Rian Qadir

Aerofit Treadmill Buyer Profile and Data Analysis Report

Table of Contents

[1 Overview 3](#_Toc180194321)

[2 Project Insights 3](#_Toc180194322)

[2.1 Product Overview 3](#_Toc180194323)

[3 Data Information 3](#_Toc180194324)

[4 Exploratory Data Analysis (EDA) 4](#_Toc180194325)

[4.1 Data Import and Overview 4](#_Toc180194326)

[4.2 Statistical Summary 5](#_Toc180194327)

[4.2.1 Categorical Features: 5](#_Toc180194328)

[4.2.2 Numerical Features: 5](#_Toc180194329)

[4.3 Analysis Without Visual Aids 6](#_Toc180194330)

[4.3.1 Value Counts: 6](#_Toc180194331)

[4.3.2 Unique Attributes: 6](#_Toc180194332)

[4.4 Visualization-Based Analysis 7](#_Toc180194333)

[4.4.1 Univariate Analysis - Numerical Features 7](#_Toc180194334)

[4.4.2 Univariate Analysis - Categorical Features 9](#_Toc180194335)

[4.4.3 Bivariate Analysis 10](#_Toc180194336)

[4.4.4 Multivariate Analysis 13](#_Toc180194337)

[4.5 Correlation Analysis 14](#_Toc180194338)

[4.5.1 Graph 14](#_Toc180194339)

[4.5.2 Observations From the Heatmap 15](#_Toc180194340)

[4.6 Outlier Detection 16](#_Toc180194341)

[4.7 Conditional Probabilities 16](#_Toc180194342)

[4.7.1 Product Purchases: 16](#_Toc180194343)

[4.7.2 Product – Gender: 16](#_Toc180194344)

[4.7.3 Product – Age: 18](#_Toc180194345)

[4.7.4 Product – Income: 18](#_Toc180194346)

[4.7.5 Product – Fitness: 19](#_Toc180194347)

[4.7.6 Product - Marital Status: 20](#_Toc180194348)

[5 Customer Demographics 21](#_Toc180194349)

[5.1 Gender Distribution 21](#_Toc180194350)

[5.2 Age Distribution 21](#_Toc180194351)

[5.3 Income Analysis 21](#_Toc180194352)

[5.3.1 Low-Income Customer Insights: 21](#_Toc180194353)

[5.3.2 High-Income Customer Insights: 21](#_Toc180194354)

[5.3.3 High-Income customer who Purchased KP781 21](#_Toc180194355)

[5.4 Fitness Level Insights 21](#_Toc180194356)

[5.4.1 Fitness Level 5 Analysis: 21](#_Toc180194357)

[5.5 Marital Status Analysis 21](#_Toc180194358)

[5.5.1 Partnered Customers: 21](#_Toc180194359)

[6 Insights and Recommendations 22](#_Toc180194360)

[6.1 Target Audience 22](#_Toc180194361)

[6.1.1 Insight: Males, especially those who are married or partnered, make up a large segment of the customer base. 22](#_Toc180194362)

[6.1.2 Recommendation: Marketing efforts should focus on these demographics, with campaigns designed to resonate with their preferences. 22](#_Toc180194363)

[6.2 Product Portfolio 22](#_Toc180194364)

[6.2.1 Insight: 22](#_Toc180194365)

[6.2.2 Recommendation: 22](#_Toc180194366)

[Continue promoting these models and consider developing similar treadmills with enhanced features. 22](#_Toc180194367)

[6.3 Pricing Strategy 22](#_Toc180194368)

[6.3.1 Insight: 22](#_Toc180194369)

[6.3.2 Recommendation: 22](#_Toc180194370)

[6.4 Marketing Strategies 22](#_Toc180194371)

[6.4.1 Insight: 22](#_Toc180194372)

[6.4.2 Recommendation: 22](#_Toc180194373)

[6.5 Product Improvements 23](#_Toc180194374)

[6.5.1 Insight: 23](#_Toc180194375)

[6.5.2 Recommendation: 23](#_Toc180194376)

[6.6 Geographic Targeting 23](#_Toc180194377)

[6.6.1 Insight: 23](#_Toc180194378)

[6.6.2 Recommendation: 23](#_Toc180194379)

[6.7 Data-Driven Decisions 23](#_Toc180194380)

[6.7.1 Insight: 23](#_Toc180194381)

[6.7.2 Recommendation: 23](#_Toc180194382)

[7 Conclusion 23](#_Toc180194383)

# Overview

This document provides an in-depth analysis of the purchasing trends, customer profiles, and product performance for Aerofit treadmills. The core aim is to extract actionable insights that will support the refinement of marketing initiatives, product evolution, and customer interaction efforts. By evaluating the target customer base for each treadmill model, Aerofit seeks to tailor recommendations for future buyers.

# Project Insights

Aerofit’s research team has set out to examine the demographic characteristics and preferences of customers who have bought treadmills from their collection. These insights will assist in aligning marketing and product strategies with customer needs.

## Product Overview

* **KP281**: Basic treadmill, retailing at $1,500.
* **KP481**: Mid-range treadmill aimed at runners, priced at $1,750.
* **KP781**: Premium treadmill equipped with high-end features, retailing at $2,500.

# Data Information

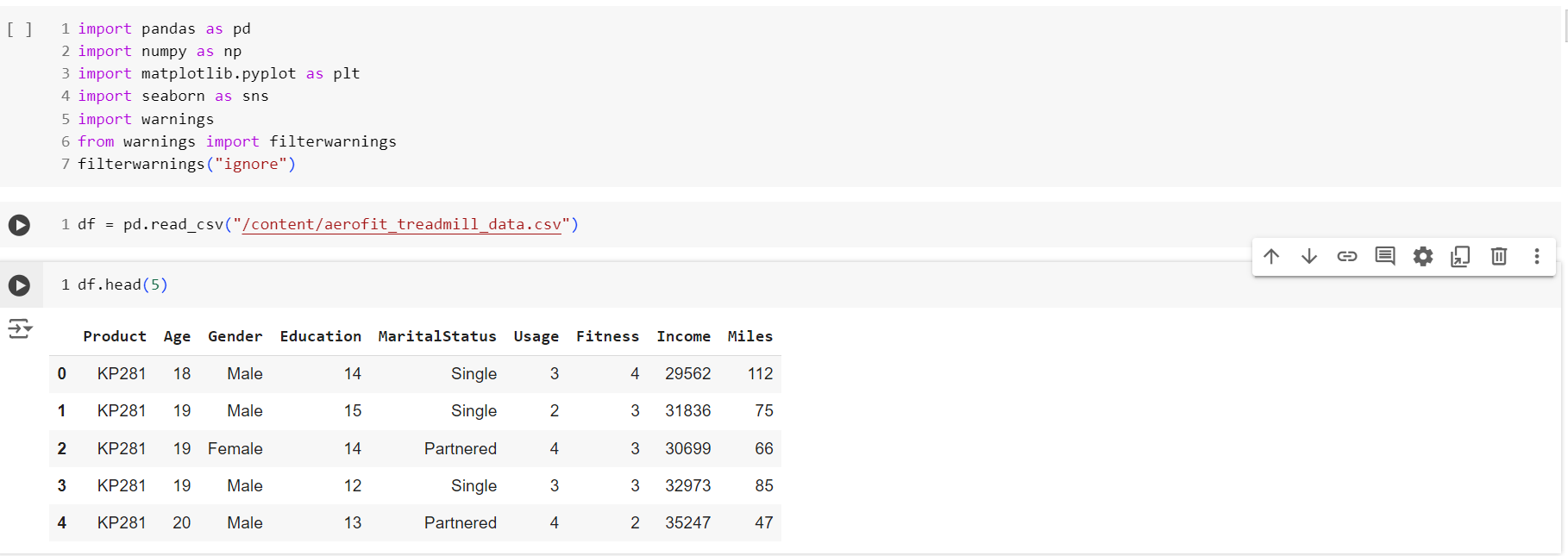
The data, sourced from the last quarter’s sales records, includes customer demographics and treadmill preferences. Key features of the dataset:

* **Product**: Model purchased (KP281, KP481, KP781).
* **Age**: Age of the buyer.
* **Gender**: Male or female customer.
* **Education**: Number of years of formal education.
* **Marital Status**: Either single or partnered.
* **Usage**: Expected weekly usage of the treadmill.
* **Fitness**: Self-reported fitness level (scale of 1-5).
* **Income**: Annual income in USD.
* **Miles**: Anticipated weekly distance (miles) covered on the treadmill.

# Exploratory Data Analysis (EDA)

## Data Import and Overview

The data was imported using the panda's library, and an initial review of its shape and types was conducted. Missing data checks and duplicate scans were also performed.



* The shape of the data frame is checked to understand the number of rows and columns.

A close-up of a computer screen

Description automatically generated

* Data types of each column are verified to ensure proper analysis.

A white background with black text

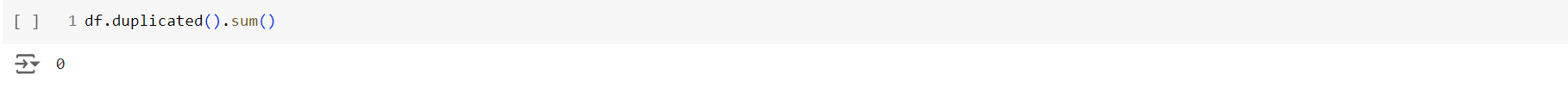
Description automatically generated

* A check for missing values is performed to assess data quality.

A white rectangular object with a black border

Description automatically generated

* The dataset is scanned for duplicate entries.



## Statistical Summary

A statistical summary, covering both qualitative and quantitative features, was compiled.

### Categorical Features:

This section presents the distribution of variables like gender, marital status, and product choice.

A white background with a black border

Description automatically generated with medium confidence

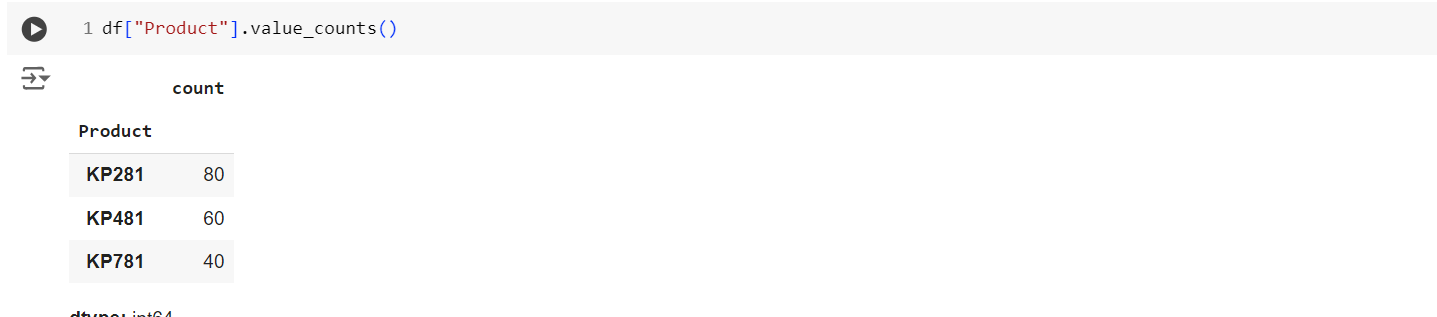
### Numerical Features:

Key statistics (mean, median, standard deviation) were computed for numerical features such as age, income, fitness, and usage.A group of numbers on a white background

Description automatically generated

## Analysis Without Visual Aids

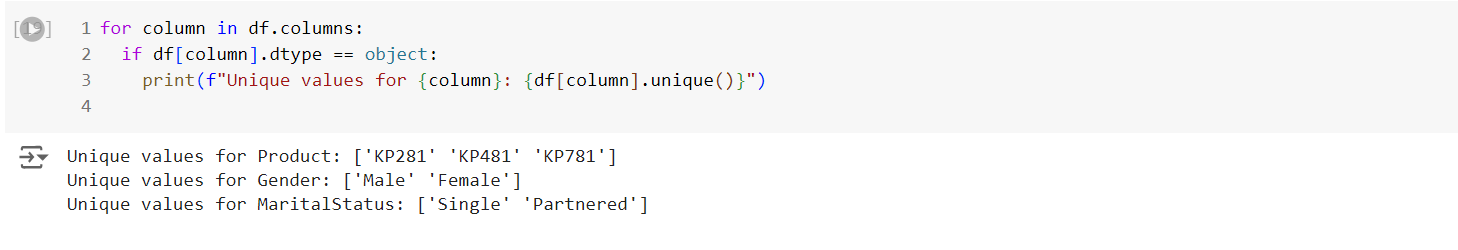
### Value Counts:

The frequency distribution of each categorical variable is outlined.

A screenshot of a computer

Description automatically generated

### Unique Attributes:

A listing of unique values for each categorical variable.

## Visualization-Based Analysis

### Univariate Analysis - Numerical Features

**Distribution Plots:** Represent the spread of numerical features like age, income, and usage.A white screen with colorful text

Description automatically generated

A group of graphs showing different types of data

Description automatically generated

A white screen with text

Description automatically generated with medium confidence

A group of graphs showing different sizes of bars

Description automatically generated with medium confidence

* **Box Plot**: Box plots to identify the spread and potential outliers in numerical features.

A computer code on a white background

Description automatically generated

A diagram of a graph

Description automatically generated with medium confidence

### Univariate Analysis - Categorical Features

* **Count Plot**: Count plots to visualize the distribution of categorical features like gender and marital status.



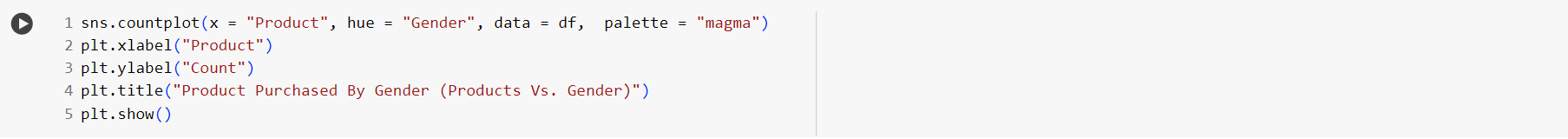
A graph of different sizes and colors

Description automatically generated with medium confidence

### Bivariate Analysis

This analysis focuses on comparing two variables, for example:

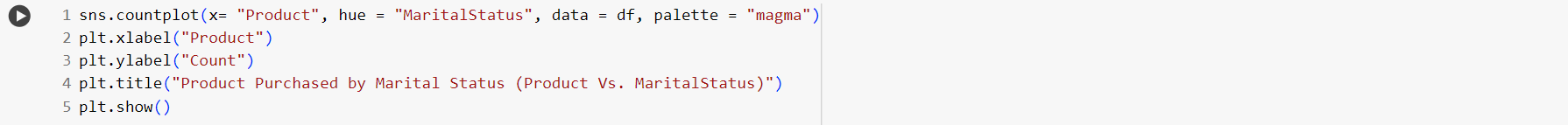
* **Product Choice by Gender**: Male and female customer trends across different treadmill models.



A graph of different colored bars

Description automatically generated

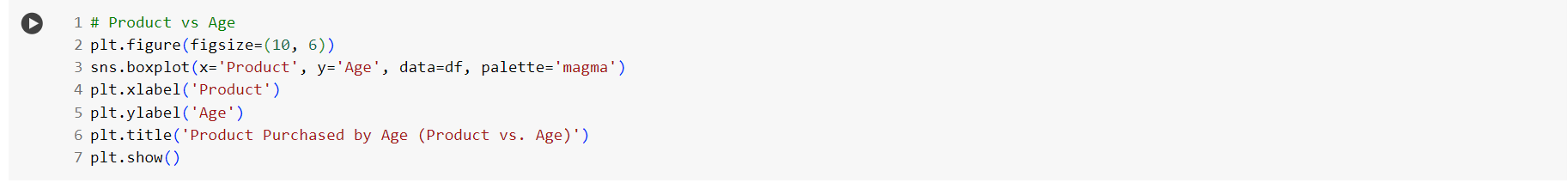
* Examining how marital status influences product choice.



A graph of a product purchased by martial status

Description automatically generated

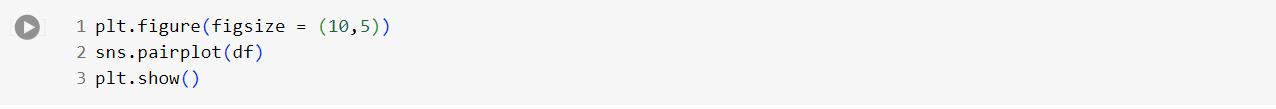
* **Product vs. Age**: Exploring the relationship between product choice and customer age.



A chart with different colored squares

Description automatically generated

### Multivariate Analysis

* **Pair Plots: These depict the interaction between several features at once.**

A screenshot of a graph

Description automatically generated

## Correlation Analysis

### Graph

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

### Observations From the Heatmap

#### Strong Positive Correlation:

* 'Miles' and 'Usage' have a strong positive correlation (0.76), which is expected as people who plan to use the treadmill more often are likely to cover more miles.
* 'Income' and 'Miles' have a moderate positive correlation (0.54), indicating that individuals with higher incomes tend to cover more miles.
* 'Income' and 'Usage' show a moderate positive correlation (0.52), suggesting that those with higher incomes tend to use the treadmill more frequently.
* 'Age' and 'Fitness' have a moderate positive correlation (0.61), which implies that older individuals tend to rate their fitness higher.
* 'Education' and 'Income' show a moderate positive correlation (0.63), which implies that users with higher education have higher income.

#### Moderate Correlation:

* 'Age' and 'Income' have a moderate positive correlation (0.51), suggesting a potential trend where older individuals tend to have higher incomes.
* 'Education' and 'Age' have a moderate positive correlation (0.63), showing that higher education individuals may be older in age.

#### Weak or No Correlation:

* 'Fitness' and 'Miles' have a very weak correlation (0.06), indicating that self-rated fitness doesn't strongly influence how many miles users plan to cover.

#### Other Observations

* Based on the correlation matrix, it's evident that 'Usage' and 'Miles' are strongly positively correlated, and 'Income' is moderately correlated with both of them.
* This indicates that users who plan to use the treadmill more and cover more miles tend to have higher incomes.
* Additionally, 'Age' and 'Fitness' are moderately positively correlated, suggesting that older individuals may rate their fitness higher.

## Outlier Detection

Outliers were flagged using the interquartile range (IQR) method, helping identify data points that could skew the results.

## Conditional Probabilities

This section includes calculations for probabilities related to treadmill purchases based on factors such as gender, age, income, and fitness level.

### Product Purchases:

The percentage of customers purchasing KP281, KP481, or KP781 is calculated.

A white rectangular object with black text

Description automatically generated

### Product – Gender:

* Percentage of male customers purchasing a treadmill.

A close-up of a computer screen

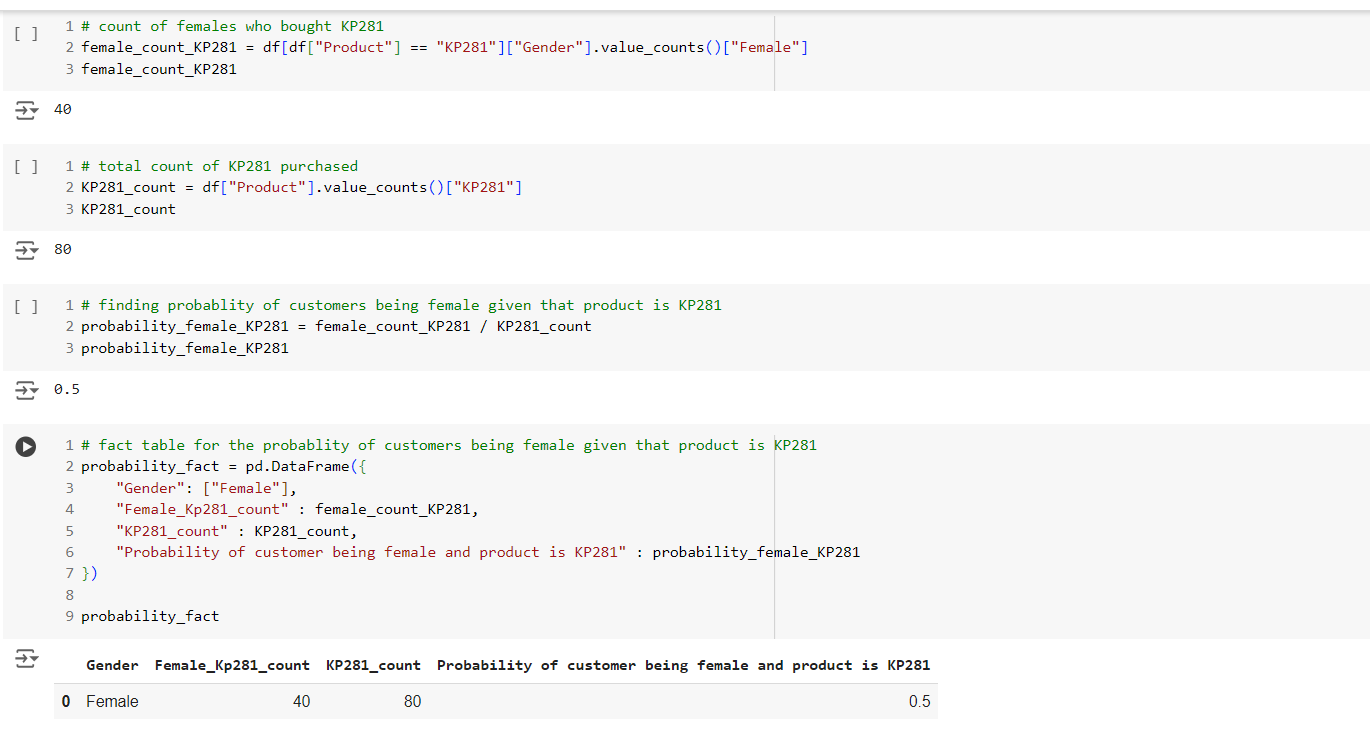
Description automatically generated

* Percentage of female customers purchasing KP781.

A screenshot of a computer

Description automatically generated

* Probability of a customer being female given that the product is KP281.



### Product – Age:

* + Percentage of customers aged between 20 and 30 among all customers.

A screenshot of a computer

Description automatically generated

### Product – Income:

* + Percentage of low-income customers purchasing a treadmill.

A screenshot of a computer

Description automatically generated

* + Percentage of high-income customers purchasing KP781.

A screenshot of a computer

Description automatically generated

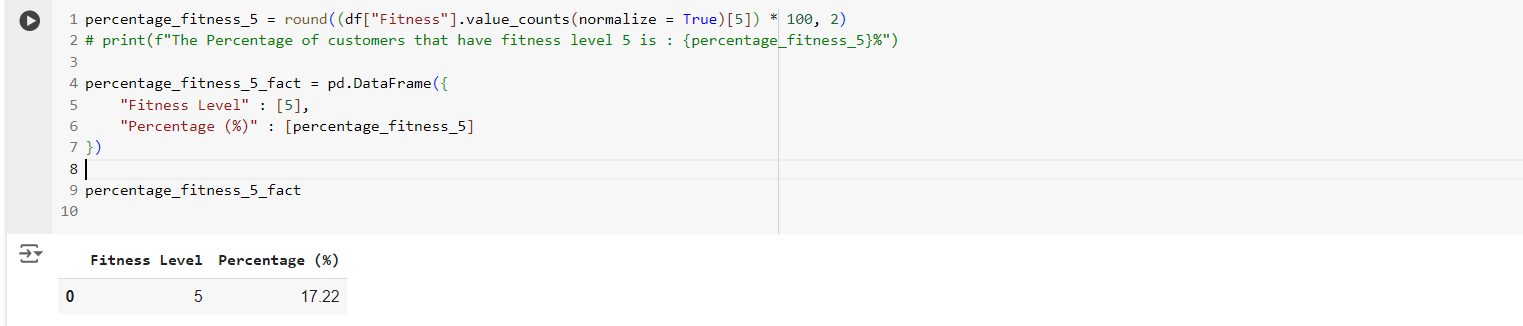
* + Percentage of high-income customers buying a treadmill given that the product is KP781.

A screenshot of a computer code

Description automatically generated

### Product – Fitness:

* + Percentage of customers with a fitness level of 5.



* + Percentage of fitness level 5 customers purchasing KP781.

A white background with colorful text

Description automatically generated

* + Percentage of customer with fitness level 5 buying KP781 treadmill

A computer screen with text

Description automatically generated

### Product - Marital Status:

* + Percentage of partnered customers using treadmills.

A screenshot of a computer

Description automatically generated

# Customer Demographics

## Gender Distribution

* Total Females Who Bought KP781: **7**
* Percentage of Females Buying KP781: **17.5%**
* Total Females Who Bought KP281: **40**
* Total Purchases of KP281: **80**
* Probability of Customer Being Female for KP281: **0.5**

## Age Distribution

* Customers Aged 20s to 30s: **67**
* Percentage of Customers Aged 20s to 30s: **37.22%**

## Income Analysis

### Low-Income Customer Insights:

* Low-Income Customers (Income ≤ $50,000): **83**
* Total Customers: **180**
* Percentage of Low-Income Customers Purchasing a Treadmill: **46.11%**

### High-Income Customer Insights:

* High-Income Customers (Income >= $80,000) : **19**
* Total KP781 Customers : **40**
* Percentage of High-income Customers Purchasing a Treadmill KP781: **47.5**

### High-Income customer who Purchased KP781

* High-Income customer who bought KP781: **19**
* Total KP781 customers : **40**
* Percentage of High-income customers purchasing KP781 given that the product is KP781: **47.5**

## Fitness Level Insights

### Fitness Level 5 Analysis:

* Percentage of Customers with Fitness Level 5: **17.22%**
* Percentage of Fitness Level 5 Customers Buying KP781: **93.55%**
* Percentage of KP781 Customers with Fitness Level 5: **72.5%**

## Marital Status Analysis

### Partnered Customers:

* Total Partnered Customers: **107**
* Total Customers: **180**
* Percentage of Partnered Customers: **59.44%**

# Insights and Recommendations

Based on the comprehensive analysis of the dataset, the following actionable insights and recommendations are provided to improve product offerings, customer targeting, and business strategies for Aerofit treadmills:

## Target Audience

### Insight: Males, especially those who are married or partnered, make up a large segment of the customer base.

### Recommendation: Marketing efforts should focus on these demographics, with campaigns designed to resonate with their preferences.

## Product Portfolio

### Insight:

KP281 and KP481 are the most popular products, appealing to customers who prefer mid-range and entry-level treadmills.

### Recommendation:

### Continue promoting these models and consider developing similar treadmills with enhanced features.

## Pricing Strategy

### Insight:

The moderate correlation between income levels and treadmill usage indicates some price sensitivity among customers.

### Recommendation:

* Analyze the relationship between price points and sales volumes for each product to optimize pricing strategies.
* Explore options like tiered pricing or bundled offers to attract customers across different income levels while maximizing revenue.

## Marketing Strategies

### Insight:

High-income customers and fitness enthusiasts are likely to be drawn to treadmills with advanced features.

### Recommendation:

* Highlight the health and fitness benefits of owning a treadmill, particularly targeting those with higher incomes or greater interest in exercise.
* Use targeted advertising and promotions to reach potential customers based on demographics (age, gender, marital status) and interests (fitness, sports, health).
* Develop content marketing strategies that focus on providing valuable information about the benefits of regular exercise, fitness tips, and how treadmills can help achieve specific goals.

## Product Improvements

### Insight:

Customers with higher fitness levels tend to purchase the KP781 treadmill, suggesting a demand for advanced features.

### Recommendation:

* Research potential improvements to product features, taking into consideration factors like age and fitness levels.
* Analyze data from customers with a fitness level of 5 and those who purchase KP781 to assess the need for additional advanced features.
* Introduce features that cater to users' preferences, such as improved console interfaces, personalized workout programs, and advanced tracking capabilities.

## Geographic Targeting

### Insight:

Certain geographical areas may exhibit a higher demand for treadmills based on regional fitness trends and preferences.

### Recommendation:

* Consider conducting market research to understand the specific geographical areas where there is a high demand for treadmills.
* Tailor marketing campaigns and product availability based on regional preferences and needs.
* Consider partnering with fitness centers or gyms in specific locations to increase product awareness and sales.

## Data-Driven Decisions

### Insight:

Data analytics offer valuable insights into customer behavior and preferences.

### Recommendation:

* Continue utilizing data analysis to monitor sales trends, customer behaviors, and marketing campaign effectiveness.
* Use insights derived from the analysis to inform decision-making, optimize pricing, and refine marketing strategies.
* Invest in building a robust data analytics infrastructure to enable efficient data collection, analysis, and reporting.

# Conclusion

This analysis of Aerofit’s treadmill buyers provides valuable insights into customer demographics and product preferences. By applying these recommendations, Aerofit can refine its marketing approach, better meet customer needs, and ultimately drive growth..