***Aerofit Business  problem :***

Identifying 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.

Differences across the products with respect to customer characteristics.

1. Descriptive analytics to create a customer profile for each Aerofit treadmill product. 2. 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.

***About Dataset :***

Aerofit Data is about 3 different products KP281 , KP481 and KP781.

Company has collected the data on individuals who purchased a treadmill from Aerofit stores during the prior three months.

Data of 180  individual Customers.

shape : 180,9

Available Columns are : 9

**Info about data-frame:**

Data columns (total 10 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

9 Fitness\_category 180 non-null object

dtypes: int64(6), object(4)

**# No null values found in dataset.**

Product 0

Age 0

Gender 0

Education 0

MaritalStatus 0

Usage 0

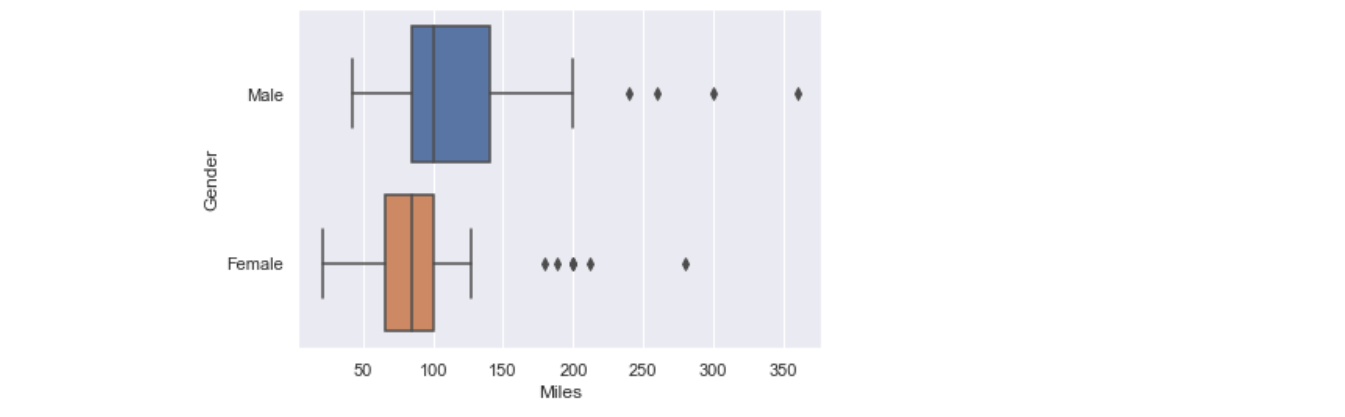
Fitness 0

Income 0

Miles 0

**# Outlier Detection :**

Below plot shows the distribution of miles run/walk by customer  , Gender wise.



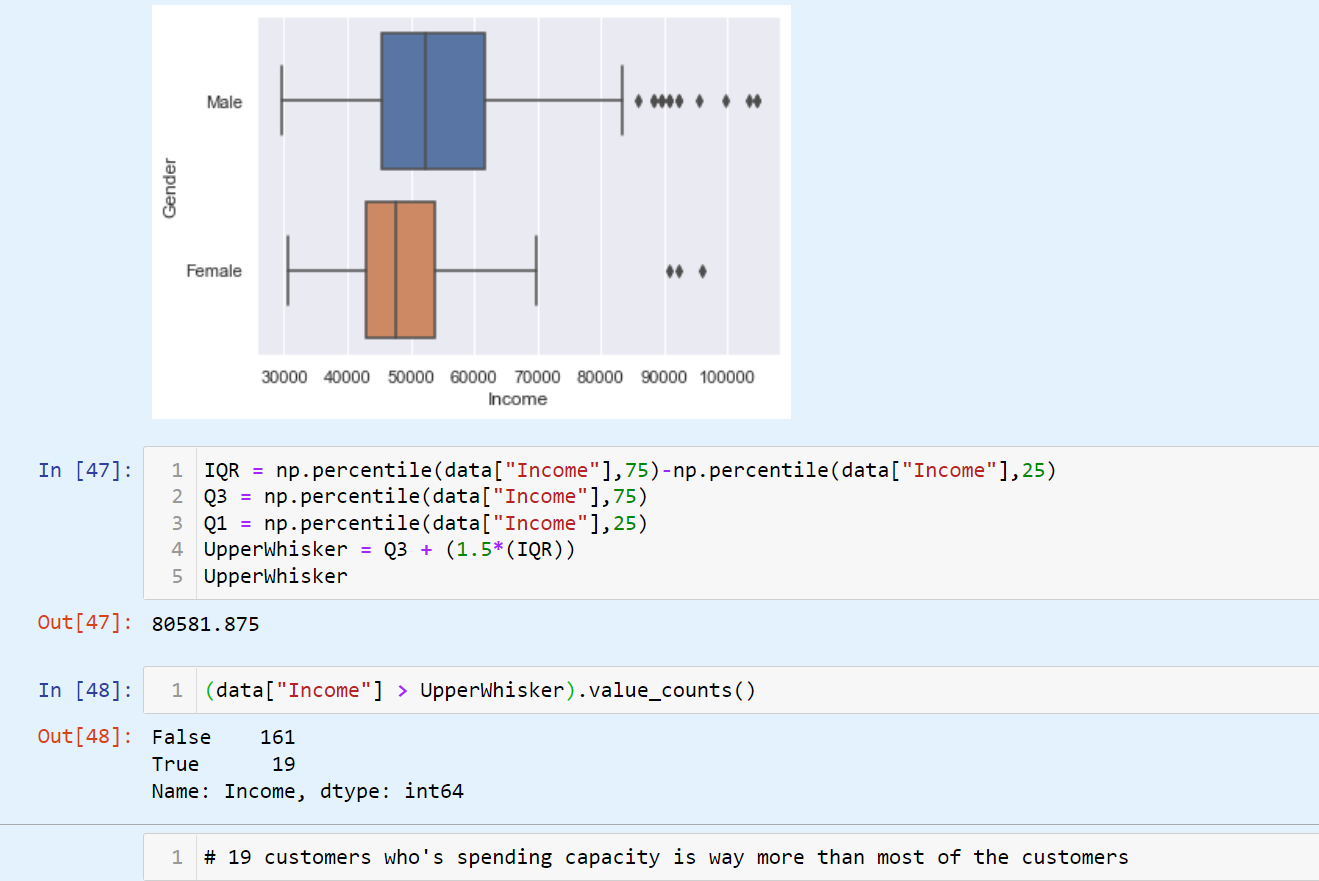


***Observations and Insights:***

*13 outlier in column "Miles".*

*Customers who fall in outliers as per their miles run/walk , uses product KP781 and are in excellent shape(from data).*

Below plot shows the distribution of income of customers  , Gender wise.



***Observations and Insights:***

19 outliers in income columns.

Probably those are the target customers who can spend more and more likely to buy expensive product.

***Pre-Processing Data for Analysis:***

categorized Fitness column :

5:"Excellent Shape"

4:"Good Shape"

3:"Average Shape"

2:"Bad Shape"

1:"Poor Shape"

Merged Product Price with data frame.



Categorized Age Column

"Teen(0-21)"

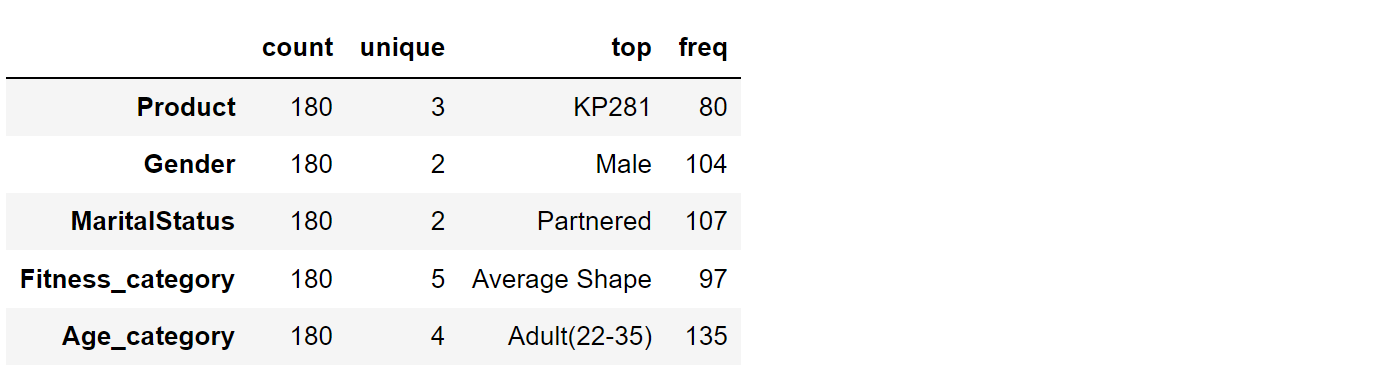
"Adult(22-35)"

"mid\_age(36-45)"

"Towards\_old-age(>46)"

***Statistical Summery :***

**Describing Categorical Features :**



***Observations and Insights:***

There 3 products : KP281,KP481 and KP781

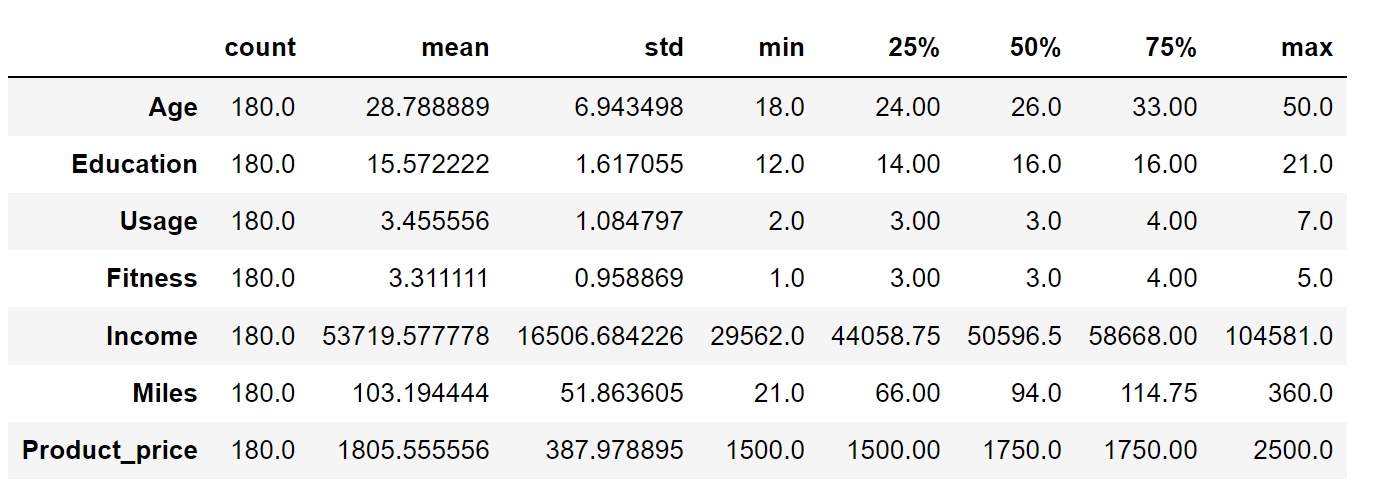
Total Customer count = 180

highest frequency data available is for Male Gender category. ( 104 / 180)

Majority of customers are Married and have average shape . (97 / 180)

Majority of customers are adults who are 22 to 35 years old. (135 / 180)

**Describing Numerical Features :**



***Observations and Insights:***

**Insights from Above Information :**

1. Median Age of Customer is 26 years.
2. Maximum users are Adults(22-35) years and are Male and Married.
3. Maximum Selling Product is KP281.
4. Maximum numbers of customers' fitness level is above average(>3 according to given data).
5. Median Miles run/walk per customer : 94 Miles

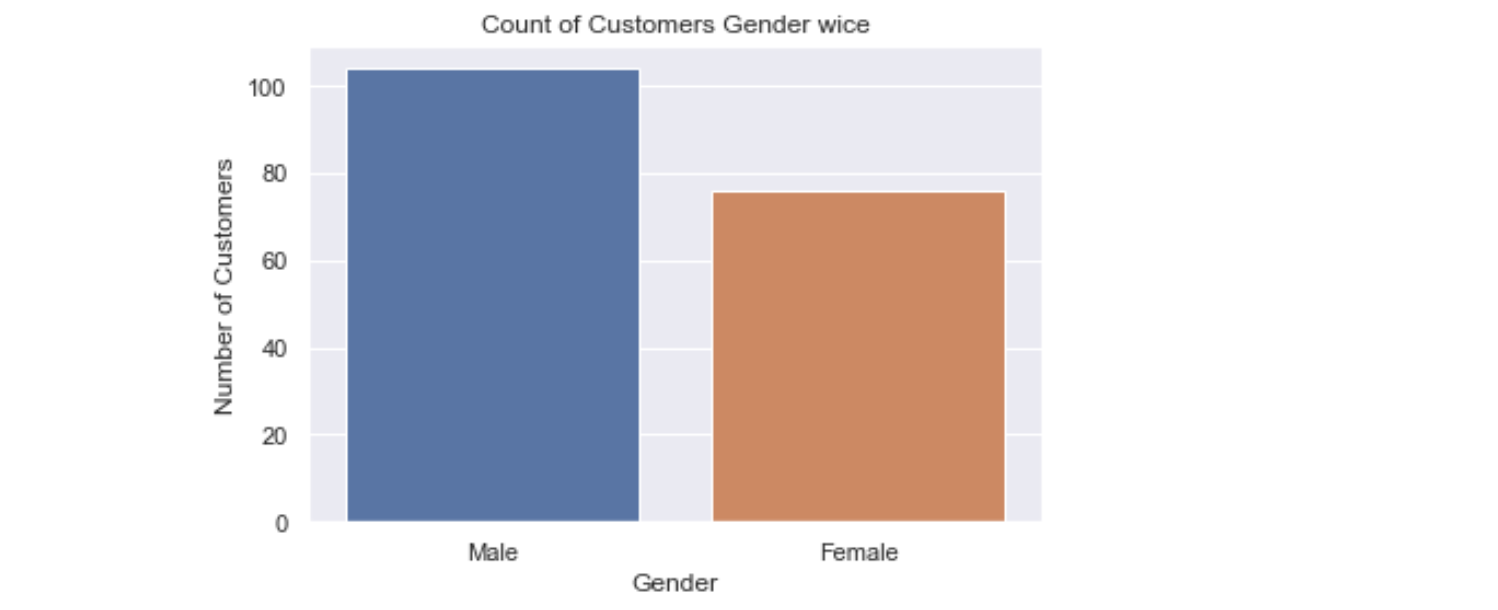
median income of the customers :50596.5 USD Median of average usage per customer : 3 days a week

Average Customer education is 15 to 16 years.

**General Sales Analysis :**

Male 104

Female 76



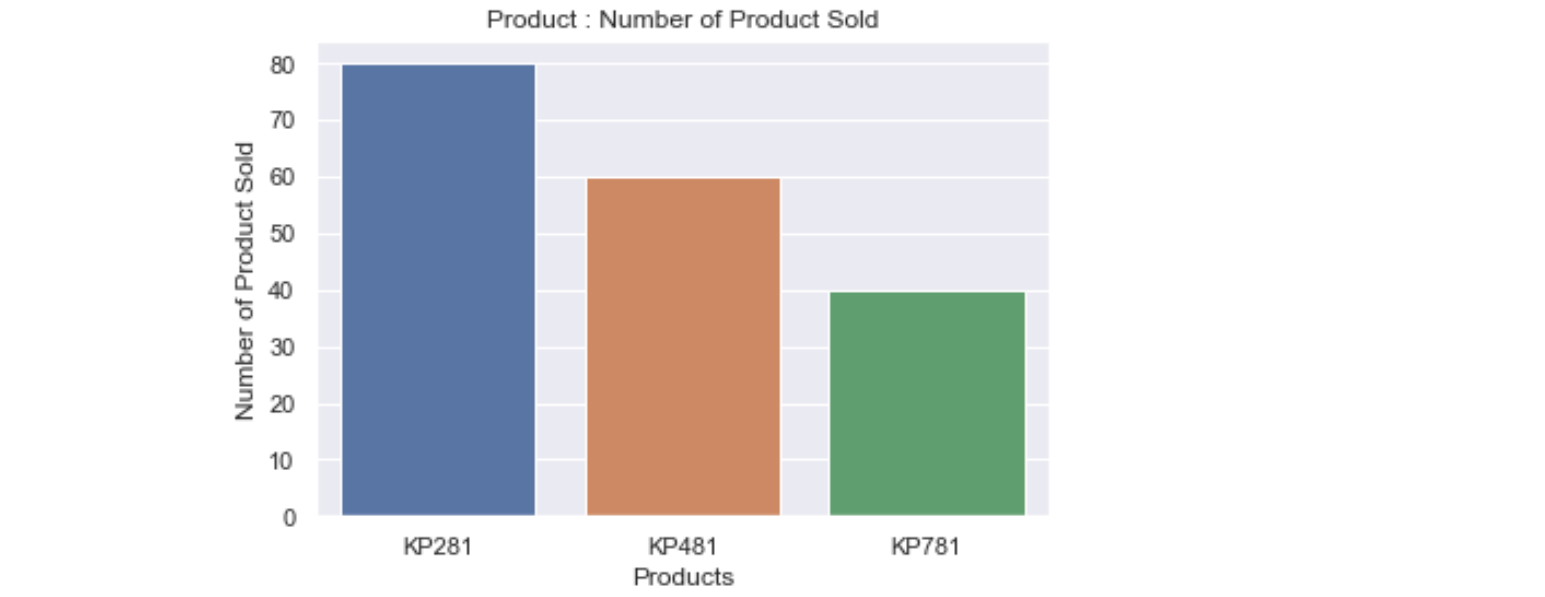
Product - Number of Product Sold

Product Count Product-Share

KP281 80 44.44 %

KP481 60 33.33 %

KP781 40 22.22 %

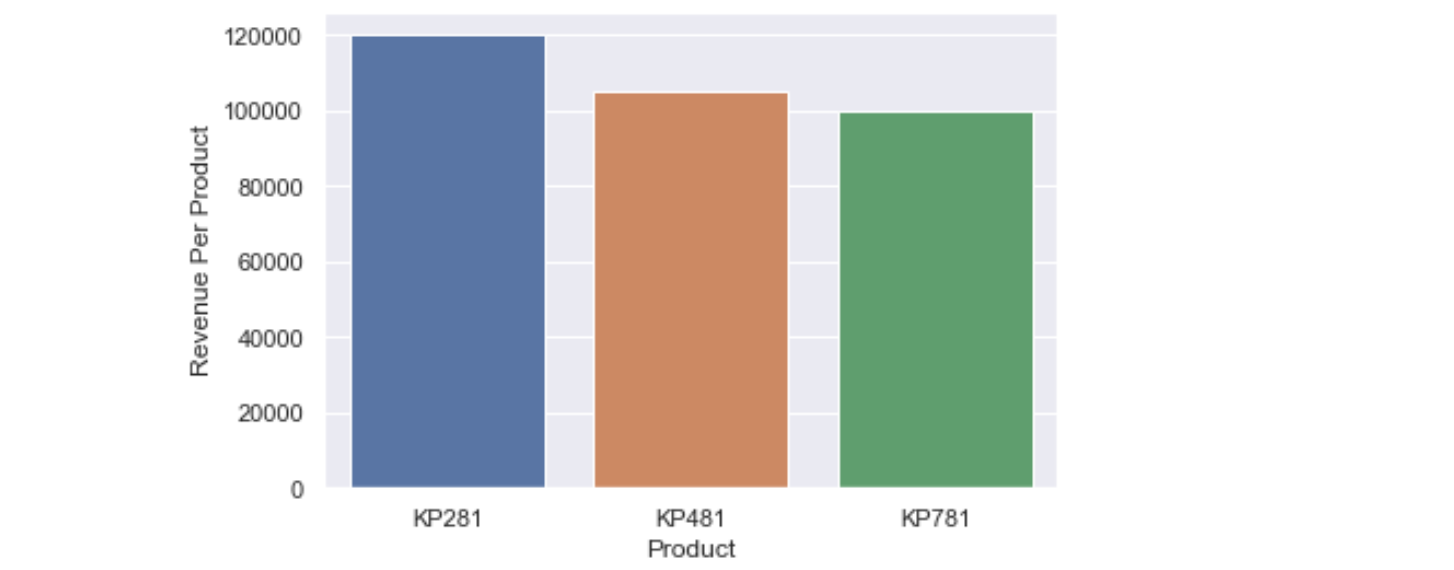


Revenue generated per Product:

Product Revenue in $

KP281 120000 USD

KP481 105000 USD  
KP781 100000 USD



***Observations and Insights:***

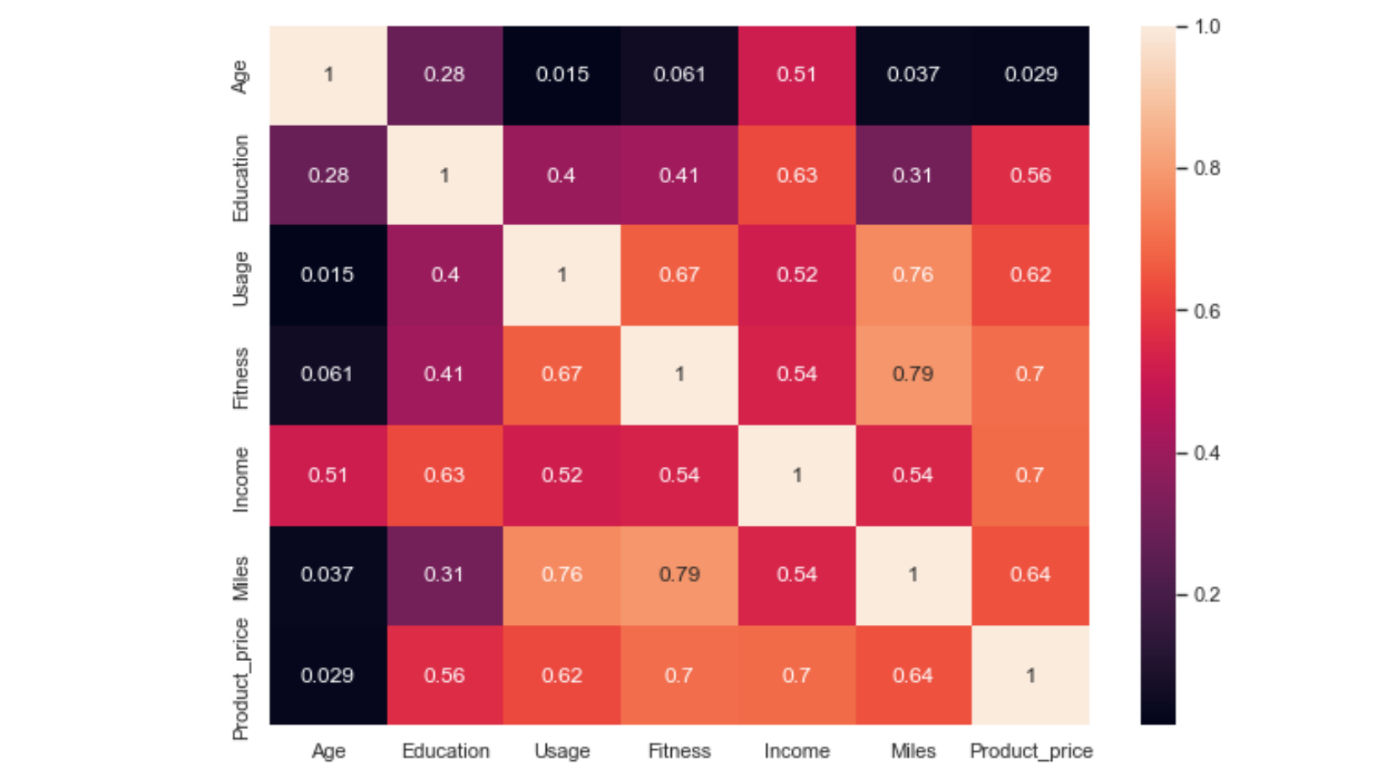
Highest  Selling Product is : KP281

KP481 and KP781 product's market share and Revenue generated by those products are significant.

Data shows KP281 as leading product which is the cheapest product and the most sold item.

Relatively KP781 is the costliest Product from Aerofit.

***Correlation Between Features :***



# Correlation  : ( features having correlation higher than 6.5 )

Usage Fitness 0.668606

Miles 0.759130

Fitness Usage 0.668606

Miles 0.785702

Product\_price 0.696616

Income Product\_price 0.695847

Miles Usage 0.759130

Fitness 0.785702

Product\_price Fitness 0.696616

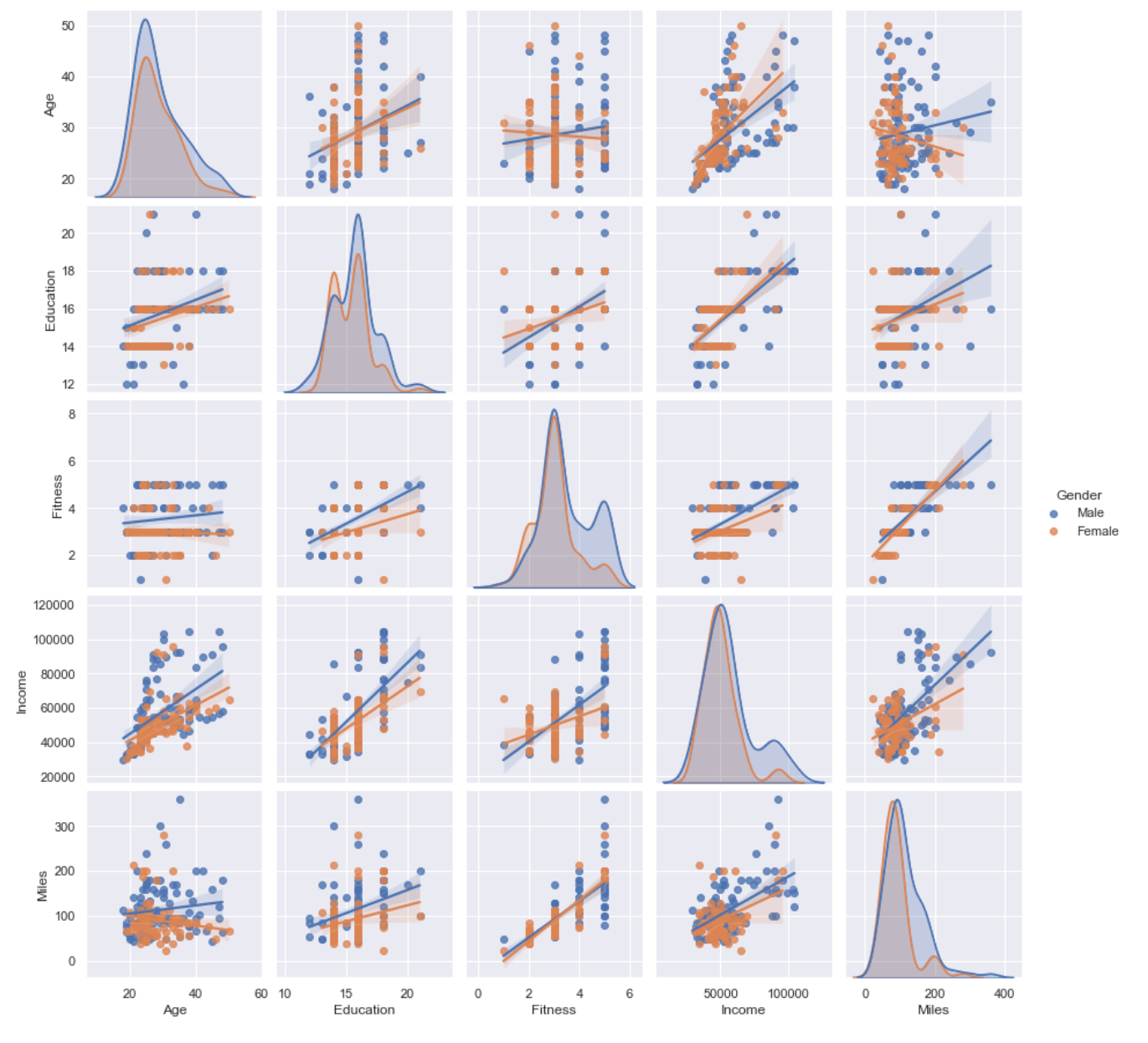
Income 0.695847

Important correlations

1. Fitness &  Miles : 0.785702

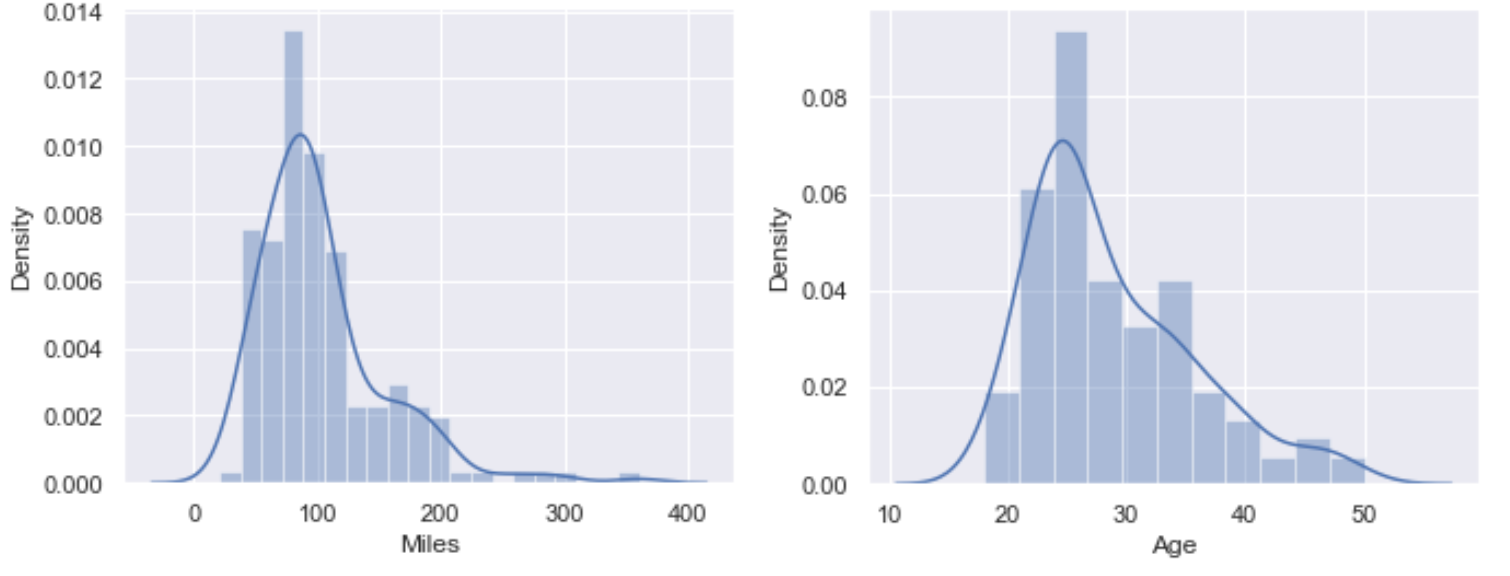
2. Product\_price & Income : 0.695847

Pairplot :



**Distribution** **Plot** of Miles Run/walk by Customers .

Distribution Plot of Age of Customers.



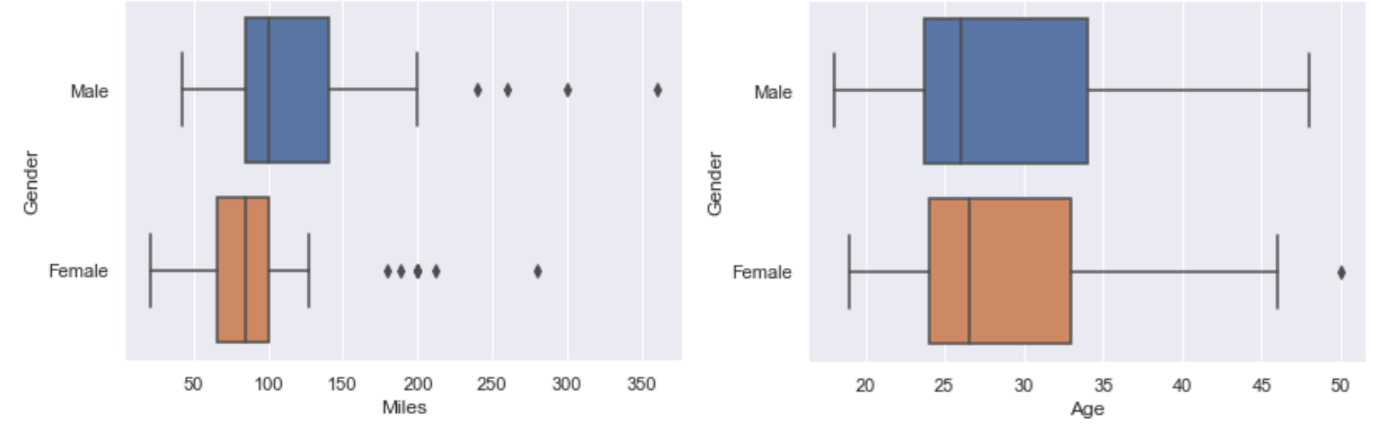
***Observations and Insights:***

Frequency of Customers run/walk from 50 to 130 is more .

and Age from 21 to 35.

**Gender Wise distribution** of Miles they walk/run  &

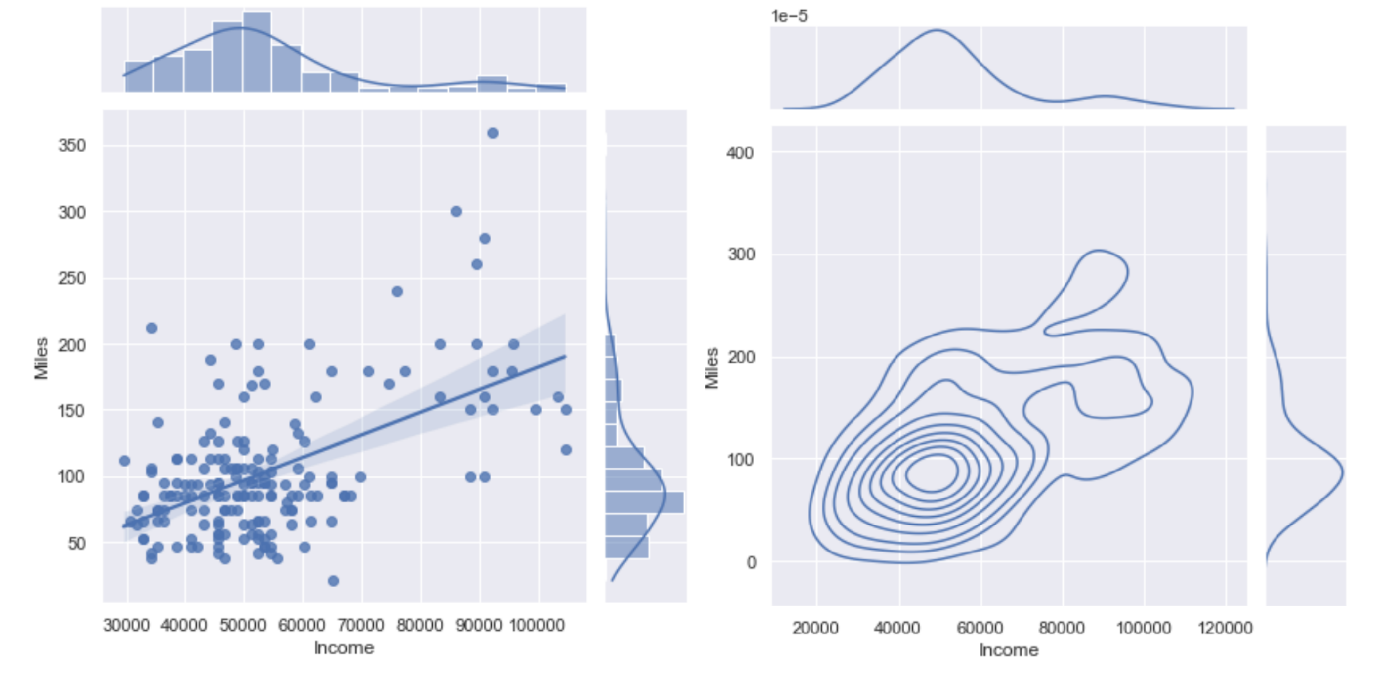
**Gender Wise distribution** of their age:



Correlation between Income and miles :

Majority Customer Information about their earning and how much they workout .

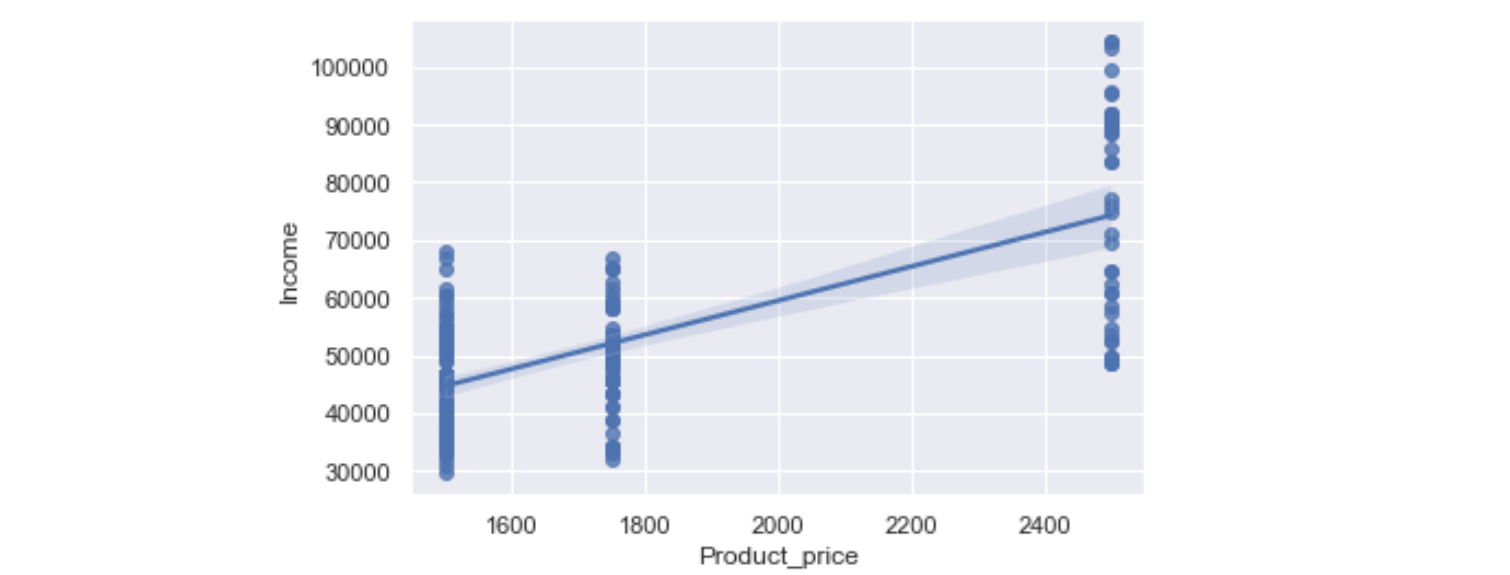
Scattered Plot : Income-Miles

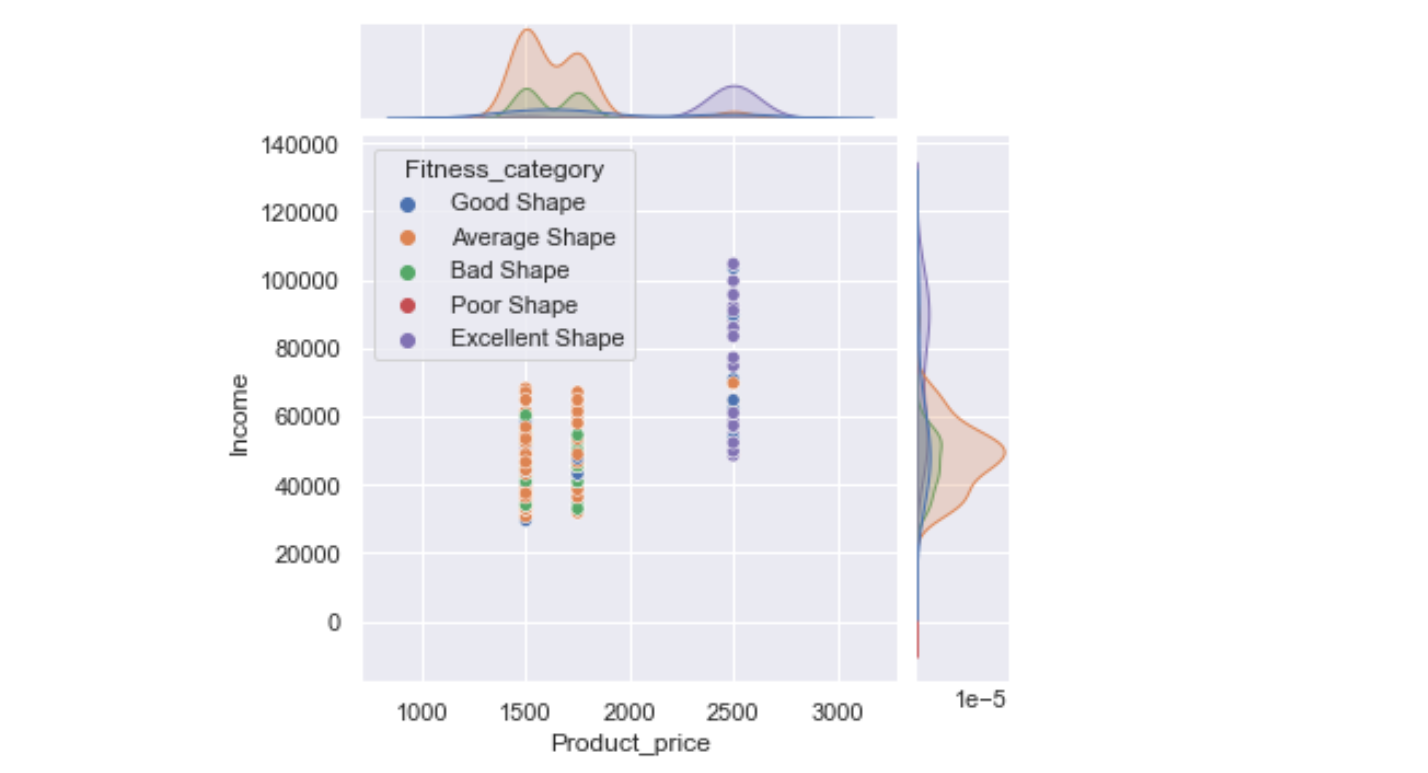


***Observations and Insights:***

Majority customer base has earning from 25,000 to 75,000USD

and prefer to exercises very less to 175 miles a week.





***Observations and Insights:***

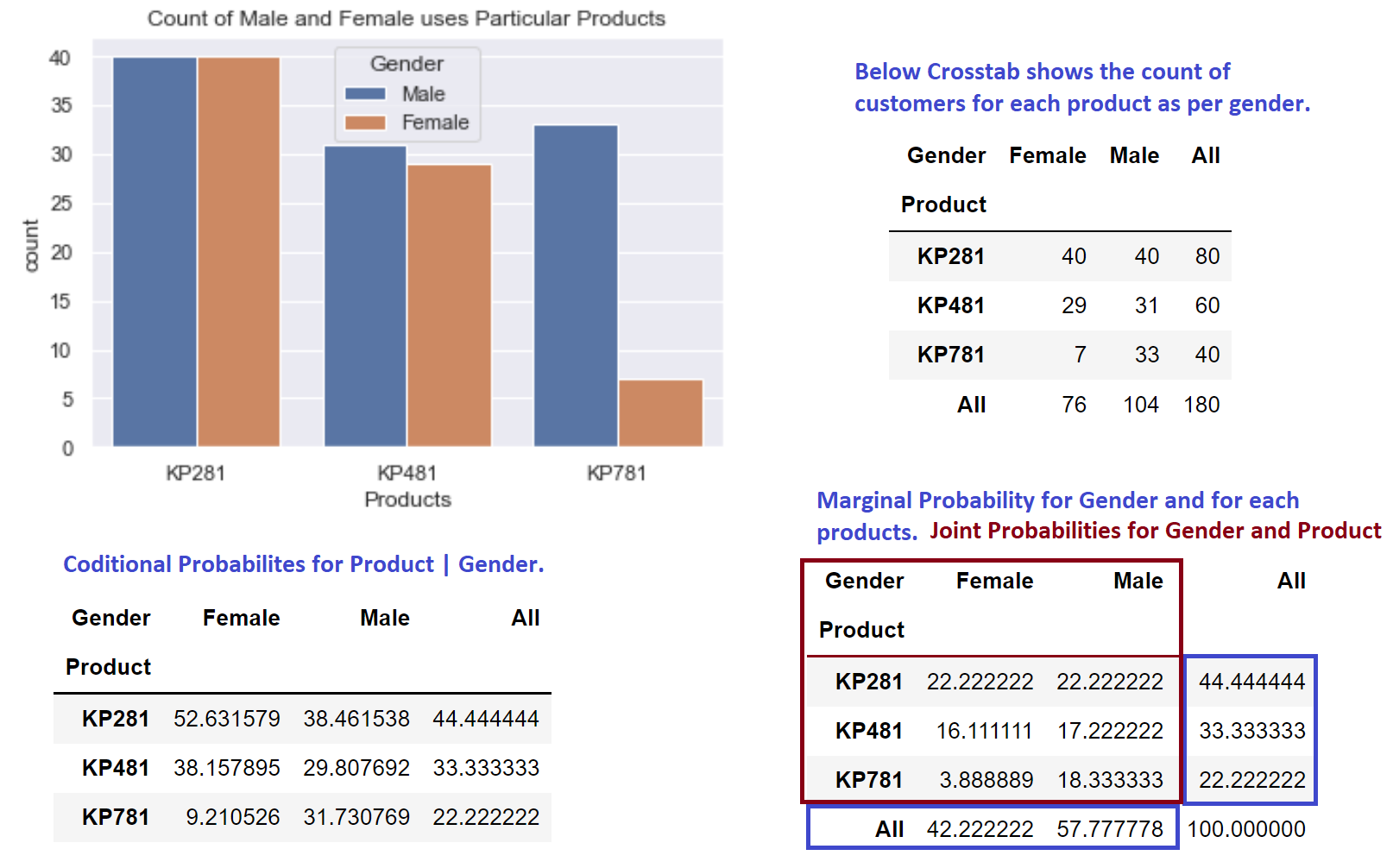
from above Plot,

we can see a positive trend , that who are earning more are likely to buy the costlier product.

and from 2nd plot , shows , people are in excellent and good shape, they are more likely spend mor amount and buy the costlier product which can be more reliable for extensive use.

***Identify the characteristics of the target audience :***

Gender wise Preference:



***Observations and Insights:***

Marginal Probability  :

(from above tables)

Probability of Male Customer Purchasing any product is : 57.77 %

Probability of Female Customer Purchasing any product is : 42.22 %

Marginal Probability of any customer buying

product KP281  is : 44.44 %  (cheapest / entry level product)

product KP481  is : 33.33 % ( for intermediate users)

product KP781  is : 22.22 % ( product for extensive use who run/walk more miles)

Conditional Probability :

Probability of Female customer buying KP281(52.63%) is more than male(38.46%).

KP281 is more recommended for female customers.

Probability of Male customer buying Product KP781(31.73%) is way more than female(9.21%).

Probability of Female customer buying Product KP481(38.15%) is significantly higher than male (29.80%. )

KP481 product is specifically recommended for Female customers who are intermediate user.

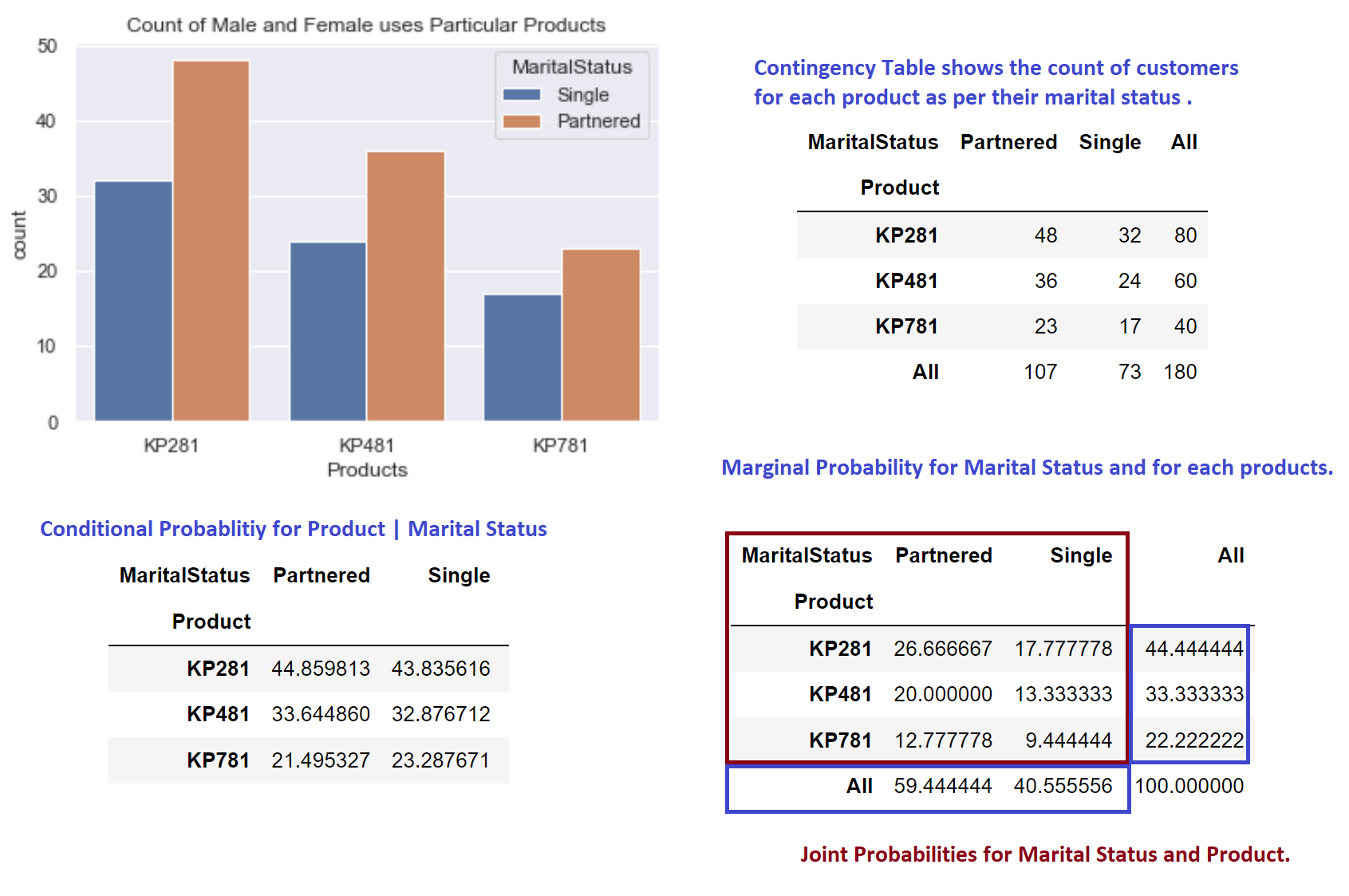
Marital Status wise Preference :

Status % share

Partnered 59.44 %

Single 40.55 %

approximately  60% customers are Married.



***Observations and Insights:***

Conditional Probabilities:

KP281 | Partnered = 44.85 %

KP481 | Partnered = 33.64 %

KP781 | Partnered = 21.49 %

KP281 | Single = 43.83 %

KP481 | Single = 32.87 %

KP781 | Single = 23.28 %

Probability of Married Person purchasing any product is 59.44 %

Probability of Single  Person purchasing any product is 40.55 %

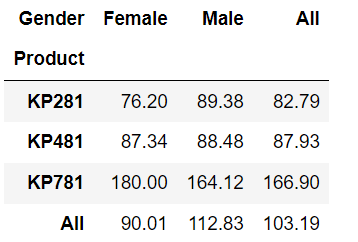
Probability of a Married person buying product KP281 and KP481 is slightly higher than the customers who are single.

Probability of a single person buying KP781 is higher than Married customers.

So , KP781 is also recommended for people who are single and exercises more.

***Product classification as per customer Usages and Miles.***

Below is the average run/walking mile per Gender for particular product.



***Observations and Insights:***

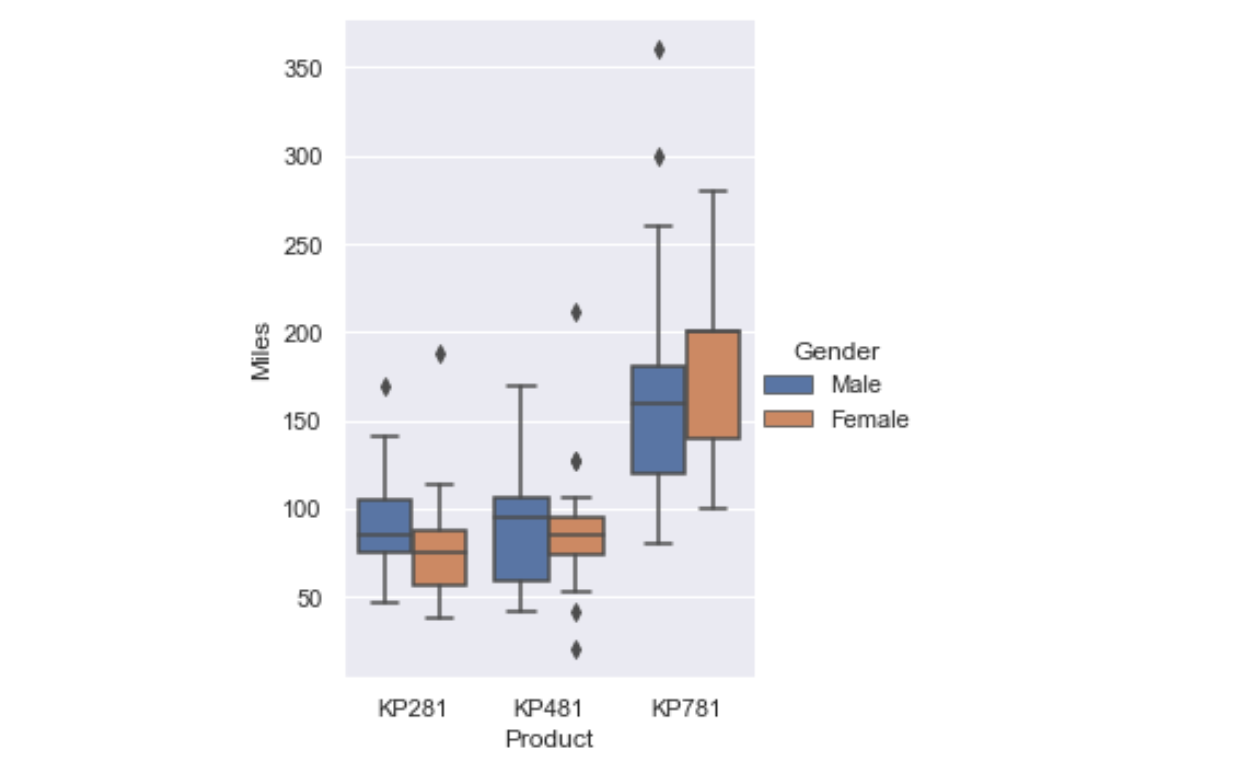
From charts and Crosstab of average miles run by customer for particular product:

Female Customers who are running average 180 miles (extensive exercise) , are using product KP781, which is higher than Male average using same product.

KP781 can be recommended for Female customers who exercises extensively.

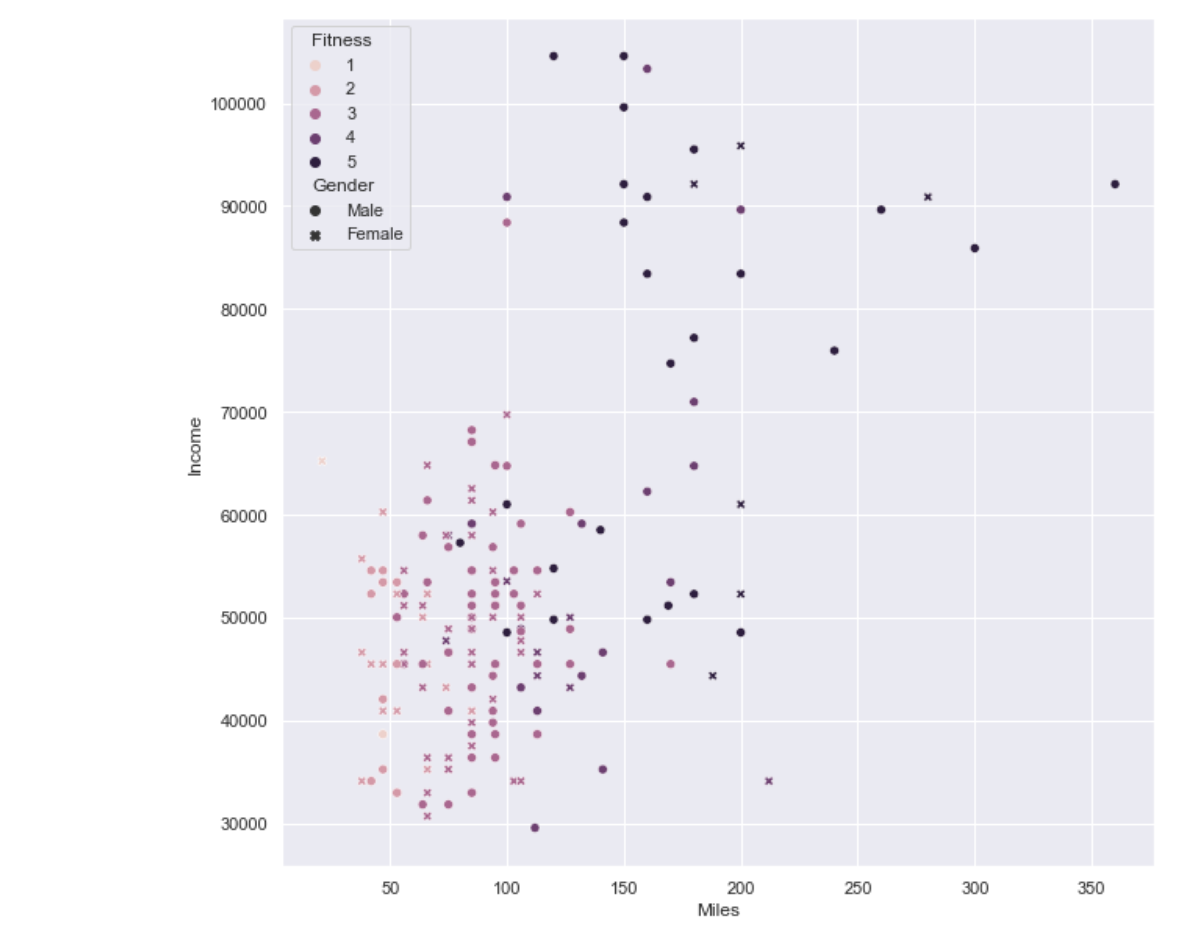
Males customers who are running average of 90 miles (average exercise) , are using product KP281 .

Males customers who are running average of 87 miles (average exercise) , are using product KP481 . and for female average running for same product is 88 miles.



***Observations and Insights:***

Since, the variation  for Product KP481 for particularly Male is more , we can say KP481 is good for people who want to run/walk for 60 to 130 miles a week. It is more a genera purpose product for intermediate use.



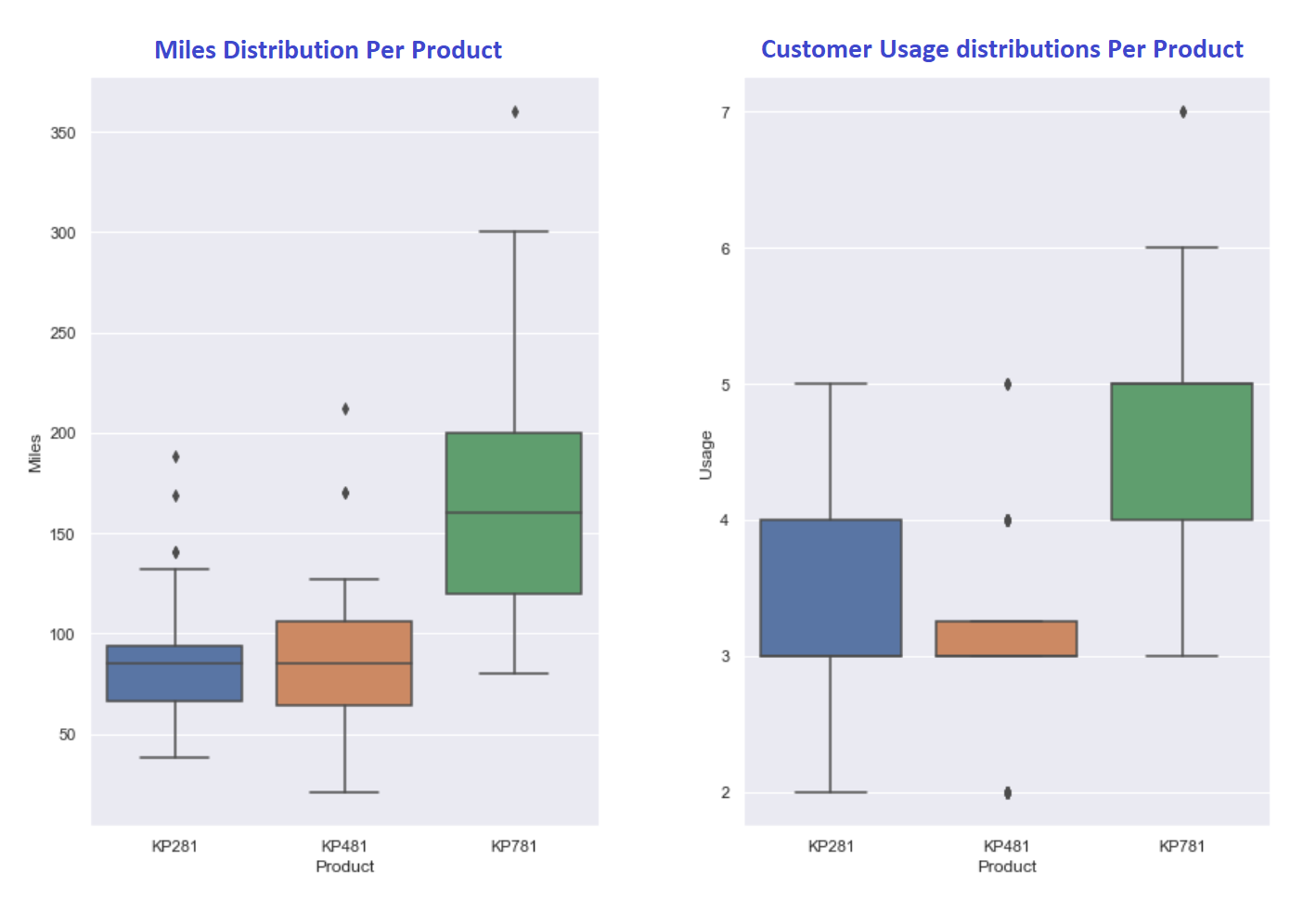
***Observations and Insights:***

Above scattered Plot shows the overall picture over customer's income, how much they exercise (run/walk miles) given their gender and their fitness level.

Most of the customer's fitness level is around 3 to 4 . and it says people who run more miles are having good fitness level.

Though there is a trend with income and miles. But there are very few customers who earn a lot and run more miles.

***Distributions across miles and usage of customers for particular product :***



***Observations and Insights:***

left categorical plot shows the miles run/walk by customer distribution across all the products.

And right plot shows the usage (weeks) distribution  across all the products.

Customers who walk/run 70-90 miles, are  using KP281

Customers who walk/run 70-130 or more miles are using  KP481.

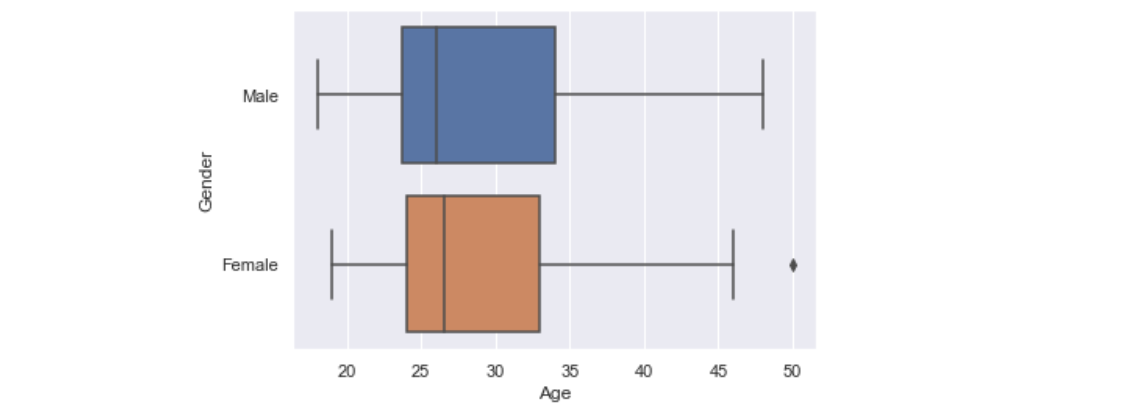
Customers who walk/run 120 to 200 or more miles uses KP781.

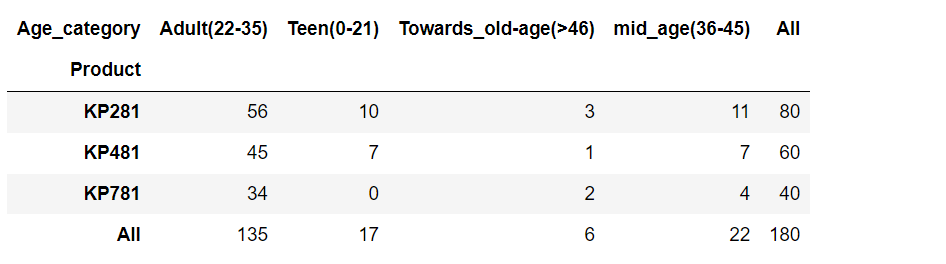
Customers who walks/run 3 to 4 days a week, are  using KP281

Customers who walk/run 3  or more days are using  KP481.

Customers who walk/run 4 to 5 or more days uses KP781.

***Product classification as per customer 's Age.***



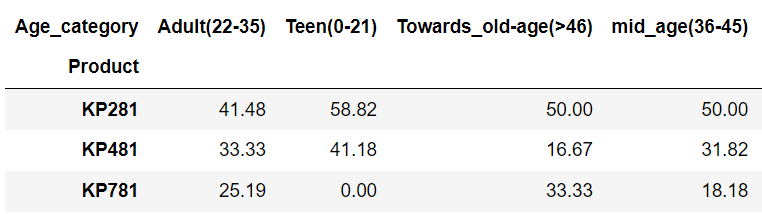


***Observations and Insights :***

from above distribution ,

Most of the customer base is from Age category Adult (22-35) : 135 customer .

customers who are in Teen and mid\_age category are 17 , 22 .



***Observations and Insights :***

Probability of Teen Age Customer buying KP281 is 58.82 % , and KP481 is 41.18 %.

Probability of Adult buying KP281 is  41.48%  , KP481 is 33.33% and KP781 is 25.19%.

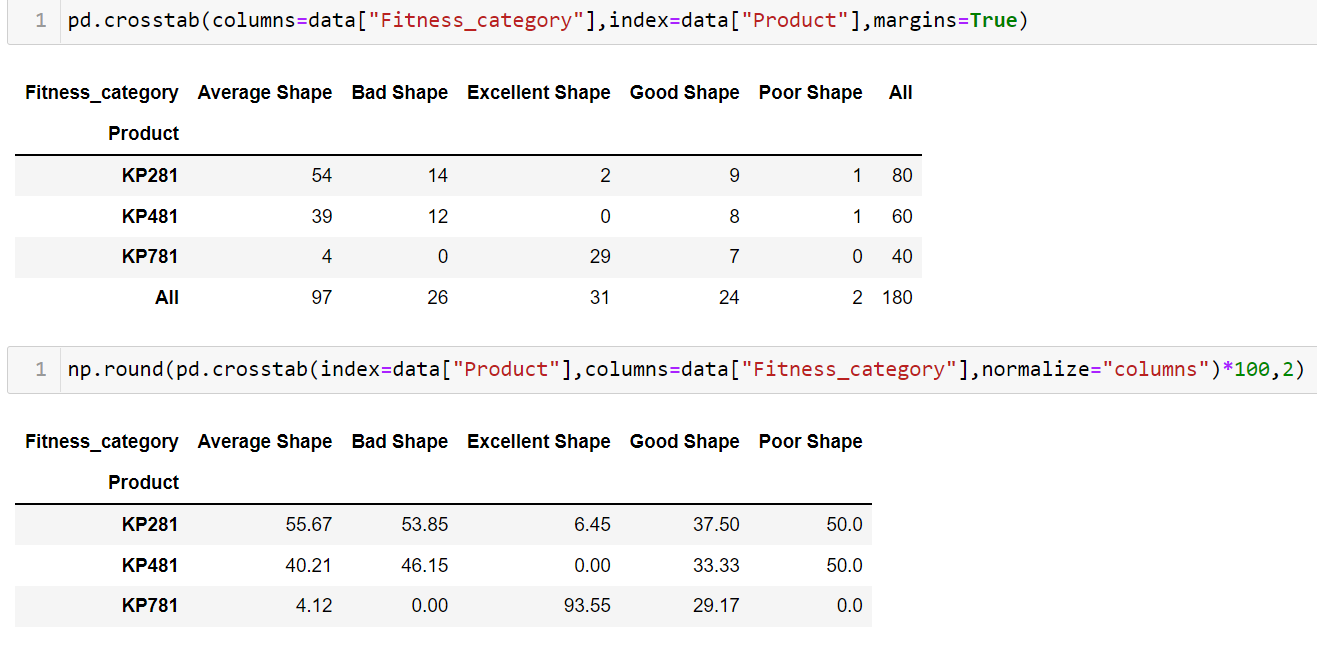
Probability of Customer age above 46 buying KP281 is 50%  , KP481 is 16.67% and KP781 is 33.33%.

Probability of Customer of mid age(36-45 years) buying KP281 is 50%  , KP481 is 31.82% and KP781 is 18.18%.

***Fitness-Categories and Product :***

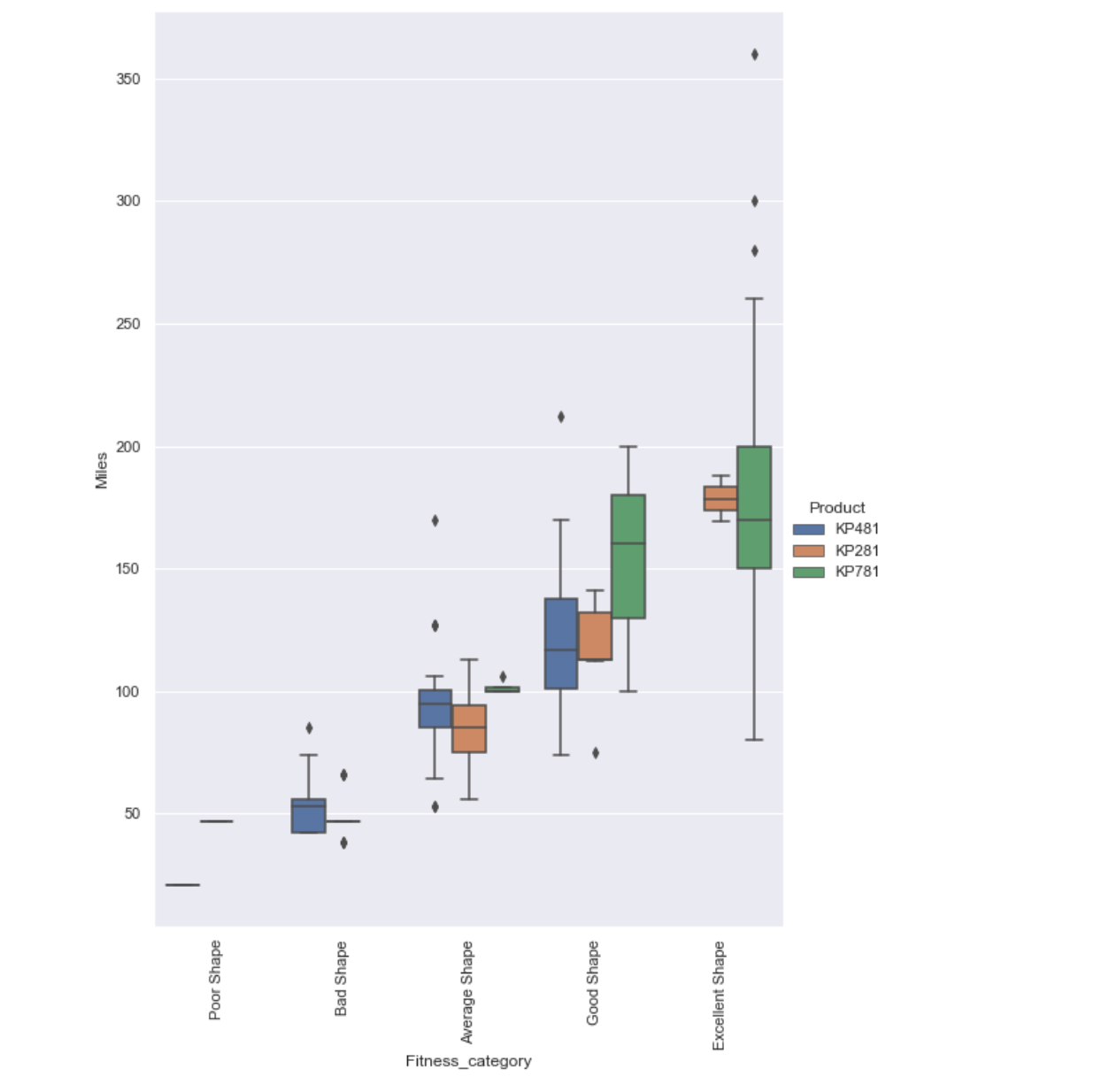
2nd table shows the conditional probabilities:

Product|FitnessLevel



***Observations and Insights :***

if person is in Excellent Shape  , the probability that he is using KP781 is more than 90%.



***Customer Profiling****- Categorization of users.*

**KP281 :**

- Most affordable and entry level and Maximum Selling Product.

-  This model popular amongst both Male and Female customers

-  Same number of  Male and Female customers.

-  Customers  walk/run average 70 to 90 miles on this product.

-  Customers use  3 to 4 times a week

-  Fitness Level of this product users is Average Shape.

- More general purpose for all age group and fitness levels.

Adult(22-35) 56

mid\_age(36-45) 11

Teen(0-21) 10

Towards\_old-age(>46) 3

**KP481 :**

- Intermediate Price Range

- Fitness Level of this product users varies from Bad to  Average Shape depending on their usage.

- Customers prefer KP481 model to use less frequent but to run more miles per week on this.

- Customer walk/run average  70 to 130 or more miles per week on his product.

- has higher probability of selling for female customers.

- Probability of Female customer buying KP481 is significantly higher than male.

 - KP481 product is specifically recommended for Female customers who are intermediate user.

- customers are from  adult, teen and mid-age  categories.

Adult(22-35) 45

Teen(0-21) 7

mid\_age(36-45) 7

Towards\_old-age(>46) 1

**KP781 :**

-  least sold product.

- high price and preferred by customers who does exercises more extensively and run more miles.

- Customer walk/run average  120 to 200 or more miles per week on his product.

-  Customers use  4 to 5 times a week at least.

- If person is in Excellent Shape  , the probability that he is using KP781 is more than 90%.

- Female Customers who are running average 180 miles (extensive exercise) , are using product  KP781, which is higher than Male average using same product.

-  KP781 can be recommended for Female customers who exercises extensively.

-  Probability of Male customer buying Product KP781(31.73%) is way more than female(9.21%).

-  Probability of a single person buying KP781 is higher than Married customers.

    So , KP781 is also recommended for people who are single and exercises more.

- most of old people who are above 45 age and adult uses this product.

Adult(22-35) 34

mid\_age(36-45) 4

Towards\_old-age(>46) 2