AEROFIT TREADMILL CUSTOMER PROFILE ANALYSIS

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.

Basic Metrics :

```
1  data.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
```

```
1  data.shape
```

```
(180, 9)
```

```
1  data.isna().sum()
```

```
Product          0
Age              0
Gender           0
Education        0
MaritalStatus    0
Usage            0
Fitness          0
Income           0
Miles            0
dtype: int64
```

- Dataset has 180 records with 9 features for each record as it can be seen in the shape of data (180,9).

- There are no missing values in any of the columns.

- We have three columns with object datatype and 6 columns with integer datatype.

```
1  data['Product'].value_counts()
```

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

```
1  data['MaritalStatus'].value_counts()
```

```
Partnered    107
Single        73
Name: MaritalStatus, dtype: int64
```

```
1  data['Gender'].value_counts()
```

```
Male      104
Female     76
Name: Gender, dtype: int64
```

```
1  data['Usage'].value_counts()
```

```
3    69
4    52
2    33
5    17
6     7
7     2
Name: Usage, dtype: int64
```

- Only three type of treadmills are sold to customers with 'KP281' the most frequently bought.

- 104 Males have purchased the treadmills compared to 76 Females.

- Partnered customers are high in number when compared to Single customers.

- It can be seen that customers use the treadmills mostly for 3 days a week and only 2 customers use it for all 7 days of a week.

```
1  data.describe()
```

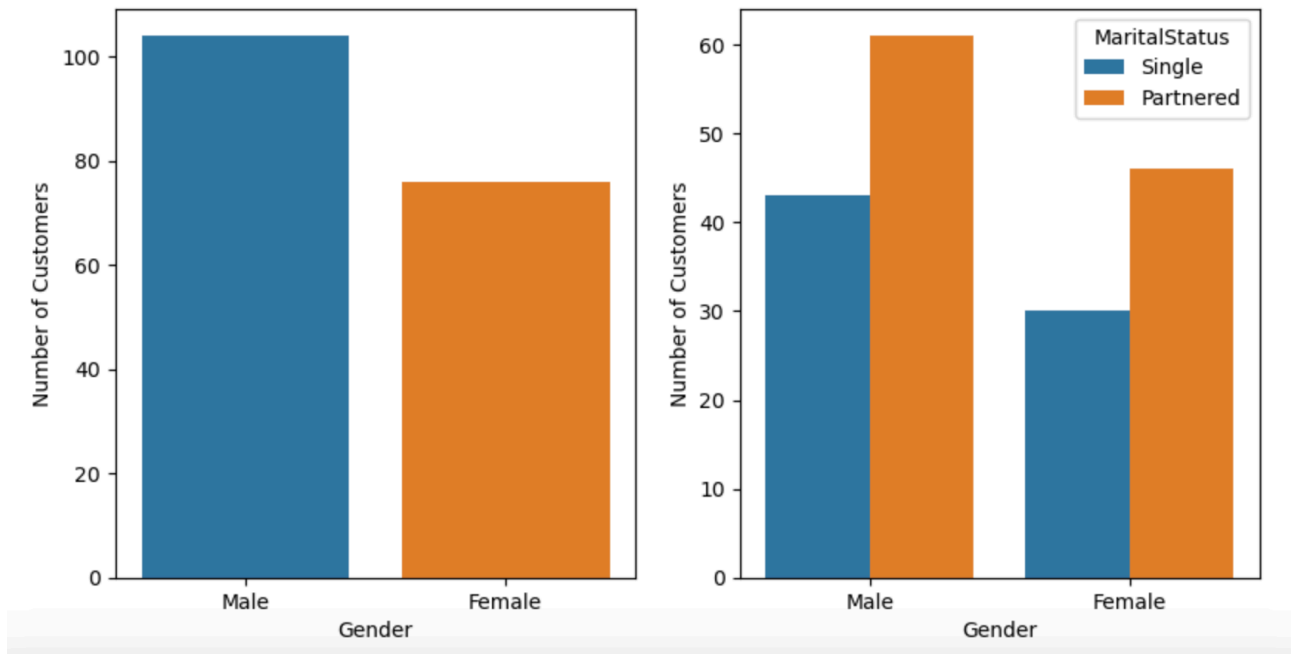|	Age	Education	Usage	Fitness	Income	Miles
count	180.000000	180.000000	180.000000	180.000000	180.000000	180.000000
mean	28.788889	15.572222	3.455556	3.311111	53719.577778	103.194444
std	6.943498	1.617055	1.084797	0.958869	16506.684226	51.863605
min	18.000000	12.000000	2.000000	1.000000	29562.000000	21.000000
25%	24.000000	14.000000	3.000000	3.000000	44058.750000	66.000000
50%	26.000000	16.000000	3.000000	3.000000	50596.500000	94.000000
75%	33.000000	16.000000	4.000000	4.000000	58668.000000	114.750000
max	50.000000	21.000000	7.000000	5.000000	104581.000000	360.000000

- data.describe() gives a descriptive Statistical analysis on the data.

This gives us the range of values in each numerical columns, the mean of the data, different percentile values , median and the standard deviation.

- Mean Age of customers is 28.7 and the median age is 26 (50 Percentile).

- Youngest customer is 18 years old whereas the oldest customer is 50 years old.

- Average fitness rating of customers is 3.3 and customers use the treadmills on an average of 3.45 days per week.

- 50 Percentile of the Total miles covered in treadmills is less than 94 miles and the mean is 103 miles.
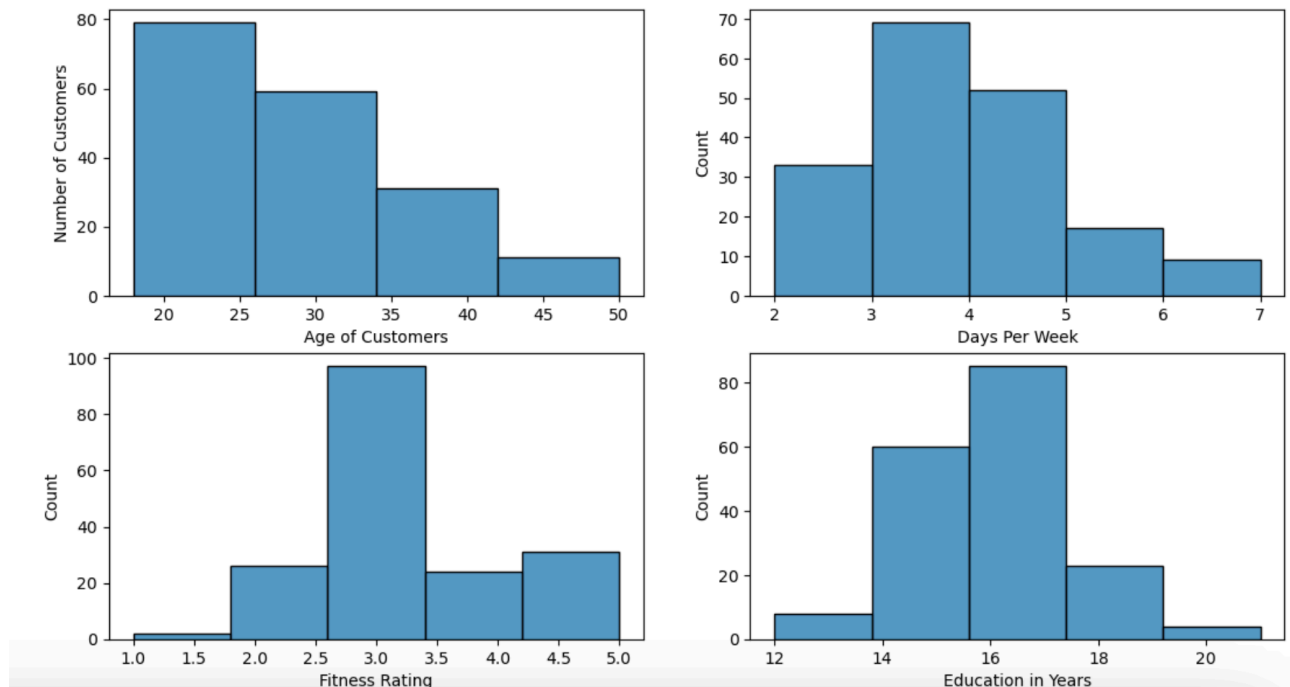
- Younger customers are more in number than older customers. Customers in the age group of 18-26 are high in number when compare to other age groups.

- Customers use the treadmills mostly for 3-5 days per week. Very few customers use it for more than 5 days.

- Very less customers have a fitness rating less than 2 and 3 is the fitness rating for high customer count.

- 14 - 17 years is the highest number of customers's education.

- Male Customers buy more treadmills than female customers.

- Also in both male and female category, the customers who are partnered prefer to buy more treadmills than single customers.


Product Category Splitup
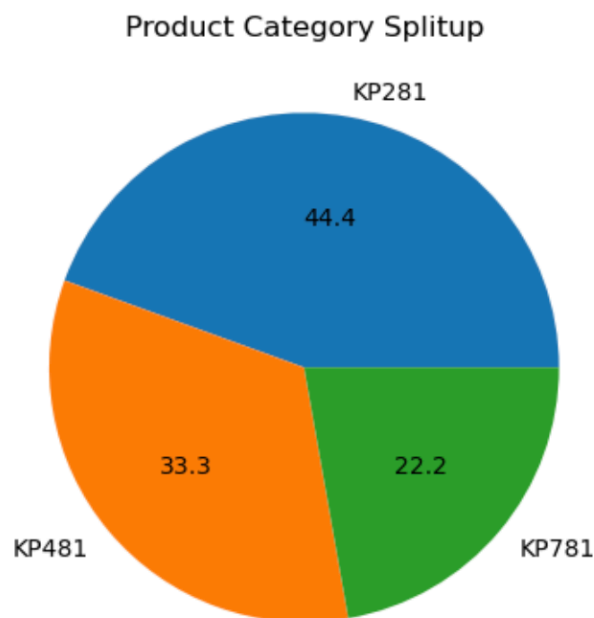
- KP 281 is the favourite product for customers closely followed by KP481.

- KP781 is the least bought treadmill type.

```
1  data.describe()
```

	Age	Education	Usage	Fitness	Income	Miles
count	180.000000	180.000000	180.000000	180.000000	180.000000	180.000000
mean	28.788889	15.572222	3.455556	3.311111	53719.577778	103.194444
std	6.943498	1.617055	1.084797	0.958869	16506.684226	51.863605
min	18.000000	12.000000	2.000000	1.000000	29562.000000	21.000000
25%	24.000000	14.000000	3.000000	3.000000	44058.750000	66.000000
50%	26.000000	16.000000	3.000000	3.000000	50596.500000	94.000000
75%	33.000000	16.000000	4.000000	4.000000	58668.000000	114.750000
max	50.000000	21.000000	7.000000	5.000000	104581.000000	360.000000

- For the total miles covered by each customer,

  - Actual Minimum : 21

  - Actual Maximum : 360

  - 25th Percentile : 66 (Q1)

  - 50th Percentile : 94 (Q2)

  - 75th Percentile : 114.75 (Q3)

  - IQR = Q3-Q1 => 48.75

  - Minimum Considered : Q1 - 1.5*IQR (or) Actual Minimum. Here Actual Minimum is considered.

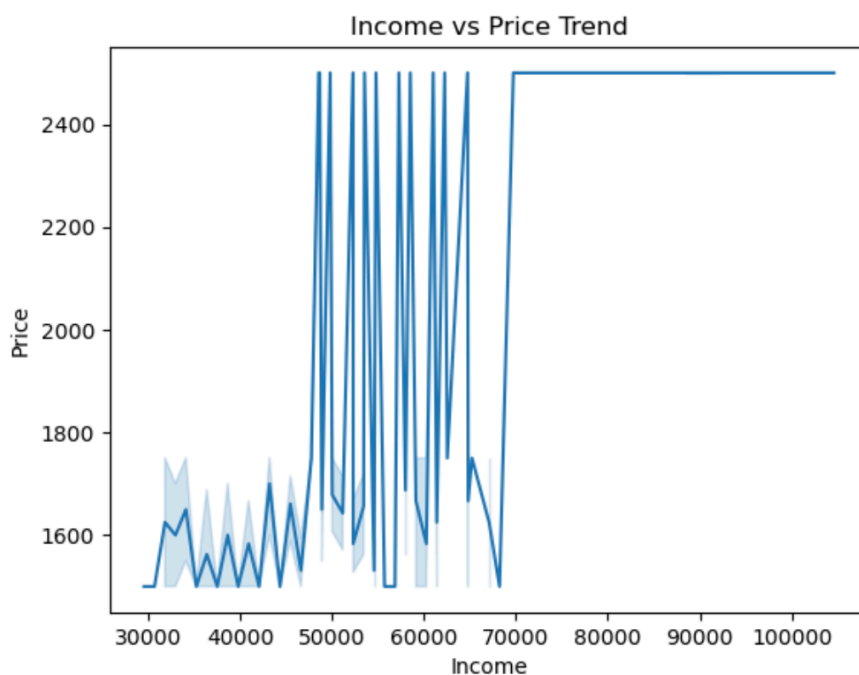Miles Outlier Detection
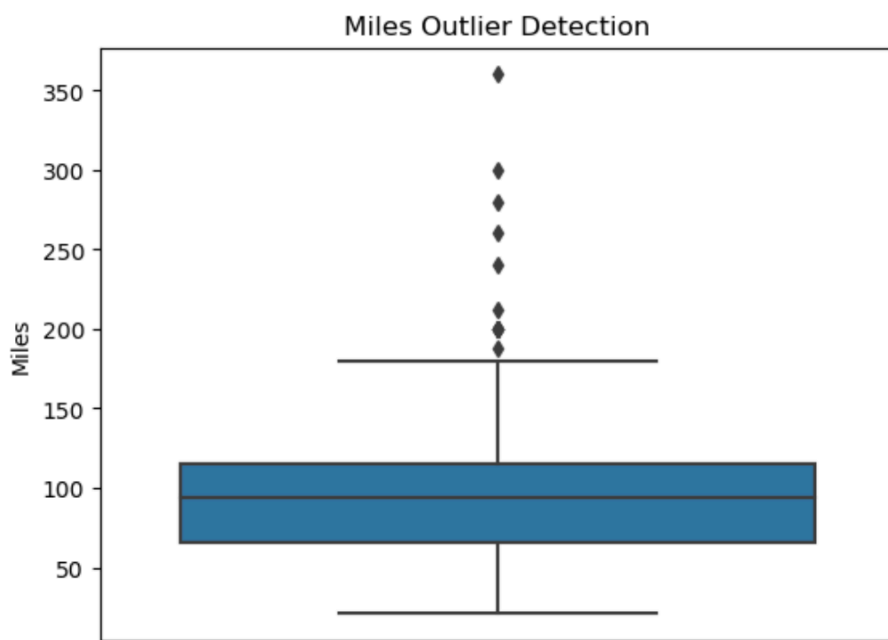
- Maximum Considered : Q3 + 1.5*IQR (or) Actual Maximum.

- Q3 + 1.5*IQR = 187.5

- Hence any value above 187.5 miles is considered as outliers in this data.
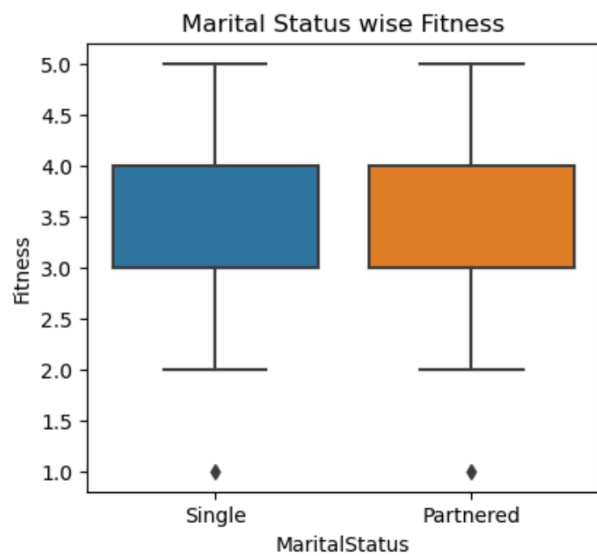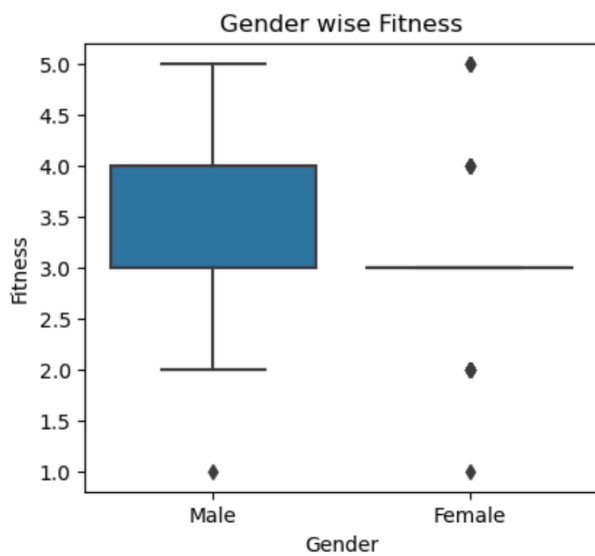
```
1  (data['Miles']>187.5).sum()
```
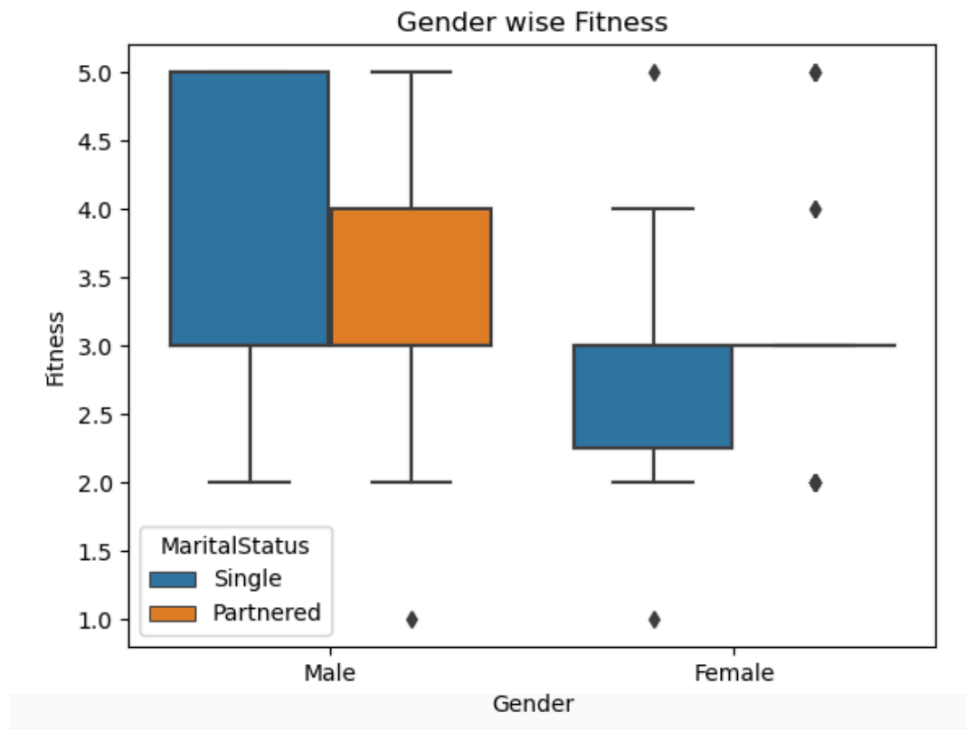
13

- 13 / 180 records are considered outliers in the miles column.


Income vs Price Trend

- As the income of the customers increase, the price of treadmills also increase which means, high earning customers prefer the treadmill with advanced features.

Gender wise Fitness

Gender wise Fitness

Marital Status wise Fitness

- Fitness rating of female customers is very low when compared to the male customers.

- Single and Partnered customers have similar Fitness rating.

Product Splitup across the Genders — Product Splitup based on Marital Status
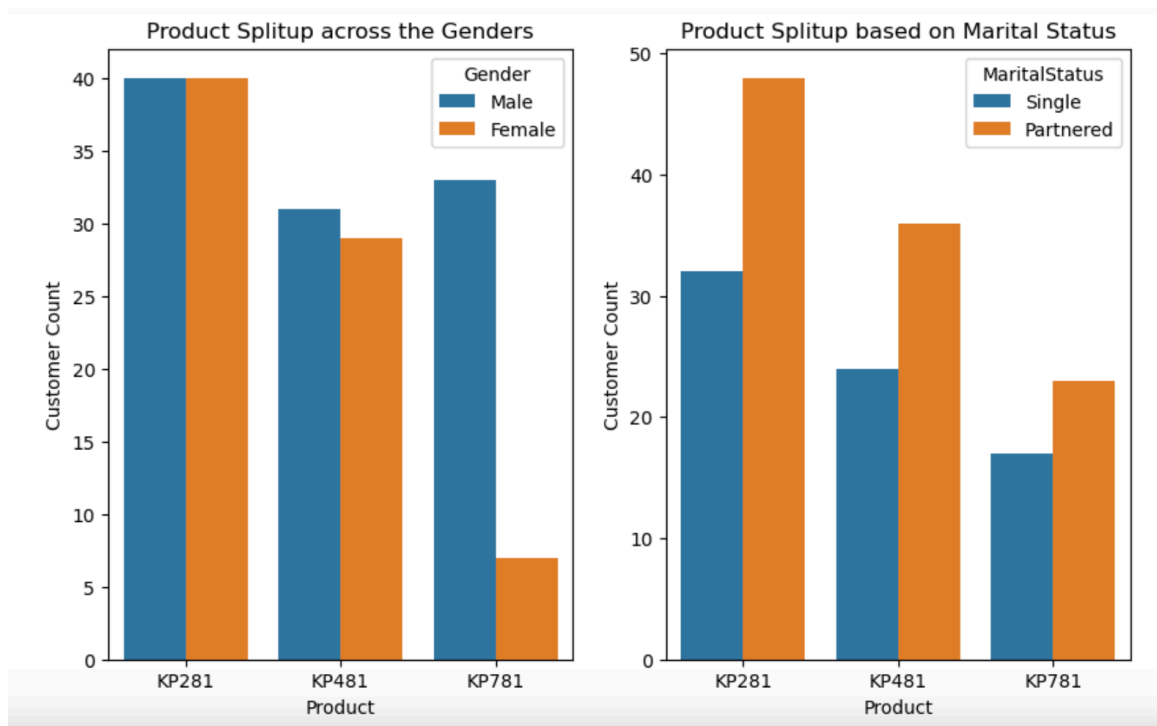
- Entry level (KP281) and Midlevel (KP481) treadmills are equally preferred by male and female customers but KP781 with advanced features is mainly preferred by male customers.

- Partnered customers are more in all the three categories of treadmills.

Miles Outlier Detection across Genders
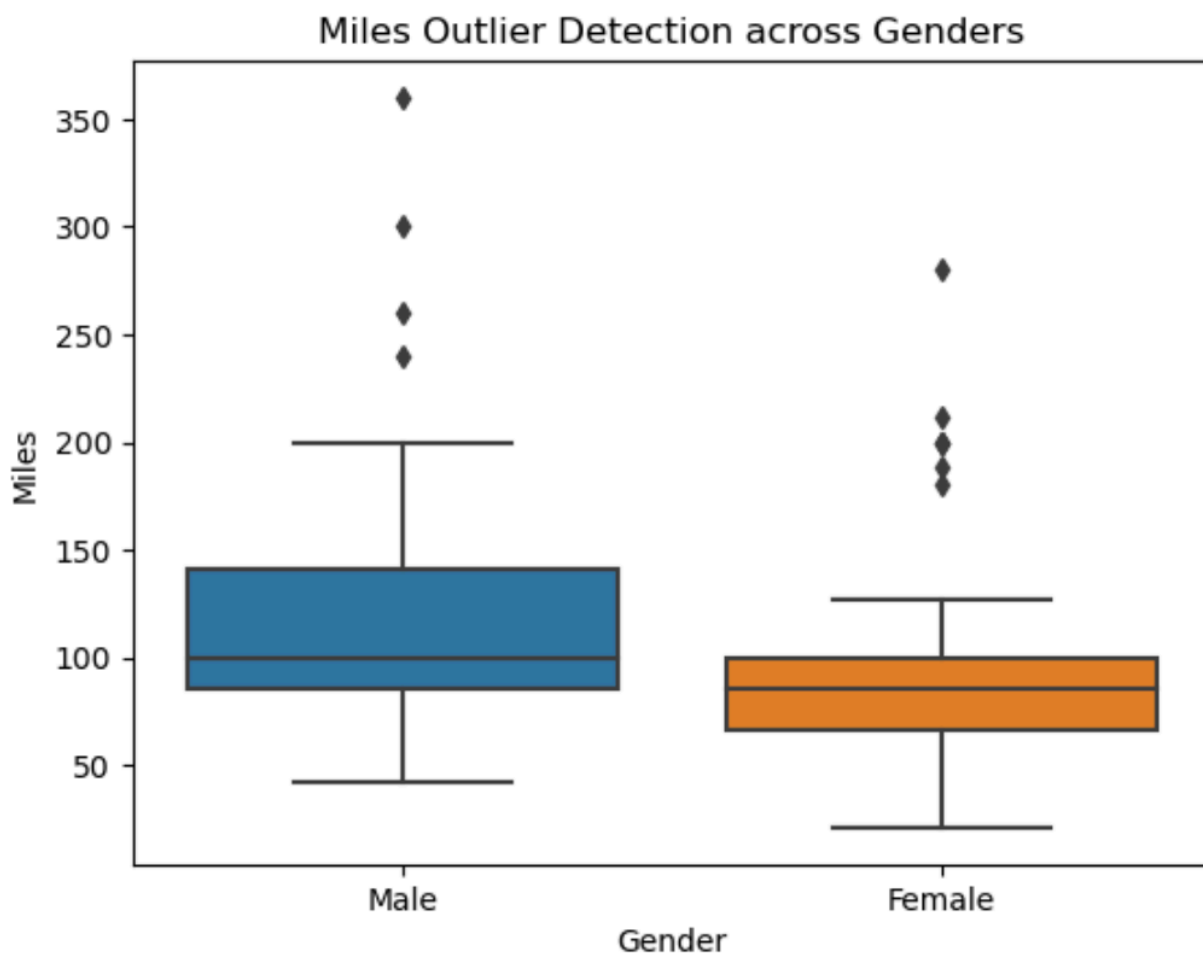
- The Median of miles covered by male customers is slightly higher than those of female customers.

- Very few female customers cover more than 150 miles and they lie in the outliers zone.

- Male customers cover distance in the range of 50 - 200 miles.
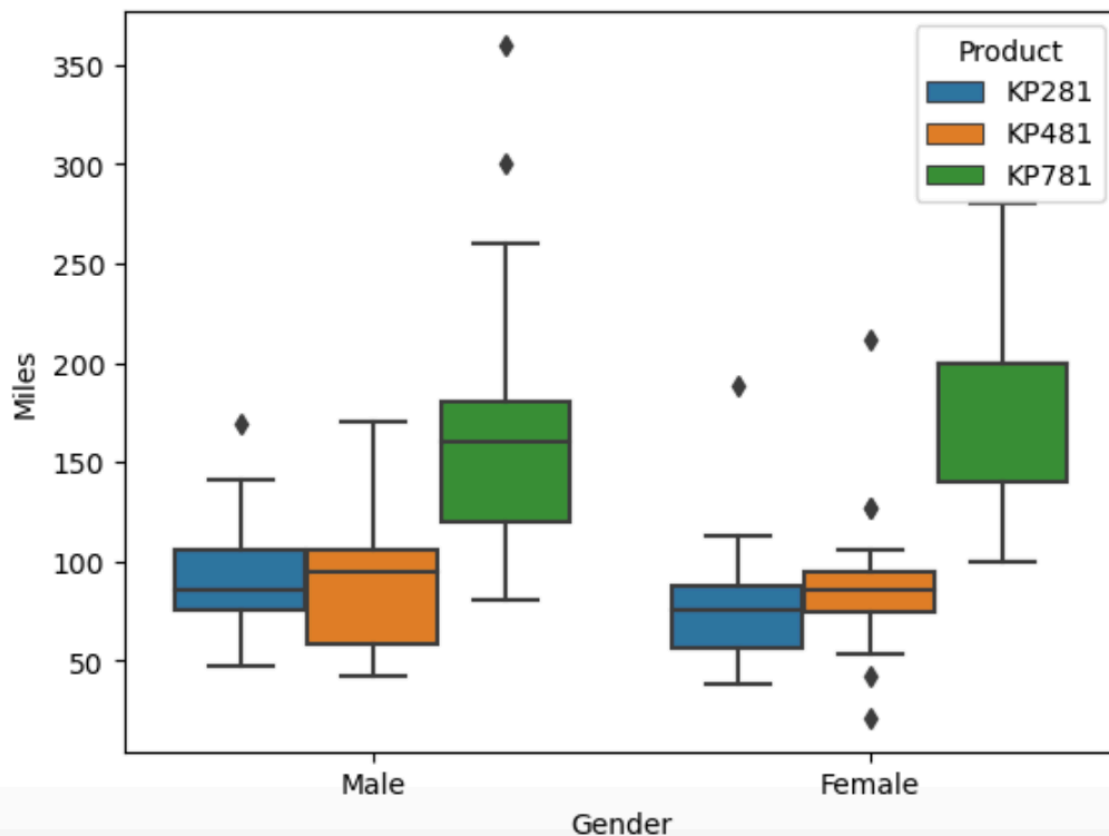

- Customers who cover more miles tend to prefer the advanced treadmill KP781.

- The Median value of miles covered by Both male and female customers increases as the level of treadmill increases.
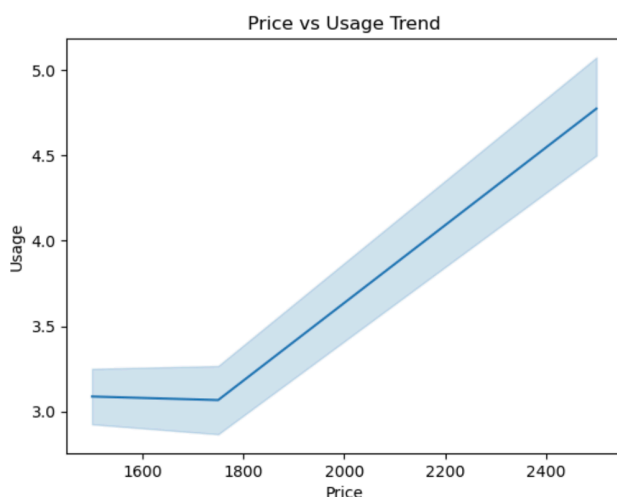
- Usage of the treadmills is increasing as the level of treadmills increases.

- Advance treadmill with higher price has more usage than Entry level treadmills with lower price.
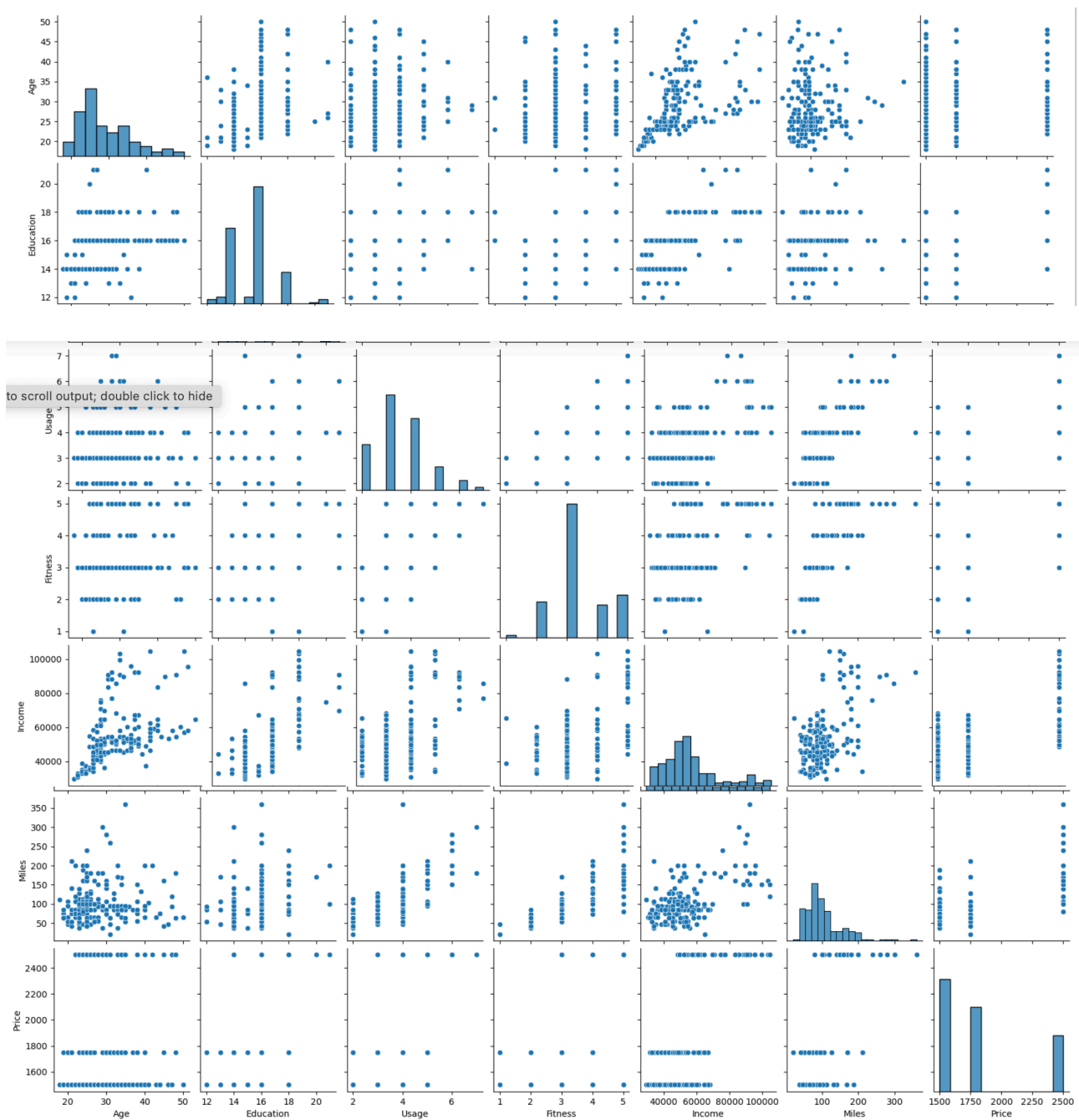
- The above pair plot shows the different relationship between the numerical columns in the data.

- Age vs Income shows, Younger people have less income when compared to older people which is logical.

- Fitness vs Miles shows, people with more fitness rating cover more miles using the treadmills.
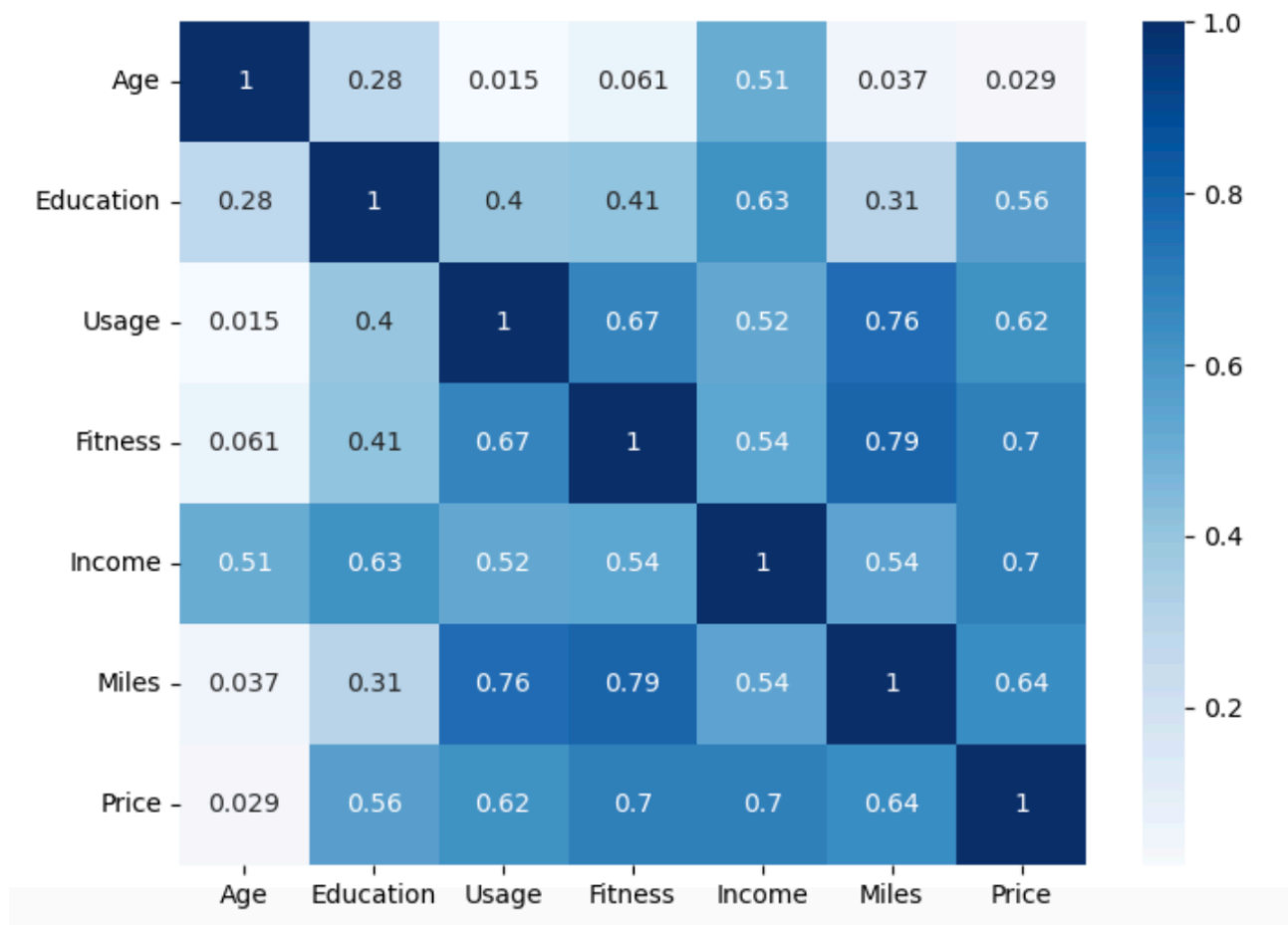
- From the above heat map, we can understand the correlation between the different Numerical Parameters in the data.

- Fitness and Miles Covered have a correlation value of 0.79 which indicates they are strongly positively correlated.

- Usage and Miles covered, fitness and Usage are also positively corelated.

- Age and Usage, Fitness and Age, Age and Miles are some of the pairs which have no correlation.

- Price and Income are positively correlated as more than income, people can buy higher price treadmills.

Gender	Female	Male	Total
MaritalStatus			
Partnered	25.555556	33.888889	59.444444
Single	16.666667	23.888889	40.555556
Total	42.222222	57.777778	100.000000

- Out of total customers, 25.56 % are partnered female, 33.89 % are partnered male, 16.67 % are single female and 23.8 % are single male.

- Therefore Partnered customers are almost 60% of total customers.

Gender	Female	Male	Total
Product			
KP281	22.222222	22.222222	44.444444
KP481	16.111111	17.222222	33.333333
KP781	3.888889	18.333333	22.222222
Total	42.222222	57.777778	100.000000

- Entry Level Treadmill - KP281 has been bought by 44.4% of customers.

- Male and female customers are almost equal in purchasing KP281.

- 33.33 % customers have purchased Mid level - KP481. Here also male and female probability is almost the same.

- KP781 has been purchased by 22.22 % of customers which is the lowest among the three types.

- Huge difference between male and female in buying the higher end treadmill with 90% of KP781 customers being male.

```
1  pd.crosstab(index = data['Product'],
2              columns = data['Income_Category'],
3              margins=True,
4              normalize=True)*100
```

Income_Category	High	Low	Medium	All
Product				
KP281	6.111111	18.888889	19.444444	44.444444
KP481	6.666667	8.333333	18.333333	33.333333
KP781	17.222222	0.000000	5.000000	22.222222
All	30.000000	27.222222	42.777778	100.000000

- Customers are categorised as low income (<=45000) , medium income (45000 to 55000) and high income (> 55000).

- Probability of KP281 being bought is :

    - Low : 0.18

    - Medium : 0.19

    - High : 0.06

- Probability of KP481 being bought is :

    - Low : 0.08

    - Medium : 0.18

    - High : 0.06

- Probability of KP681 being bought is :

    - Low : 0

    - Medium : 0.05

    - High : 0.17

Income_Category	High	Low	Medium	All
Gender				
Female	9.444444	13.333333	19.444444	42.222222
Male	20.555556	13.888889	23.333333	57.777778
All	30.000000	27.222222	42.777778	100.000000

- Male customers have higher percentage in high income category and medium category where as low income category the male and female split up is almost equal.
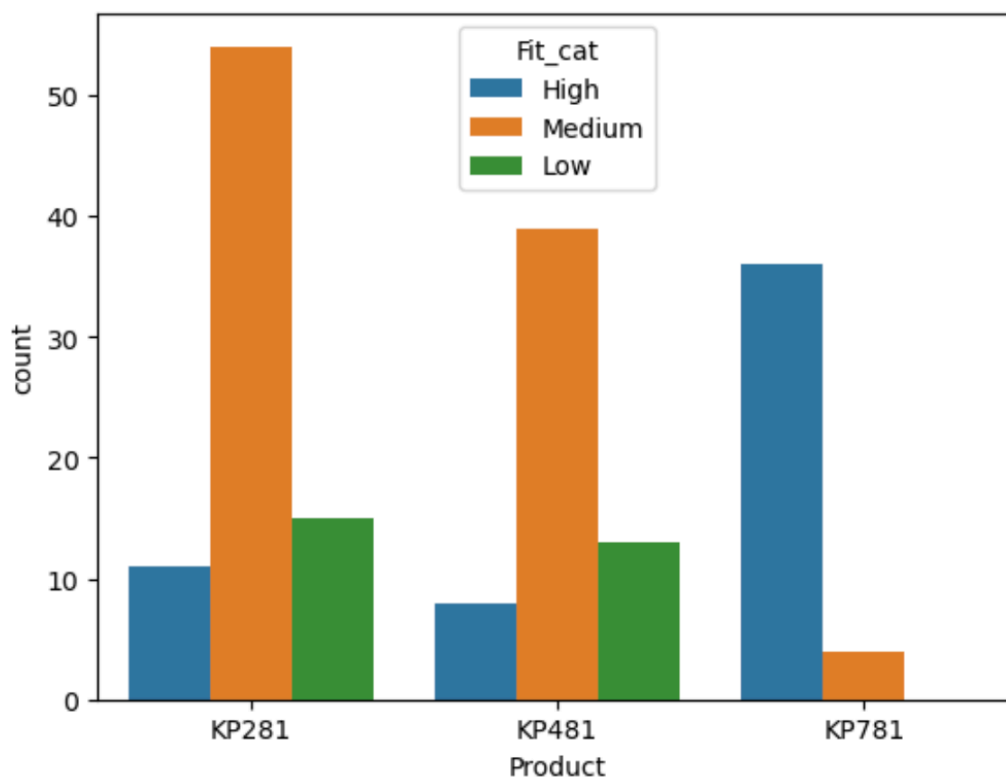
- Low Fitness : 1,2 Medium Fitness : 3 High Fitness: 4,5

- High Fitness customers prefer high level treadmills.

KP281 - Entry Level Treadmill - $1500 : 80 Customers

• Suitable for Low Income Category Individuals.

• Both Male and Female equally prefer this type.

• Low and Medium Fit customers also prefer this type of treadmill.

KP481 - Middle Level Treadmill - $1750 : 60 Customers

• Suitable for Low and Medium Income Category Individuals.

• Male customers are more in this type of treadmill than female customers.

• Low and Medium Fit customers also prefer this type of treadmill.

KP781 - High Level Treadmill - $2500 : 40 Customers

• Suitable for High Income Category Individuals.

• Male customers prefer this type of treadmill.

• High Fit customers prefer this type of treadmill.