

## Case Study: AeroFit Treadmill Customer Profiling

### 1. Executive Summary

This project focuses on analyzing customer data for AeroFit, a manufacturer of treadmills. The primary goal is to understand the characteristics of customers who purchase different models of AeroFit treadmills. By performing descriptive analytics and probability analysis, we aim to create detailed customer profiles for each product. These profiles will enable AeroFit to enhance its marketing strategies, improve product recommendations for new customers, and ultimately align its product offerings with specific market segments.

### 2. Business Problem

The market research team at AeroFit wants to identify the characteristics of the target audience for each type of treadmill offered by the company. This is to provide better, more personalized recommendations to new customers. The team has decided to investigate whether there are significant differences across the product line with respect to customer characteristics. The core task is to perform a descriptive analysis to create a distinct customer profile for each AeroFit treadmill product.

### 3. Product Portfolio

AeroFit's product line consists of three distinct treadmill models, catering to different fitness levels and budgets:

- **The KP281:** An entry-level treadmill that sells for **\$1,500**.
- **The KP481:** A mid-level treadmill designed for intermediate runners, which sells for **\$1,750**.
- **The KP781:** A heavy-duty treadmill with advanced features, which sells for **\$2,500**.

### 4. Dataset Description

The analysis is based on a dataset collected from individuals who purchased an AeroFit treadmill from a retail store during the prior three months.

**Dataset Source:** [Aerofit treadmill.csv](#)

The dataset contains the following features:

Feature	Description
Product	The model of the treadmill purchased (KP281, KP481, or KP781).
Age	Age of the customer in years.
Gender	Gender of the customer (Male/Female).

Education      Education level of the customer in years.

MaritalStatus      Marital status of the customer (Single/Partnered).

Usage      The average number of times the customer plans to use the treadmill each week.

Income      Annual income of the customer (in \$).

Fitness      Self-rated fitness level on a 1-to-5 scale (1: poor shape, 5: excellent shape).

Miles      The average number of miles the customer expects to walk/run each week.

## 5. Proposed Plan of Analysis

The analysis will be conducted in several stages to ensure a comprehensive understanding of the customer data.

- **Data Exploration and Cleaning:**
  - Import the dataset and examine its structure, characteristics, and data types.
  - Check for missing values and inconsistencies.
  - Detect potential outliers using descriptive statistics (.describe()) and boxplots.
- **Descriptive Analytics & Customer Profiling:**
  - Analyze the distribution of individual customer attributes (Age, Gender, Income, etc.).
  - Create customer profiles for each treadmill model by analyzing the characteristics of buyers for each product using visualizations like histograms, count plots, and boxplots.
- **Bivariate and Probability Analysis:**
  - Construct **two-way contingency tables** (using pandas.crosstab) to explore the relationship between categorical variables (e.g., Product vs. Gender, Product vs. MaritalStatus).
  - Calculate **marginal probabilities** to understand the overall distribution of product purchases (e.g.,  $P(KP281)$ ,  $P(KP481)$ ,  $P(KP781)$ ).
  - Calculate **conditional probabilities** to answer specific business questions, such as:
    - *What is the probability that a customer will purchase a KP781, given they are male?* (i.e.,  $P(KP781 \mid \text{Male})$ )

- *What is the probability that a customer has a high income, given they purchased a KP481?*

- **Correlation Analysis:**

- Generate a heatmap of the correlation matrix or a pair plot to visualize the relationships between numerical features (Age, Income, Usage, Miles, etc.). This will help identify which factors are strongly associated with each other.

## **6. Expected Outcomes & Business Impact**

The successful completion of this analysis will yield the following outcomes:

1. **Clear Customer Profiles:** A detailed profile for the typical buyer of each treadmill model (KP281, KP481, and KP781), outlining their demographic, fitness, and income characteristics.
2. **Data-Driven Insights:** Actionable insights derived from probability analysis that quantify the likelihood of a purchase based on customer attributes.
3. **Strategic Recommendations:** Based on the findings, we will provide recommendations for:
  - **Targeted Marketing:** Tailoring advertising campaigns to the specific profiles of customers most likely to buy each model.
  - **Sales Strategy:** Equipping the sales team with insights to better guide customers to the right product.
  - **Product Development:** Informing future product enhancements or new product designs based on the needs of different customer segments.