greatlearning
Power Ahead

Cardio Good Fitness Project

Contents

Business Problem Overview and Solution Approach

- **Core business idea**

Cardio Good Fitness is a retail store that sells high-quality cardiovascular exercise equipments such as treadmills, bikes, ellipticals, rowers etc. There are three different models of Treadmill sold here which are TM195, TM498 and TM798. A file containing the data of customers who purchased these Treadmills over a year has been provided. Various Customer factors such as Age, Gender, Marital Status, Income, Education, Expected Usage, Fitness etc. can affect the sales of these Treadmill models in positive or negative ways. We need to detect the patterns connecting these factors and come up with thoughtful insights in order to improve the Treadmill sales in the Retail Store.

- **Problem to tackle**

To extract actionable insights from the data that we have collected over the past 1 year to optimize resources and identify area of growth and improvement. We will be majorly focusing on the following problems :

- Come up with a customer profile (characteristics of a customer) of the different products
- Perform univariate and multivariate analyses
- Generate a set of insights and recommendations that will help the company in targeting new customers

- **Financial Implications**

There are three different models of Treadmill sold here which are TM195, TM498 and TM798. Once we evaluate the file containing the data of customers who purchased these Treadmills over one year, we would know how the various Customer factors such as Age, Gender, Marital Status, Income, Education, Expected Usage, Fitness etc. can affect the sales of these Treadmill models. From these Analysis we can come up with solutions to increase the sale of the Treadmill products and thus benefit the Retail store.

Data Overview

- **Data Description:**

- Product - The model no. of the treadmill
- Age - Age of the customer in no of years
- Gender - Gender of the customer
- Education - Education of the customer in no. of years
- Marital Status - Marital status of the customer
- Usage - Avg. # times the customer wants to use the treadmill every week
- Fitness - Self rated fitness score of the customer (5 - very fit, 1 - very unfit)
- Income - Income of the customer
- Miles- Miles that a customer expects to run

Observations	Variables	Duration
180	9	1 year

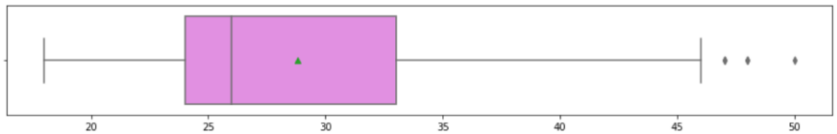
Significant manipulations made to raw data

- All columns have 180 observations. Hence there are no missing values.
- Product, Gender and MaritalStatus are object datatype i.e. String data type in Python. They were changed to Categorical variables as they have limited set of values.
- All other variables are Numerical variables with Integer datatype.

EDA

- **Univariate Analysis of Numerical Variables.**

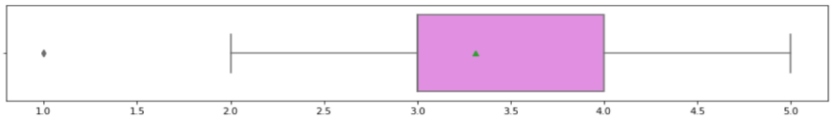
➢ **Age**



Observations

- Age is skewed towards the right.
- Mean age is 28, Median is 26 and Mode is 25. This indicated that Young customers are buying more treadmills.
- The oldest person buying the treadmill is aged 50.
- Very few customers with age less than 20 and greater then 40 are buying treadmills.

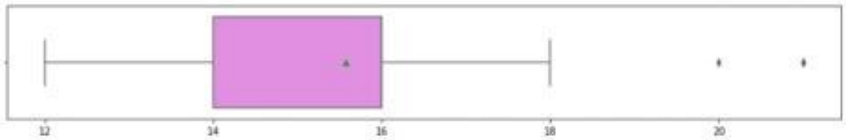
➢ **Fitness**



Observations

- Mode and Median is 3 for Fitness. Mean is slightly more than 3. It means most customers have Fitness level 3 which is a Moderately Fit level. The graph is slightly skewed to the right.
- Only few customers consider themselves unfit or very fit as most of them consider themselves as Moderately Fit.

➢ Education



Observations

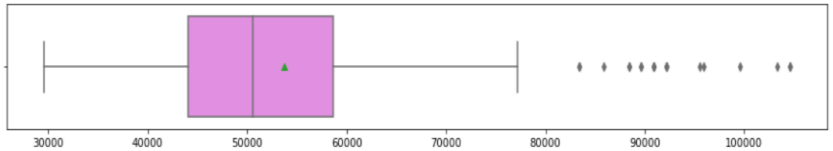
- Mode and Median for Education is the same which is 16 years.It means most of the cutomers are University Bachelor graduates. Mean is around 15.5 which is slightly less than the Median indicating the graph is slightly skewed to the left.
- There are only very few outliers to the right indicating very few customers who are University post graduates with more than 18 years of education are buying the Products.

➢ Usage

Observations

- Mode and Median is 3 for Usage. Mean is slightly more than 3.
- The graph is slightly skewed to the right and there are few outliers to the right.
- Very few customers use the product more than 5 times a week.

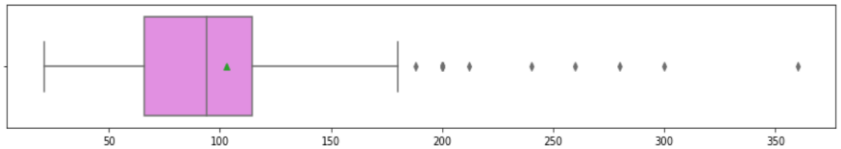
➢ Income



Observations

- Most of the customers have income in the range of 30000 to 70000.
- Mean is around 53000 and median is around 50000.
- The graph is slightly skewed to the right and there are few outliers to the right.
- Only few customers buying the product have income more than 80000 and less than 30000.

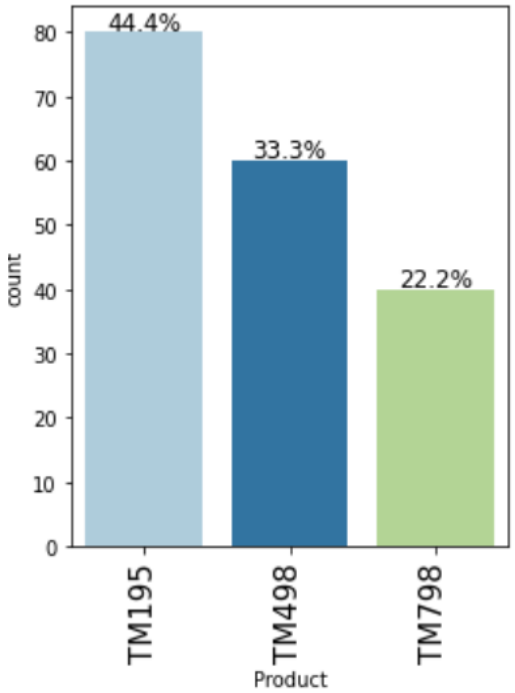
➢ Miles

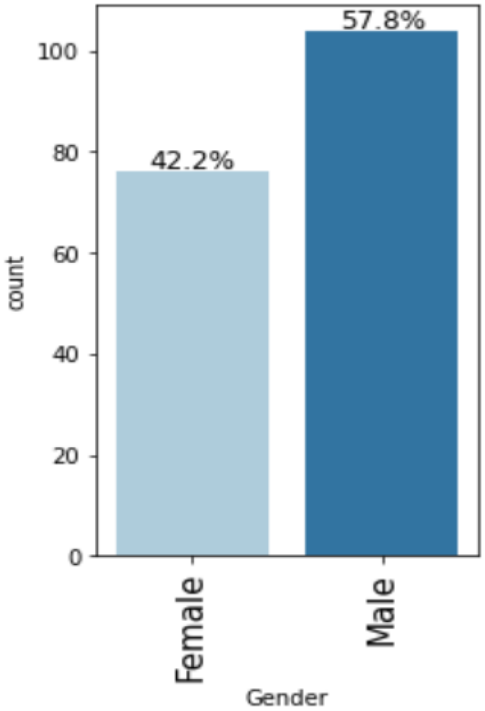


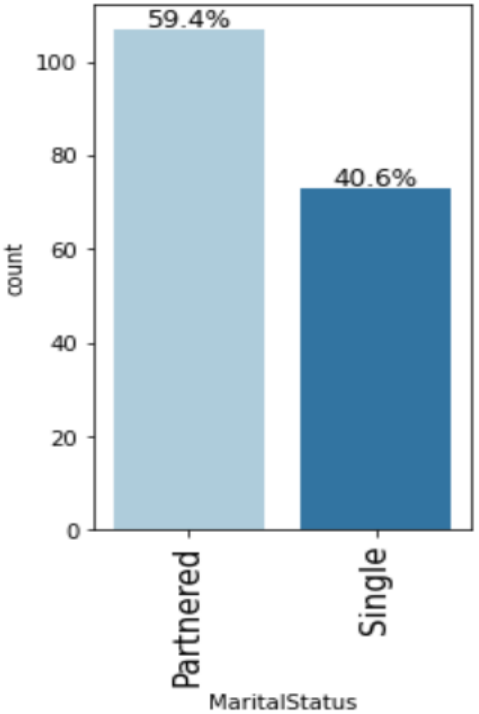
Observations

- Majority of the customers are expecting to run 40 to 120 miles per week. Most of them are targetting 80 miles per week.
- Mean is around 103 and median is around 94.
- The graph is slightly skewed to the right and there are some outliers to the right.
- Only few customers expecting to run more than 200 miles per week.

Univariate Analysis of Categorical Variables.







Observations

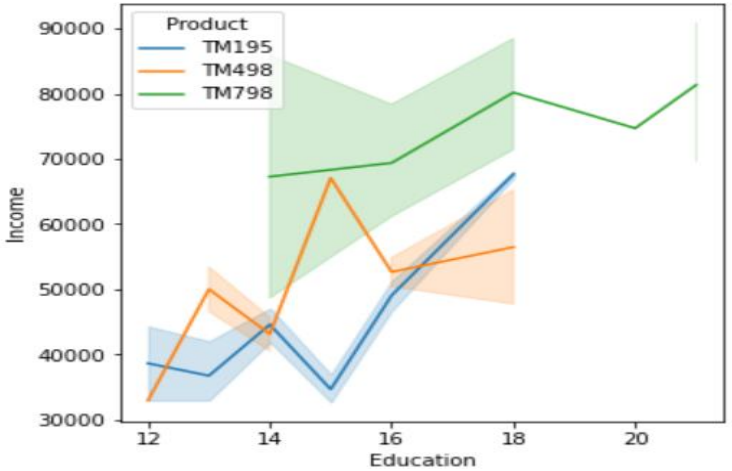
- The most purchased model is TM195 and the least purchased model is TM798.
  Males are buying the product more than females.
  Partnered customers are buying the product more than Single customers.

Bivariate Analysis – Correlation Matrix

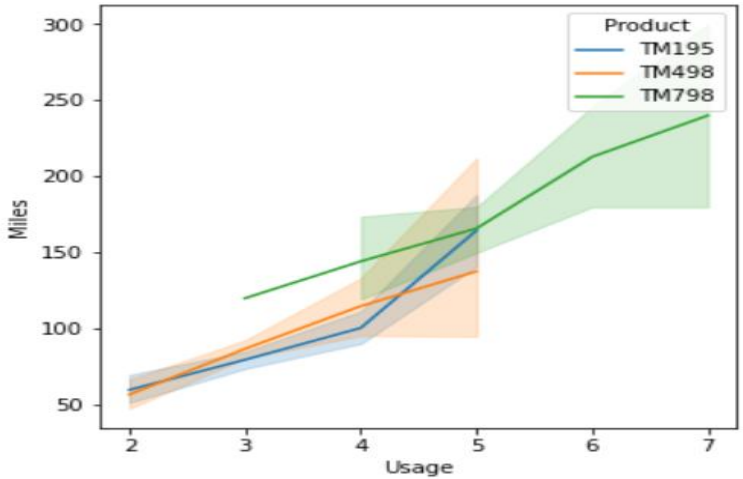
Observations

- Miles, Usage and Fitness have strong correlation with each other. As more number of miles are covered, there is more Usage and Fitness level will also increase. Similarly with more Usage , more miles are covered. And with more Fitness there is more Usage.
- Age and income have significant correlation.
- Age has almost no correlation with Usage, Fitness and Miles.
- Age has some correlation with Education.
- Education has significant correlation with income. Higher education makes way to higher income.
- Education has some correlation with Usage, Fitness and Miles.
- Income has almost the same amount of correlation with Age, Usage, Fitness, Miles and a bit more strong correlation with Income.

Multivariate Analysis



Education, Income and Product

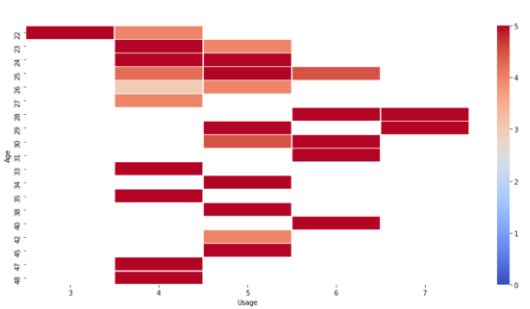


Usage, Miles and Product

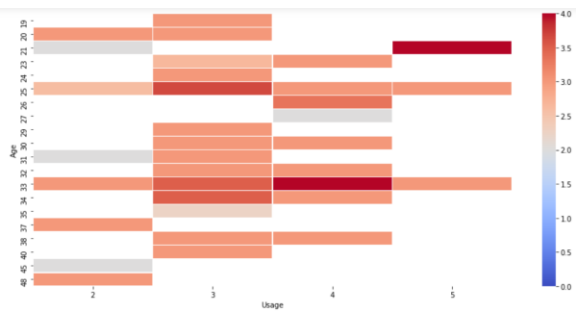
Observations

- For all three products , Income increases with increase in years of Education.
- Customers with less Education have less Income and they prefer to buy the Product TM195 as it might be less costly.
- Customers with a middle range of income prefer to buy TM498.
- Customers with higher Education have high income and they prefer to buy the Product TM798 as they can afford it.
- For all three products, Usage and Miles increase together in a linear way.
- TM195 and TM498 have almost the same usage but TM195 surpasses TM498 in the Miles expected to be covered.
- TM798 has the most Usage and higher number of miles expected to be covered.

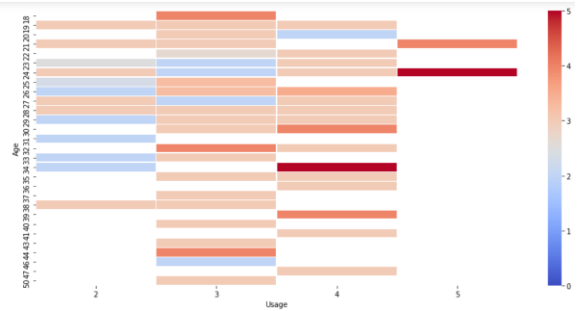
Product Heat Maps



TM798



TM498



TM195

Observations

- **TM798** has high Fitness levels throughout with all age groups.
- Most Usage is for Level 4 and 5 . And the least one is for Level 3.
- Customers with age 28 and 29 have highest usage of 6 and 7 times per week.
- Customers with Age 31 to 48 have high Fitness levels and Usage between 4 and 6.
- **TM498** does not have as high Fitness level as TM798 with the highest level being 4.
- Customers aged 21 had highest usage of 5 with Fitness Level 4.
- There are many customers with Usage as low as 2 times per week and Fitness levels less than 3.
- Most customers have Usage 3 times per week.
- **TM195** most customers have Usage levels 3 and 4.
- Customers with Usage 2 times per week have very less fitness levels.
- Customers of all ages enjoy this Product more compared to TM798 and TM498.

Business Insights and Recommendations

Conclusion

We analyzed a dataset of 180 Customers who have bought the three differnt Models of TreadMills TM195,TM498 and TM798. The data spanned over one year. The main factors affecting the sales of Products are the Customers Age, Income, Marital Status, Gender, Usage, Fitness, Miles expected to run and Education. We have been able to conclude that :

- TM195 is the most bought Treadmill. It is also the cheapest and attracts customers with low income and less Usage.
- TM798 is the least bought Treadmill. It is bought by customers with High Education, Good income and who have High Fitnes Levels.They also expect to cover longer miles.
- Male Customers are buying Treadmills more than the female customers.
- Customers with Partners are buying Treadmills more than the Single customers.
- All three treadmills are mainly used by the younger generation of age 20-35.
- Miles, Usage and Fitness have strong correlation with each other. As more number of miles are covered, there is more Usage and Fitness level will also increase. Similary with more Usage , more miles are covered. And with more Fitness there is more Usage.
- Age has almost no corelation with Usage, Fitness and Miles.
- Education has significant correlation with income. Higher education makes way to higher income.

Recommendation to business

- As young generation are the main customers, attract more young customers by using a youth icon as Ambassador.
- As there are less Female customers when compared to Males, organize and sponsor Women Fitness Awareness Programmes so they become more aware of the Products.
- As there are less customers above the age of 35, more ads should be designed in a way that should reach out to the older customers as well.
- TM195 being the cheaper model should be able to attract wide range of customers with less income.
- As TM798 is used by Very Fit customers , it should be projected to be the Treadmill King of Fitness and more fitness enthusiasts should be able to buy them.

greatlearning
Power Ahead

Happy Learning !