

AEROFIT CASE STUDY

https://colab.research.google.com/drive/1Px3_J8LOcAbs3TXgEpZJ1WVZbKYjdA8C?usp=sharing

Step-1 :- Defining the problem statement & analyzing basic metrics

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.

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.

Step-2:- Observation on data

```
[2] import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

[3] from google.colab import drive
drive.mount('/content/drive')

Mounted at /content/drive

[4] df=pd.read_csv('/content/drive/MyDrive/Netflix-case study/aerofit_treadmill.csv')

[ ] 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

[ ] df.info()

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

unique_values=df.unique()
unique_values.name='count_of_values'
unique_values

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

[ ] df.describe().transpose()

count mean std min 25% 50% 75% max
Age 180.0 28.788889 6.943498 18.0 24.00 26.0 33.00 50.0
Education 180.0 15.572222 1.617055 12.0 14.00 16.0 16.00 21.0
Usage 180.0 3.455556 1.084797 2.0 3.00 3.0 4.00 7.0
Fitness 180.0 3.311111 0.958869 1.0 3.00 3.0 4.00 5.0
Income 180.0 53719.577778 16506.684228 29562.0 44058.75 50599.5 58698.00 104581.0
Miles 180.0 103.194444 51.863605 21.0 66.00 94.0 114.75 360.0

[ ] df.isna()

Product Age Gender Education MaritalStatus Usage Fitness Income Miles
0 False False False False False False False False False
1 False False False False False False False False False
2 False False False False False False False False False
3 False False False False False False False False False
4 False False False False False False False False False
...
175 False False False False False False False False False
176 False False False False False False False False False
177 False False False False False False False False False
178 False False False False False False False False False
179 False False False False False False False False False
180 rows x 9 columns
```

. Observation (Data Type)

1. Data type of Product is object (string).
2. Data type of Age is int64.
3. Data type of Gender is object (string).
4. Data type of Education is int64.
5. Data type of MaritalStatus is object (string).
6. Data type of Usage is int64.
7. Data type of Fitness is int64.
8. Data type of Income is int64.
9. Data type of Miles is int64.

Observations from Descriptive Statistics (Numerical)

1. Age: Median age of the customer(s) is 26 years, having maximum age of 50 years and minimum age of 18 years.
2. Education (Years): Median education years of the customer(s) is 16 years, with maximum education years is 21, and minimum years is 12.
3. Usage (Per week) : Median usage of treadmill is 3 times per week, with maximum 7 times per week and minimum 2 times per week.
4. Fitness (1-5) : Median fitness rating of customer(s) is 3 (moderately fit) and mean fitness roughly lies around the median.
5. Income (\$): Median income of customer(s) is 50.5K annually. Maximum income is 104K annually, and minimum income is 29.5K.
6. Miles: Median distance travelled (walk/run) by customer(s) is 94. Maximum distance travelled is 114.75 and minimum is 21.

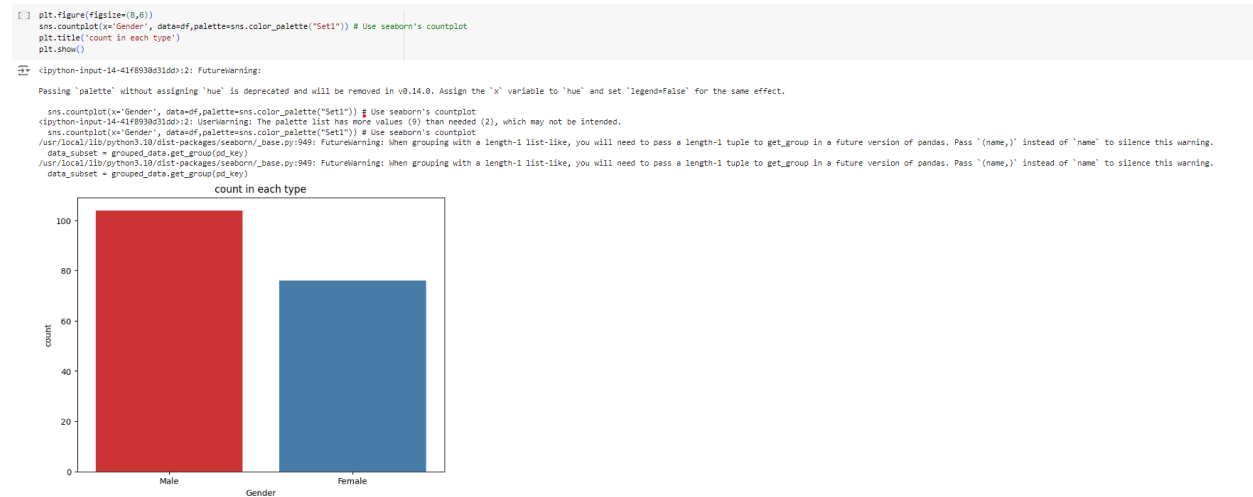
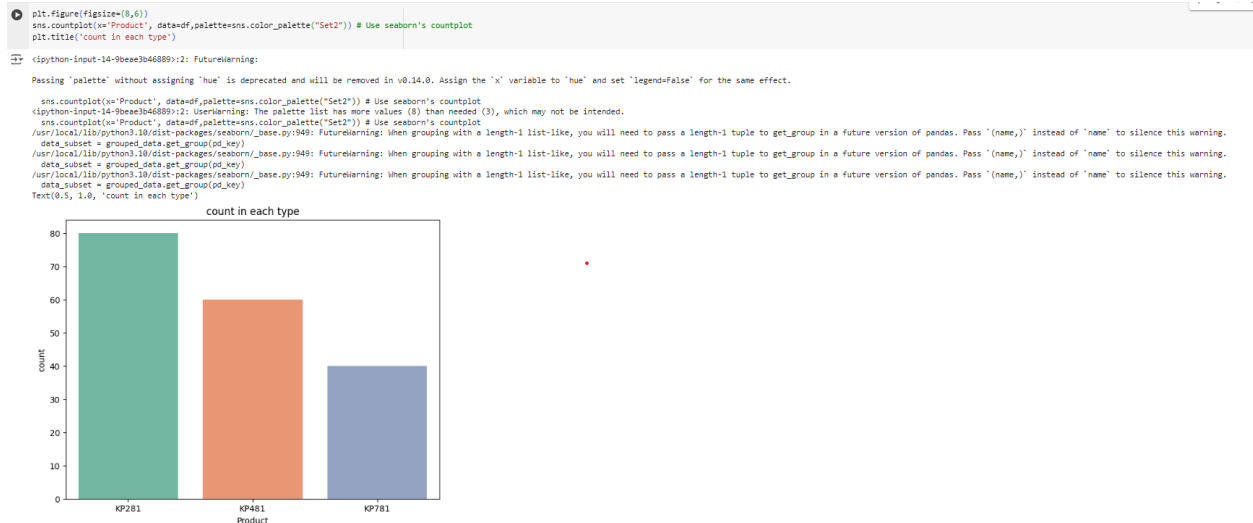
General Observations

1. Fitness: More than 50% customers have given self-rating of 3, followed by 5 and 2.
2. Usage: Fair amount (38.3%) of people have reported usage of treadmills 3 times per week, followed by 4 times per week, and 2 times per week respectively.
3. Product: Most commonly product purchased is KP281, followed by KP481, and KP781 respectively.

- Gender: Male is the most common gender who purchased more of the aerofit products.
- MaritalStatus: Couples purchased more products compare to Single people. Maybe Couple Goals!

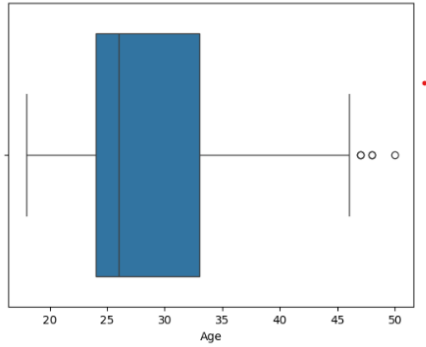
Step-3 Visual Analysis-univariate & Bivariate

Univariate Analysis :-



```
[ ] sns.boxplot(data=df,x='Age')
```

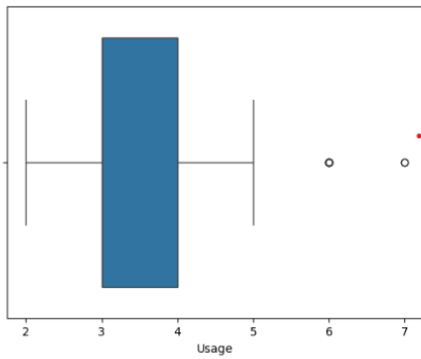
```
/usr/local/lib/python3.10/dist-packages/seaborn/categorical.py:640: FutureWarning: SeriesGroupBy.grouper is deprecated and will be removed in a future version of pandas.  
positions = grouped.grouper.result_index.to_numpy(dtype=float)  
<Axes: xlabel='Age'>
```



1. Most of the people who are working out is age between 24 to 33
2. 46-50 age group people workout very less(Outliers)

```
[ ] sns.boxplot(data=df,x='Usage')
```

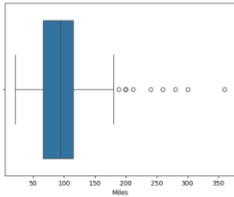
```
/usr/local/lib/python3.10/dist-packages/seaborn/categorical.py:640: FutureWarning: SeriesGroupBy.grouper is deprecated and will be removed in a future version of pandas.  
positions = grouped.grouper.result_index.to_numpy(dtype=float)  
<Axes: xlabel='Usage'>
```



1. Most of the customers use treadmill 3-4 times per week
2. Very few people walk run on treadmill 6-7 times per week(Outliers).
3. Difference between 25th and 75th percentile is 1.0
4. Overall, it looks like very few pwople are regular in their workouts while others are working out casually.

```
[ ] sns.boxplot(data=df,x='Miles')
```

```
/usr/local/lib/python3.10/dist-packages/seaborn/categorical.py:640: FutureWarning: SeriesGroupBy.grouper is deprecated and will be removed in a future version of pandas.  
positions = grouped.grouper.result_index.to_numpy(dtype=float)  
<Axes: xlabel='Miles'>
```



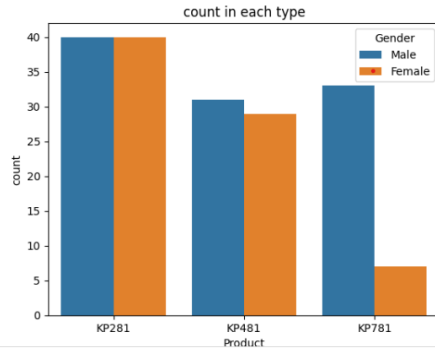
observation

1. Most of the distance travelled by the customer on the treadmill is roughly between 75-125 Miles.
2. Very few people have travelled more than roughly 200 miles(Outliers)
3. Difference between 25th and 75th percentile is 48.75 miles(running/walking)

▼ Bivariate

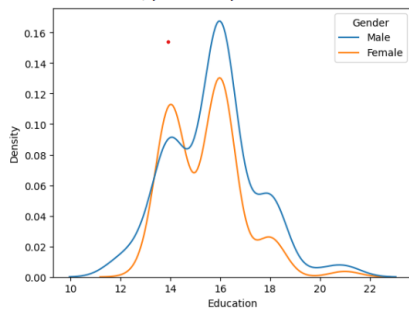
```
sns.countplot(data=df, x='Product', hue='Gender')
plt.title('count in each type')
```

```
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
Text(0.5, 1.0, 'count in each type')
```



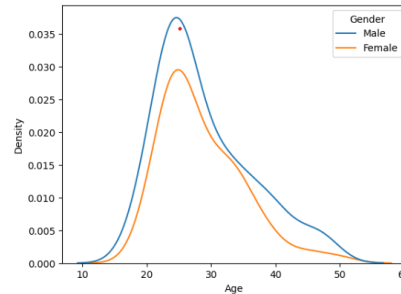
```
sns.kdeplot(data=df, x='Education', hue='Gender')
```

```
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
<Axes: xlabel='Education', ylabel='Density'>
```



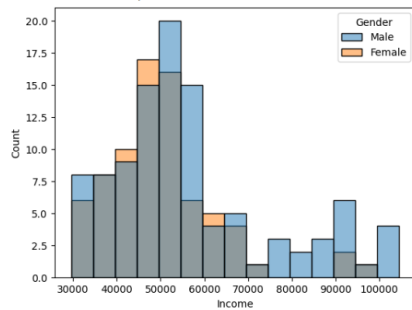
```
sns.kdeplot(data=df, x='Age', hue='Gender')
```

```
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
<Axes: xlabel='Age', ylabel='Density'>
```



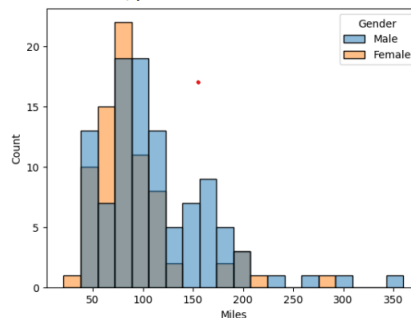
```
sns.histplot(data=df, x='Income', hue='Gender')
```

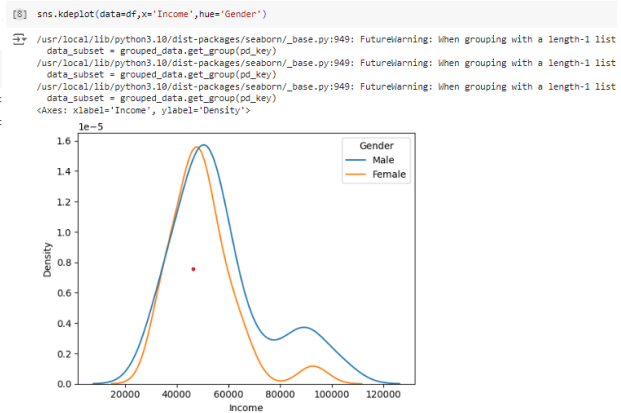
```
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
<Axes: xlabel='Income', ylabel='Count'>
```



```
sns.histplot(data=df, x='Miles', hue='Gender')
```

```
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
/usr/local/lib/python3.10/dist-packages/seaborn/_base.py:949: FutureWarning: When grouping with a length-1 list-like, you will need to pass a length-1
data_subset = grouped_data.get_group(pd_key)
<Axes: xlabel='Miles', ylabel='Count'>
```





-Distplot, histogram, countplot, boxplot, pairplot has been used here

Step-4&5 Business insights based on non-graphical and visual analysis

Conclusions

Business Insights(based on Non-Graphical and Visual Analysis):

1. The top three purchased treadmill models are KP281, KP481, and KP781, in that order.
2. There is a higher proportion of male buyers compared to female buyers.
3. More customers are in a partnered marital status compared to single.
4. The average age of customers is 28, with a range between 18 to 50 years and a median of 26 years.
5. The average education level of customers is 15.5 years, with a range between 2 to 21 years and a median of 16 years.
6. On average, customers plan to use the treadmill three times per week, with a range between 2 to 7 times per week and a median of three times per week.
7. The average self-fitness rating of customers is 3, with a range between 1 to 5 and a median of 3.
8. Customers' average annual income is 53.7K dollars, with a range between 29.5K dollars to 104K dollars and a median income of 50.5K dollars.
9. The average distance traveled by customers on the treadmill is 103 miles, with a range between 21 to 360 miles and a median of 94 miles.
10. The relationship between fitness and distance traveled on the treadmill is strong.
11. Similarly, there is a strong relationship between usage frequency and distance traveled on the treadmill.
12. Most customers use the treadmill 3-4 times per week, with very few using it 6-7 times per week.
13. The majority of customers rate themselves as moderately fit.
14. The mean income for KP281 buyers is 46.4K dollars, for KP481 buyers is 48.9K dollars, and for KP781 buyers is 75.4K dollars.
15. KP281 and KP481 have the same mean usage of 3, while KP781 has a mean usage of 4.
16. The mean fitness rating for KP281 and KP481 buyers is 3, while for KP781 buyers, it is 4.6.
17. KP781 is the most preferred treadmill among male customers, while females show the least preference for it.
18. Overall, male customers tend to use treadmills more frequently than females.
19. The income distribution between both genders is roughly similar.
20. Males tend to have a higher fitness level compared to females.

21. The distance traveled on the treadmill is roughly the same for both genders, but men tend to cover longer distances, with some going beyond 320 miles.
22. Partnered customers tend to have a higher fitness level compared to singles.

Step-6 Recommendations

1. Promote KP281 and KP481 treadmills as budget-friendly options, especially targeting customers with annual incomes in the range of 39K - 53K Dollars.
2. Market KP781 treadmill as a premium product with advanced features, targeting professionals and athletes.
3. Enhance the marketing strategy for KP781 by associating it with renowned athletes, leveraging their achievements for better outreach.
4. Run special marketing campaigns on Women's Day and Mother's Day to encourage more women to adopt an exercise routine, highlighting the benefits of using our treadmills.
5. Conduct research to expand the customer base beyond 50 years of age. Offer basic treadmill models (KP281/KP481) as suitable options for beginners in this age group.
6. Encourage existing customers to upgrade their treadmills to high-end models as their usage increases over time, leading to increased revenue for the business.