

---

## ▼ \*AeroFit \*

---

AeroFit is a leading brand in the field of fitness equipment. AeroFit provides a product range including machines such as treadmills, exercise bikes, gym equipment, and fitness accessories to cater to the needs of all categories of people.

---

## 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, to provide a better recommendation of the treadmills to the new customers. The team decides to investigate whether there are differences across the product with respect to customer characteristics. Perform descriptive analytics to create a customer profile for each AeroFit treadmill product by developing appropriate tables and charts. For each AeroFit treadmill product, construct two-way contingency tables and compute all conditional and marginal probabilities along with their insights/impact on the business.

---

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import math
from scipy import stats
from scipy.stats import binom
from scipy.stats import norm
```

---

## ▼ What good looks like?

---

- 1.Import the dataset and do usual data analysis steps like checking the structure & characteristics of the dataset
  - 2.Detect Outliers (using boxplot, "describe" method by checking the difference between mean and median)
  - 3.Check if features like marital status, age have any effect on the product purchased (using countplot, histplots, boxplots etc)
  - 4.Representing the marginal probability like - what percent of customers have purchased KP281, KP481, or KP781 in a table (can use pandas.crosstab here)
  - 5.Check correlation among different factors using heat maps or pair plots.
  - 6.With all the above steps you can answer questions like: What is the probability of a male customer buying a KP781 treadmill?
  - 7.Customer Profiling - Categorization of users.
  - 8.Probability- marginal, conditional probability.
  - 9.Some recommendations and actionable insights, based on the inferences.
-

```
df = pd.read_csv('https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/001/125/original/aerofit_treadmill.csv?1639992749')
df
```

	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
...	...	...	...	...	...	...	...	...	...
175	KP781	40	Male	21	Single	6	5	83416	200
176	KP781	42	Male	18	Single	5	4	89641	200
177	KP781	45	Male	16	Single	5	5	90886	160
178	KP781	47	Male	18	Partnered	4	5	104581	120
179	KP781	48	Male	18	Partnered	4	5	95508	180

180 rows x 9 columns



1. Import the dataset and do usual data analysis steps like checking the structure & characteristics of the dataset

```
df.head()
```

	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

```
len(df)
```

180

```
df.dtypes
```

Product	object
Age	int64
Gender	object
Education	int64

```

MaritalStatus    object
Usage            int64
Fitness          int64
Income           int64
Miles            int64
dtype: object

```

```
df.describe(include="all")
```

	Product	Age	Gender	Education	MaritalStatus	Usage	Fitness
<b>count</b>	180	180.000000	180	180.000000	180	180.000000	180.000000
<b>unique</b>	3	NaN	2	NaN	2	NaN	NaN
<b>top</b>	KP281	NaN	Male	NaN	Partnered	NaN	NaN
<b>freq</b>	80	NaN	104	NaN	107	NaN	NaN
<b>mean</b>	NaN	28.788889	NaN	15.572222	NaN	3.455556	3.311111
<b>std</b>	NaN	6.943498	NaN	1.617055	NaN	1.084797	0.958869
<b>min</b>	NaN	18.000000	NaN	12.000000	NaN	2.000000	1.000000
<b>25%</b>	NaN	24.000000	NaN	14.000000	NaN	3.000000	3.000000
<b>50%</b>	NaN	26.000000	NaN	16.000000	NaN	3.000000	3.000000
<b>75%</b>	NaN	33.000000	NaN	16.000000	NaN	4.000000	4.000000
<b>max</b>	NaN	50.000000	NaN	21.000000	NaN	7.000000	5.000000

```
df.info()
```

```

<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

```

```
df["Product"].value_counts()
```

```

KP281    80
KP481    60
KP781    40
Name: Product, dtype: int64

```

**Insights-**

1. There are 180 rows and 9 columns.
2. There are no null values in the provided data.
3. Minimum and Maximum age of the person is 18 & 50 mean is 28.79 and 75 % of the persons have age less than or equal to 33
4. Out of 180 data of gender, 104 persons are Male and rest are Female.
5. Most of the people are having 16 years of education i.e. 75% of persons are having education  $\leq 16$  years.
6. Product name KP281 is the most frequent product with values 80.
7. Frequency of Marital Status 'Partnered' is 107 out of 180.
8. There must be outliers in column Income & Miles as the standard deviation of these data are very high.

**Recommendations-**

1. Age and Education: Since the majority of individuals fall in the age range of 18 to 33 and have 16 years of education, it would be beneficial to design fitness programs and marketing strategies targeting this age group. Offering specialized fitness programs that cater to their needs and preferences can attract more customers.
2. Gender: With 104 males and the rest being females, it indicates a potential opportunity to explore ways to encourage more female participation. This could involve creating female-focused fitness classes, promoting a supportive and inclusive environment, and offering targeted promotional offers.
3. Product Strategy: The product "KP281" is the most frequently purchased one, with a value of 80. This product seems to be popular among customers. Analyzing the reasons behind its popularity can provide valuable insights into designing new products or improving existing ones to meet customer demands.
4. Marital Status: Since the frequency of "Partnered" individuals is 107 out of 180, consider offering family-oriented fitness packages or programs. Providing incentives for couples or families to work out together can increase customer retention and loyalty.
5. Outliers in Income and Miles: Given the high standard deviation in Income and Miles columns, it's likely that there are outliers in the data. These outliers can significantly impact statistical analysis and predictions. It is recommended to carefully examine the data to identify and handle outliers appropriately. Depending on the specific use case, these outliers may need to be removed, transformed, or analyzed separately.
6. Data Quality and Collection: Ensure that data collection methods are accurate and consistent to maintain the integrity of the dataset. Regularly check for missing or erroneous entries and address any issues promptly.

2. Detect Outliers (using boxplot, "describe" method by checking the difference between mean and median)

## ▼ Univariate Analysis

- Understanding the distribution of data for the quantitative attributes.

1. Age
  2. Education
  3. Usage
  4. Fitness
  5. Income
  6. Miles
- 

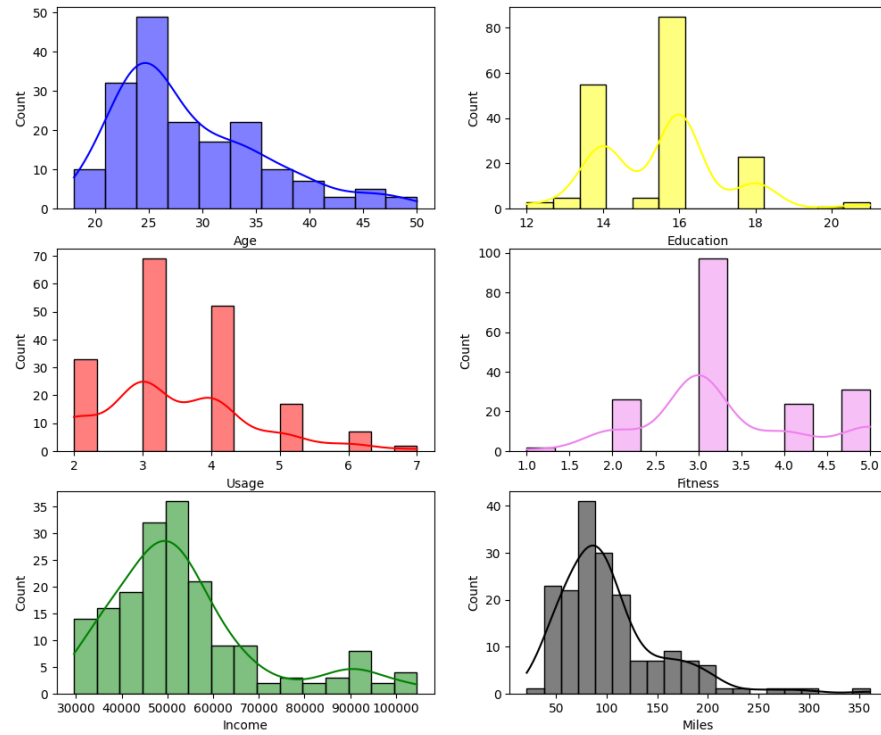
```
df.nunique()
```

```
Product      3
Age          32
Gender        2
Education     8
MaritalStatus 2
Usage         6
Fitness       5
Income       62
Miles        37
dtype: int64
```

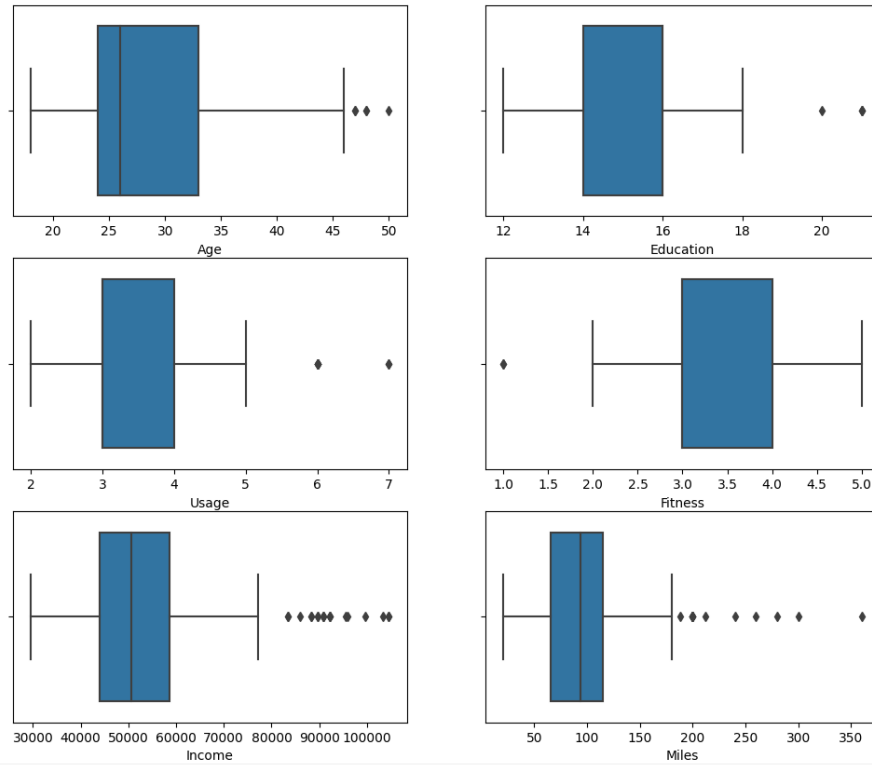
```
df["Product"].value_counts()
```

```
KP281    80
KP481    60
KP781    40
Name: Product, dtype: int64
```

```
fig, axis = plt.subplots(3, 2, figsize=(12, 10))
sns.histplot(data=df, x="Age", kde=True, ax=axis[0, 0], color="blue")
sns.histplot(data=df, x="Education", kde=True, ax=axis[0, 1], color="yellow")
sns.histplot(data=df, x="Usage", kde=True, ax=axis[1, 0], color="Red")
sns.histplot(data=df, x="Fitness", kde=True, ax=axis[1, 1], color="violet")
sns.histplot(data=df, x="Income", kde=True, ax=axis[2, 0], color="Green")
sns.histplot(data=df, x="Miles", kde=True, ax=axis[2, 1], color="black")
plt.show()
```



```
fig,axis = plt.subplots(3,2, figsize=(12,10))
sns.boxplot(data=df,x="Age",ax=axis[0,0])
sns.boxplot(data=df,x="Education",ax=axis[0,1])
sns.boxplot(data=df,x="Usage",ax=axis[1,0])
sns.boxplot(data=df,x="Fitness",ax=axis[1,1])
sns.boxplot(data=df,x="Income",ax=axis[2,0])
sns.boxplot(data=df,x="Miles",ax=axis[2,1])
plt.show()
```



```
df.describe(include="all")
```

	Product	Age	Gender	Education	MaritalStatus	Usage	Fitness
count	180	180.000000	180	180.000000	180	180.000000	180.000000
unique	3	NaN	2	NaN	2	NaN	NaN
top	KP281	NaN	Male	NaN	Partnered	NaN	NaN
freq	80	NaN	104	NaN	107	NaN	NaN
mean	NaN	28.788889	NaN	15.572222	NaN	3.455556	3.311111
std	NaN	6.943498	NaN	1.617055	NaN	1.084797	0.958869
min	NaN	18.000000	NaN	12.000000	NaN	2.000000	1.000000
25%	NaN	24.000000	NaN	14.000000	NaN	3.000000	3.000000
50%	NaN	26.000000	NaN	16.000000	NaN	3.000000	3.000000
75%	NaN	33.000000	NaN	16.000000	NaN	4.000000	4.000000
max	NaN	50.000000	NaN	21.000000	NaN	7.000000	5.000000

---

## Insights

---

1. From the boxplot, we can see that Income and Miles have more outliers and after that age has some.
  2. From the countplot,
    - we can see that maximum concentrated age of persons is 25.
    - we can see that maximum concentrated Education of persons is 16 years.
    - we can see that maximum concentrated income of persons is 50000.
  3.
    - Mean age-28.788889
    - Mean Education-15.57 years
    - Mean Miles-103.194444
    - Mean Income-53719.57 INR
  4.
    - std dev.age-6.94
    - std dev. Education-1.61
    - std dev. Miles-51.86
    - std dev. Income-16506.68
- 

## Recommendation

---

1. Handling Outliers: Since Income and Miles have a significant number of outliers, it's important to handle them appropriately. Outliers can greatly influence the statistical analysis and model predictions. Consider using robust statistical methods or transformations to mitigate the impact of outliers or consider removing them if they are data entry errors. However, be cautious when removing outliers, as they could represent genuine and valuable information.
2. Targeted Marketing Strategies: With the maximum concentrated age being 25 and maximum concentrated education being 16 years, consider designing targeted marketing campaigns to attract individuals in this age and education group. Offering promotions, discounts, or specialized fitness programs that cater to the preferences of this demographic can be effective.
3. Income-based Offers: Since the mean income is approximately 53719.57 INR, consider creating membership plans or offers that align with different income levels. For instance, you could have different tiers of membership with varying price points and benefits to accommodate individuals with different financial capacities.
4. Improving Education Levels: Although the mean education level is 15.57 years, consider offering educational workshops or classes to encourage individuals to pursue higher education. Promote the benefits of continuous learning and how it can positively impact their overall health and well-being.
5. Customer Retention: Given the standard deviation of age and income, there might be diverse customer segments. Focus on customer retention strategies to ensure that you cater to the varying needs of different age groups and income levels. Providing personalized fitness plans and excellent customer service can foster loyalty and encourage repeat visits.



6. Market Research: Conduct further market research to understand the reasons behind the concentration of individuals around certain age, education, income, and mileage levels. This can help in refining the offerings and tailoring services to better meet the specific needs of the target audience.
7. Data Consistency: Ensure data consistency and accuracy in the collection process to maintain reliable insights. Regularly update and analyze the data to stay informed about any changes in customer behavior and preferences.
8. Competitor Analysis: Monitor the offerings and strategies of competitors to identify gaps in the market and differentiate AeroFit from other fitness centers. Understanding the strengths and weaknesses of competitors can inform improvements and marketing approaches.

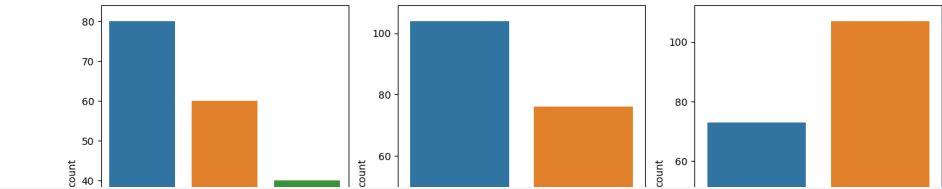
3. Check if features like marital status, age have any effect on the product purchased (using countplot, histplots, boxplots etc)

4. Representing the marginal probability like - what percent of customers have purchased KP281, KP481, or KP781 in a table (can use `pandas.crosstab` here)

- Understanding the distribution of the data for the qualitative attributes.

1. Product
2. Gender
3. Marital Status

```
Fig, axs= plt.subplots(1,3 , figsize=(15,6))
sns.countplot(data=df,x='Product',ax=axs[0])
sns.countplot(data=df,x='Gender',ax=axs[1])
sns.countplot(data=df,x='MaritalStatus',ax=axs[2])
plt.show()
```



```
df1= df[["Product", "Gender", "MaritalStatus"]].melt()  
df1.groupby(['variable', 'value'])['value'].count()/len(df)
```

		value	
variable	value		
Gender	Female	0.422222	
	Male	0.577778	
MaritalStatus	Partnered	0.594444	
	Single	0.405556	
Product	KP281	0.444444	
	KP481	0.333333	
	KP781	0.222222	

Insights-

1. Product-
- 44.44% of the customers have purchased KP281 product.
  - 33.33% of the customers have purchased KP481 product.
  - 22.22% of the customers have purchased KP781 product.
2. Gender- 57.78% of the customers are male and rest are female.
3. Marital status-59.44% of the customers are partnered.

Recommendations-

1. Product Promotion: Since KP281 is the most popular product with 44.44% of customers purchasing it, focus on promoting this product even further. Create targeted marketing campaigns that highlight the unique benefits and features of KP281. Consider offering special discounts or bundle deals for KP281 to encourage more customers to try it.
2. Product Diversification: While KP281 is popular, it's essential to diversify the product offerings to appeal to a broader customer base. Introduce new products or variations of existing ones to cater to different fitness preferences and goals. Conduct market research to identify potential gaps in the product lineup and innovate accordingly.
3. Gender-Specific Programs: With 57.78% of customers being male, consider creating gender-specific fitness programs and marketing them accordingly. For instance, offer male-focused fitness classes or workshops that address the specific health and fitness needs of men. Similarly, create female-focused programs to attract and retain more female customers.

4. Inclusive Marketing: While gender-specific programs can be beneficial, ensure that overall marketing efforts remain inclusive. Promote a diverse and inclusive atmosphere in the fitness center, where people of all genders feel welcome and comfortable. Avoid gender stereotypes in marketing materials and focus on promoting health and fitness as a universal goal.
5. Couples Fitness Classes: With 59.44% of customers being partnered, introduce couples fitness classes or partner workout sessions. These classes can enhance the fitness journey for couples, providing a fun and motivating way to exercise together. Offer incentives for couples to join, such as discounted rates or exclusive offers for partners.

---

## ▼ Bivariate Analysis

Checking if features- Gender or Marital status have any effect on product purchased.

---

```
sns.set_style(style='whitegrid')
fig, axs= plt.subplots(1,2,figsize=(15,6.5))
sns.countplot(data=df,x="Product",hue="Gender",palette=["#7fcdbb","#edf8b1"],ax=axs[0])
axs[0].set_title("Product vs Gender")
sns.countplot(data=df,x="Product",hue="MaritalStatus",palette=['#edf8b1','#FAAE7B'],ax=axs[1])
axs[1].set_title("Product vs MaritalStatus")
plt.show()
```

Product vs Gender

Product vs MaritalStatus

**Insights-****Product vs Gender**

1. Equal number of Males and Females have purchased KP281 product and almost same for the product KP481.
2. Most of the male customers have purchased the KP781.

**Product vs Marital Status**

1. Customers who is Partnered, is more likely to purchase the product and it is true for all the products

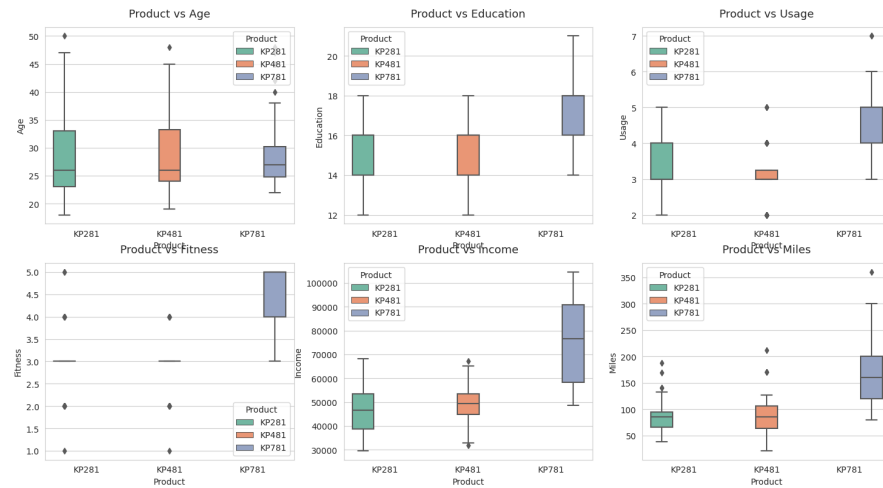
**Recommendations-****Product vs Gender:**

1. Gender-Specific Marketing: While the number of males and females purchasing KP281 and KP481 products seems to be equal, consider tailoring marketing efforts based on the specific preferences and needs of each gender.
2. Understanding Product Appeal: Analyze why KP281 and KP481 products have an equal number of male and female customers. Conduct surveys or focus groups to understand what aspects of these products appeal to both genders.
3. Promoting KP781 to Males: As most male customers have purchased KP781, leverage this trend to design targeted marketing campaigns to further boost sales of this product among male clientele.

**Product vs Marital Status:**

1. Partnered Customer Programs: Since customers who are partnered are more likely to purchase products, consider offering exclusive partner programs or incentives. Create fitness packages designed for couples that provide a shared fitness experience.
2. Promote Health as a Shared Goal: In marketing materials and messaging, emphasize the benefits of working out and staying healthy together as a couple.
3. Family-Oriented Programs: Extend the focus beyond partnered individuals to include families. Design fitness programs that cater to families and promote a healthy lifestyle for all members.

```
var= ['Age', 'Education', 'Usage', 'Fitness', 'Income', 'Miles']
sns.set_style("whitegrid")
fig,axs=plt.subplots(2,3,figsize=(18,6))
fig.subplots_adjust(top=1.3)
count=0
for i in range(2):
    for j in range(3):
        sns.boxplot(data=df,x='Product',y=var[count],ax=axs[i,j],hue='Product',palette="Set2")
        axs[i,j].set_title(f"Product vs {var[count]}",pad=12,fontsize=13)
        count+=1
```



## Insights-

### 1. Product vs Age

- Customers purchasing products KP281 & KP481 are having same age median value.
- Customers whose age lies between 25-30, are more likely to buy KP781 product.

### 2. Product vs Education

- Customers whose education is greater than 16, have more chances to purchase the kp781 product.
- While the customers with education less than 16 have equal chances of purchasing kp281 or kp481.

### 3. Product vs Usage

- Customers who are planning to use the treadmill greater than 4 times a week, are more likely to purchase the kp781 product

b. While the other customers are likely to purchasing kp281 or kp481.

#### 4.Product vs Fitness

a. The more the customer is fit (fitness $\geq$  3), higher the chances of the customer to purchase the kp781 product

#### 5.Product vs Income

a.Higher the income of the customer (income  $\geq$  60000), higher the chances of the customer to purchase the kp781 product

#### 6.Product vs Miles

aif the customer expects to walk/run greater than 120 miles per week, it is more likely that the customer will buy kp781 product.

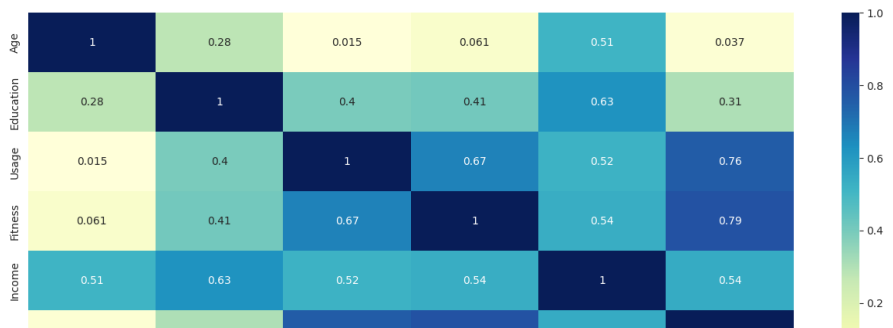
---

#### Recommendations-

1. Product vs Age: Since customers purchasing KP281 and KP481 have the same age median value, consider creating joint marketing campaigns that appeal to this specific age group. Highlight how both products can cater to their fitness goals and preferences.
  2. Product vs Education: To target customers with education greater than 16, promote the unique benefits of KP781 that align with their higher educational background. For customers with education less than 16, emphasize the versatility and suitability of both KP281 and KP481 products.
  3. Product vs Usage: Create marketing materials that showcase how KP781 is suitable for customers who plan to use the treadmill more than 4 times a week. Highlight features that cater to frequent usage and durability to attract this segment.
  4. Product vs Fitness: Focus on promoting KP781 to customers with a fitness level of 3 or higher. Highlight how this product can enhance their fitness routine and assist them in achieving their fitness goals more effectively.
  5. Product vs Income: Target customers with income greater than or equal to 60000 and emphasize the value and benefits of KP781. Showcase how this product provides excellent features and functionalities that justify its price.
  6. Product vs Miles: Highlight the advantages of KP781 for customers who expect to walk/run more than 120 miles per week. Showcase its durability, performance, and suitability for intensive workout routines.
- 

```
plt.figure(figsize=(16, 6))
sns.heatmap(df.corr(), cmap="YlGnBu", annot=True)
plt.show()
```

```
<ipython-input-32-46cee23e128a>:2: FutureWarning: The default value of numeric_
sns.heatmap(df.corr(), cmap="YlGnBu", annot=True)
```



Insights

- 1. Miles and Fitness and Miles and Usage are highly correlated, which means if a customer's fitness level is high they use more treadmills.
- 2. Income and education show a strong correlation. High-income and highly educated people prefer high-end models (KP798), as mentioned during Bivariant analysis of Categorical variables.
- 3. There is no corelation between Usage & Age or Fitness & Age which mean Age should not be barrier to use treadmills or specific model of treadmills.

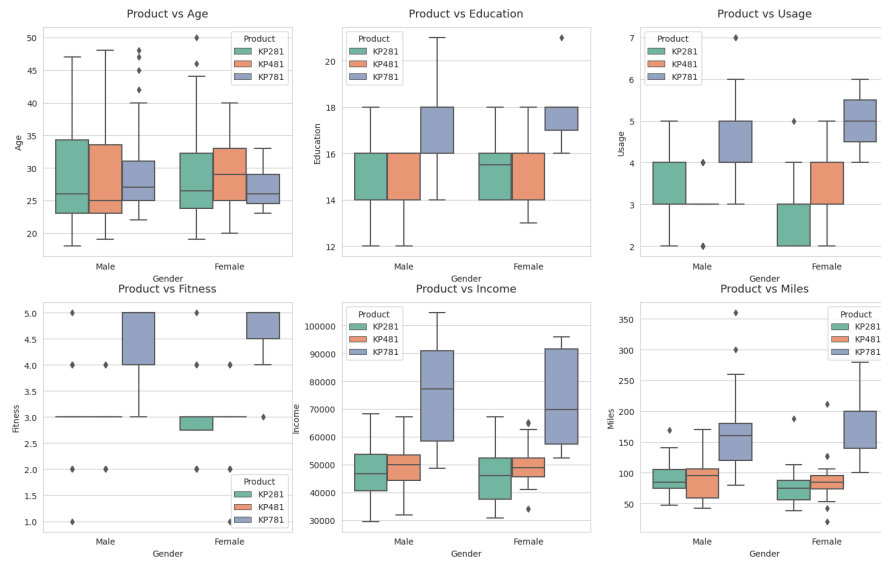
Recommendations

- 1. Promote Treadmill Usage Across Age Groups: Highlight the correlation between Age and Treadmill Usage to dispel age-related barriers. Encourage customers of all age groups to utilize treadmills as a suitable fitness option.
- 2. Target High-Income, Highly Educated Customers: Capitalize on the strong correlation between Income and Education by promoting high-end treadmill models like KP798 to this specific customer segment. Showcase the advanced features and benefits that cater to their preferences.
- 3. Diversify Fitness Programs: Offer a diverse range of fitness programs and classes to cater to customers with varying fitness levels and goals. Create a welcoming and inclusive environment that motivates individuals of all backgrounds to achieve their fitness aspirations with AeroFit.

▼ Multivariate Analysis

```
var= ['Age', 'Education', 'Usage', 'Fitness', 'Income', 'Miles']
sns.set_style("whitegrid")
fig,axs=plt.subplots(2,3,figsize=(18,7))
fig.subplots_adjust(top=1.3)
count=0
for i in range(2):
```

```
for j in range(3):
    sns.boxplot(data=df, x='Gender', y=var[count], hue='Product', ax=axis[i, j], palette="Set2")
    axis[i, j].set_title(f"Product vs {var[count]}", pad=12, fontsize=13)
    count+=1
```





## ▼ Marginal Probability

```
pd.concat( [ df.Product.value_counts(), df.Product.value_counts(normalize=True) ], keys=['counts', 'Marginal_Prob'], axis=1, )
```

	counts	Marginal_Prob
KP281	80	0.444444
KP481	60	0.333333
KP781	40	0.222222

## ▼ Conditional Probability

Probability of each product given gender.

```
def p_prod_given_gender(gender, print_marginal=False):
    if gender!= "Female" and gender!= "Male":
        return "Invalid gender value."
    df1= pd.crosstab(index=df['Gender'],columns=[df['Product']])
    p_781= df1['KP781'][gender] / df1.loc[gender].sum()
    p_481= df1['KP481'][gender] / df1.loc[gender].sum()
    p_281= df1['KP281'][gender] / df1.loc[gender].sum()
    if print_marginal:
        print(f"P(Male): {df1.loc['Male'].sum()/len(df):.2f}")
        print(f"P(Female): {df1.loc['Female'].sum()/len(df):.2f}")
        print(f"P(KP781/{gender}):{p_781:.2f}")
        print(f"P(KP481/{gender}):{p_481:.2f}")
        print(f"P(KP281/{gender}):{p_281:.2f}\n")
    p_prod_given_gender('Male',True)
    p_prod_given_gender('Female')
```

```
P(Male): 0.58
P(Female): 0.42
P(KP781/Male):0.32
P(KP481/Male):0.30
P(KP281/Male):0.38
```

Probability of each product given marital status

```
def p_prod_given_MaritalStatus(status, print_marginal=False):
    if status!= "Single" and status!= "Partnered":
        return "Invalid MaritalStatus value."
    df1= pd.crosstab(index=df['MaritalStatus'],columns=[df['Product']])
    p_781= df1['KP781'][status] / df1.loc[status].sum()
    p_481= df1['KP481'][status] / df1.loc[status].sum()
    p_281= df1['KP281'][status] / df1.loc[status].sum()
    if print_marginal:
```

```
print(f"P(Single): {df1.loc['Single'].sum()/len(df):.2f}")
print(f"P(Partnered): {df1.loc['Partnered'].sum()/len(df):.2f}\n")
print(f"P(KP781/{status}):{p_781:.2f}")
print(f"P(KP481/{status}):{p_481:.2f}")
print(f"P(KP281/{status}):{p_281:.2f}\n")
p_prod_given_MaritalStatus('Single', True)
p_prod_given_MaritalStatus('Partnered')
```

```
P(Single): 0.41
P(Partnered): 0.59
```

```
P(KP781/Single):0.23
P(KP481/Single):0.33
P(KP281/Single):0.44
```

---

## Insights

1. In both Gender customers whose education is greater than 16(Education>= 16) prefer to buy KP781 product.
2. In both Gender customer who are planning to use treadmill more than four times (Usage>=4), prefer to buy KP781 product
3. Females who are planning to use treadmill 3-4 times a week, are more likely to buy KP481 product.
4. In both Gender, customer whose Income is more than 55000 are more likely to buy KP781 product.
5. The probability of a male customer buying a KP781 treadmill is 0.32.
6. The probability of a male customer buying a KP481 treadmill is 0.30.
7. The probability of a male customer buying a KP281 treadmill is 0.38.
8. The probability of a customer buying a KP281 treadmill and are single is 0.44.
9. The probability of a customer buying a KP481 treadmill and are single is 0.33.
10. The probability of a customer buying a KP781 treadmill and are single is 0.23.

---

## Recommendations

1. Product vs Gender and Education: Target customers of both genders who have an education level greater than 16 (Education >= 16) with promotions and marketing campaigns that emphasize the benefits of KP781. Highlight how this product aligns with their higher educational background and offers advanced features to support their fitness goals.
  2. Product vs Gender and Usage: Focus on customers of both genders who plan to use the treadmill more than four times a week (Usage >= 4). Showcase how KP781 is designed to withstand frequent usage and offers the desired functionalities that can enhance their workout experience.
  3. Product vs Gender and Usage for Females: For female customers who are planning to use the treadmill 3-4 times a week, target marketing efforts towards promoting KP481. Highlight the specific features of KP481 that cater to their fitness routines and preferences.
  4. Product vs Gender and Income: Attract customers of both genders with an income higher than 55000 by promoting KP781. Emphasize how this product provides excellent value for their investment and offers top-of-the-line features to enhance their fitness journey.
-

---

✓ 1s completed at 23:04

