significantly improve the conversion rates.
- **H1**: The new ads campaign does significantly improve the conversion rates.

Test Group	Conversion Rate	Metrics Different	P-Value	Z Statistic	Confidence Interval
PSAs	1.63 %	⬆ 1.15 %	0.005	-2.78	[0.011, 0.021]
ADs	2.78 %				[0.021, 0.034]

P-Value < 0.05

"Reject the null hypothesis"

It statistically validates that the new ads campaign **does significantly improve** the conversion rates.

CODE

Conclusion

New Design Page

The A/B test results indicate that **there is no statistically significant evidence to support the claim that the new design page improves conversion rates**. With a Z statistic of 1.31 and a p-value of 0.190, which is greater than the significance level of 0.05, we fail to reject the null hypothesis.

- **Failure to Reject the Null Hypothesis**: A p-value of 0.190 suggests that there is a 19% chance of observing the obtained results (or more extreme results) if there is truly no difference in conversion rates between the old and new design pages. This probability is higher than the commonly used significance level of 0.05, leading to the failure to reject the null hypothesis.
- **Confidence Intervals**: The 95% confidence intervals for the conversion rates of both the control group (old page) and the treatment group (new page) overlap. This overlap further supports the conclusion that there is no significant difference between the two designs.
- **Conversion Rates**: The observed conversion rates for the control group (12.04%) and the treatment group (11.88%) are very similar. The slight difference is not statistically significant, as indicated by the p-value and confidence intervals.

New Ads Campaign

The A/B test provides **statistically significant evidence to conclude that the new ad campaign improves conversion rates** compared to the public service announcements (PSA).

- **Rejecting the Null Hypothesis**: With a p-value of 0.005, which is less than the significance level of 0.05, we reject the null hypothesis. This indicates a statistically significant difference in conversion rates between the two groups.
- **Confidence Intervals**: The 95% confidence intervals for the ad campaign (0.021, 0.034) and the PSA group (0.011, 0.021) do not overlap, supporting the conclusion of a significant difference.
- **Improved Conversion Rates**: The new ad campaign demonstrates a higher conversion rate (2.78 %) than the PSA group (1.63 %), representing an improvement of 1.15 %.
- **Z Statistic**: The Z statistic of -2.78 further indicates the significance of the difference observed.
- **Potential Revenue Gain**: In this scenario, the potential revenue gain from the new ad campaign would be $**1.445.55***.

*assume an AOV (Average Order Value) of $50 and traffic of user 2514

Recommendation

New Design Page

Based on these findings, it is recommended to **retain the existing design page or explore alternative design modifications for further testing.** The current new design page does not demonstrate a statistically significant improvement in conversion rates, and therefore, implementing it might not yield the desired positive impact on business outcomes.

New Ads Campaign

Based on these findings, it is recommended to **implement the new ad campaign to potentially improve overall conversion rates and drive positive business outcomes**. The statistical evidence strongly supports its effectiveness in driving conversions compared to the previous approach using public service announcements.

Thanks!

Let’s get in touch!


Email


LinkedIn


Github