Project Title:

Marketing Spend Effectiveness Analysis Using Statistical Techniques
_____

1. Project Overview

This project aims to evaluate the performance of marketing campaigns by conducting descriptive and inferential statistical analysis on key marketing metrics—Click-Through Rate (CTR) and Cost Per Acquisition (CPA). The focus is to uncover insights on data distribution, detect outliers, estimate population parameters, and test hypotheses to support data-driven decision-making.

2. Objective

- To understand the distribution and behavior of key marketing performance indicators.
- To identify outliers using the Empirical Rule and Z-scores.
- To compute confidence intervals to estimate population mean for CTR and CPA.
- To perform hypothesis testing to validate assumptions and support strategic decisions.
- To enhance data readiness for statistical analysis through log transformation and normalization.

3. Data Description

- Source: www.kaggle.com
- Size: 200,000 records
- Key Variables:
  1. Click_Through_Rate (CTR)
  2. Cost_Per_Acquisition (CPA)
  3. Age_Group
  4. Channel_Name
  5. ROI
  6. Conversion_Rate

4. Scope of Work

In Scope:

- Exploratory Data Analysis (EDA)
- Visualization of distribution patterns (histograms, KDE plots)
- Log transformation of right-skewed variables
- Shapiro-Wilk Normality Test
- Application of the Empirical Rule
- Z-score based Outlier Detection

- Confidence Interval Estimation (95% and 99%)
- One-sample T-tests on log-transformed CTR and CPA

5. Tools & Technologies

- Python (Pandas, NumPy, SciPy, Seaborn, Matplotlib)
- Jupyter Notebook
- Statistical principles (Normality, Confidence Intervals, T-tests)

6. Expected Deliverables

- Cleaned and pre-processed dataset with transformations
- Visual distribution reports and outlier analysis
- Confidence intervals for CTR and CPA
- Summary report with hypothesis testing conclusions

7. Business Impact

- Equip marketing teams with a data-backed understanding of campaign performance.
- Detect anomalies in CPA that could signal campaign inefficiencies or outliers.
- Validate performance benchmarks using hypothesis testing and confidence intervals.
- Showcase analytical capabilities for data analyst portfolio visibility.

Data Analyst: Shijin Ramesh