

EDA Portfolio Project Report AeroFit treadmills

This project analyzes customer purchasing trends for AeroFit treadmills using Exploratory Data Analysis (EDA). The objective is to identify the characteristics of the target audience for each treadmill model and provide recommendations to improve product marketing and sales. The analysis explores key demographic attributes such as age, income, fitness level, and marital status. Findings reveal correlations between product preference and various customer traits, highlighting insights that can enhance sales strategies.

Prepared by
Muhammad TAHA Zaman

Date
March 9, 2025

Prepared for
AeroFit treadmills



Table of Contents:

Abstract.....	02
Introduction	03
Methodology:	03
Key Findings and Insights	04
Actionable Recommendations:	07
Conclusion.....	08
References:	08

Abstract:

This project analyzes customer purchasing trends for AeroFit treadmills using **Exploratory Data Analysis (EDA)**. The objective is to identify the characteristics of the target audience for each treadmill model and provide recommendations to improve product marketing and sales. The analysis explores key demographic attributes such as age, income, fitness level, and marital status. Findings reveal correlations between product preference and various customer traits, highlighting insights that can enhance sales strategies.

Introduction:

This report investigates customer demographics and purchasing behavior at **AeroFit**. The company offers three treadmill models:

- **KP281**: Entry-level treadmill (\$1,500)
- **KP481**: Mid-level treadmill (\$1,750)
- **KP781**: High-end treadmill (\$2,500)

The primary goal is to analyze **customer attributes and trends** to determine how factors such as **age, income, fitness level, and marital status** influence treadmill selection. This EDA-based study will provide actionable insights for marketing and sales improvements.

The analysis is structured into the following sections:

1. **Data Exploration & Preprocessing**
2. **Statistical & Non-Graphical Analysis**
3. **Graphical Analysis (Univariate, Bivariate, and Multivariate)**
4. **Correlation & Outlier Detection**
5. **Conditional Probability Analysis**
6. **Actionable Insights & Recommendations**

Methodology:

Tools Used:

- **Programming Language**: Python
- **Libraries**: Pandas, NumPy, Seaborn, Matplotlib, SciPy
- **Dataset**: AeroFit Treadmill Sales Data

Steps Involved:

1. **Data Exploration & Preprocessing:**
 - Loaded the dataset and checked for missing values & duplicates.
 - Examined the data structure, column types, and basic statistics.
2. **Statistical & Non-Graphical Analysis:**
 - Analyzed numerical and categorical variables.
 - Computed summary statistics for age, income, fitness level, and treadmill usage.
3. **Graphical Analysis:**
 - **Univariate Analysis**: Histograms, boxplots, and count plots for numerical and categorical data.
 - **Bivariate Analysis**: Analyzed relationships between product purchase and customer attributes (Gender, Age, Marital Status, and Income).
 - **Multivariate Analysis**: Created pair plots to visualize feature relationships.

4. Correlation & Outlier Detection:

- Generated a correlation heatmap to understand feature relationships.
- Identified outliers using the IQR method.

5. Conditional Probability Analysis:

- Calculated probability distributions for product selection based on gender, income, age, fitness level, and marital status.

6. Insights & Recommendations:

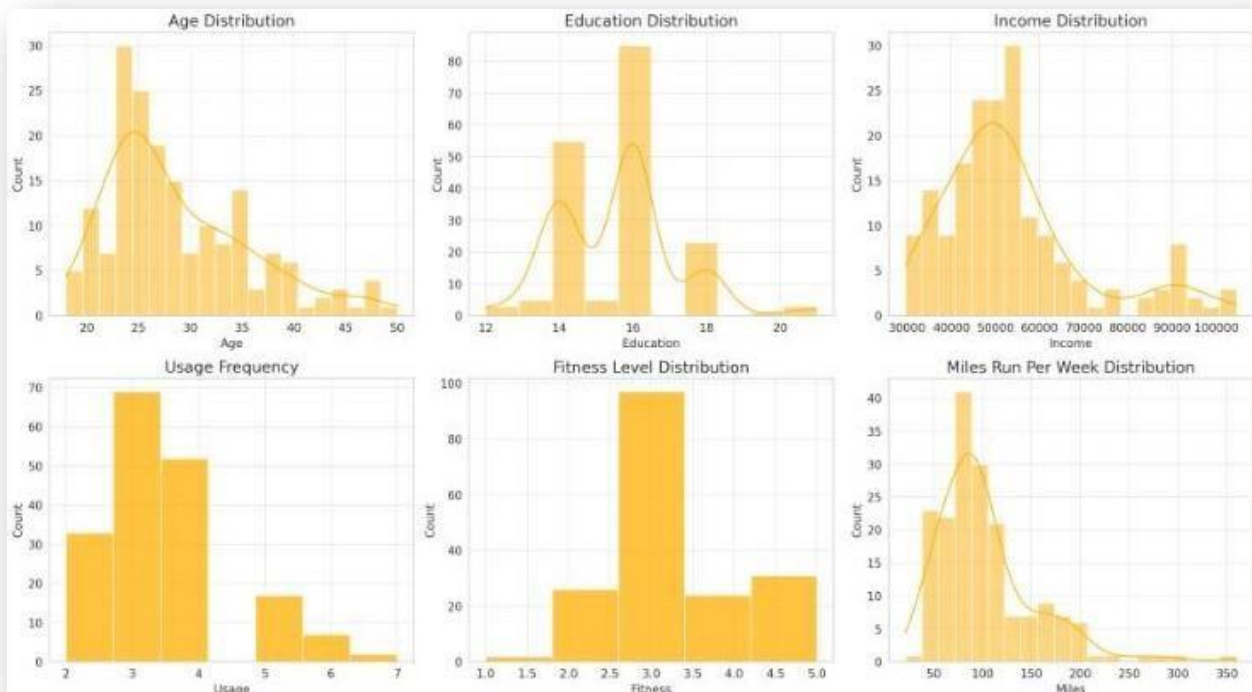
- Summarized findings and suggested marketing strategies based on data trends.

Key Findings & Insights:

1. Customer Demographics & Buying Trends:

- Most buyers fall within the **20s to 30s age range**.
- **KP281 has the highest sales**, while **KP781 lags in demand**.
- High-income customers (earning \$70,000+) do not necessarily buy premium models like **KP781**.

Data Visualization:

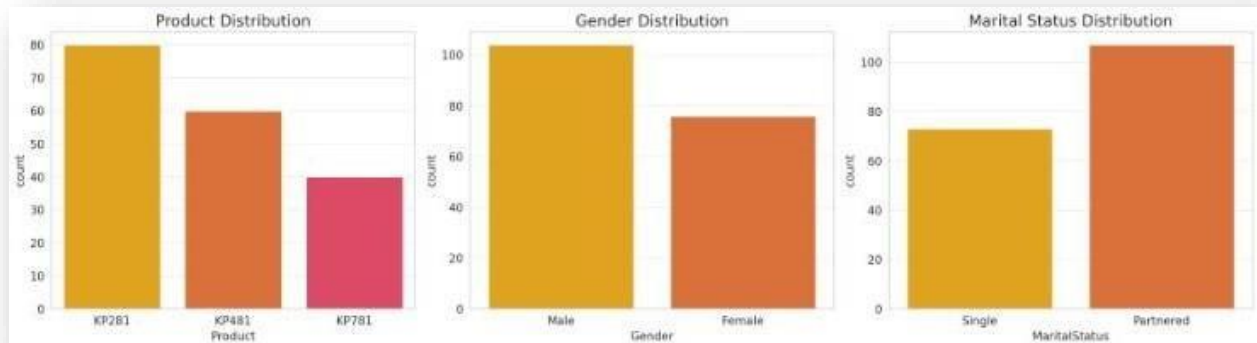


Customer Demographics & Buying Trends

2. Product Preferences & Correlation Analysis:

- **Age vs. Product Preference:** Younger customers tend to buy more budget-friendly models.
- **Income vs. Purchase Behavior:** No strong correlation between income and treadmill price selection.
- **Fitness Level vs. KP781 Sales:** **Highly fit customers prefer KP781**, but the overall sales remain low.

Data Visualization:

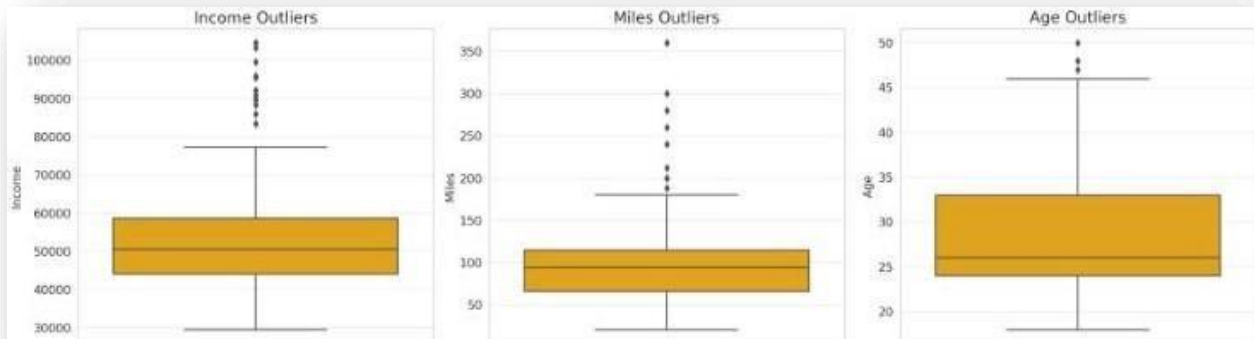


Product Preferences & Correlation Analysis

3. Outlier Detection & Data Quality:

- Income and mileage had **significant outliers**, suggesting either **niche customers** or **data entry errors**.
- Some **high-salary customers** still opted for entry-level treadmills.

Data Visualization:

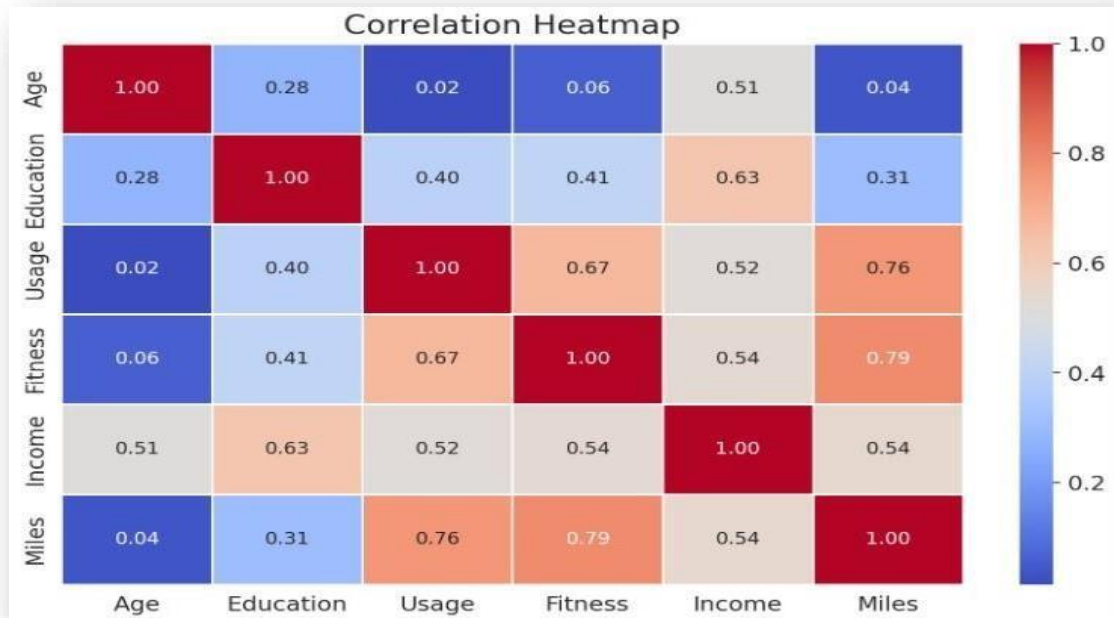


Outlier Detection & Data Quality

4. Gender & Marital Status Trends:

- Males purchase treadmills more frequently than females.
- Partnered customers make up a majority of treadmill buyers.

Data Visualization:



Gender & Marital Status Trends

Actionable Recommendations:

1. Targeted Marketing Strategies:

- Focus advertising on **young buyers (20s-30s)** with fitness-related messaging.
- Emphasize KP781's **premium features** to high-income customers.
- Increase engagement among **female buyers** through specialized promotions.

2. Sales Promotions & Pricing Adjustments:

- Introduce **discounts, EMI plans, and bundle deals** for KP281, KP481, and KP781.
- Offer **premium incentives** (e.g., **VIP fitness plans**) to high-income customers to boost KP781 sales.

3. Data Quality & Customer Segmentation:

- Perform **data cleaning** to remove incorrect outliers.
- Use **AI-based recommendations** to provide personalized product suggestions.

4. Product Positioning & Branding:

- Improve **KP781's market appeal** by partnering with **influencers & fitness trainers**.
- Introduce **limited-time offers & loyalty programs** to encourage repeat purchases.

Conclusion:

This project provided an in-depth analysis of AeroFit's treadmill sales data using **EDA techniques**. The study identified trends in customer demographics, income levels, and fitness preferences, leading to actionable recommendations. By **refining marketing strategies, improving product positioning, and optimizing data quality**, AeroFit can enhance customer engagement and drive higher sales. Future analysis can incorporate **predictive modeling** to better forecast customer purchasing behavior.

References:

- AeroFit Treadmill Sales Dataset (aerofit_treadmill_data.xlsx)
- Python Libraries: Pandas, NumPy, Seaborn, Matplotlib
- Exploratory Data Analysis Techniques
- Data Source: AeroFit customer purchase records analyzed through EDA techniques.
- Python Libraries: Pandas, NumPy, Seaborn, Matplotlib
- Exploratory Data Analysis Techniques