Process Documentation – Marketing A/B Testing

This document outlines the step-by-step process used to handle the Marketing A/B Testing dataset, including data acquisition, cleaning, analysis, and visualization. The goal was to evaluate the performance of ads versus public service announcements (PSAs) and derive actionable insights.

1 - Data Acquisition

• The dataset was sourced from Kaggle, ensuring anonymity while preserving essential insights.

• The file format was Excel (.xlsx), which was downloaded and stored locally before processing.

• Python was used to read and manipulate the dataset using the Pandas library.

2 - Data Cleaning & Preparation

• The dataset was loaded into Pandas for initial examination, identifying missing values and inconsistencies.

• Cleaning steps included:

- Removing duplicate records to avoid bias.

- Converting column types where necessary (e.g., ensuring `user_id` is an integer).

- Handling missing values by either imputing or removing incomplete records.

- Standardizing categorical values to ensure consistency (e.g., ensuring 'ad' and 'psa' labels are uniform).

3 - Data Analysis

• Performed Exploratory Data Analysis (EDA) using Python to identify trends in conversion rates, ad effectiveness, and user engagement.

• Key insights derived included:

- Conversion rates for ads vs. PSAs.

- The most effective days and hours for ad engagement.

- User behavior based on the frequency of ads displayed.

• Data was aggregated to create summary tables for better visualization in Power BI.

4 - Data Export & Transformation

• Once cleaned and analyzed, the processed dataset was exported as a CSV file for easy integration into Power BI.

• Additional transformations performed included:

- Creating calculated fields (e.g., conversion rate percentages).

- Aggregating ad impressions by time of day and day of the week.

5 - Data Visualization in Power BI

• The cleaned dataset was imported into Power BI to create an interactive dashboard.

• Visualizations included:

- Conversion Rate Comparison (Ads vs. PSAs).

- Ad Performance Trends by Day and Hour.

- Heatmaps for Most Ads Seen by Users.

- User Engagement Metrics.

6 - Key Learnings & Actionable Insights

• Peak Conversion Times: Ads were most effective on Fridays, while PSAs had better engagement midweek.

• Time-Based Optimization: Ads were seen the most between 10 AM - 6 PM, suggesting an ideal window for targeting users.

• Ad Frequency Analysis: Users exposed to more ads had a higher likelihood of conversion, highlighting the need for frequency optimization.

• Actionable Recommendations:

- Increase ad spend during peak engagement hours.

- Improve PSA content strategies to boost conversions.

- Use frequency capping to avoid ad fatigue.