In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
df=pd.read_csv("aerofit_analysis.csv")
df
```

Out[2]:

	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 × 9 columns

In [3]:
```
df.head()
```

Out[3]:

	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

In [4]:
```
# Shape of Data
df.shape
```

Out[4]:
```
(180, 9)
```

In [5]:
```
# Describe of Data
df.describe()
```

Out[5]:

	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

In [6]:
```
# NULL Values
df.isnull().sum()
```

Out[6]:
```
Product          0
Age              0
Gender           0
Education        0
MaritalStatus    0
Usage            0
Fitness          0
Income           0
Miles            0
dtype: int64
```

In [7]:
```
# Value_Counts
df['Product'].value_counts(normalize=True)
```

Out[7]:
```
Product
KP281    0.444444
KP481    0.333333
KP781    0.222222
Name: proportion, dtype: float64
```

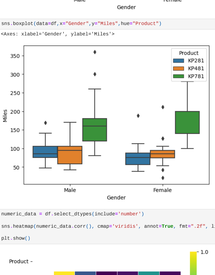
In [8]:
```
# Unique Values
df.nunique()
```

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

In [9]:
```
# counting Male and Female
ax=sns.countplot(data=df,x="Gender")
for i in ax.containers:
    ax.bar_label(i)
```

In [10]:
```
# Males Have More Numbers of Peoples
```

In [11]:
```
ax=sns.histplot(data=df,x="Age",kde=True)
for i in ax.containers:
    ax.bar_label(i)
```



In [12]:
```
# Between 23 To 26 Have More Customers
```

In [13]:
```
ax=sns.barplot(data=df,x="Product",y="Income")
for i in ax.containers:
    ax.bar_label(i)
```

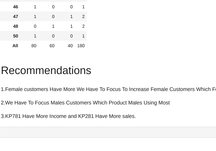
In [14]:
```
# product KP781 More Income
```

In [15]:
```
sns.boxplot(data=df,x="Gender",y="Income",hue="Product")
```

Out[15]:
```
<Axes: xlabel='Gender', ylabel='Income'>
```

In [16]:
```
sns.boxplot(data=df,x="Gender",y="Miles",hue="Product")
```

Out[16]:
```
<Axes: xlabel='Gender', ylabel='Miles'>
```



In [20]:
```
numeric_data = df.select_dtypes(include='number')

sns.heatmap(numeric_data.corr(), cmap='viridis', annot=True, fmt=".2f", linewidths=.5)

plt.show()
```

In [17]:
```
pd.crosstab(index=df['Gender'], columns=df['Product'], margins=True, normalize="index")*100
```

Out[17]:

Product	KP281	KP481	KP781
Gender			
Female	52.631579	38.157895	9.210526
Male	38.461538	29.807692	31.730769
All	44.444444	33.333333	22.222222

In [ ]:
```
# KP281 Have High Percenatge Compare to Other Product
# Female Have High Percentage In KP281 and KP481 and Male Have More In KP781
```

In [18]:
```
pd.crosstab(index=df["Income"],columns=df["Product"],margins=True)
```

Out[18]:

Product	KP281	KP481	KP781	All
Income				
29562	1	0	0	1
30699	1	0	0	1
31836	1	1	0	2
32973	3	2	0	5
34110	2	3	0	5
...	...	...	...	...
95866	0	0	1	1
99601	0	0	1	1
103336	0	0	1	1
104581	0	0	2	2
All	80	60	40	180

63 rows × 4 columns

In [ ]:
```
# KP281 Have More Sales
```

In [19]:
```
pd.crosstab(index=df["Age"],columns=df["Product"],margins=True)
```

Out[19]:

Product	KP281	KP481	KP781	All
Age				
18	1	0	0	1
19	3	1	0	4
20	2	3	0	5
21	4	3	0	7
22	4	0	3	7
23	8	7	3	18
24	5	3	4	12
25	7	11	7	25
26	7	3	2	12
27	3	1	3	7
28	6	0	3	9
29	3	1	2	6
30	2	2	3	7
31	2	3	1	6
32	2	2	0	4
33	2	5	1	8
34	2	3	1	6
35	3	4	1	8
36	1	0	0	1
37	1	1	0	2
38	4	2	1	7
39	1	0	0	1
40	1	3	1	5
41	1	0	0	1
42	0	0	1	1
43	1	0	0	1
44	1	0	0	1
45	0	1	1	2
46	1	0	0	1
47	1	0	1	2
48	0	1	1	2
50	1	0	0	1
All	80	60	40	180

Recommendations

1.Female customers Have More We Have To Focus To Increase Female Customers Which Females Like The Most Female Customers Using KP281 Have Most

2.We Have To Focus Males Customers Which Product Males Using Most

3.KP781 Have More Income and KP281 Have More sales.

In [ ]: