

# Product Pricing Challenge - Amazon ML Challenge 2025

## Problem Statement

In e-commerce, determining the optimal price for products is essential for business success and customer satisfaction. The goal of this challenge is to develop a machine learning model that predicts the price of a product based on its attributes such as title, description, specifications, and quantity.

Product pricing is influenced by various factors including brand, specifications, and packaging details. This challenge aims to analyze these complex relationships and build an ML solution capable of recommending accurate product prices.

## Data Description

Column Name	Description
sample_id	Unique identifier for each sample
catalog_content	Concatenated text field containing product title, description, and Item Pack Quantity (IPQ)
image_link	Public URL for the product image
price	Target variable (only available in the training set)

## Dataset Split

- Training Set: 75,000 products with labeled prices
- Test Set: 75,000 products for evaluation

## Evaluation Metric

Submissions are evaluated using **Symmetric Mean Absolute Percentage Error (SMAPE)**.

### Formula:

$$\text{SMAPE} = (1/n) * \sum | \text{Predicted} - \text{Actual} | / ((| \text{Predicted} | + | \text{Actual} |) / 2)$$

Lower SMAPE indicates better performance and it is bounded between 0% (perfect) and 200% (worst).