|  |
| --- |
| Running on treadmill |
| Aerofit Treadmill Buyer Profile and Data Analysis Report |
| |  |  |  | | --- | --- | --- | | Khola Shams | 10/14/24 | Business Analysis / EDA | |

Table of Contents

[1 Introduction 1](#_Toc179819507)

[2 Project Details 1](#_Toc179819508)

[2.1 Product Portfolio 1](#_Toc179819509)

[3 Data Description 1](#_Toc179819510)

[4 Data Exploration and Processing 2](#_Toc179819511)

[4.1 Data Import and Overview 2](#_Toc179819512)

[4.2 Statistical Summary 3](#_Toc179819513)

[4.2.1 Categorical Features: 3](#_Toc179819514)

[4.2.2 Numerical Features: 4](#_Toc179819515)

[4.3 Non-Graphical Analysis 4](#_Toc179819516)

[4.3.1 Value Counts: 4](#_Toc179819517)

[4.3.2 Unique Attributes: 5](#_Toc179819518)

[4.4 Graphical Analysis 5](#_Toc179819519)

[4.4.1 Univariate Analysis - Numerical Features 5](#_Toc179819520)

[4.4.2 Univariate Analysis - Categorical Features 8](#_Toc179819521)

[4.4.3 Bivariate Analysis 9](#_Toc179819522)

[4.4.4 Multivariate Analysis 12](#_Toc179819523)

[4.5 Correlation Analysis 13](#_Toc179819524)

[4.5.1 Graph 13](#_Toc179819525)

[4.5.2 Observations From the Heatmap 14](#_Toc179819526)

[4.6 Outlier Detection 15](#_Toc179819527)

[4.7 Conditional Probabilities 15](#_Toc179819528)

[4.7.1 Product Purchases: 15](#_Toc179819529)

[4.7.2 Product – Gender: 15](#_Toc179819530)

[4.7.3 Product – Age: 17](#_Toc179819531)

[4.7.4 Product – Income: 17](#_Toc179819532)

[4.7.5 Product – Fitness: 18](#_Toc179819533)

[4.7.6 Product - Marital Status: 19](#_Toc179819534)

[5 Customer Demographics 20](#_Toc179819535)

[5.1 Gender Distribution 20](#_Toc179819536)

[5.2 Age Distribution 20](#_Toc179819537)

[5.3 Income Analysis 20](#_Toc179819538)

[5.3.1 Low-Income Customer Insights: 20](#_Toc179819539)

[5.3.2 High-Income Customer Insights: 20](#_Toc179819540)

[5.3.3 High-Income customer who Purchased KP781 20](#_Toc179819541)

[5.4 Fitness Level Insights 20](#_Toc179819542)

[5.4.1 Fitness Level 5 Analysis: 20](#_Toc179819543)

[5.5 Marital Status Analysis 20](#_Toc179819544)

[5.5.1 Partnered Customers: 20](#_Toc179819545)

[6 Insights and Recommendations 21](#_Toc179819546)

[6.1 Target Audience 21](#_Toc179819547)

[6.1.1 Insight: 21](#_Toc179819548)

[6.1.2 Recommendation: 21](#_Toc179819549)

[6.2 Product Portfolio 21](#_Toc179819550)

[6.2.1 Insight: 21](#_Toc179819551)

[6.2.2 Recommendation: 21](#_Toc179819552)

[6.3 Pricing Strategy 21](#_Toc179819553)

[6.3.1 Insight: 21](#_Toc179819554)

[6.3.2 Recommendation: 21](#_Toc179819555)

[6.4 Marketing Strategies 21](#_Toc179819556)

[6.4.1 Insight: 21](#_Toc179819557)

[6.4.2 Recommendation: 22](#_Toc179819558)

[6.5 Customer Retention 22](#_Toc179819559)

[6.5.1 Insight: 22](#_Toc179819560)

[6.5.2 Recommendation: 22](#_Toc179819561)

[6.6 Product Improvements 22](#_Toc179819562)

[6.6.1 Insight: 22](#_Toc179819563)

[6.6.2 Recommendation: 22](#_Toc179819564)

[6.7 Geographic Targeting 22](#_Toc179819565)

[6.7.1 Insight: 22](#_Toc179819566)

[6.7.2 Recommendation: 22](#_Toc179819567)

[6.8 Data-Driven Decisions 23](#_Toc179819568)

[6.8.1 Insight: 23](#_Toc179819569)

[6.8.2 Recommendation: 23](#_Toc179819570)

[7 Conclusion 23](#_Toc179819571)

# Introduction

This report presents a thorough analysis of customer demographics, purchasing behavior, and product performance related to Aerofit treadmills. The primary objective is to derive actionable insights and recommendations that can guide marketing strategies, product development, and customer engagement initiatives. By investigating the characteristics of the target audience for each type of treadmill, Aerofit aims to enhance its recommendations to new customers.

# Project Details

The market research team at Aerofit seeks to identify the characteristics of the target audience for each type of treadmill offered by the company. The insights gathered will inform marketing strategies and product development.

## Product Portfolio

* **KP281**: Entry-level treadmill, priced at **$1,500**.
* **KP481**: Mid-level treadmill designed for runners, priced at **$1,750**.
* **KP781**: Advanced treadmill with premium features, priced at **$2,500**.

# Data Description

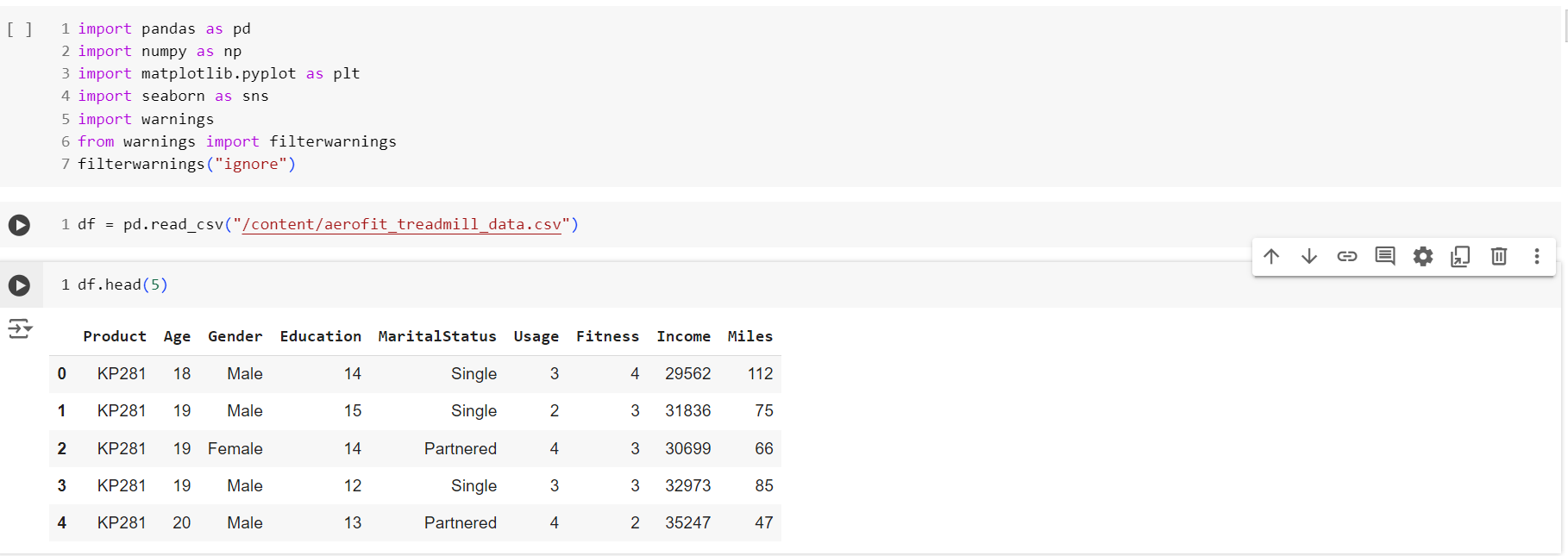
The dataset, **aerofit\_treadmill\_data.csv**, contains information on individuals who purchased a treadmill from Aerofit stores over the past three months. The features included are:

* **Product**: Type of product purchased (KP281, KP481, or KP781).
* **Age**: Customer age in years.
* **Gender**: Customer gender (male/female).
* **Education**: Education level in years.
* **MaritalStatus**: Customer marital status (single or partnered).
* **Usage**: Average number of times the customer plans to use the treadmill each week.
* **Fitness**: Self-rated fitness level on a scale of 1-5.
* **Income**: Annual income in US dollars.
* **Miles**: Average number of miles the customer expects to walk/run each week.

# Data Exploration and Processing

## Data Import and Overview

* The data is imported using pandas for analysis.



* The shape of the DataFrame 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

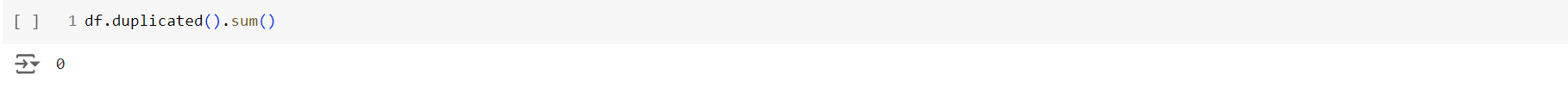
Description automatically generated

* 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 is generated for both categorical and numerical features. Key observations include:

### Categorical Features:

Distribution of gender, marital status, and product types.

A white background with a black border

Description automatically generated with medium confidence

### Numerical Features:

Summary statistics including mean, median, and standard deviation for age, income, usage, fitness, and miles.

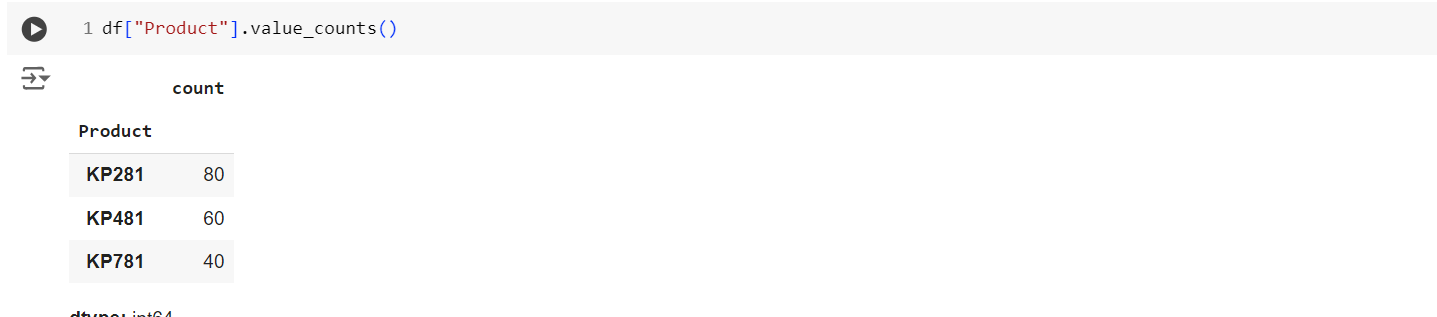
A group of numbers on a white background

Description automatically generated

## Non-Graphical Analysis

### Value Counts:

Counts of unique values for all categorical features are documented.

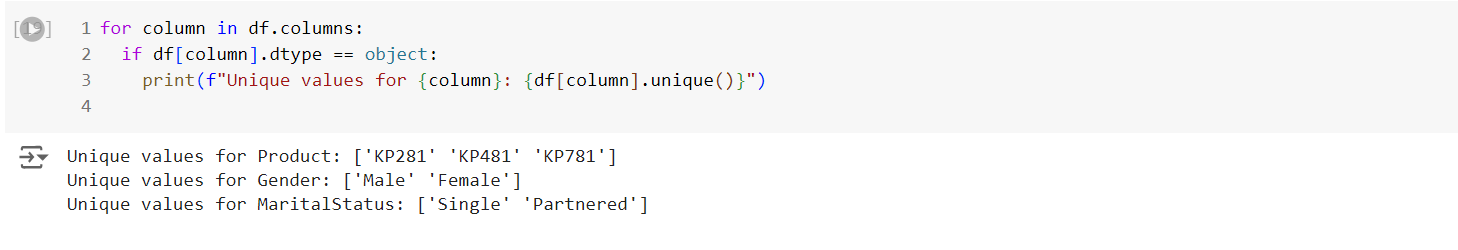


A screenshot of a computer

Description automatically generated

### Unique Attributes:

The unique attributes for each categorical feature are listed.



## Graphical Analysis

### Univariate Analysis - Numerical Features

* **Distribution Plot**: Plots to visualize the distribution of numerical features such as 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

* **Count Plot**: Count plots for the frequency of different fitness levels and product types.

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.

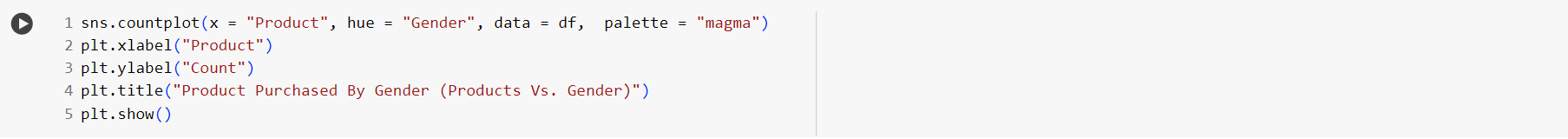


A graph of different sizes and colors

Description automatically generated with medium confidence

### Bivariate Analysis

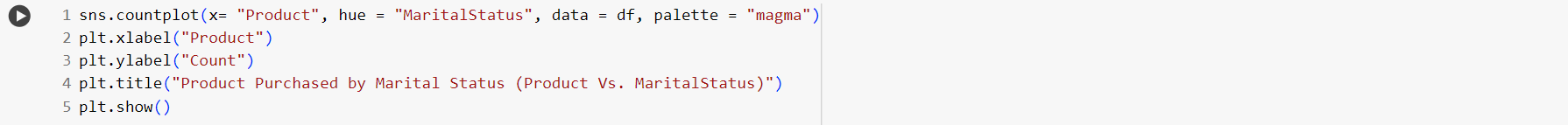
* **Product vs. Gender**: Analyzing the relationship between the product purchased and the customer's gender.



A graph of different colored bars

Description automatically generated

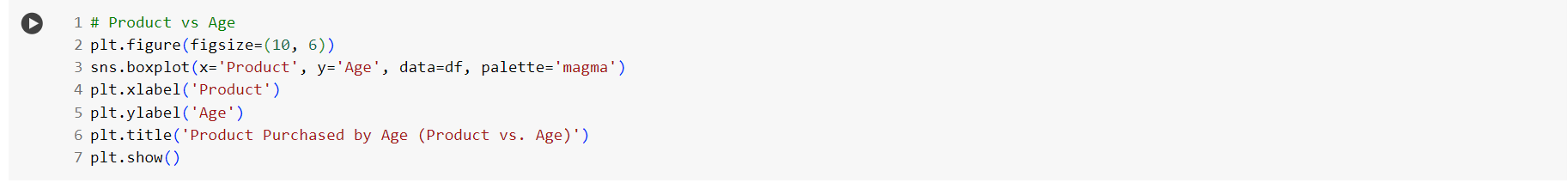
* **Product vs. Marital Status**: Examining how marital status influences product choice.



A graph of a product purchased by martial status

Description automatically generated

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

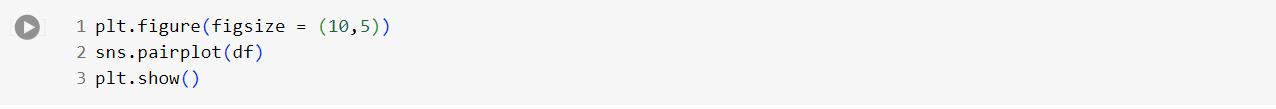


A chart with different colored squares

Description automatically generated

### Multivariate Analysis

* **Pair Plots**: Creating pair plots to show relationships among multiple features simultaneously.



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.
* 'Fitness' and 'Usage' have a very weak correlation (0.05), indicating that self-rated fitness doesn't strongly influence how many times users plan to use the treadmill per week.

#### 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 are detected using the Interquartile Range (IQR) method, allowing identification of extreme values that may affect analysis.



## Conditional Probabilities

### 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

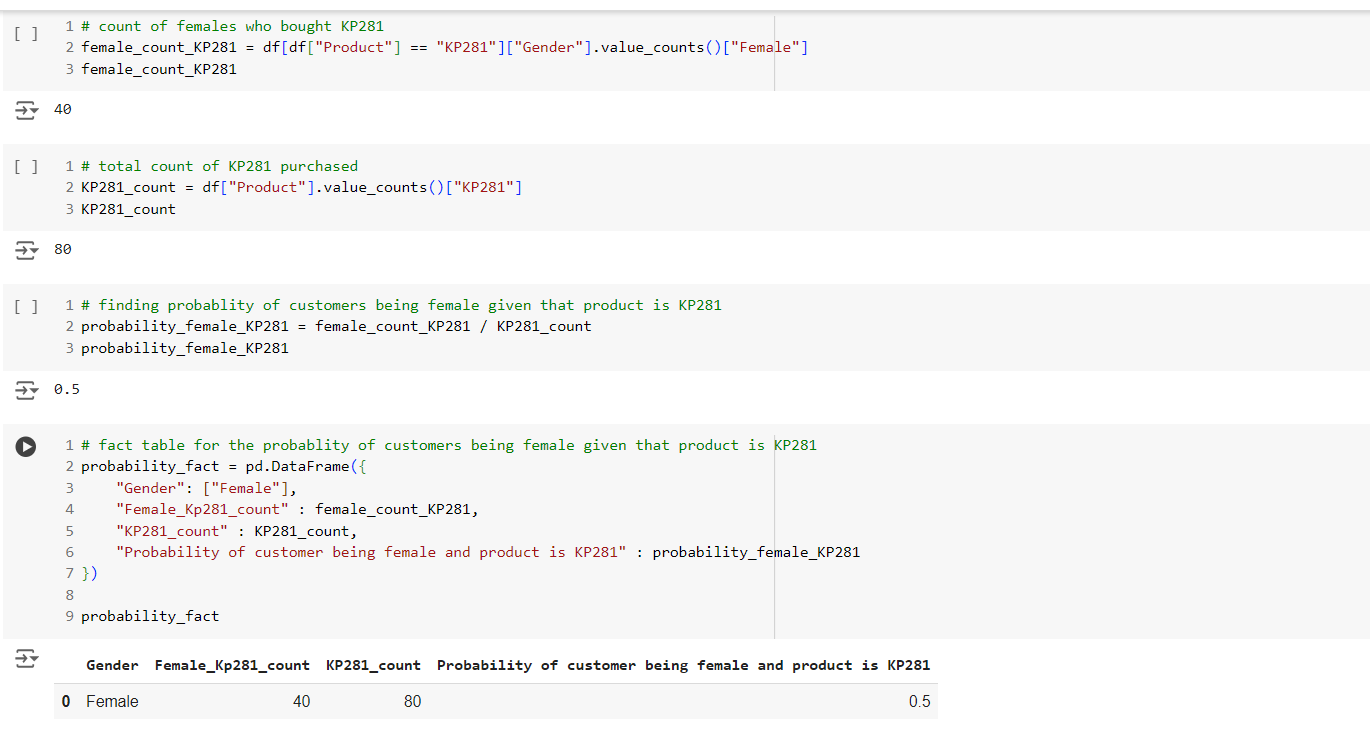
Description automatically generated

* 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

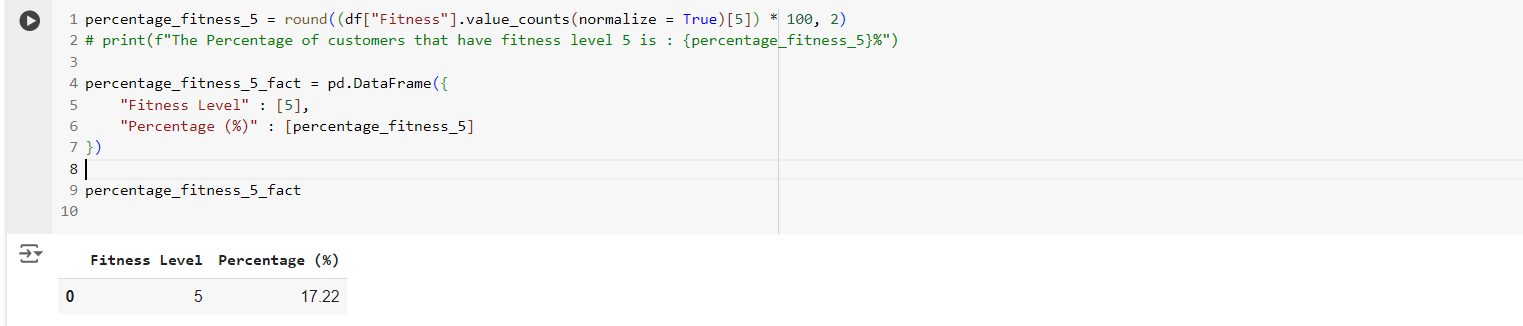
* + Percentage of high-income customers buying a treadmill given that the product is KP781.

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:

A significant portion of customers purchasing treadmills are males, particularly those who are partnered/married.

### Recommendation:

* Focus marketing efforts on these segments, highlighting the benefits and features that cater to their specific needs.
* Consider tailoring marketing campaigns and promotional offers to attract and retain male customers.

## Product Portfolio

### Insight:

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

### Recommendation:

* Continue investing in developing and promoting these products to capitalize on their market demand.
* Consider introducing new models that fall within these price ranges and cater to specific customer needs, such as features related to fitness tracking, entertainment options, and customization.

## 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.

## Customer Retention

### Insight:

Customer engagement can be enhanced through personalized experiences and exclusive offers.

### Recommendation:

* Encourage customer loyalty by building a strong relationship with users, providing excellent after-sales support, and offering exclusive benefits.
* Consider launching programs like fitness challenges, user communities, or educational resources to increase engagement and retain existing customers.
* Collect customer feedback through surveys or online reviews to understand their experiences and address any pain points.

## 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 comprehensive analysis of Aerofit treadmill purchases provides valuable insights into customer demographics, preferences, and behaviors. By implementing the recommended strategies and continuously monitoring market trends, Aerofit can strengthen its market position, enhance customer satisfaction, and drive sales growth.