Data Dictionary – Marketing A/B Testing

This Data Dictionary provides definitions and descriptions of all key fields used in the dataset, ensuring clarity in understanding the data structure.

1 - Data Fields

Column Name	Data Type	Description
user_id	Integer	Unique identifier assigned to each user.
test_group	String	The group the user was assigned to: 'ad' or 'psa'.
converted	Boolean	Indicates whether the user converted (TRUE) or not (FALSE).
total_ads	Integer	Total number of ads the user was shown.
most_ads_day	String	The day of the week when the user saw the most ads.
most_ads_hour	Integer	The hour of the day when the user saw the most ads.

2 - Data Source

The dataset was sourced from **Kaggle**, a well-known platform for open-source datasets. It has been structured to maintain anonymity while preserving its analytical value. The dataset was used for A/B testing analysis of marketing ads versus public service announcements (PSAs).

3 - How the Data Was Used

The dataset underwent several stages of processing and transformation to create meaningful insights:

• **Data Uploading & Cleaning:** The dataset was imported into Python using Pandas. Initial cleaning involved removing duplicates, handling missing values, and standardizing formats.

• **Data Analysis:** Exploratory Data Analysis (EDA) was performed in Python to uncover trends in conversion rates, ad effectiveness, and user engagement.

• **Data Transformation:** Aggregation and feature engineering were performed to create additional insights before exporting the dataset to Power BI.

• **Data Visualization:** The cleaned and processed data was imported into **Power BI**, where an interactive dashboard was built to visualize A/B test results and key performance indicators (KPIs).