Aerofit®

HELPING YOU FROM FIT-LESS TO FITNESS

SINCE 1986

Business Case Solution

By Munmun Mishra

Introduction

In the ever-evolving fitness equipment industry, Aerofit stands out by seamlessly blending data science into its core strategy to enhance the customer experience. The company's commitment to innovation is evident as it leverages analytics to redefine how treadmills are recommended, tailoring suggestions based on individual customer characteristics.

At the intersection of fitness and data-driven decision-making, Aerofit's approach departs from the one-size-fits-all model. By utilising advanced analytics, the company ensures that its products align with the unique needs of its diverse customer base. This personalised methodology sets Aerofit apart as a customer-centric leader and positions it to adapt to the market's ever-changing demands.

Aerofit's exploration into data-driven customisation demonstrates its dedication to technological advancements and highlights its ability to anticipate trends in the fitness landscape. Through a thorough customer data analysis, the company understands individual preferences and uses this insight to shape future product developments.

This business case solution delves into the specifics of Aerofit's data science-driven strategy, showcasing how it propels the company forward in a competitive market. As we unpack the layers of Aerofit's approach, it becomes clear that the fusion of fitness and analytics is not just a technological upgrade but a transformative shift with the potential to reshape the industry.

Objectives

In this analysis, our primary objectives are

- **Customer Characterization:**
    - Analyse the Aerofit treadmill dataset comprehensively to gain insights into various customer characteristics.
    - Uncover patterns and trends within the dataset to understand the diverse preferences of Aerofit's customer base.
- **Customer Profiles:**
    - Develop distinct customer profiles for each treadmill product based on the dataset.
    - Identify key attributes that define customer segments and tailor profiles to each product category.
- **Probability Computation:**
    - Computing conditional and marginal probabilities to deepen our understanding of customer behaviours
    - Utilise probabilities to inform and enhance product recommendations, ensuring a personalised approach for each customer.
- **Actionable Insights:**
    - Extract actionable insights from the analysis that can guide informed business decisions.
    - Translate statistical findings into practical strategies for product development, marketing, and customer engagement.

Objectives and Scope

The scope of this analysis encompasses

- **Demographic Exploration:**
    - Investigate customer demographics within the dataset to uncover trends related to age, gender, location, and other relevant factors.
    - Understand how demographic variables influence customer preferences for Aerofit treadmills.
- **Usage Patterns:**
    - Explore patterns in treadmill usage, including frequency, duration, and preferred features.
    - Correlate usage patterns with customer characteristics to identify key factors driving product choices.
- **Product Preferences:**
    - Assess the impact of customer demographics and usage patterns on the preference for specific treadmill products.
    - Formulate insights that guide product positioning and marketing strategies.

Data Acquisition and Preprocessing

We kick-started our analysis by seamlessly bringing the Aerofit treadmill dataset into our environment. To ensure a smooth analytical journey, we took the following streamlined preprocessing steps:

- **Categorical Attributes Upgrade:**
  Optimising our analytical processes, we converted categorical attributes like 'Gender' and 'MaritalStatus' to the 'category' data type. This tweak streamlines computations and opens doors to a more nuanced exploration of these attributes.

- **Data Integrity Assurance:**
  Our dataset inspection brought good news — no missing values were detected. This clean slate ensures our dataset's reliability, forming a sturdy foundation for precise and meaningful analyses.

With this straightforward yet effective approach to data preparation, we're geared up for a thorough exploration of Aerofit's treadmill dataset. These behind-the-scenes steps set the tone for a confident and reliable analysis, ensuring our findings are as insightful and accurate.

Importing the CSV file

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
# Load the Aerofit treadmill dataset
url =
 "https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/001/125/origi
 nal/aerofit_treadmill.csv?1639992749"
# Replace with the actual dataset link
aerofit_data = pd.read_csv(url)
```

Data Summary

Structure and Characteristics of the dataset

```
print(aerofit_data.info())
print(aerofit_data.head())

# Descriptive statistics for numerical variables
print(aerofit_data.describe())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 180 entries, 0 to 179
Data columns (total 9 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Product        180 non-null    object
 1   Age            180 non-null    int64
 2   Gender         180 non-null    object
 3   Education      180 non-null    int64
 4   MaritalStatus  180 non-null    object
 5   Usage          180 non-null    int64
 6   Fitness        180 non-null    int64
 7   Income         180 non-null    int64
 8   Miles          180 non-null    int64
dtypes: int64(6), object(3)
memory usage: 12.8+ KB
None
  Product   Age   Gender   Education MaritalStatus   Usage   Fitness   Income
Miles
0   KP281    18     Male          14         Single       3         4   29562
112
1   KP281    19     Male          15         Single       2         3   31836
75
2   KP281    19   Female          14      Partnered       4         3   30699
66
3   KP281    19     Male          12         Single       3         3   32973
85
4   KP281    20     Male          13      Partnered       4         2   35247
47
              Age   Education       Usage     Fitness         Income  \
count  180.000000  180.000000  180.000000  180.000000     180.000000
mean    28.788889   15.572222    3.455556    3.311111   53719.577778
```

	Age	Education	Usage	Fitness	Income
std	6.943498	1.617055	1.084797	0.958869	16506.684226
min	18.000000	12.000000	2.000000	1.000000	29562.000000
25%	24.000000	14.000000	3.000000	3.000000	44058.750000
50%	26.000000	16.000000	3.000000	3.000000	50596.500000
75%	33.000000	16.000000	4.000000	4.000000	58668.000000
max	50.000000	21.000000	7.000000	5.000000	104581.000000

	Miles
count	180.000000
mean	103.194444
std	51.863605
min	21.000000
25%	66.000000
50%	94.000000
75%	114.750000
max	360.000000

- **DataFrame Overview:**
  - The data frame has 180 entries (rows) and 9 columns.
  - No missing values are in any of the columns (non-null counts are all 180).
- **Columns:**
  - **Product:** Categorical variable representing the product used.
  - **Age:** Numerical variable representing the age of the individuals.
  - **Gender:** Categorical variable indicating the gender of the individuals.
  - **Education:** Numerical variable indicating the education level of the individuals.
  - **MaritalStatus:** Categorical variable indicating the marital status of the individuals.
  - **Usage:** Numerical variable indicating product usage.
  - **Fitness:** Numerical variable indicating self-rated fitness level.
  - **Income:** Numerical variable representing the income of the individuals.
  - **Miles:** Numerical variable representing the number of miles the individuals are willing to travel.
- **Descriptive Statistics:**
  - For numerical columns (Age, Education, Usage, Fitness, Income, Miles):
  - Mean (average) values are provided.
  - Standard deviation (std) indicates the amount of variation or dispersion.
  - Minimum (min) and Maximum (max) values give the range of the data.
  - Quartiles (25%, 50%, 75%) provide insights into the distribution.

- **Interpretations:**
    - **Age:** The average age is around 28.8, with a minimum of 18 and a maximum of 50.
    - **Education:** On average, the individuals have a level of education of around 15.6.
    - **Usage and Fitness:** On average, the product is used 3.46 times, and the self-rated fitness level is 3.31.
    - **Income:** The average income is approximately 53,719, with a significant standard deviation, indicating a range of income levels.
    - **Miles:** On average, individuals are willing to travel about 103 miles, with a minimum of 21 and a maximum of 360.
- **Categorical Variables:**
    - For categorical variables (Gender, MaritalStatus), we can analyse the distribution of categories, but the summary statistics mainly apply to numerical variables.
- **Usage Patterns:**
    - We can observe usage patterns, income distribution, and preferences.

These meticulously chosen features form the backbone of our analysis, providing a comprehensive view of customer characteristics and preferences. By delving into the intricate interplay of these variables, we aim to extract accurate and actionable insights, aligning our findings with Aerofit's commitment to a personalised and customer-centric approach. As we navigate through the dataset, these key elements will serve as critical touchpoints, guiding our exploration of the multifaceted landscape of Aerofit's customer interactions.

Data Visualisation

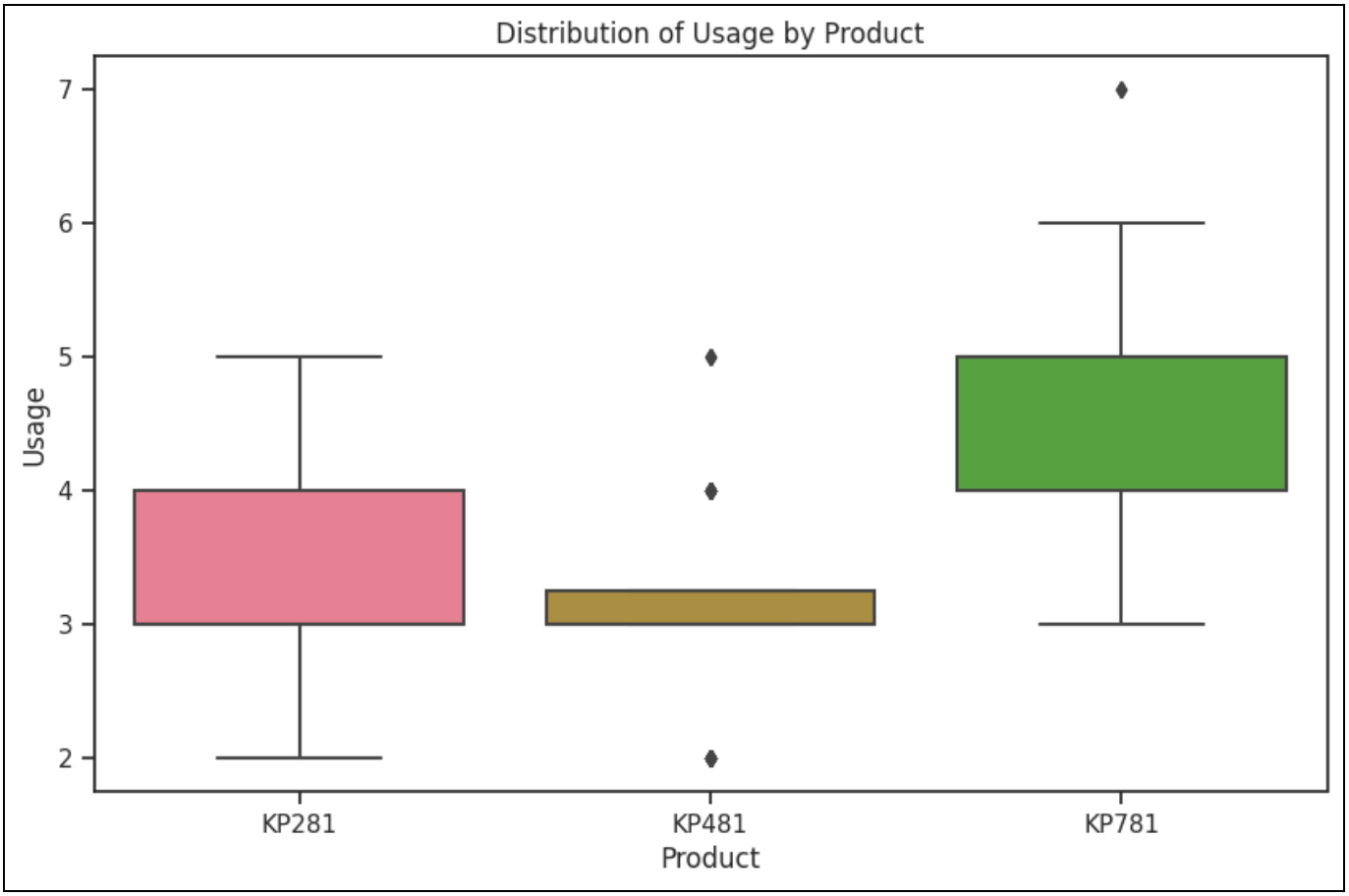
- - - - X

Pair Plots for Numerical Variables:

```
import seaborn as sns
import matplotlib.pyplot as plt

# Load your data into 'aerofit_data'

# Set up the pair plot with Seaborn
sns.set(style="ticks")
sns.set_palette("husl")  # Choose an appropriate color palette

# Define the pair plot
pair_plot = sns.pairplot(aerofit_data, hue='Product')

# Set title and adjust layout for better visualization
pair_plot.fig.suptitle("Pair Plots for Numerical Variables by Product
Purchased", y=1.02)

# Show the plot
plt.show()
```

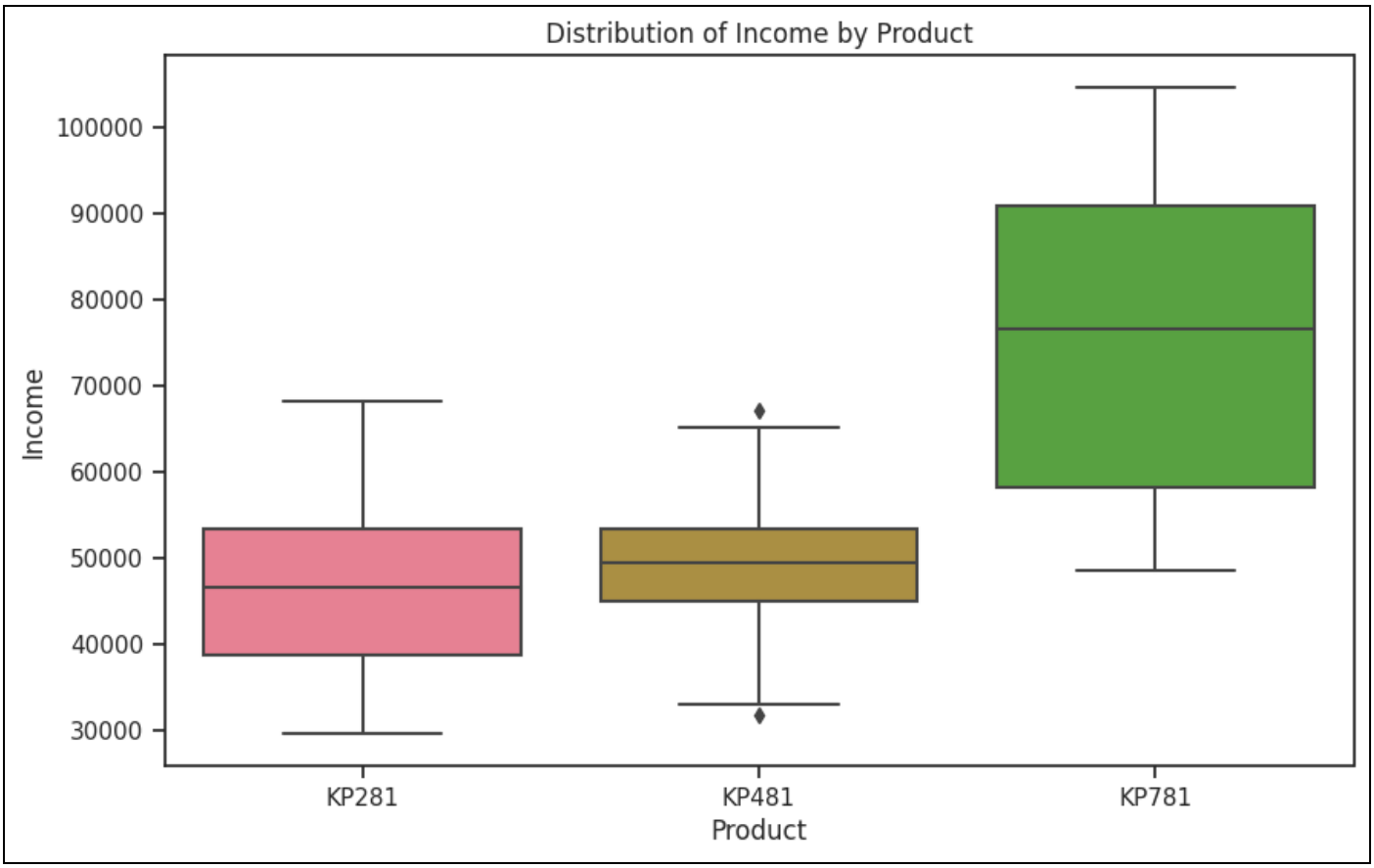
Distribution of Usage by Product Purchased:

```
# Visualize the distribution of usage by product
plt.figure(figsize=(10, 6))
sns.boxplot(x='Product', y='Usage', data=aerofit_data)
plt.title('Distribution of Usage by Product')
plt.show()
```



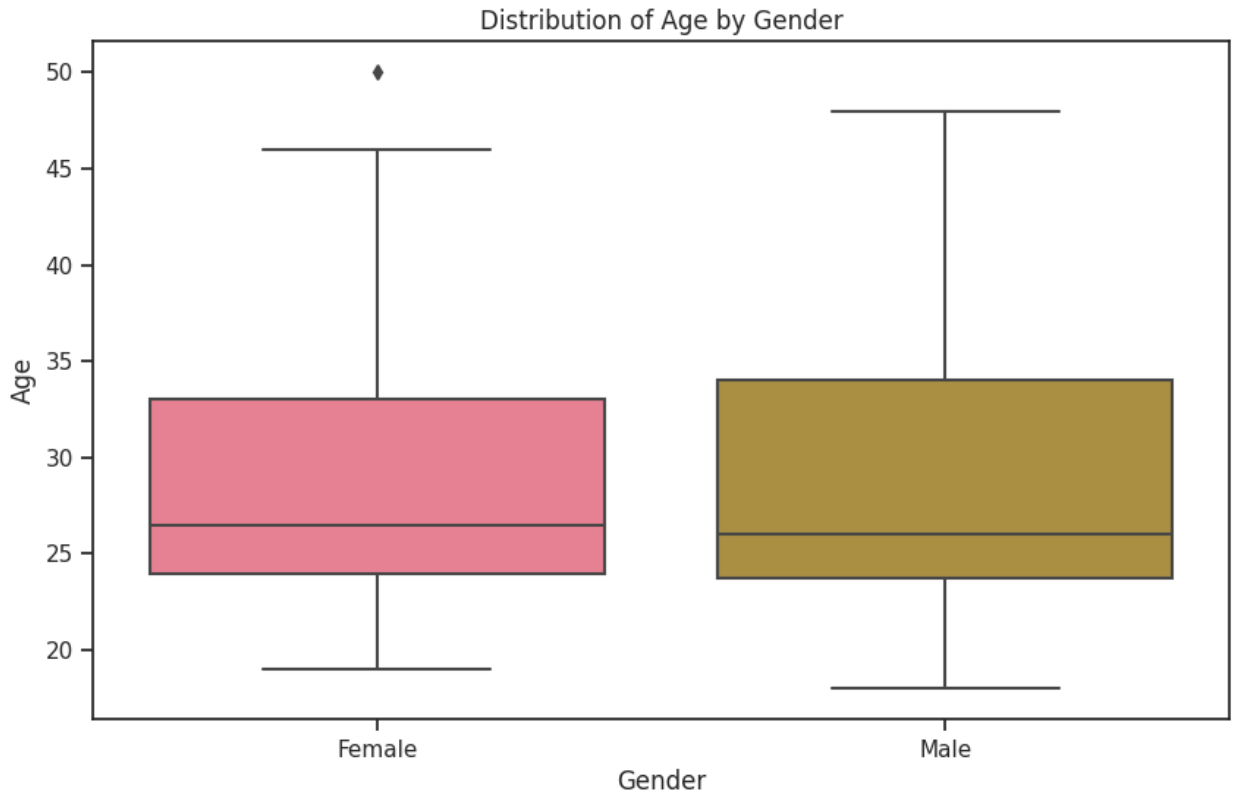
Income Distribution by Product Purchased:

```
# Visualize the distribution of income by product
plt.figure(figsize=(10, 6))
sns.boxplot(x='Product', y='Income', data=aerofit_data)
plt.title('Distribution of Income by Product')
plt.show()
```



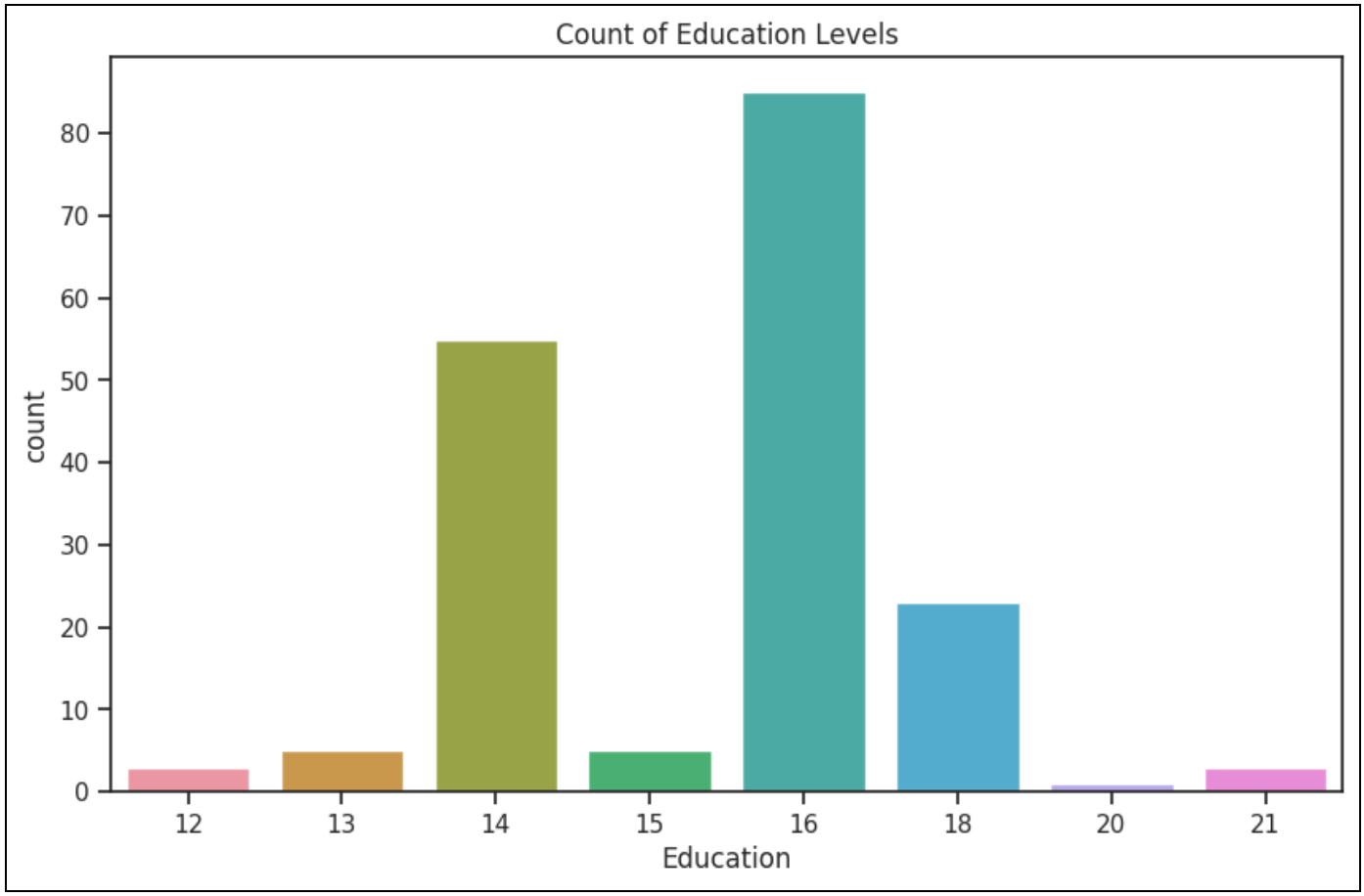
Age Distribution by Gender:

```
# Visualize the distribution of age by gender
plt.figure(figsize=(10, 6))
sns.boxplot(x='Gender', y='Age', data=aerofit_data)
plt.title('Distribution of Age by Gender')
plt.show()
```



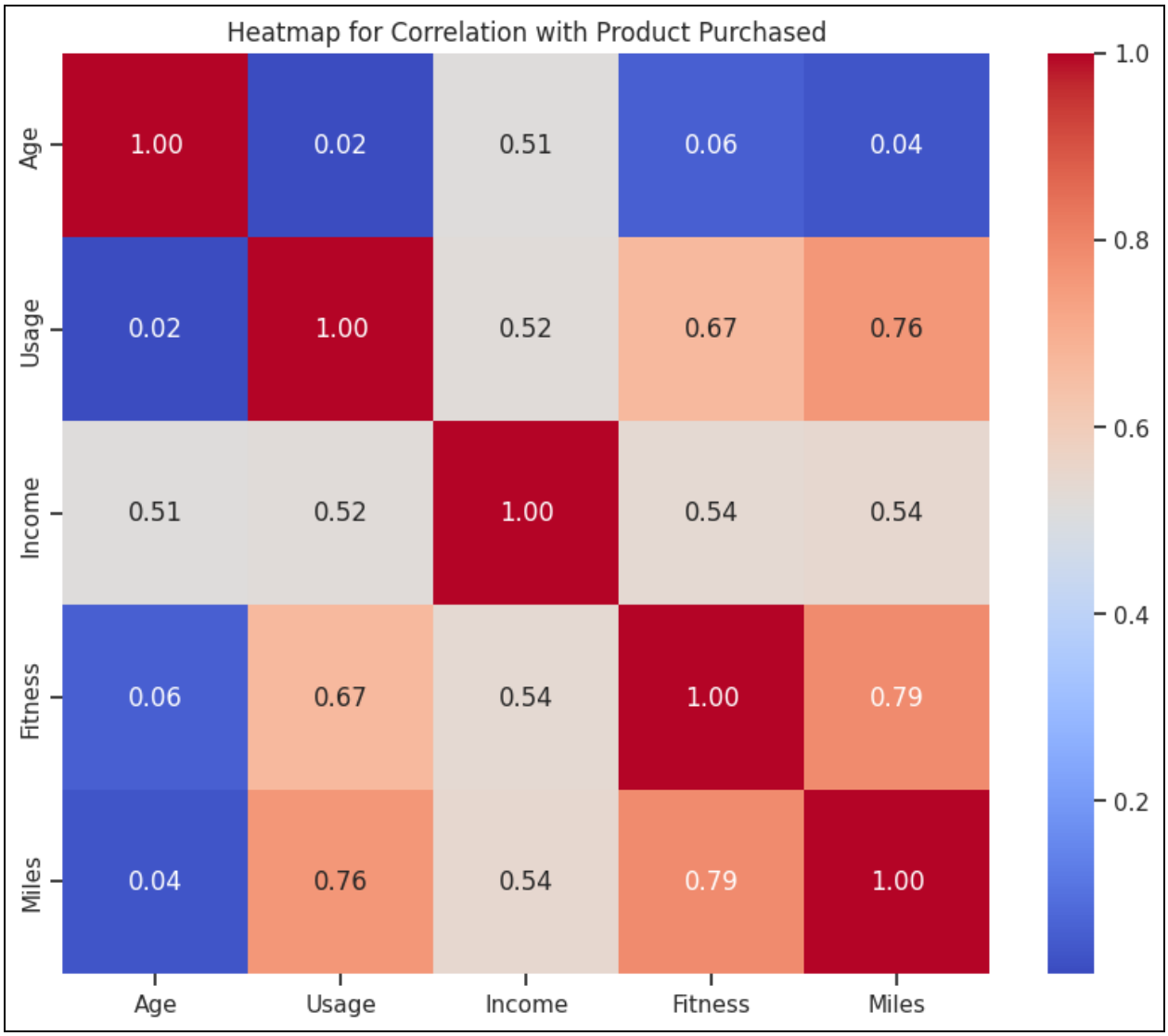
Education Level Distribution:

```
# Visualize the count of education levels
plt.figure(figsize=(10, 6))
sns.countplot(x='Education', data=aerofit_data)
plt.title('Count of Education Levels')
plt.show()
```



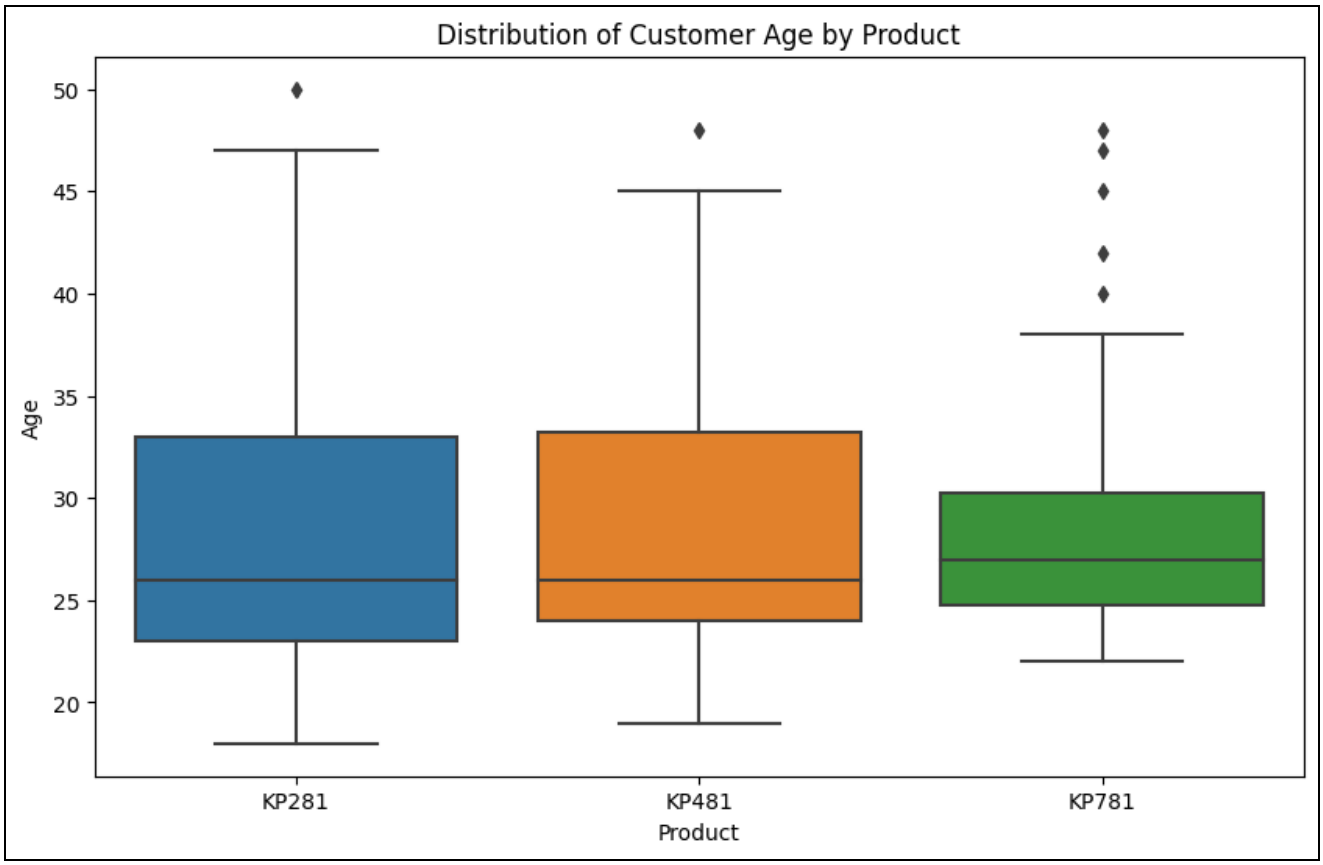
Heatmap for Product Purchased and Numerical Variables:

```
# Visualize a heatmap for correlation between numerical variables and product purchased
plt.figure(figsize=(10, 8))
heatmap_data = aerofit_data[['Age', 'Usage', 'Income', 'Fitness', 'Miles']]
heatmap_data['Product'] = aerofit_data['Product']
sns.heatmap(heatmap_data.corr(), annot=True, cmap='coolwarm', fmt=".2f")
plt.title('Heatmap for Correlation with Product Purchased')
plt.show()
```



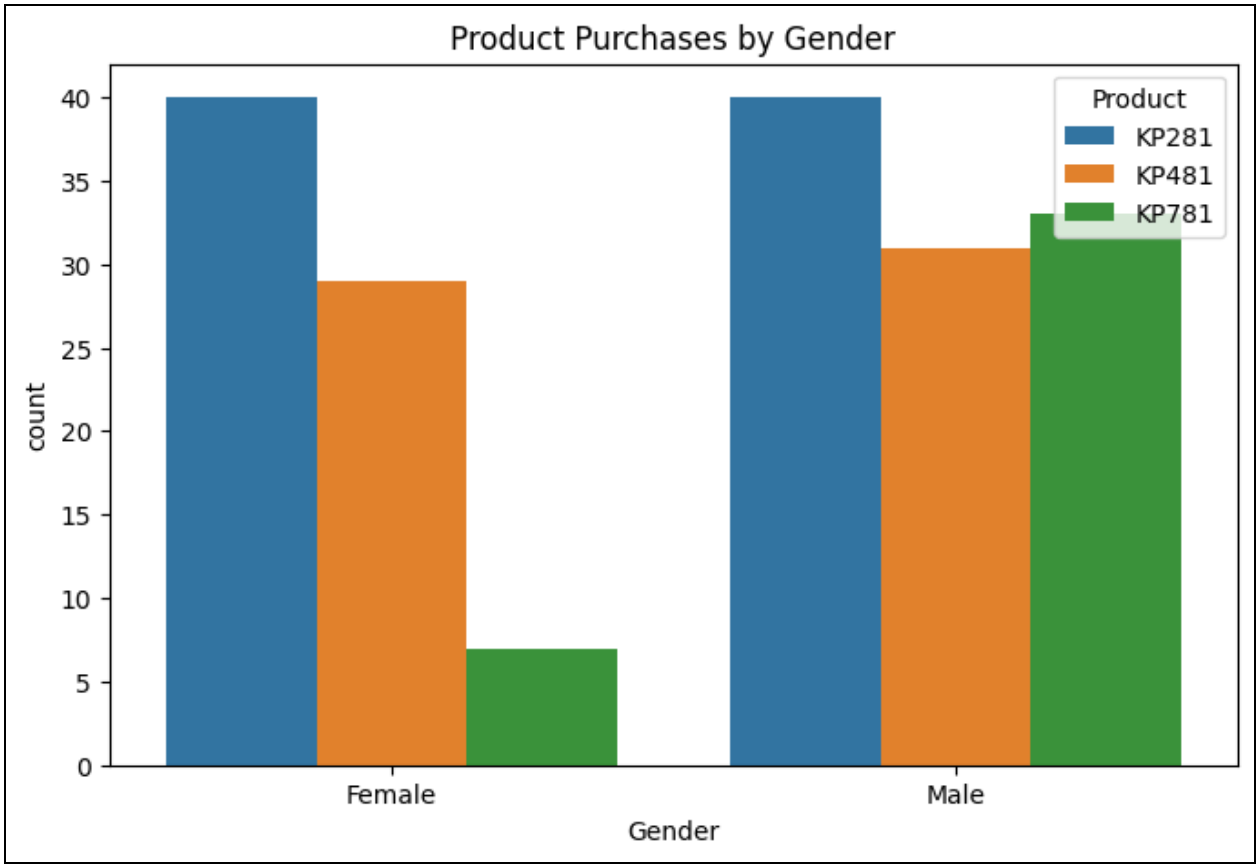
Distribution of Customer age by product

```
# Visualize the distribution of customer age by product
plt.figure(figsize=(10, 6))
sns.boxplot(x='Product', y='Age', data=aerofit_data)
plt.title('Distribution of Customer Age by Product')
plt.show()
```



Count of products purchased by gender

```
#Visualize the count of products purchased by gender
plt.figure(figsize=(8, 5))
sns.countplot(x='Gender', hue='Product', data=aerofit_data)
plt.title('Product Purchases by Gender')
plt.show()
```



Two-way contingency and conditional probabilities for Product vs Marital Status

```
# Create two-way contingency table
contingency_table_marital = pd.crosstab(aerofit_data['Product'],
aerofit_data['MaritalStatus'], margins=True, margins_name="Total")

# Compute conditional probabilities
conditional_prob_marital =
contingency_table_marital.div(contingency_table_marital['Total'], axis=0)

# Display contingency table and conditional probabilities
print("Two-way Contingency Table: Product vs. MaritalStatus")
print(contingency_table_marital)
print("\nConditional Probabilities: Product vs. MaritalStatus")
print(conditional_prob_marital)
```

```
Two-way Contingency Table: Product vs. MaritalStatus
MaritalStatus  Partnered  Single  Total
Product
KP281                 48      32     80
KP481                 36      24     60
KP781                 23      17     40
Total                107      73    180
```

```
Conditional Probabilities: Product vs. MaritalStatus
MaritalStatus  Partnered    Single  Total
Product
KP281           0.600000  0.400000    1.0
KP481           0.600000  0.400000    1.0
KP781           0.575000  0.425000    1.0
Total           0.594444  0.405556    1.0
```

Two-way contingency and conditional probabilities for Product vs Education

```
#Education vs. Product Purchased
#Create two-way contingency table
contingency_table_education = pd.crosstab(aerofit_data['Product'],
aerofit_data['Education'], margins=True, margins_name="Total")

# Compute conditional probabilities
conditional_prob_education =
contingency_table_education.div(contingency_table_education['Total'],
axis=0)

# Display contingency table and conditional probabilities
print("\nTwo-way Contingency Table: Product vs. Education")
print(contingency_table_education)
print("\nConditional Probabilities: Product vs. Education")
print(conditional_prob_education)
```

```
Two-way Contingency Table: Product vs. Education
Education  12  13  14  15  16  18  20  21  Total
Product
KP281       2   3  30   4  39   2   0   0     80
KP481       1   2  23   1  31   2   0   0     60
KP781       0   0   2   0  15  19   1   3     40
Total       3   5  55   5  85  23   1   3    180

Conditional Probabilities: Product vs. Education
Education         12         13         14        15         16
18  \
Product
KP281       0.025000   0.037500   0.375000  0.050000  0.487500
0.025000
KP481       0.016667   0.033333   0.383333  0.016667  0.516667
0.033333
KP781       0.000000   0.000000   0.050000  0.000000  0.375000
0.475000
```

```
Total        0.016667   0.027778   0.305556   0.027778   0.472222
0.127778
```

```
Education         20         21  Total
Product
KP281       0.000000   0.000000    1.0
KP481       0.000000   0.000000    1.0
KP781       0.025000   0.075000    1.0
Total       0.005556   0.016667    1.0
```

Two-way contingency and conditional probabilities for Product vs Gender

```
# Create two-way contingency table for Product Purchased and Gender
contingency_table_gender = pd.crosstab(aerofit_data['Product'],
aerofit_data['Gender'], margins=True, margins_name="Total")

# Compute conditional probabilities
conditional_prob_gender =
contingency_table_gender.div(contingency_table_gender['Total'], axis=0)

# Display contingency table and conditional probabilities
print("\nTwo-way Contingency Table: Product vs. Gender")
print(contingency_table_gender)
print("\nConditional Probabilities: Product vs. Gender")
print(conditional_prob_gender)
```

```
Two-way Contingency Table: Product vs. Gender
Gender   Female  Male  Total
Product
KP281        40    40     80
KP481        29    31     60
KP781         7    33     40
Total        76   104    180
```

Business Insights and Recommendations for AeroFit Treadmill Products

1. **Targeted Marketing Strategies:**
    a. **KP281:** Given that KP281 attracts a younger demographic (21-26 years old), marketing efforts should focus on platforms and channels that resonate with this age group, such as social media influencers, fitness apps, and college/university partnerships. Consider incorporating vibrant and youthful imagery in advertising campaigns.
    b. **KP481:** Targeting an older demographic with higher education and income levels, emphasising the premium features of KP481 in marketing materials. Consider partnerships with fitness experts or wellness influencers who cater to a more mature audience. Highlighting the durability and advanced functionalities of the product could enhance its appeal.
    c. **KP781:** With a predominantly male customer base, marketing for KP781 should highlight its professional-grade features. Collaborations with fitness trainers, sports professionals, or athletes could strengthen the product's image. Emphasise the benefits of high-intensity workouts and advanced training capabilities.

2. **Upselling Opportunities:**
    a. **KP481 Upselling:** Given that KP481 customers have higher income levels, there's an opportunity for upselling accessories, extended warranties, or premium subscription services. Highlight the long-term value and additional benefits of investing in a higher-end product.
    b. **KP781 Upselling:** Upsell opportunities for KP781 can focus on premium training programs, virtual coaching services, or accessories designed for serious runners. Consider bundling these with the product to enhance perceived value.

3. **Female Customer Engagement:**
    a. **Targeted Campaigns:** All products show potential for growing the female customer base. Develop marketing campaigns featuring female athletes using AeroFit treadmills to appeal to a broader audience. Highlight the versatility of the products for users of all fitness levels.

b. **Inclusive Imagery:** Ensure marketing materials feature a diverse range of individuals to create an inclusive image. Showcase success stories, testimonials, and transformational journeys of female customers to inspire and connect with prospective buyers.

4. **Product Differentiation:**
    a. Positioning KP781: Given its high usage and fitness levels among customers, position KP781 as a professional-grade treadmill for serious runners and fitness enthusiasts. Highlight its robust build, advanced features, and suitability for intensive training sessions.
    b. Highlighting Unique Features: Emphasize unique features for each product that cater to specific customer needs. For example, stress the compact design and user-friendly interface of KP281, the advanced technology and connectivity options of KP481, and the performance-enhancing capabilities of KP781.

5. **Geographic Targeting:**
    a. **Regional Preferences:** Explore regional customer preferences and demographic variations. Tailor marketing strategies and campaigns based on regional insights. Consider partnerships with local fitness influencers or events to enhance brand visibility.
    b. **Localised Promotions:** Implement localised promotions or discounts based on regional trends. This approach can help capture specific market segments and enhance customer engagement.

6. **Customer Loyalty Programs:**
    a. **Incentivize Repeat Purchases:** Introduce loyalty programs that offer discounts, exclusive content, or early access to new products for repeat customers. Encourage brand loyalty by creating a sense of community among AeroFit treadmill users.
    b. **Feedback Mechanism:** Establish a feedback mechanism to understand customer satisfaction and identify areas for improvement. Use customer feedback to enhance product features, customer service, and overall brand experience.

These insights and recommendations aim to guide AeroFit in refining its marketing strategies, appealing to a broader customer base, and maximising upselling opportunities for sustained business growth.

Conclusion

In decoding the data mosaic, AeroFit's customer landscape comes into sharp focus. The analysis illuminates distinct personas for each treadmill product, offering a roadmap for strategic initiatives.

Tailoring marketing for each demographic, seizing upselling opportunities, and championing inclusivity emerge as pivotal strategies. Female empowerment takes centre stage, reflecting the brand's commitment to diversity.

As AeroFit charts its course, these insights serve as beacons. By embracing regional nuances, nurturing loyalty, and staying innovative, AeroFit is poised to meet and exceed customer expectations.

In this symphony of data and insights, the crescendo signals boundless potential. May these insights propel AeroFit toward new heights, where each treadmill resonates with an ever-expanding community of fitness enthusiasts.