# **Trade Promo Analysis Project - Case Study**

## **Project Overview**

This project focuses on analyzing trade promotions using a variety of models and algorithms. The project involves processing raw data, running experiments, and generating insightful visualizations and reports.

## **Calculator for Decomposition of Sales.xlsx**

This excel sheet contains the calculations for decomposing a given sales value into various categories across all the PPG x Retailer variants.

## **TPO Case Study Presentation - Fazil & Upasana.ppt**

This ppt contains the various plots from the exploratory data analysis of all the models, followed by model performance and the interpretation of the models across all the PPG x Retailer variants.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

### **Data**

The Data directory contains the source data for the project. It is essential to note that this data is in its original, unprocessed form as received initially.

### **Final Experiment**

The Final\_Experiment directory represents the experiment run. It contains various subfolders (outputs) generated by the model during the experiment. The contents of these subfolders will be explained in detail below.

### **Project\_Notebook\_Main.ipynb**

The Project\_Notebook\_Main.ipynb is the main project notebook. Ensure that the suitable virtual environment (in this case, we use ta-lib-dev) is activated before running the notebook. There are three variables to be configured in the second cell:

* promo\_path: The path where the original provided data resides.
* scan\_path: The path where the original provided data resides.
* experiment\_folder: The name of the experiment. All outputs generated during this run will be saved inside this folder.

Once these variables are set, the notebook will run smoothly end-to-end.

Upon successful execution, the notebook generates the following folders:

* **Raw\_Data**: Contains 12 unprocessed datasets, each representing the product promotion group (ppg) by retailer.
* **Processed\_Data**: Holds the processed versions of the 12 datasets mentioned above.
* **EDA\_Plots**: Consists of graphs illustrating the exploratory data analysis segment.
* **Metrics\_Weights**: Contains files for 84 model variants, including weights, bias, train\_r2, and test\_r2 for each of the 12 datasets.
* **Contribution**: Provides a data-wise sales breakup for all 84 model variants.
* **Contribution\_Percent**: Expresses the raw contribution numbers in terms of percentages, with 12 CSV files representing contribution percent information for the 7 models run.
* **Final\_Report**: Includes the following files:
  + **All\_Models\_Performance.csv**: Offers the overall performance of all 84 models.
  + **All\_Models\_Relative\_Contribution.csv**: Provides insight into contributions among promotional activity features alone.
  + **Best\_Performing\_Models.csv**: Filters out the best-performing models for each dataset.
  + **Best\_Models\_Relative\_Contribution.csv**: Highlights the relative contribution among promotional activities for the filtered best-performing models.
  + **Decomposition\_of\_Sales.csv**: Shows how 100% of sales are attributed to various explanatory features.
  + **Decomposition\_of\_Sales.png**: A visual representation of the sales decomposition.

This README serves as a guide to understanding the structure of the project and the purpose of each directory and file. Feel free to explore and analyze the generated outputs for deeper insights into the trade promotion analysis.