**Project Title: Retail Price Optimization Using Machine Learning**

**Link:** <https://www.kaggle.com/datasets/bhanupratapbiswas/retail-price-optimization-case-study>

**Problem & Dataset description:** In the competitive world of retail, setting optimal prices for products is crucial for attracting customers while maximizing profits. Retailers need to dynamically adjust prices based on various factors such as competition, demand, and market trends. However, manual price optimization can be time-consuming and error-prone. This project aims to leverage machine learning techniques to develop a price optimization solution that automates the process and improves overall profitability. The dataset from Kaggle provides valuable insights into the retail domain. It consists of historical sales and pricing data for various products. The dataset contains approximately 678 rows and 29 columns, including product attributes, prices, and sales volume. This data will serve as the foundation for training and evaluating the price optimization model.

**Aim:** The primary goal of this project is to develop a machine-learning model that can predict optimal prices for products in a retail environment.

* To analyze the data with spark processing and identify the important features and visualize the findings in Tableau
* To create a machine learning model for price optimization with the best accuracy
* To Evaluate the developed model using evaluation metrics after hyper-parameter tuning

**Tasks and Workflow:**

**Week 1: Data Preprocessing and Exploration**

Explore the dataset statistically and visually to understand feature distributions and correlation and create a dashboard in the tableau by creating multiple smaller charts

**Week 2: Model Development and Training**

Feature selection to Identify the most relevant features and choose the appropriate machine learning algorithms followed by Tune hyperparameters using techniques like grid search or random search to optimize model performance.

**Week 3: Model Evaluation and Document Preparation**

Evaluate model performance using the test dataset and the by Comparing the performance of different models to select the best-performing model and finally document the project