# Applied Data Science with Python

## Project Title: Marketing Campaigns Analysis

## Project Overview

This project involves analyzing customer behavior and identifying patterns and trends to assist in optimizing marketing campaigns. The primary objective is to gain insights into the effectiveness of marketing strategies by exploring and understanding various customer and sales-related factors. The analysis is aligned with the concept of the Marketing Mix (4Ps: Product, Price, Place, Promotion) and leverages data science techniques to address the problem scenario.

## Problem Scenario

The Marketing Mix framework revolves around the 4Ps:  
- People: Characteristics like birth year, education, and income.  
- Product: Spending behavior on items like wine, fruits, gold, etc.  
- Place: Information on sales channels like websites and stores.  
- Promotion: Results from various campaigns and customer responses.

The objective is to use exploratory data analysis (EDA), hypothesis testing, and data visualization to uncover factors contributing to customer acquisition.

## Steps Undertaken

### 1. Data Cleaning and Preparation

- Imported the dataset and verified the integrity of variables such as Dt\_Customer and Income.  
- Addressed missing income data by imputing values based on customers with similar education and marital status.  
- Created new variables:  
 - Total Children: Derived from existing children-related columns.  
 - Age: Computed from birth-year.  
 - Total Spending: Aggregated the spending across categories like wine, fruits, gold, etc.

### 2. Exploratory Data Analysis (EDA)

- Generated box plots and histograms to analyze distributions and detect outliers.  
- Performed outlier treatment to ensure the data was ready for statistical analysis.  
- Created a heatmap to visualize correlations between variables, aiding in hypothesis testing and feature selection.

### 3. Feature Engineering

- Encoded categorical variables:  
 - Used ordinal encoding for ordered categories like education levels.  
 - Applied one-hot encoding for nominal variables like marital status along with Standardization technique for model training.

### 4. Hypothesis Testing

- Tested hypotheses related to customer behavior:  
 - Older customers may prefer in-store shopping over online channels.  
 - Customers with children may prefer online shopping due to time constraints.  
 - Sales through different channels might cannibalize in-store sales.  
 - Compared US performance against other regions in terms of total purchases.

### 5. Data Visualization and Insights

- Analyzed product performance in terms of revenue to identify best and least-performing categories.  
- Examined relationships between customer age and campaign acceptance rates.  
- Visualized customer distribution by country for campaign acceptance.  
- Investigated patterns between the number of children at home and total spending.  
- Studied the education background of customers who lodged complaints in the past two years.

## Key Takeaways

This project demonstrates the application of data science techniques in a real-world marketing context. By combining statistical analysis, EDA, and feature engineering, we gained actionable insights to improve marketing strategies and drive customer acquisition.

**Technologies Used:**  
- Python: Core language for analysis and modeling.  
- Pandas: Data cleaning, transformation, and management.  
- Matplotlib & Seaborn: Data visualization.  
- Scipy: Hypothesis testing.

**Author**: Ravi Ranjan Lal

**Email**: lalraviranjan@gmail.com