Ad Impact: Quantitative Evaluation of Advertising Campaigns through A/B Testing

1. Background and Problem

Background and Overview of the problem

This research examines the effectiveness of targeted advertisements (ads) compared to public service announcements (PSAs) by analyzing data from 588,101 participants exposed to a varying number of ads. Using techniques such as Propensity Score Matching (PSM) and hypothesis testing, the study aims to understand how ad exposure influences conversion rates, and to ensure that comparisons between the ad and PSA groups are fair by accounting for potential biases.

Data Collection

The dataset utilized is sourced from Kaggle and simulates a typical marketing A/B testing scenario. It comprises data collected from a randomized control trial where participants are exposed to either targeted advertisements (ads) or a Public Service Announcement (PSA). The data includes variables such as the group exposed to (ad or PSA), conversion status, total ads seen, and the day and hour when most ads were viewed.

Research Objectives

The primary objectives of this research are to:

- Determine the overall success of the ad campaign in driving conversions.
- Quantify the impact of ads on the likelihood of conversion, distinguishing it from other factors.

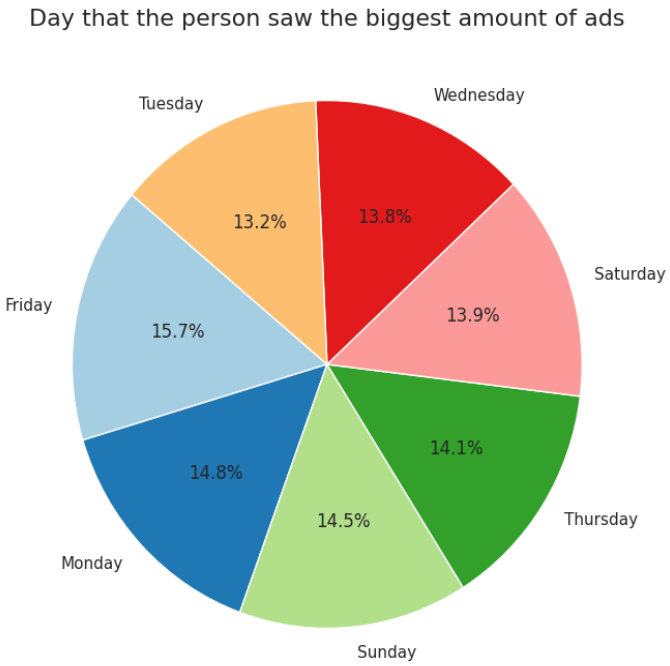
2. Data Summary and Exploratory Analysis

The dataset, consisting of observations from 588,101 participants, provides a rich foundation for understanding the dynamics of advertising exposure and effectiveness. Descriptive statistics indicate that participants are exposed to an average of 25 ads, with a substantial range in individual ad exposure, varying from 1 to 2,065 ads. Notably, most ads are viewed around 2 PM, highlighting peak viewing times in the early afternoon to evening.

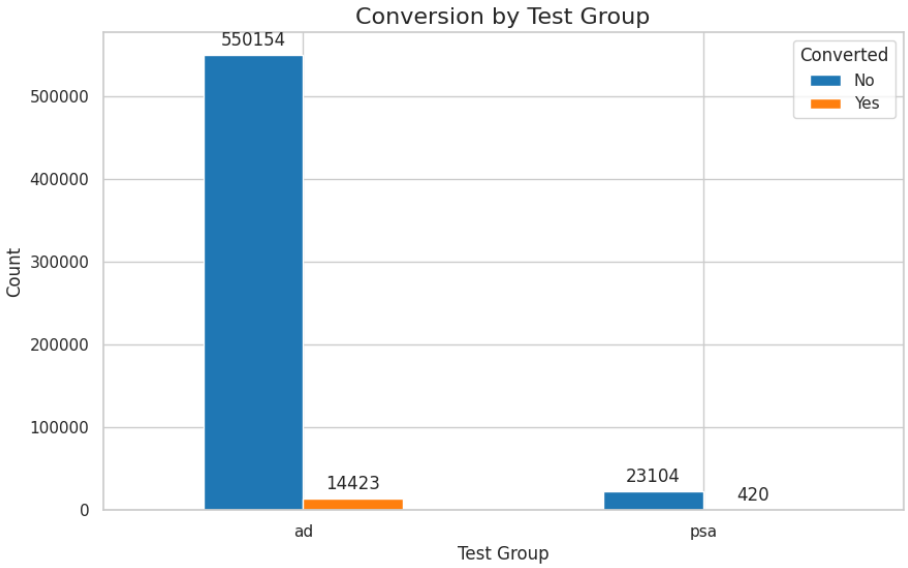
Descriptive Statistics

	total_ads	most_ads_hour
count	588101	588101
mean	25	14
std	44	5
min	1	0
25%	4	11
50%	13	14
75%	27	18
max	2065	23

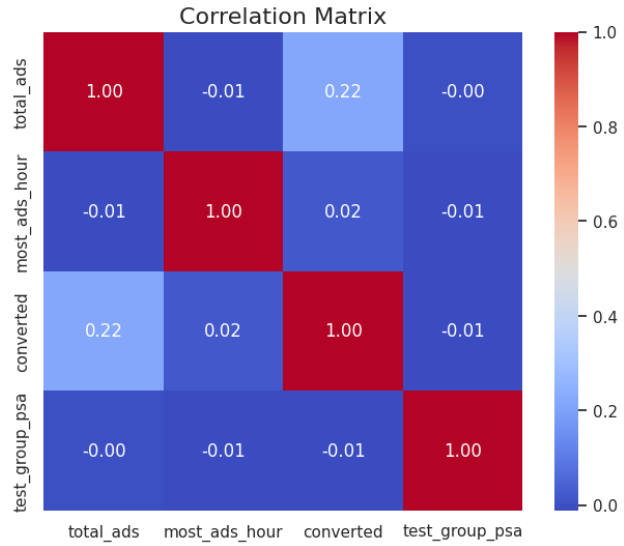
The next pie chart illustrates the distribution of ad views throughout the week. Friday stands out as the peak day (15.7% of total views), signaling a surge in engagement or online activity. Though Sunday sees fewer views, the weekend collectively shows robust interaction. This spread is relatively consistent, with daily percentages ranging narrowly from 13.2% to 15.7%, suggesting a stable viewer engagement throughout the week.



The conversion analysis via a bar chart reveals distinct differences between two test groups: 'ad' and 'psa'. The 'ad' group encompasses a large segment of participants but records a low conversion rate with 550,154 non-converters and 14,423 converters (3%). In contrast, the smaller 'psa' group consists of 23,104 non-converters and 420 converters, reflecting an even lower conversion rate of about 2%. This contrast in group sizes and conversion rates underscores the need for a nuanced analysis of advertising effectiveness.

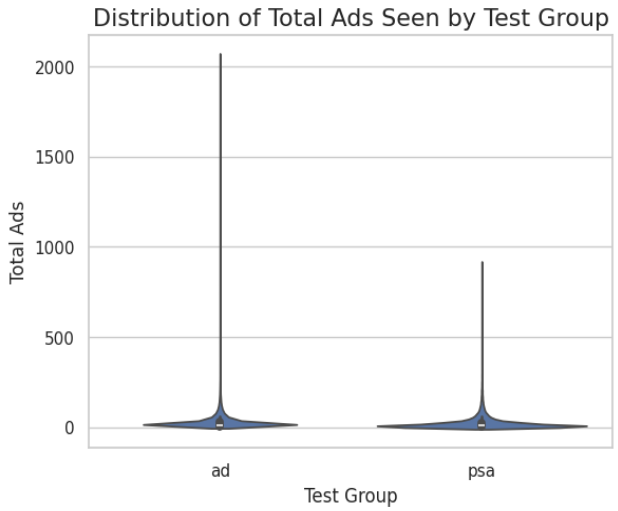


A correlation matrix provides further insight, identifying a moderate positive correlation (0.22) between the total ads seen and conversions. This suggests that increased ad exposure might potentially enhance conversion rates. However, other factors such as the peak ad-viewing hour and group membership show minimal impact on conversions, highlighting the complex interplay of variables affecting ad effectiveness.

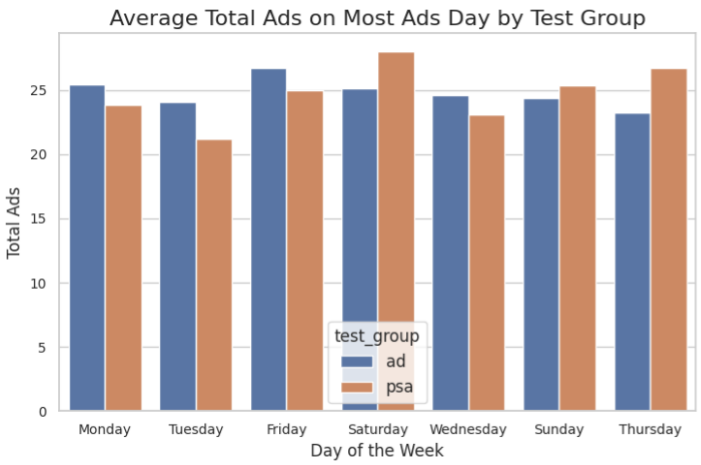


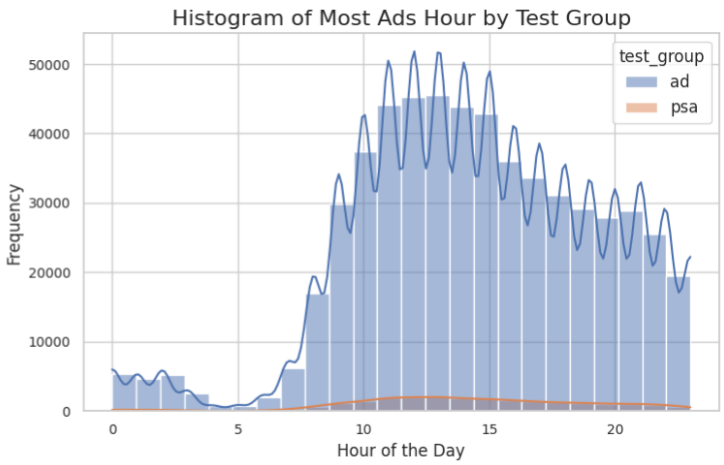
.

The violin plot and bar plot add another layer to our analysis, depicting the distribution and average number of ads seen by each test group. The violin plot shows that both groups have a similar distribution shape in terms of total ads seen, although the 'ad' group exhibits more variability.



The bar plot corroborates this, indicating that on their most active days, the 'ad' group tends to see more ads than the 'psa' group. Meanwhile, the histogram of the most ads hour by test group reveals a broader distribution for the 'ad' group across the day, suggesting that their ad exposure is not only higher but also more spread out during the day.





Given these insights, there is a need to employ PSM to adjust for potential biases arising from unequal sample sizes and variable exposure levels among the test groups. PSM will ensure a more equitable comparison of the effectiveness of targeted ads versus public service announcements by standardizing participant characteristics across groups, ultimately leading to more reliable conclusions about the impact of advertising strategies.

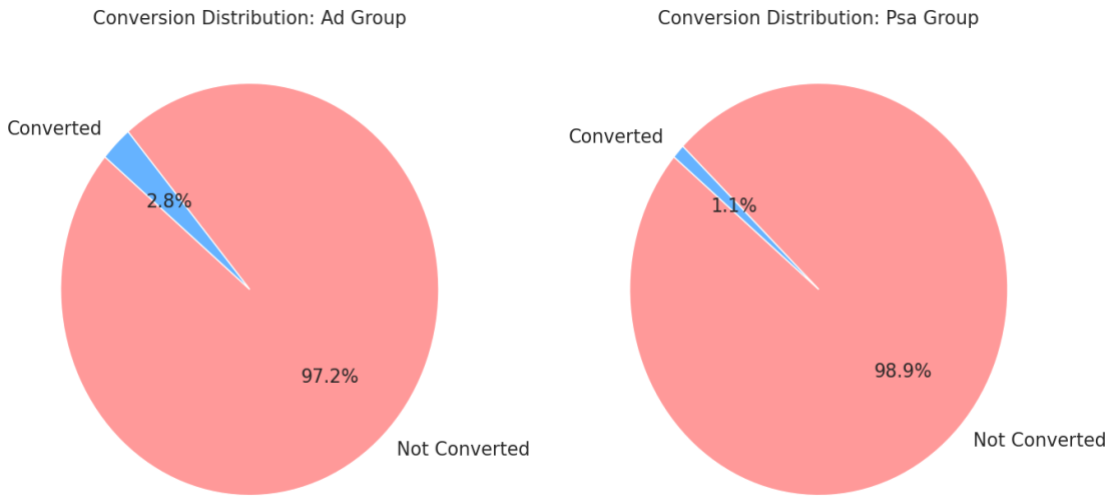
3. Propensity Score Matching (PSM)

Following detailed descriptive analysis and statistical testing, I have closely examined the relationships and differences between the ad and PSA groups using both the Chi-Squared Test and T-test. The Chi-squared tests revealed strong statistical associations in ad viewership patterns by day (p-value: 4.85e-48) and hour (p-value: 1.09e-28), indicating significant differences in how ads are viewed across these groups. However, the T-test, which compared the average number of ads seen by each group, reported no significant difference (p-value: 0.831), suggesting uniform ad exposure levels across both groups despite these categorical variances.

The findings highlight that variations in ad viewing times do not necessarily correspond to differences in ad volume, illustrating the complex nature of advertising impact. To address these complexities and accurately assess the effects of ads versus PSAs, PSM was utilized. PSM involved using a logistic regression model to calculate propensity scores, which determined the likelihood of individuals belonging to either the 'ad' group or the 'psa' group, based on categorical data like the days and hours when ads were most viewed. This model excluded identifiers and direct outcome measures such as conversions. After calculating these scores, we standardized them along with related covariates to maintain consistency in the matching process, which was meticulously executed using a Nearest Neighbors algorithm. This ensured that each ad-exposed individual was paired with a PSA-exposed counterpart based on their propensity scores.

Post-matching, a dataset of these pairs was created to facilitate comparisons, assessing match quality by measuring the distances between matched pairs and quantifying the similarity of their characteristics. We then verified the balance of covariates across treatment and control groups in this new dataset, confirming that observed differences in outcomes, such as conversion rates, are attributable to ad exposure rather than any inherent biases between groups. This validation underscored the effectiveness of the matching process.

After achieving covariate balance, individuals in the treatment group were found to have been exposed to more ads (25.17 ads compared to 19.35 ads in the control group). Additionally, the treatment group exhibited a higher average conversion rate of 2.8%, significantly contrasting with 1.1% in the control group. This difference emphasizes the effectiveness of ad exposure in enhancing consumer behavior, affirming the superior impact of targeted advertising over PSAs. Thus, the post-matching analysis clearly attributes the observed differences in conversion rates to the type of ad exposure, rather than to any pre-existing biases between the groups.



4. Hypothesis Testing

After achieving a balanced distribution of covariates between the treatment and control groups through matching, I conducted further hypothesis testing to assess the impact of ad exposure on conversion rates. A Chi-squared test was performed using a contingency table of conversions by test group. The resulting Chi-squared statistic of 346.909, along with an extremely low p-value, decisively indicates a significant association between group membership and conversion rates, suggesting that the observed differences are not due to chance.

This analysis confirms that exposure to ads is significantly correlated with increased conversion rates, demonstrating the effectiveness of targeted advertising in influencing consumer behavior. Specifically, the data shows that ad exposure enhances conversion rates by approximately 1.65%, underscoring the success of the advertising campaign. These findings reinforce the conclusion that the differential in conversion rates between the groups is directly attributable to the type of ad exposure received, validating the strategic use of targeted advertisements to boost consumer engagement and conversions.

5. Conclusion

- Effectiveness of Targeted Ads: The data consistently shows that targeted advertisements significantly enhance conversion rates compared to public service announcements. Specifically, individuals in the ad group had higher exposure and conversion rates, demonstrating the effectiveness of targeted ads in influencing consumer behavior.
- The analysis highlighted that not only the quantity but also the timing and frequency of ads play a crucial role in maximizing consumer engagement. Ads shown more frequently and during optimal times (early afternoon), have a higher impact, as evidenced by the peak viewing times and higher conversion rates.
- Employing PSM helped ensure that differences in conversion rates between the groups were attributable to the ads themselves and not pre-existing differences. This method proved essential for validating the direct impact of advertising strategies on consumer actions.
- The findings provide insights that can guide businesses on how to allocate marketing budgets more efficiently. By understanding which aspects of the advertising campaigns yield higher conversions, companies can optimize their marketing efforts for better ROI.
- The study also underscores the need for transparency and ethical considerations in digital advertising. The clear impact of targeted advertising suggests that companies must consider consumer privacy and consent when designing such campaigns, potentially informing regulatory policies on digital advertising practices.

These conclusions affirm the value of rigorous A/B testing and sophisticated analytical techniques like PSM in quantifying the impact of different advertising strategies, enabling more informed business decisions and policy formulations in the digital marketing realm.