## **Clothing Price Prediction**

**Problem Statement:**

Effective pricing plays a crucial role in determining the success of a product, particularly for clothing items, where numerous factors like brand value, style trends, color, material and consumer behavior influence the final price. The challenge lies in developing an AI model capable of accurately predicting clothing prices, enabling retailers to optimize pricing strategies and enhance overall profitability. The primary objective is to create a predictive model that can analyze various features of a garment and provide an estimate of its price with high

accuracy.

**Dataset Selected:**

It includes various features related to clothing items along with their corresponding prices. This dataset is intended for use in regression analysis, machine learning model training, and statistical modeling aimed at clothes price prediction based on their attributes. Researchers, data scientists, and analysts can explore patterns and relationships between clothing attributes and prices to build predictive models for estimating the price of clothes based on their characteristics.

**Dataset Description:**

The dataset comprises information about apparel items across various categories, brands, colors, sizes, materials, and their corresponding prices. It consists of 1000 entries with six columns: Brand, Category, Color, Size, Material, and Price. Each entry represents a unique apparel item, where 'Brand' denotes the manufacturer or brand name, 'Category' signifies the type of apparel (e.g., dress, jeans, shoes), 'Color' represents the color of the item, 'Size' indicates the available sizes (e.g., XS, M, L), 'Material' specifies the fabric or material used in manufacturing, and 'Price' denotes the cost of the item in the local currency. The dataset is rich in categorical features, making it suitable for exploring patterns and relationships across different apparel attributes.

The dataset has undergone several preprocessing steps, including handling missing values and duplicate entries. There are no missing values in any of the columns, ensuring completeness and consistency in the data. Additionally, duplicate entries have been removed, resulting in a clean dataset with 1000 unique apparel items. Furthermore, one-hot encoding has been applied to categorical variables like 'Category,' 'Color,' 'Size,' 'Material,' and 'Brand' to convert them into numerical format, enabling machine learning algorithms to process the data effectively. This processed dataset is ready for analysis, modeling, and prediction tasks related to apparel pricing and market trends.