

Cardio Good Fitness

New Customer Targeting Exploratory Data Analysis

April 2021

Brian Hall – DSBA – Project 1

03 | Objective

04 | Data Overview

05 | Customer Overview

06 | Product Overview

07 | Customer Profiles – TM195

08 | Customer Profiles – TM485

09 | Customer Profiles – TM795

10 | Recommended Next Steps

11 | Additional Data Recommendations

Appendix

13 | Uni-Variate Data Analysis

19 | Multi-Variate Data Analysis

Objective

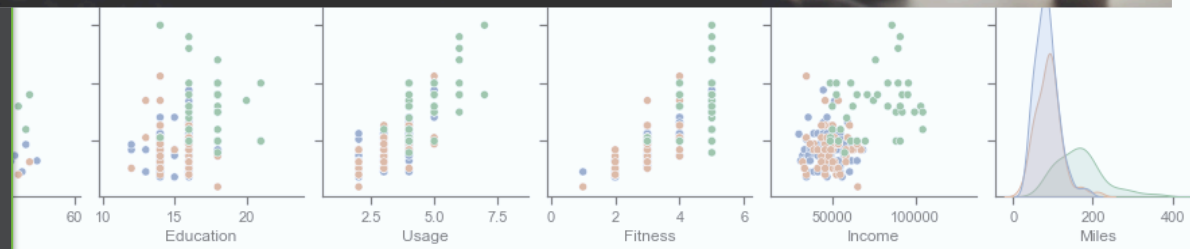
Preliminary Data Analysis for the Cardio Good Fitness dataset, which contains data of the customers of the treadmill products of a retail store called Cardio Good Fitness

1. Produce a customer profile (characteristics of a customer) of the different products
2. Perform uni-variate and multi-variate analyses
3. Generate a set of insights and recommendations that will help the company in targeting new customers



April 2021

Product
● TM195
● TM498
● TM798



Data Overview

The data is for customers of the treadmill products of a retail store called Cardio Good Fitness (period of time unspecified)

Data has been provided via CSV (CardioGoodFitness.csv | 7.28k)

Variable

Description

Product	The model # of the treadmill
Age	In # of years, of the customer
Gender	Of the customer
Education	In # of years, of the customer
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	Of the customer
Miles	Expected to run

The Data Contains:

- 180 rows and 9 columns
- No missing data & no duplicate rows
- Column Keys match data description provided
- 6 numerical attributes of type int64
- 3 Categorical attributes of type object

Customer age groups will be generated and added to the DataFrame as "Age Group" with bins = [1,12,19,30,64, np.inf]

Miles - No time period has been given for expected miles to run - observations range from 21 to 360 with many high outliers.

Customer Overview

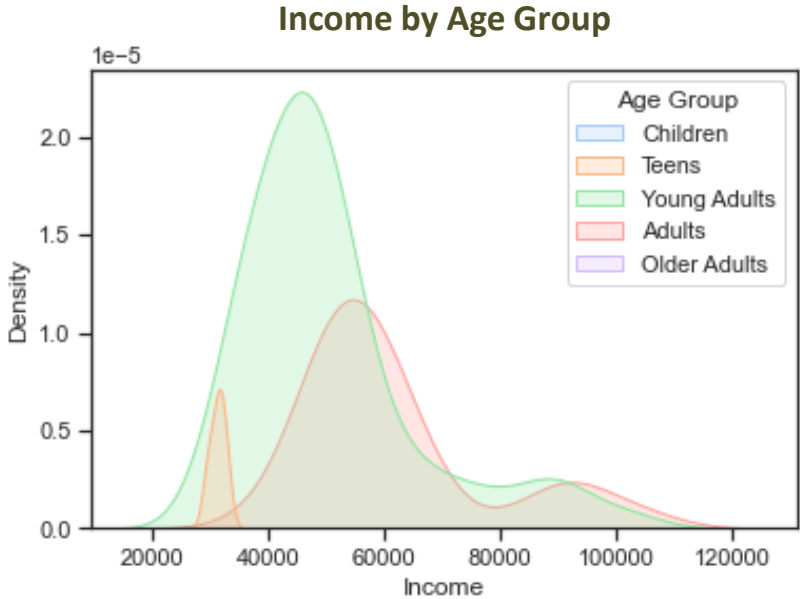
180 observations

Young Adult (age 20 - 30)	64%
Adult (age 30 - 64)	33%
Teen (age 13 - 19)	2%
Older Adult (age 65+)	0%
Mean Age	29
Age Range	18 - 50

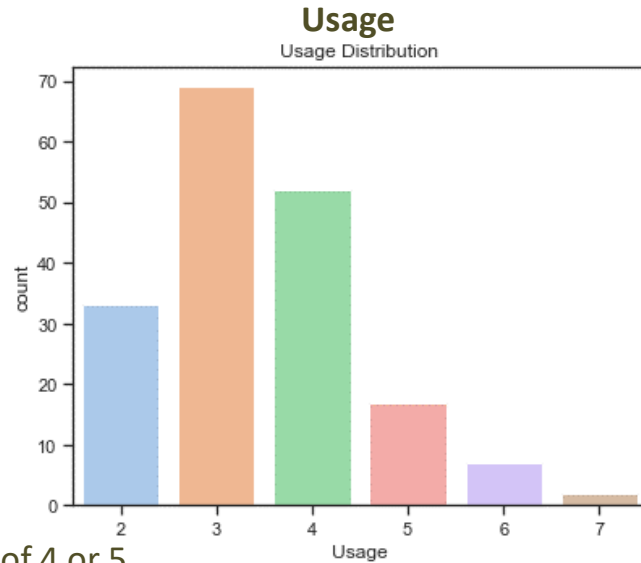
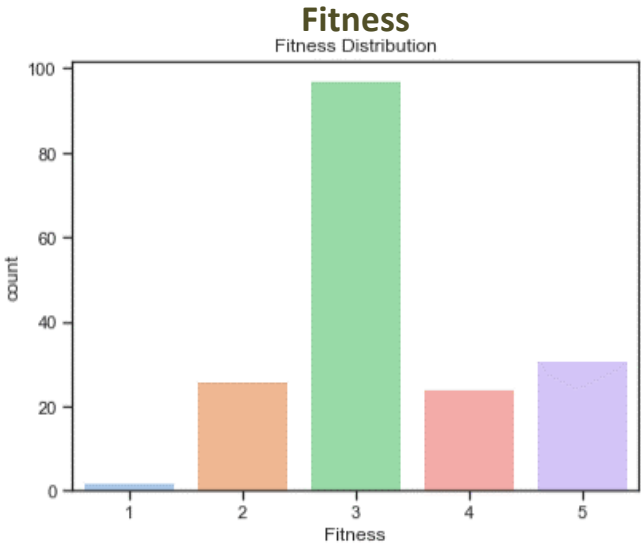
Males	58%
Females	42%

Average Income	\$53,719
Income Range	\$29,562 - \$104,581

See appendix for data analysis and distributions

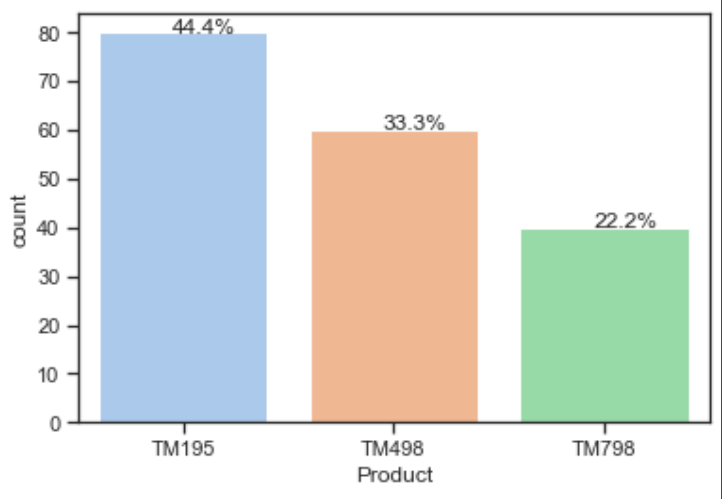


- Sales to Young Adults are higher across all products
- 68% of customers are partnered
- 47% of customers have at least a 4 year degree, 31% have at least 2 years of post secondary & 15% of customers have higher than a 4 year degree
- Nearly 54% of users rate their fitness at a 3 (scale 1-5) and 32% of users rate as a 4 or 5
- 66% of customers want to use the product 3 to 4 times a week, 18% want to use 2 times a week
- All customers with an income > \$70,000 report a usage of 4 and above
- Most customers with an income > \$70,000 report a fitness of 4 or 5
- Income increases by about 20% for partnered customers starting around \$45,000
- High correlations for Fitness, Usage & Miles - 0.76 - 0.79
- Mild correlations for Income, Miles, Fitness & Usage - 0.5
- Very low correlation for Age, Miles, Fitness & Usage - 0.015 to 0.03

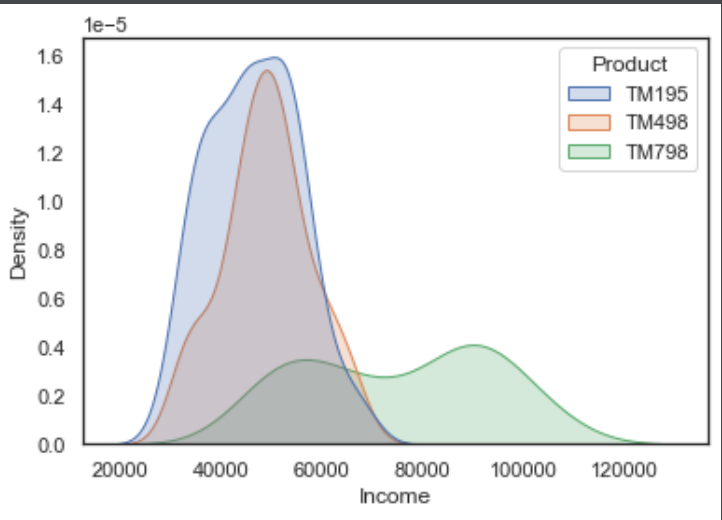


Product Overview

Product Sales



Income of buyers



See appendix for data analysis and distributions

TM195

44% of sales and is primarily purchased by customers with **\$30,000 to \$60,000** income with low to mid level fitness.



TM498

33% of sales with most of it's purchases by customers with incomes in the **\$40,000 to \$70,000** range with mid level fitness.



TM798

22% of sales to primarily males across most incomes with a high level of fitness. It is the **only model that has higher income customers, starting around \$70,000.**



Actual products not shown

Customer Profiles by Product

TM195

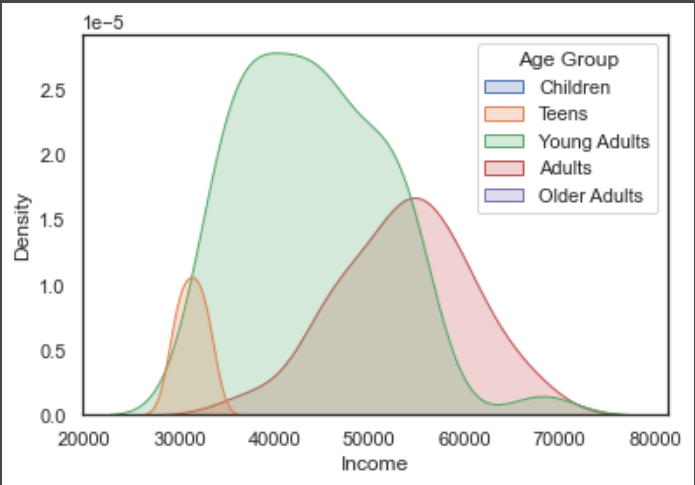
Educated
Young Adults & Adults,
'entry level' product



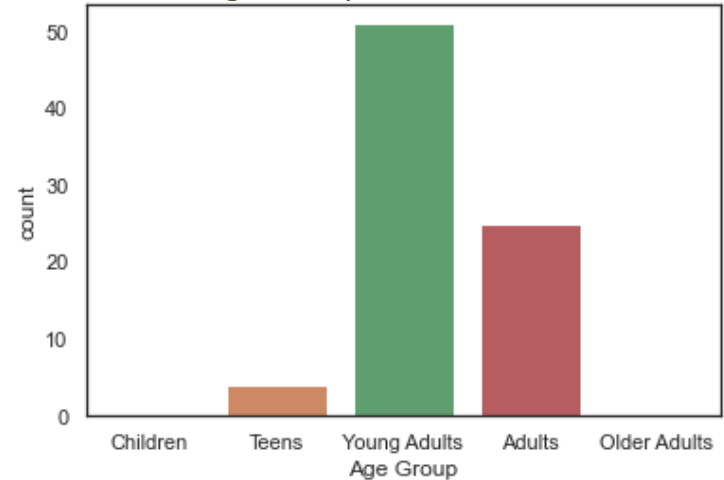
44% of sales and is primarily purchased by customers with **\$30,000 to \$60,000** income with low to mid level fitness.

- Young Adults account for about 63% of all sales
- Customer gender is roughly even
- There are roughly 25% more partnered customers
- TM195 is the only model purchased by Teens
- 51% of customers have 16+ years of education and 38% have 14
- Average income of customers is \$46,400
- Young Adult income distribution between \$30,000 & \$60,000
- Adult income distribution primarily between \$40,000 & \$70,000
- Those who purchase the TM195 report a mid level of fitness
- Customers want to primarily use the TM195 2 - 4 days a week with 46% at 3 days

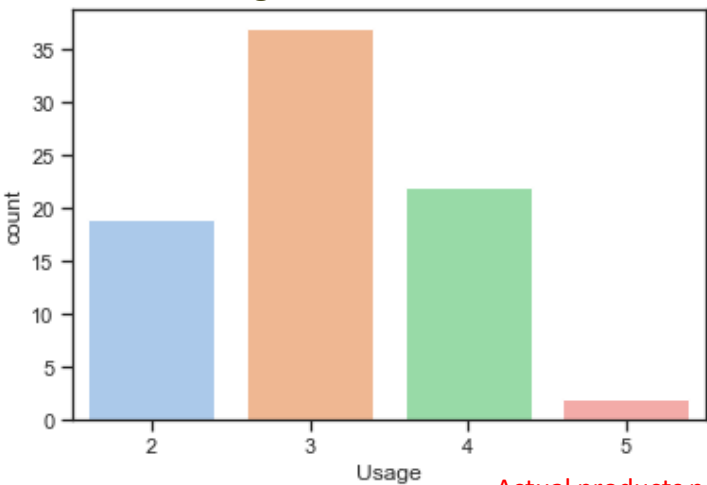
Income of buyers of TM195



TM195 Age Groups



TM195 Usage



Customer Profiles by Product

TM498

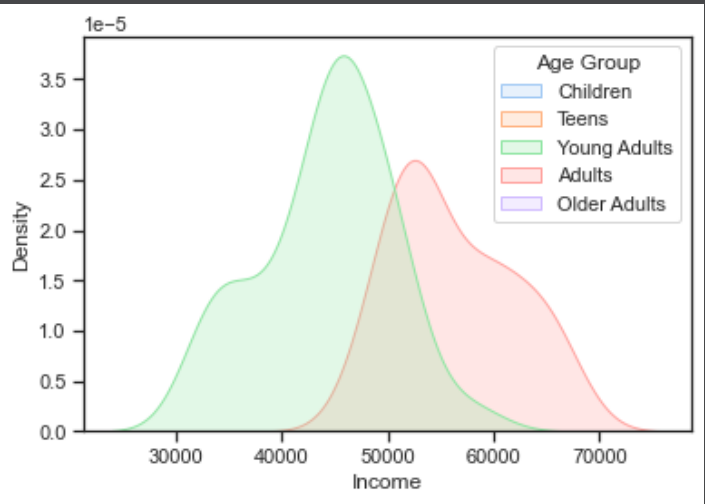
Educated
Young Adults & Adults,
‘mid-level’ product

33% of sales with most of it's purchases by customers with incomes in the \$40,000 to \$70,000 range with mid level fitness.

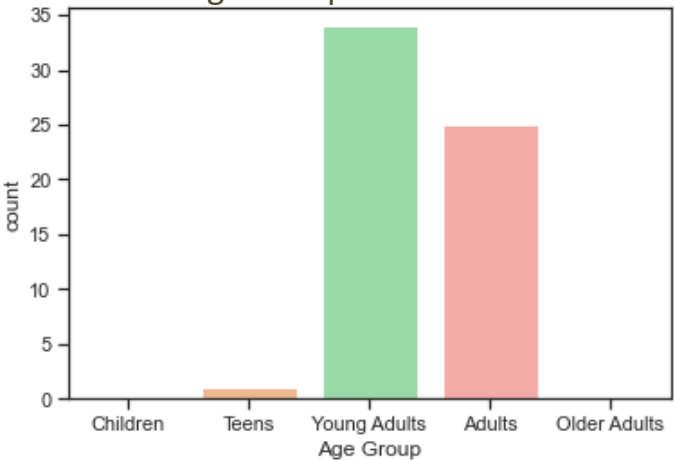


- Young Adults account for 57% of all sales and Adults 42%
- Customers gender is roughly even
- There are roughly 25% more partnered customers
- 95% of customer have 14+ years of education, with 55% having 16+ years
- Average income of customers is \$49,000
- Young Adult income distribution primarily between \$30,000 and \$60,000
- Adult income distribution primarily between \$45,000 & \$70,000
- Those who purchase the TM498 report a mid level of fitness
- Customers want to primarily use the TM498 2 - 4 days a week with 52% at 3 days

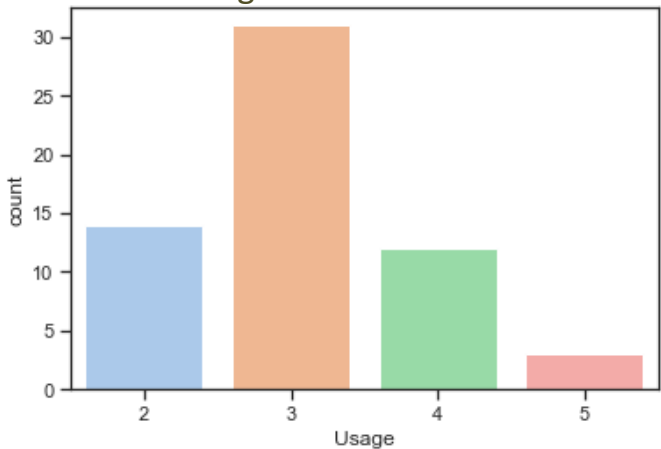
Income TM498 Buyers



TM498 Age Groups



TM498 Usage

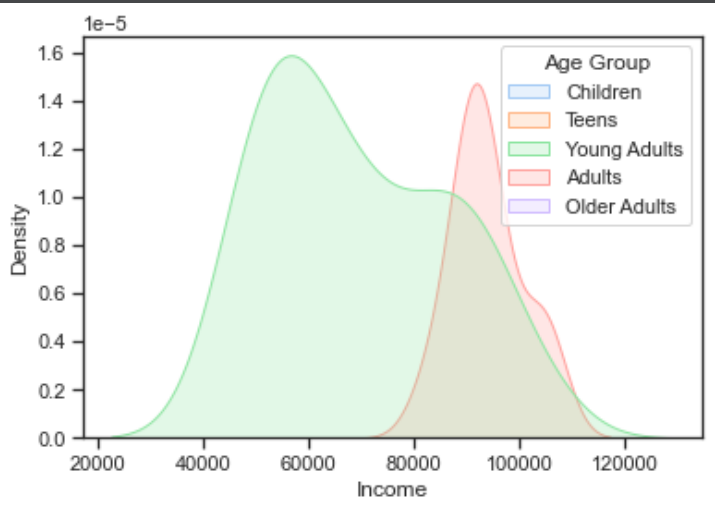


Customer Profiles by Product

TM798

Highly Educated
Young Adult & Adult
Male Fitness
Enthusiasts

Income TM798 Buyers



See appendix for data analysis and distributions

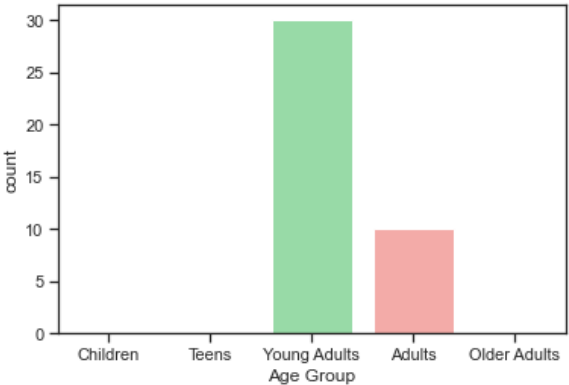
22% of sales to primarily males across most incomes with a high level of fitness. It is the only model that has higher income customers, starting around \$70,000



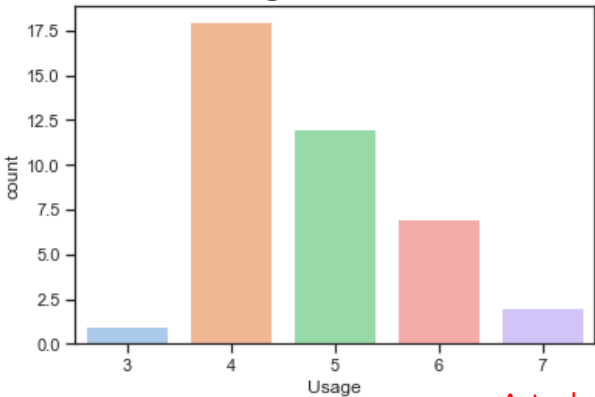
April 2021

- Young Adults account for about 75% of all sales
- Males account for 80% of all sales
- All customers have 14+ years of education, with 58% having 18+ years
- Average income of customers is \$75,500
- Young Adult customers have fairly evenly distributed incomes
- All Adult customers make \$80,000 and above
- Those who purchase the TM798 all report high fitness, mostly 5
- 92% of customers want to use the TM798 4+ times a week
- Speculation - 798 additional features may primarily be attractive to fitness enthusiasts
- Speculation - 798 customers may purchase for status & perception of quality

TM798 Age Groups



TM798 Usage



Actual products not shown

Recommended Next Steps



Increase marketing for TM798 to adult males with > 50,000 income

Message - Status & "Better fitness" features

- Goal is to increase the attention of a wider income range of adult fitness enthusiasts
- Goal is to increase the attention of more adults with mid level fitness to "better features", etc.



Increase marketing for TM798 to Young Adult & Adult female fitness enthusiasts

Message – "Better fitness" features

- Goal is to increase the attention of female fitness enthusiasts



Increase marketing for TM498 to Partnered Young Adults with > 50,000 income and possibly newer home owners

Message - Long lasting, mid-range product for the family

- Goal is to increase the attention of Partnered Young Adults with mid level and above income, possibly starting a family



Increase marketing for TM195 to adults with < 45,000 income

Message - Less expensive option to get or stay fit

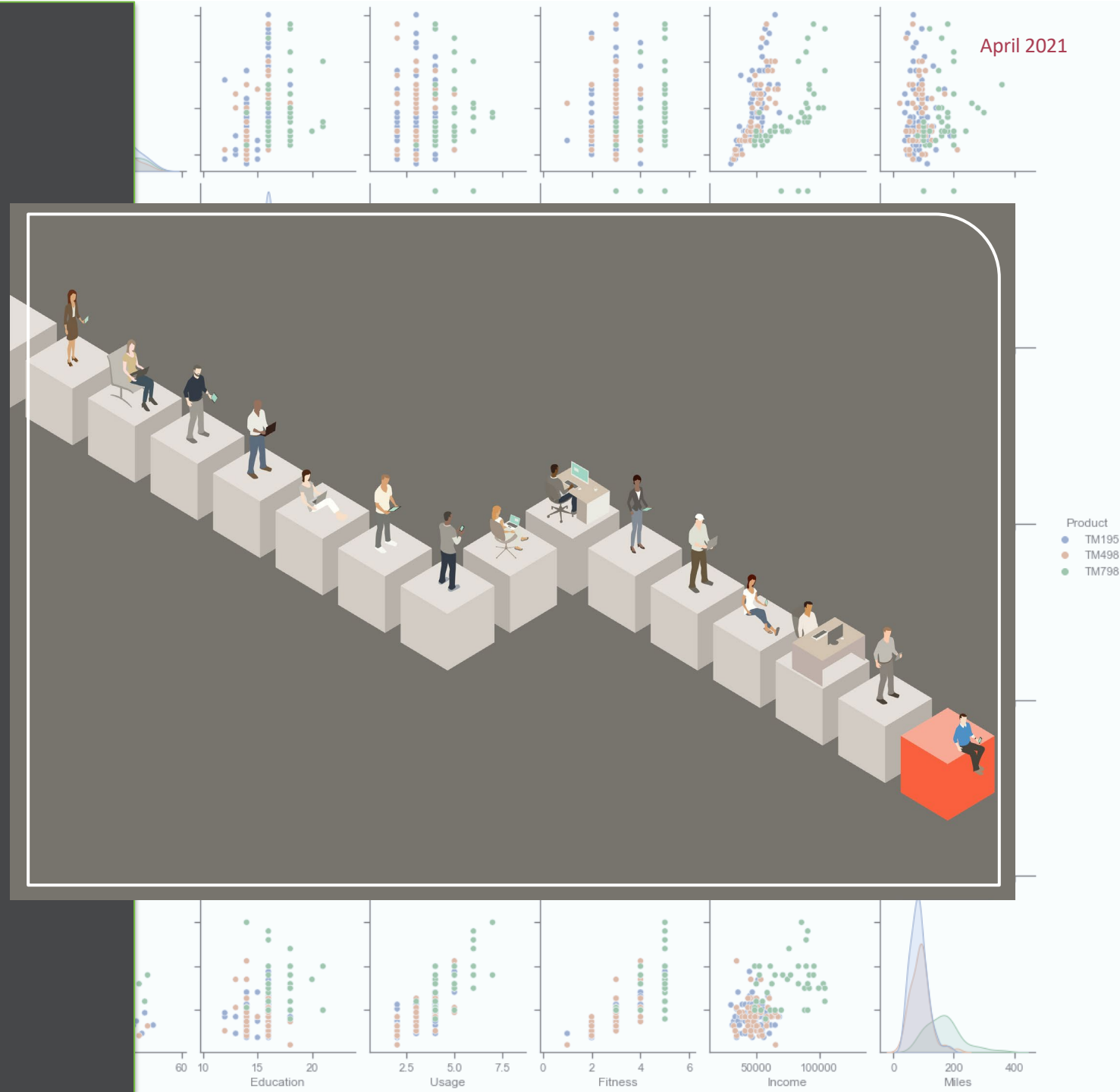
- Goal is to increase the attention of lower income adults

Additional Data Recommendations

Obtaining additional data corresponding to customers & products would significantly increase the insights needed to accurately analyze and recommend marketing, production, design & feature needs.

Potential Data to Acquire

1. Primary user surveys
2. Actual usage metrics
3. Rent vs Own (home)
4. Student vs Full Time
5. Geo-spatial data
6. Product specifications & features
7. Existing marketing campaigns and outcomes
8. Product cost, sales & production
9. Competitive product data & usage



Appendix

Cardio Good Fitness – Customer Targeting

April 2021

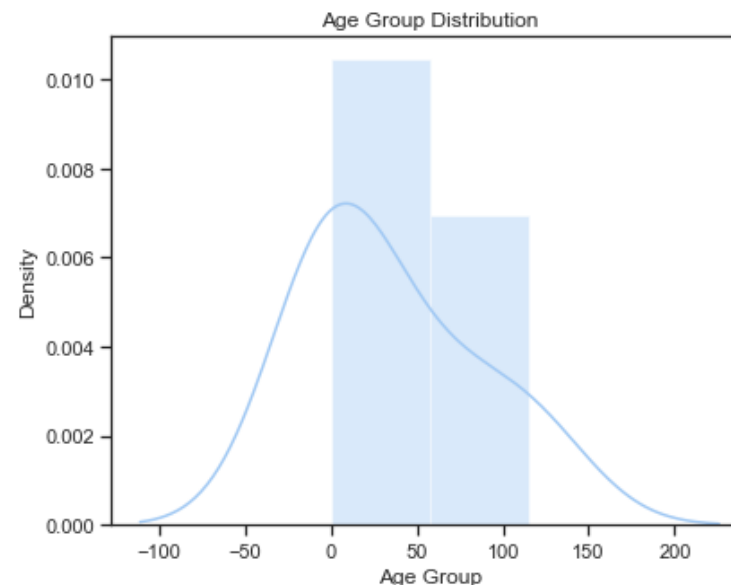
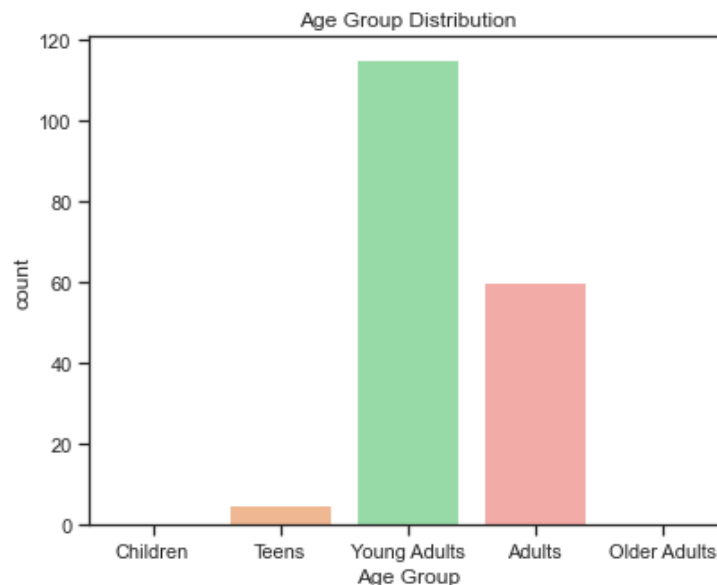
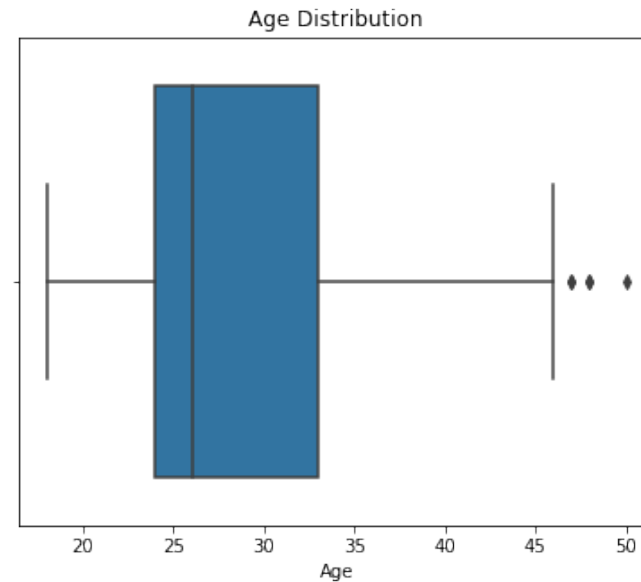
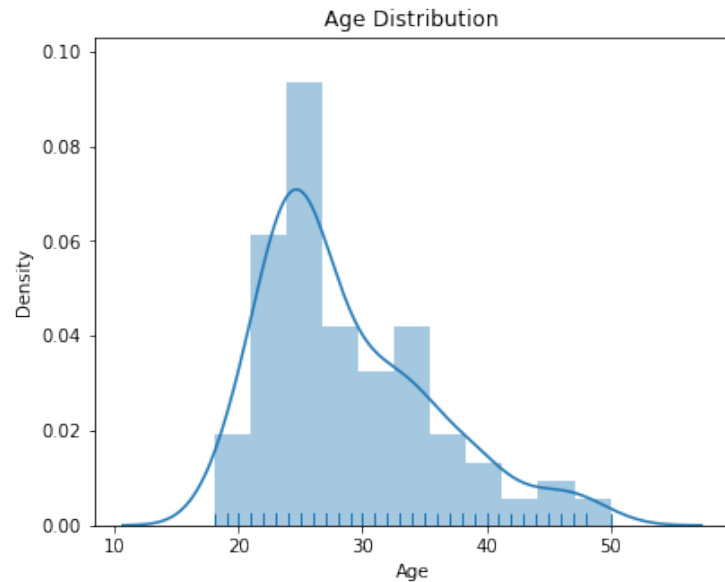
Uni-Variate Analysis

Cardio Good Fitness – Uni-Variate Analysis

April 2021

Univariate Analysis

Age & Age Group Distribution



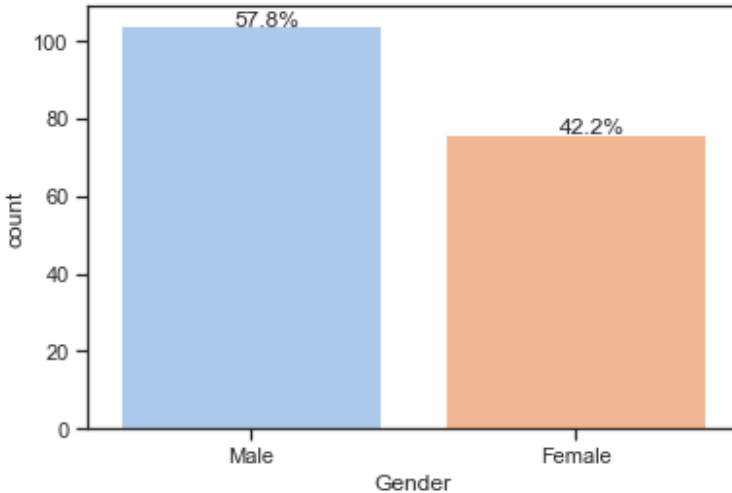
Observations

- Ages are right skewed with few outliers
- Mean age is 29
- Age range is 18 – 50
- Standard Deviation is 6.94
- Age groups will be created and primarily used in further observations

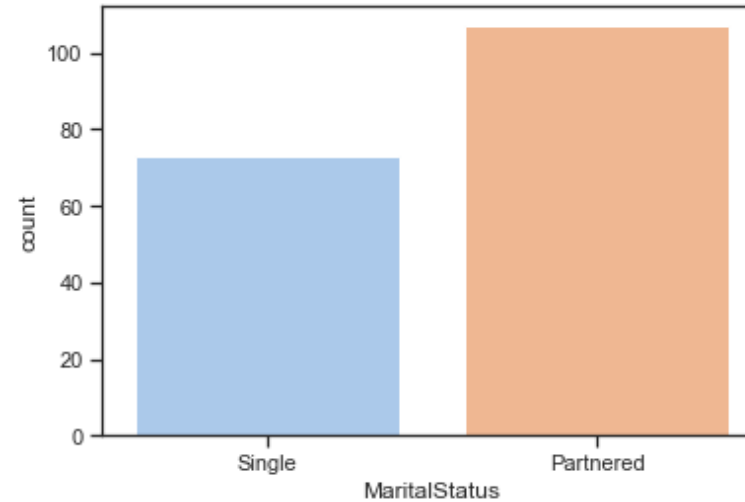
- Young Adult (age 20 - 30) 64%
- Adult (age 30 - 64) 33%
- Teen - (age 13 - 19) 3%
- Older Adult (age 65+) 0%
- Children (age under 13) 0%

Univariate Analysis

Gender Distribution



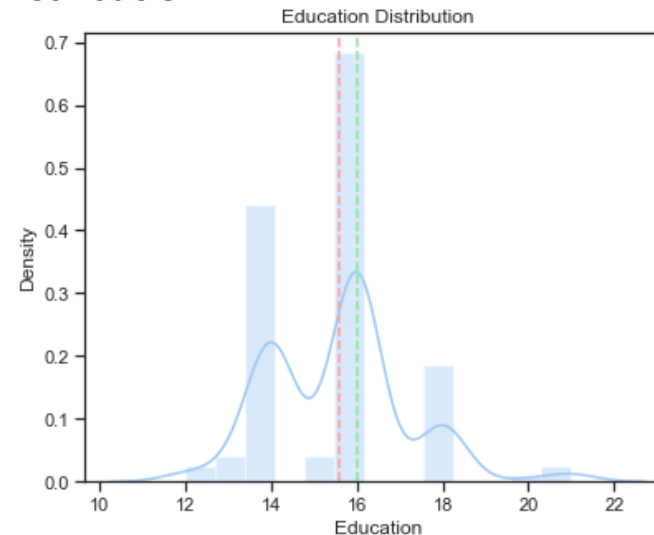
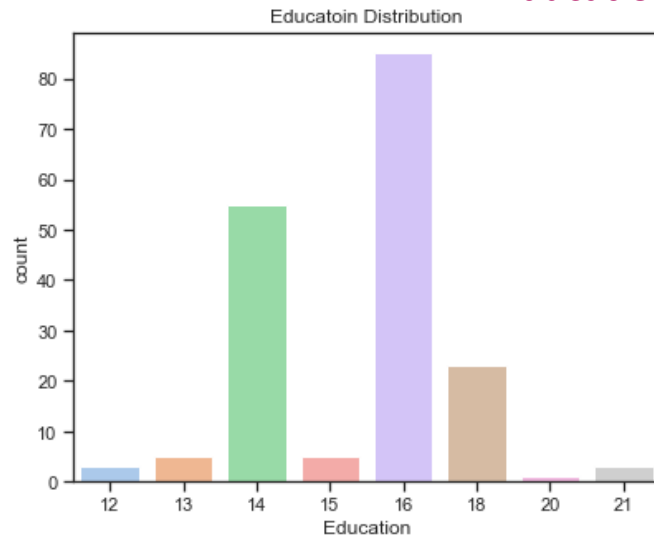
Marital Status



Observations

- 58% of customers are male
- 42% of customers are female
- 68% of customers are partnered

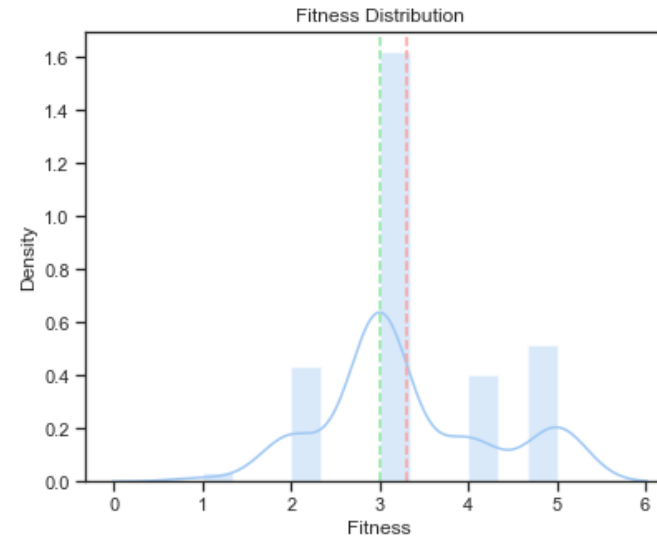
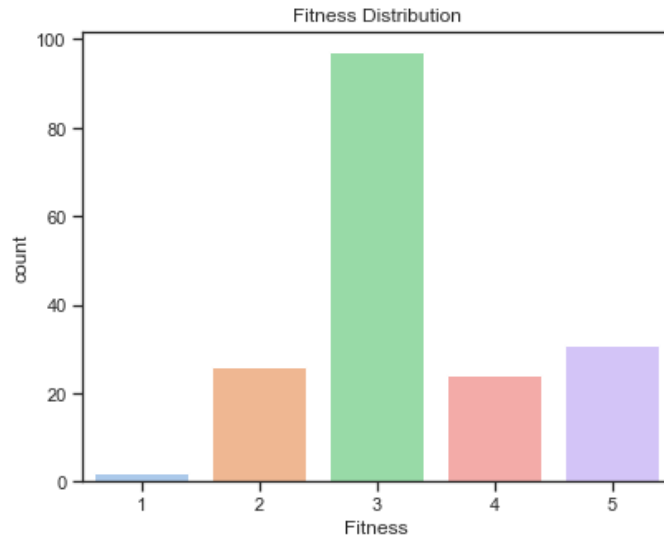
Education Distribution



- Majority of customers have at least 16 years of education
- Mean Education is 16.12 years
- 47% of customers have at least a 16 years of education
- 31% of customers have at least 14 years of education
- 15% of customers have more than 18 years of education

Univariate Analysis

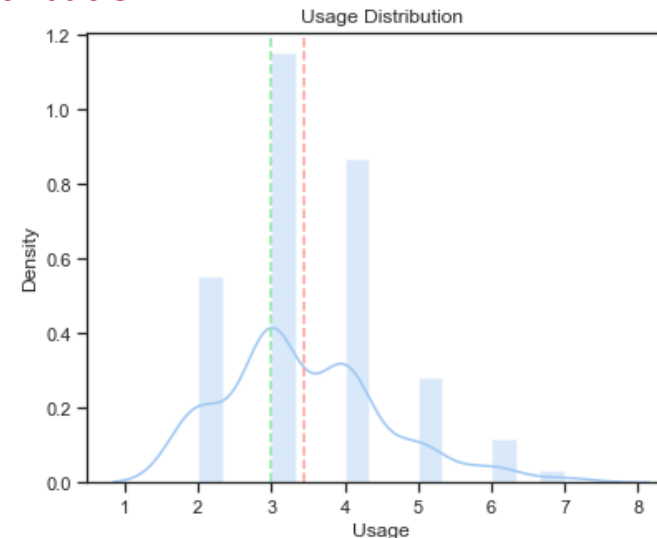
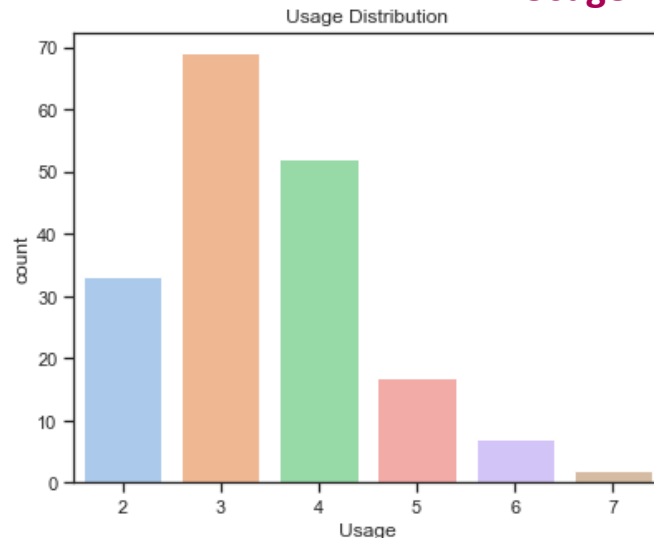
Fitness Distribution



Observations

- Fitness is a self-reported attribute
- Nearly 54% of users rate their fitness at a 3 (scale 1-5)
- 32% of users rate as a 4 or 5
- Mean fitness - 3.31
- Very few people report themselves with a fitness of 1

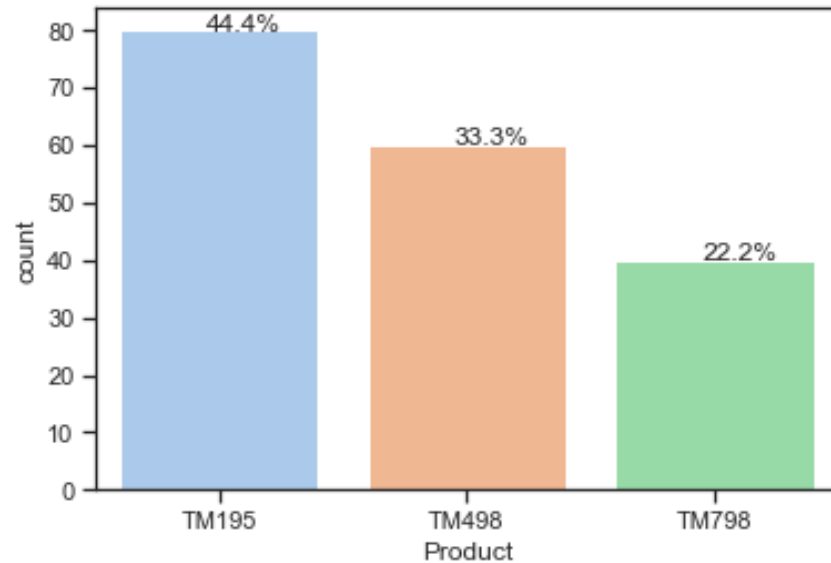
Usage Distribution



- Usage is a self-reported attribute
- Usage is fairly right skewed
- Majority of customers, 66% wants to use the product 3 to 4 times a week
- 18% want to use 2 times a week

Univariate Analysis

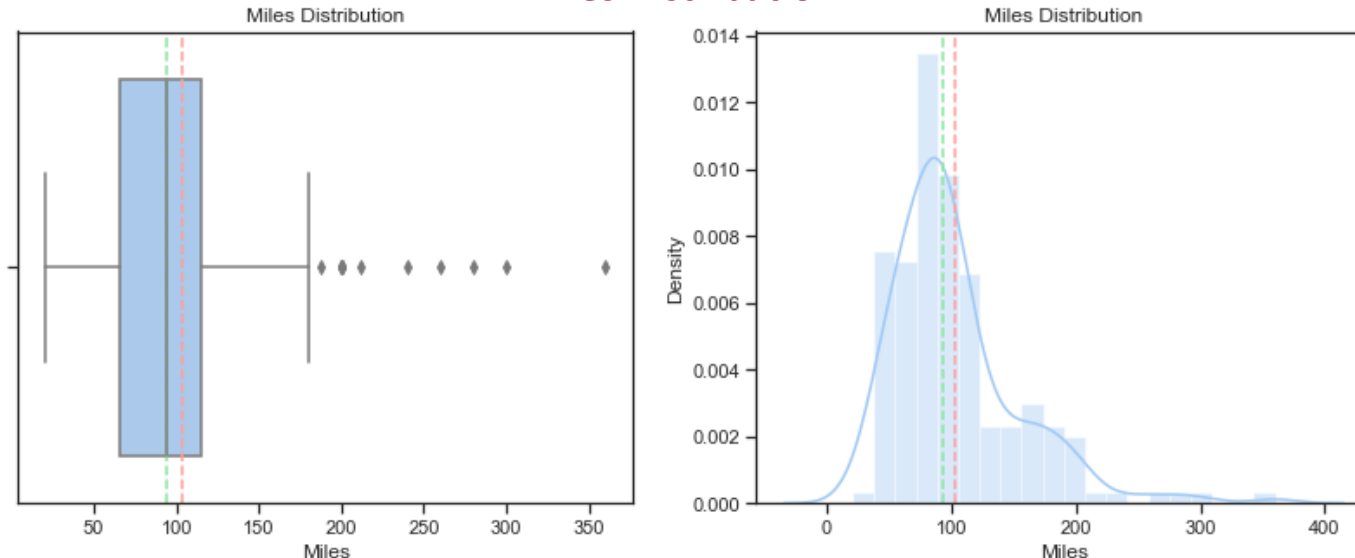
Product Distribution



Observations

- TM195 has the highest sales at 44.4%
- TM498 accounts for 33.3% of sales
- TM798 accounts for 22.2% of sales

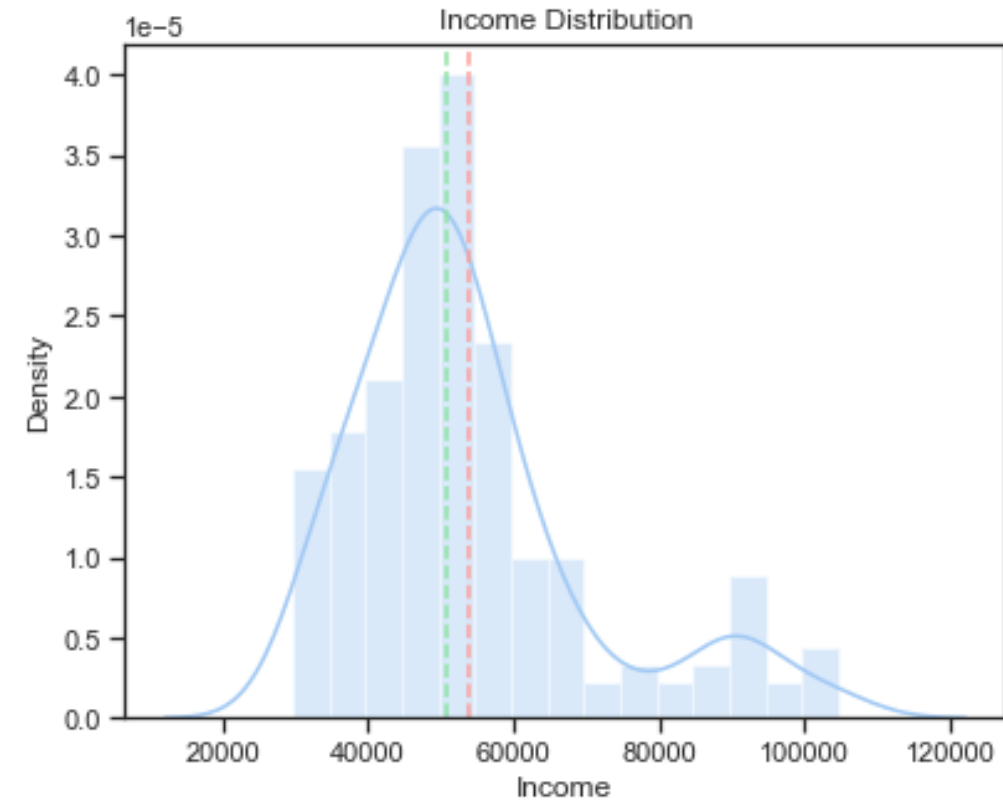
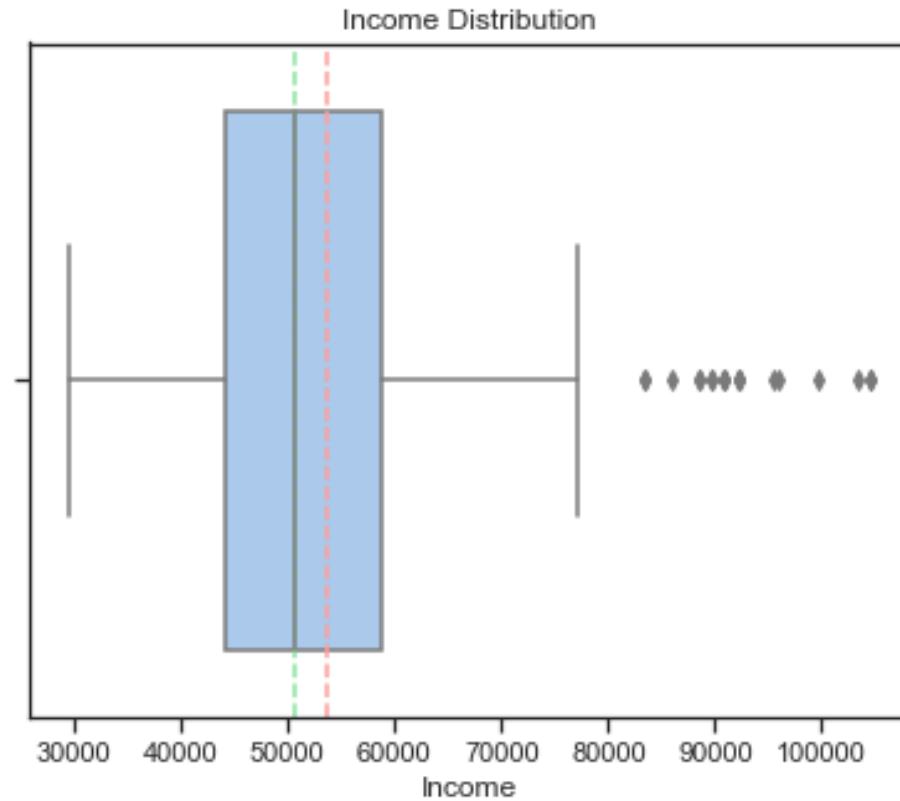
Miles Distribution



- Miles is a self-reported attribute
- Mean - 103
- Range - 21 – 360
- Standard Deviation - 52
- Many extreme outliers
- Highly skewed to the right

Univariate Analysis

Income Distribution



Observations

- Most income between \$40,000 - \$60,000
- Many extreme outliers
- Fairly skewed to the right
- Average income is \$53,719
- Range \$29,562 - \$104,581

Multi-Variate Analysis

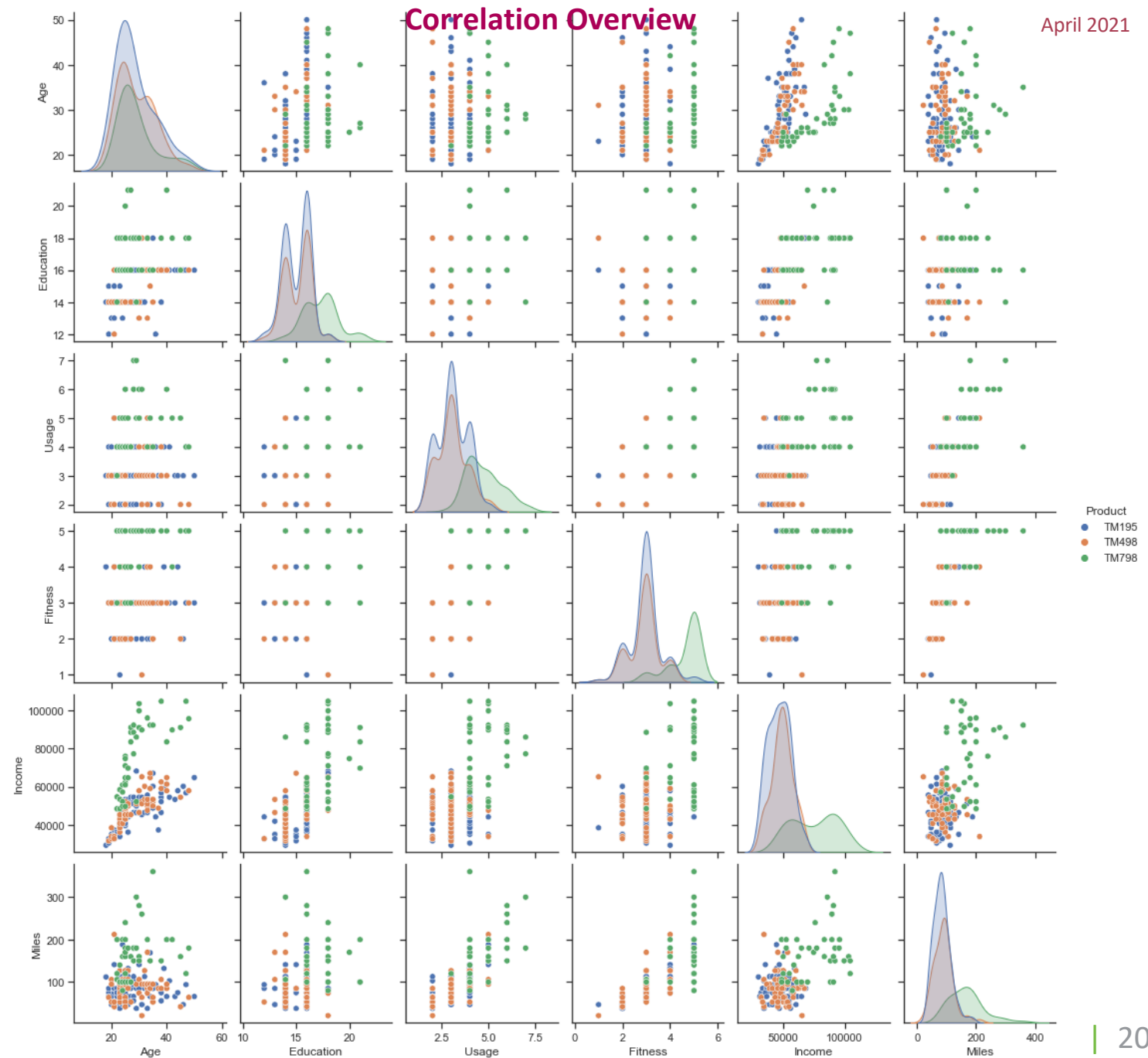
Cardio Good Fitness – Multi-Variate Analysis

April 2021

Multi-Variate Analysis

Observations

- Highest correlations are between Fitness, Usage & Miles - 0.67 to 0.79
- Moderate correlations for Income & Education - 0.63
- Mild correlations for Income, Miles, Fitness & Usage - 0.5
- Very little correlation for Age, Miles, Fitness & Usage - 0.015 to 0.037



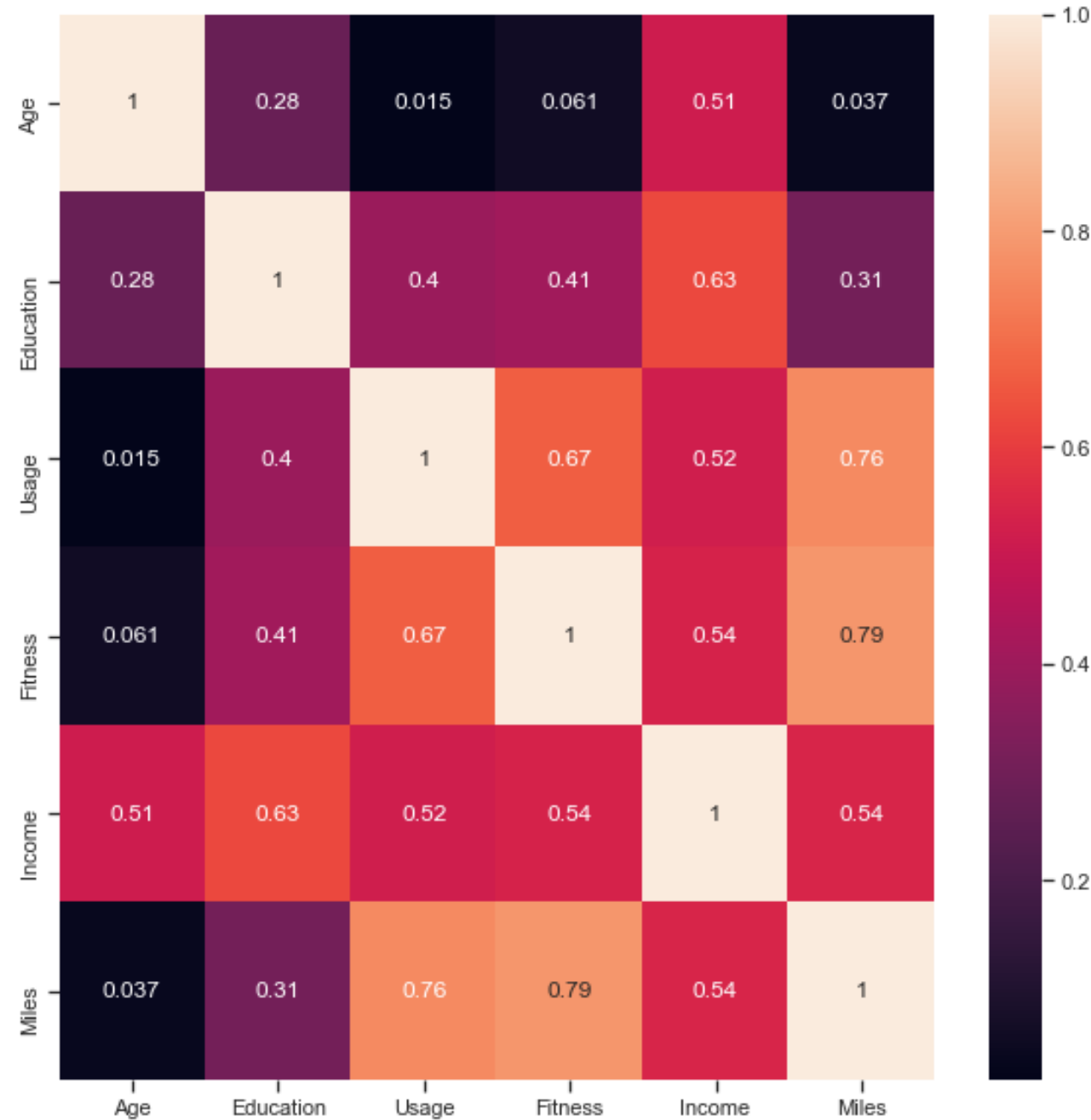
Multi-Variate Analysis

Observations

- Highest correlations are between Fitness, Usage & Miles - 0.67 to 0.79
- Moderate correlations for Income & Education - 0.63
- Moderate correlation for age & income – 0.51
- Mild correlations for Income, Miles, Fitness & Usage - 0.5
- Very little correlation for Age, Miles, Fitness & Usage - 0.015 to 0.037

Correlation Overview

April 2021



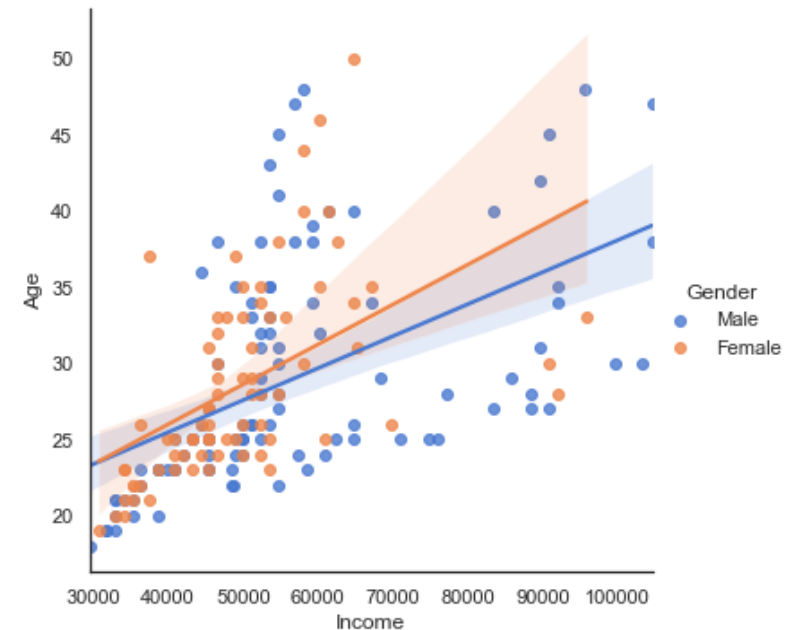
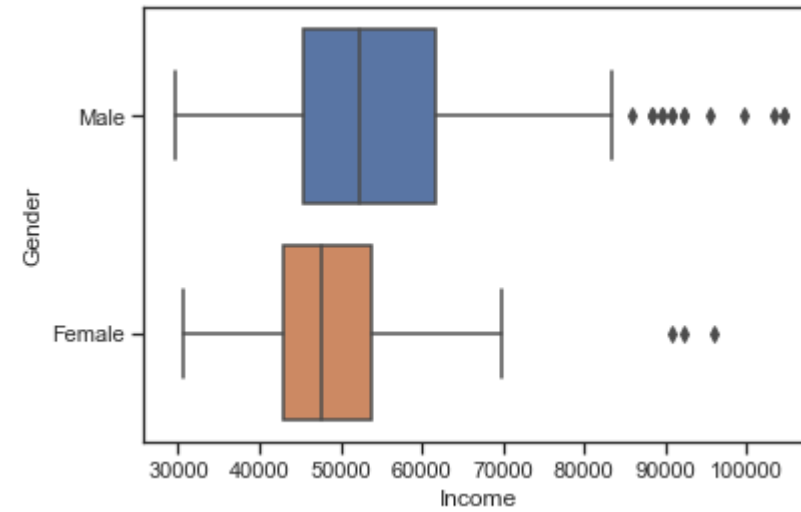
Multi-Variate Analysis

Observations

- Male customers have a higher average income
- Males have a greater number of higher income outliers
- Income and age are moderately correlated
- Males have a wider distribution of income

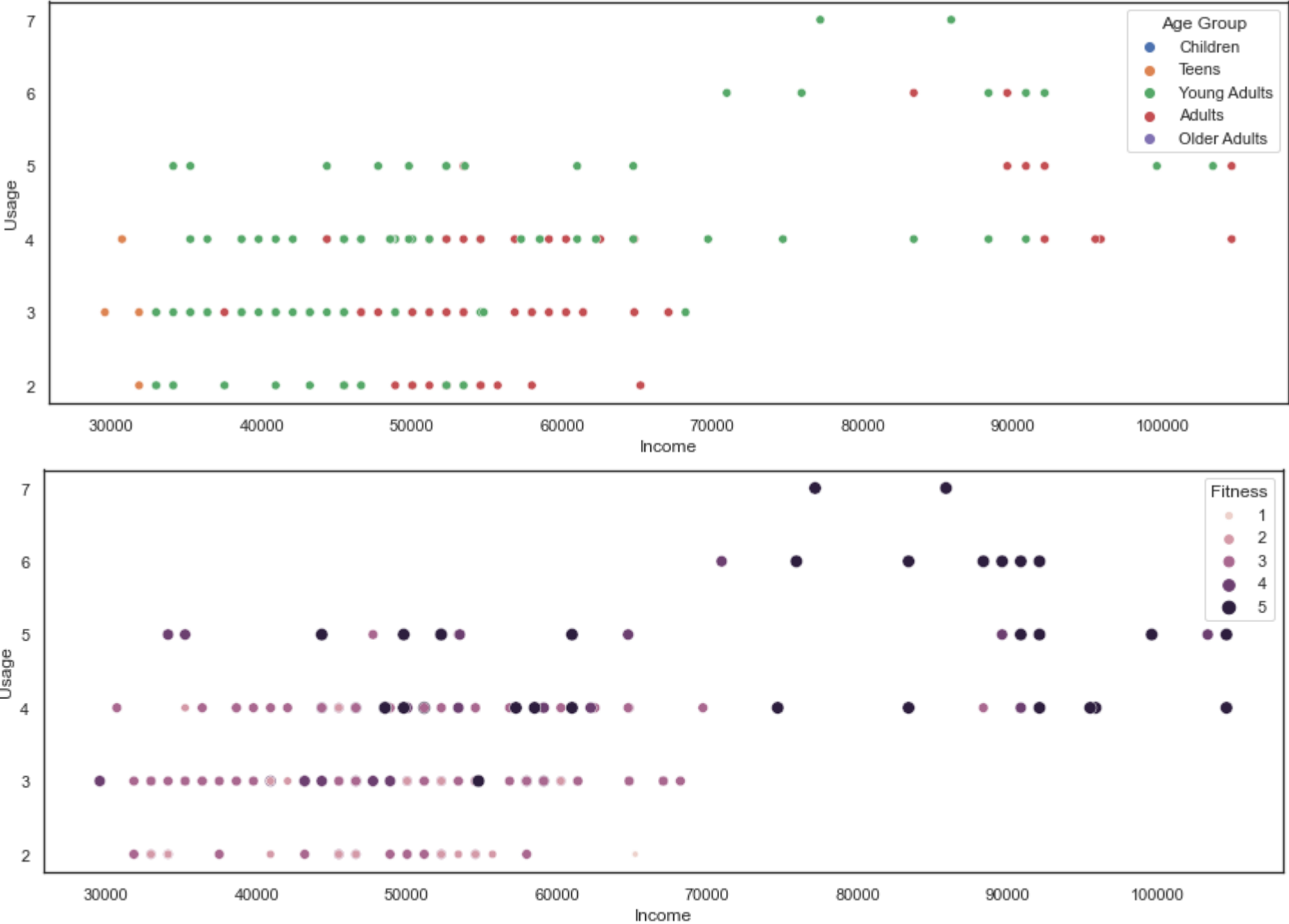
Income by Gender & Age

April 2021



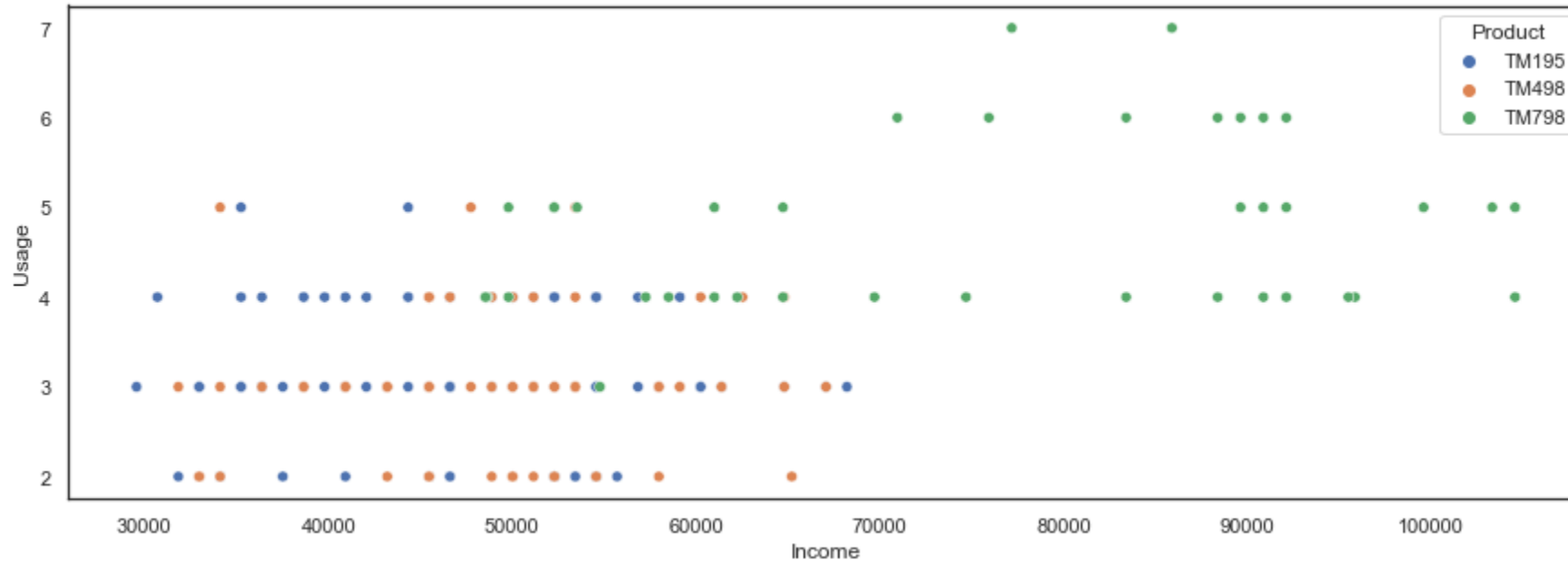
Multi-Variate Analysis

Income by Usage, Age Group, Fitness & Product



Multi-Variate Analysis

Income by Usage, Age Group, Fitness & Product

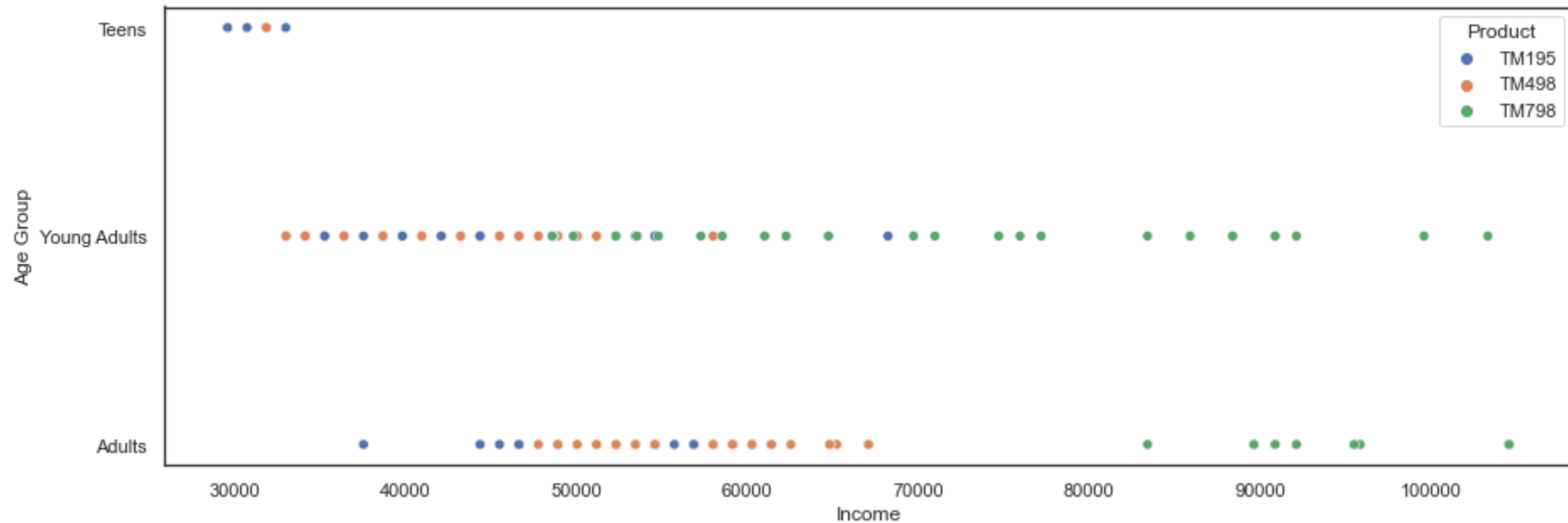


Observations

- As would be expected those with higher fitness expect to use the product 4 or more times a week
- The majority of TM798 product users expect to use the product 4 or more times a week
- Usage partially clustered for young adults & teens between 32k & 48k
- Usage partially clustered for Adults between 50k & 65k
- Most usage above 3 are young adults
- Nearly all lower to mid income customers report usage of 4 and below
- All higher income customers report usage of 4 and above
- The majority of TM798 product users have an income higher then 70k
- Only TM798 products are use by individuals with income higher then 70k
- The majority of individuals with income higher than 70k report a higher fitness (4 or 5)
- Usage gap between 70k and 85k

Multi-Variate Analysis

Income by Usage, Age Group, Fitness & Product

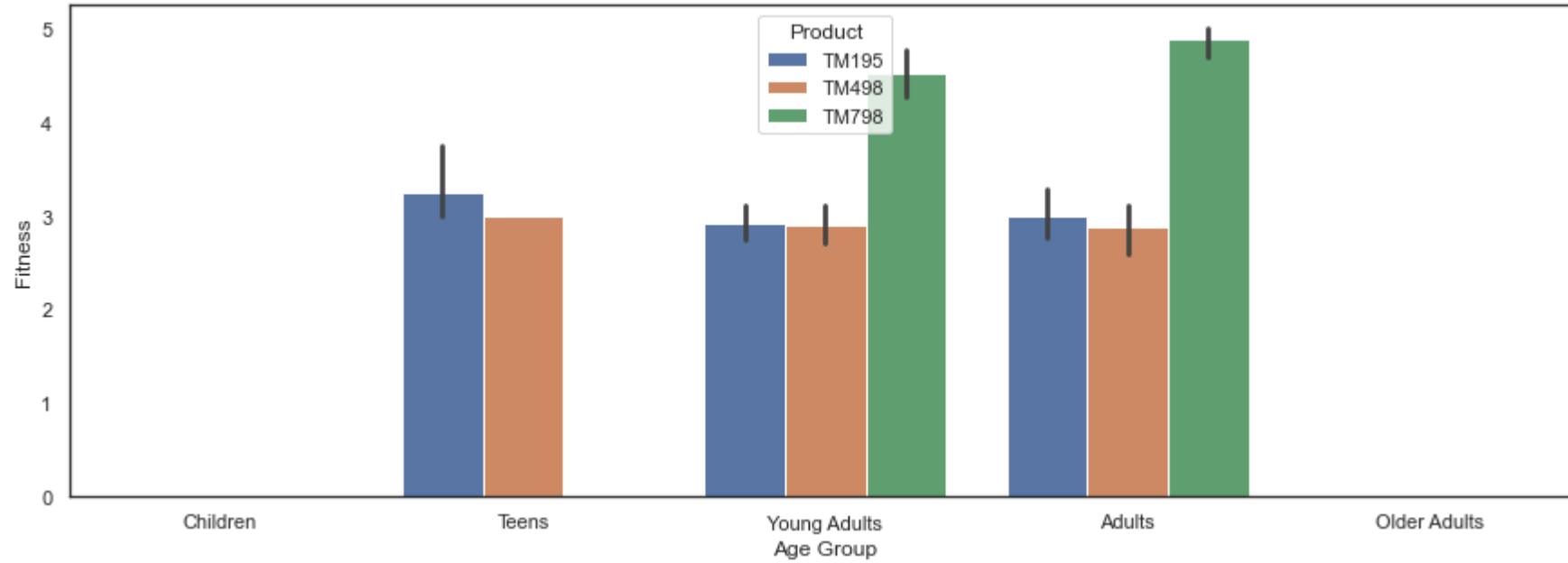


Observations

- All Teens have an income below 40k, most use the TM195, none use the TM798
- The majority of Adults have an income between 45k and 70k and primarily use the TM498
- The vast majority of Young Adults with an income above 60k use the TM798
- There is a gap for Adults with an income below 45k and between 70k & 90k

Multi-Variate Analysis

Age Group Fitness

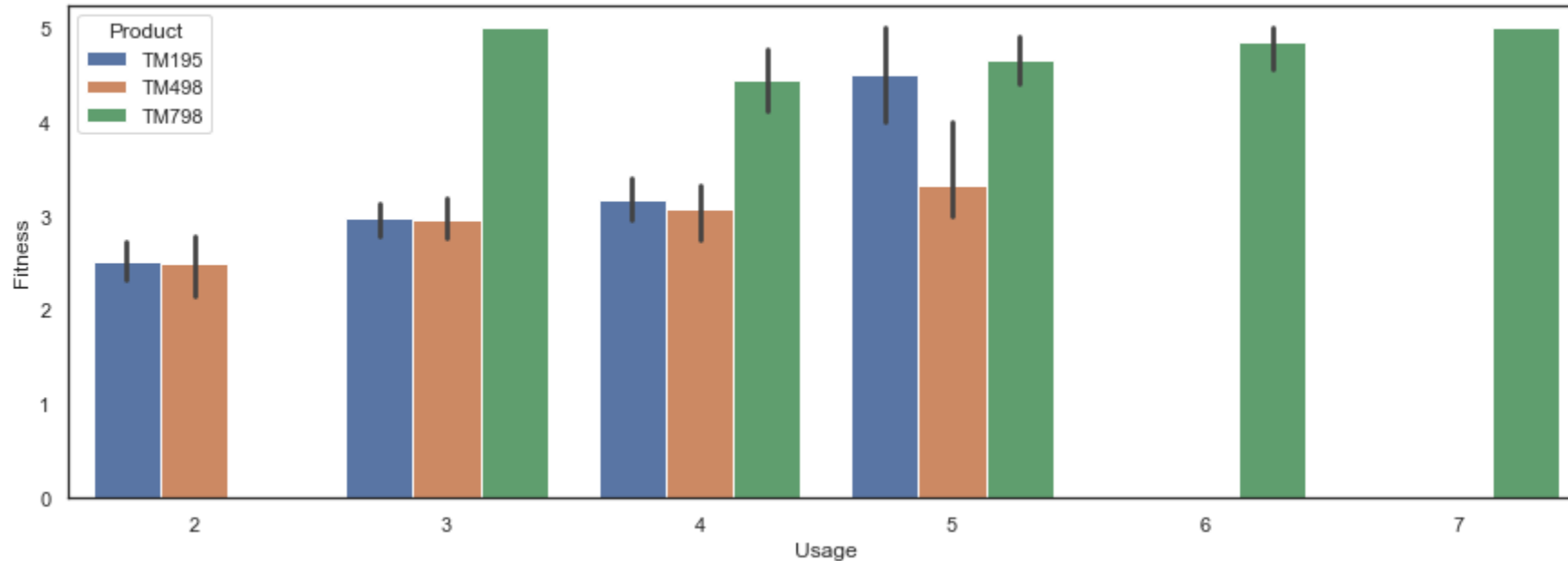


Observations

- Fitness among age groups and products are roughly the same
- Fitness for TM798 Customers is about 30% to 40% more than other products

Multi-Variate Analysis

Usage, Fitness & Products

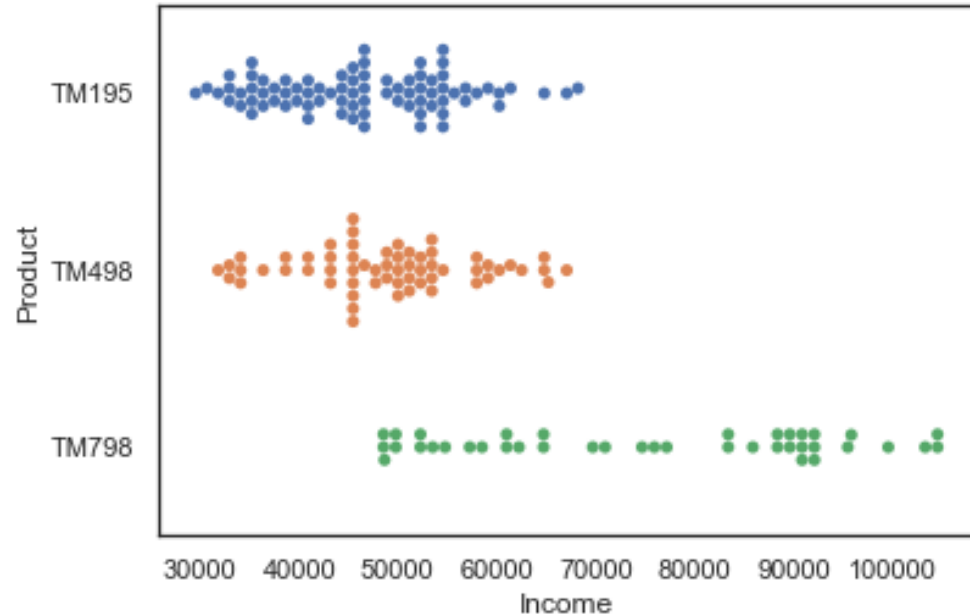
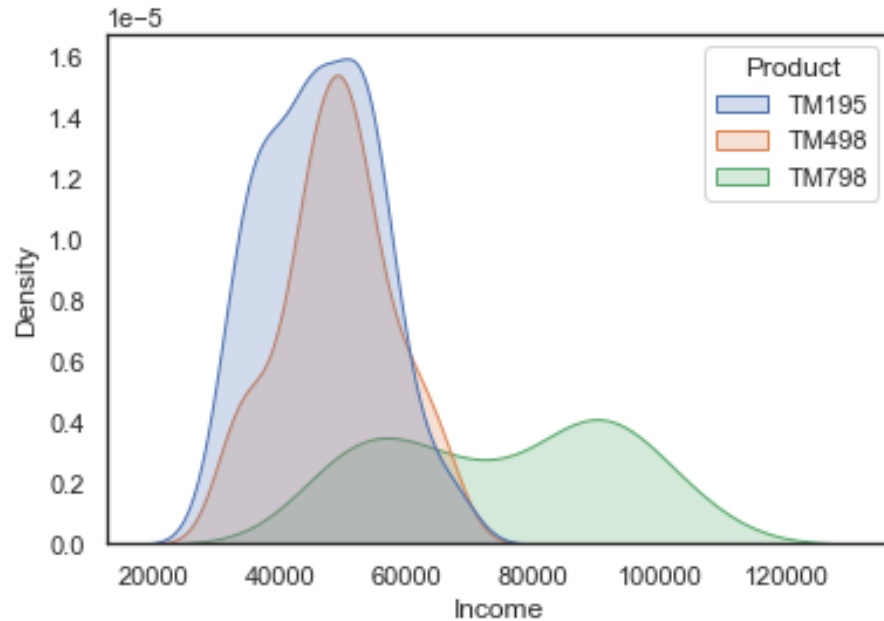


Observations

- There are no purchases for 795 below a usage of 3
- Fitness for 798 customers is roughly equal for usage above 2
- Those who purchase the 798 all report high fitness, mostly 5
(798 additional features may primarily be attractive to fitness enthusiasts)
- TM195 customers with a usage of 5 is about 25% to 50% higher than for other usages
- Fitness for 195 & 498 are roughly even except for usage of 5 where 195 is about 25% higher
- 498 seems to appeal to customers with primarily mid range fitness

Multi-Variate Analysis

Income and Product

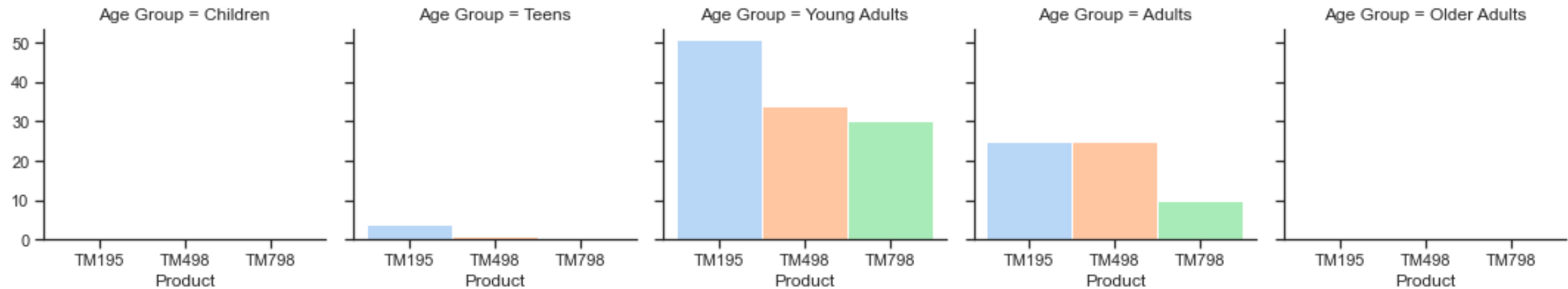


Observations

- TM798 - while this model has sales across most incomes and accounts for 22% of all sales
- it is the only model that has higher income customers, starting around 70k
- TM195 - has 44% of sales and is primarily purchased by customers in the lower to mid income ranges.
- Most sales around 30k to 60k earning customers
- TM498 - has 33% of sales with most of its purchases by customers with incomes in the 40k to 65k range
- TM195 and TM498 have a significant overlap for customers with customers with incomes in the 20k to 70k range
- TODO Need to determine if there are characteristic differences for TM 195 & 495.
- TODO Need to create a profile for TM798 customers
- TODO Determine if income is the only significant difference for TM798 purchases

Multi-Variate Analysis

Product by Age Group

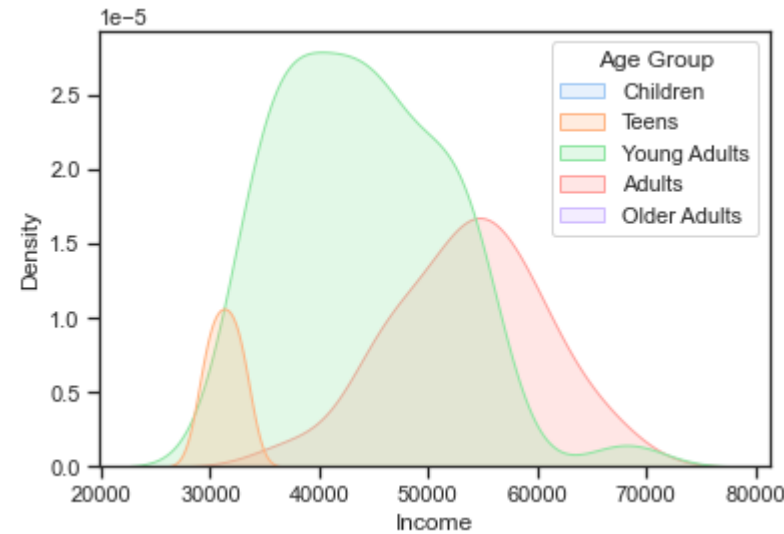


Observations

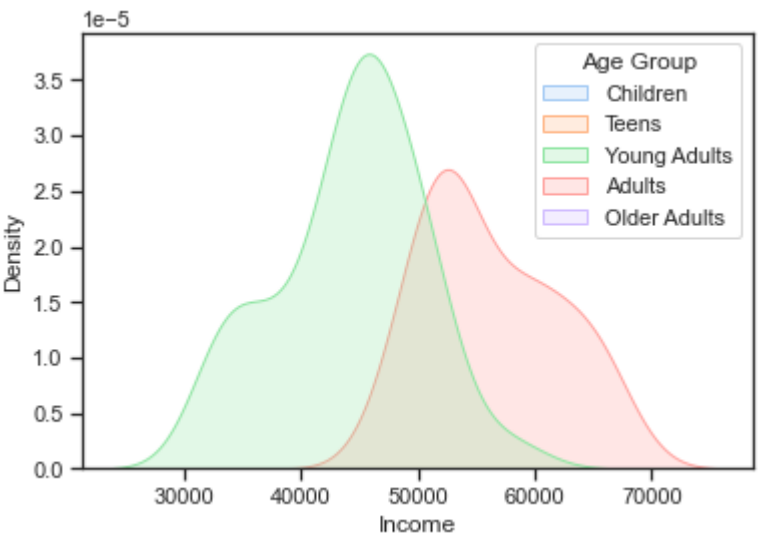
- Sales to Young Adults are higher across all products
- Sales of TM195 are about 50% higher for young adults than Adults
- Sales of TM498 are about 10% higher for young adults than Adults
- Sales of TM798 are about 75% higher for Young Adults than Adults

Multi-Variate Analysis

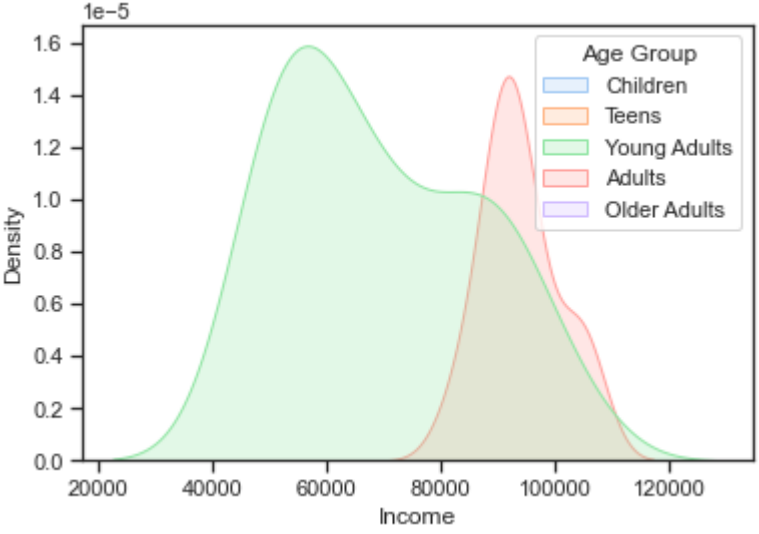
TM195 Income by Age Group



TM498 Income by Age Group



TM798 Income by Age Group

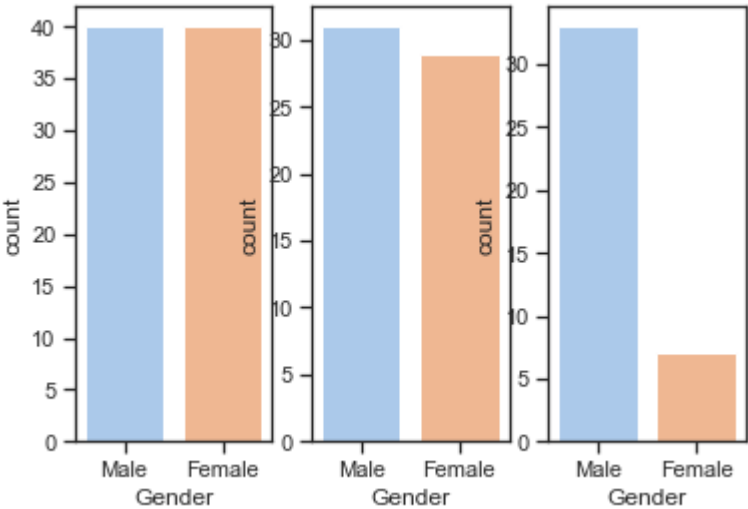


Observations

- RECOMMEND increase marketing to adults < 45k
- RECOMMEND - Increase marketing to adults with higher incomes > 50k for the 798

Multi-Variate Analysis

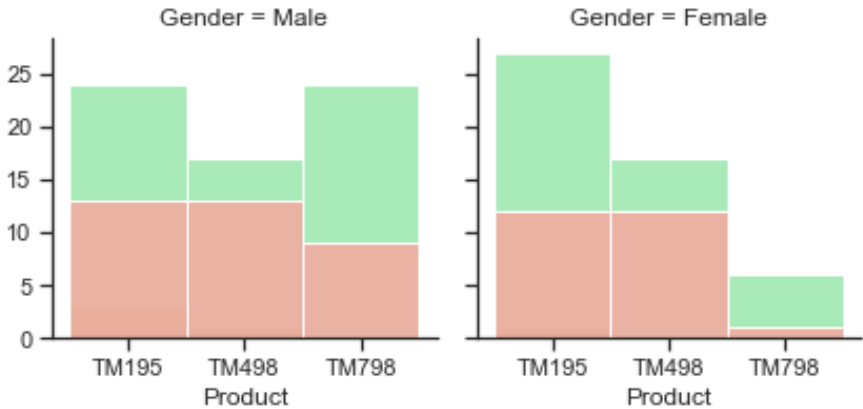
Gender by Product



Observations

- Genders for the 195 and 495 customers are primarily equal
- Gender for the 798 is 82% male

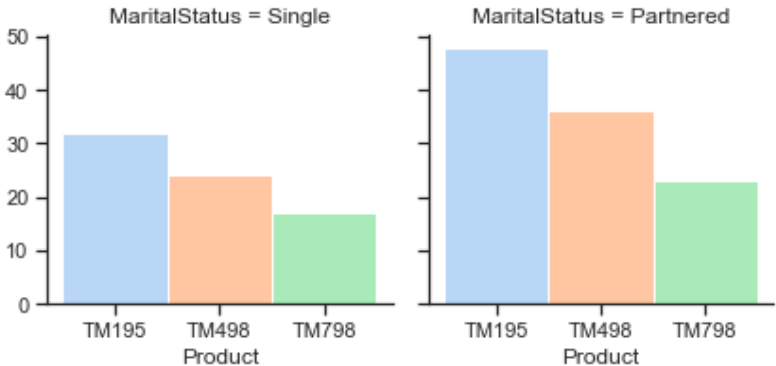
Gender by Product with Age Group



Observations

- Recommend - increase marketing of the TM798 to female fitness enthusiasts

Product by Marital Status



Observations

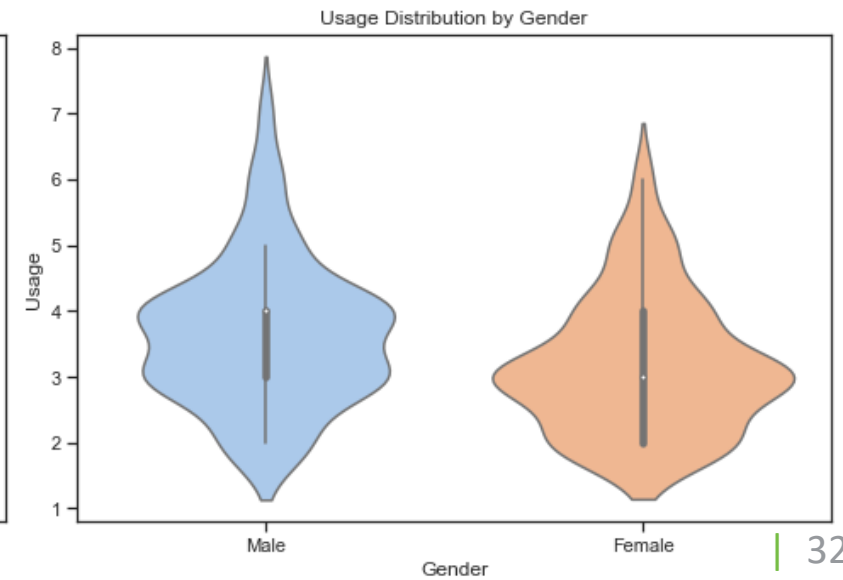
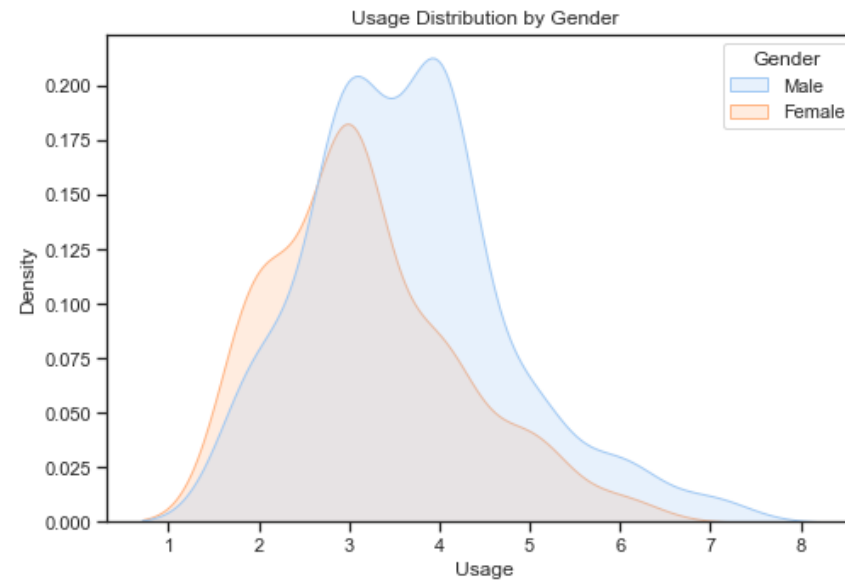
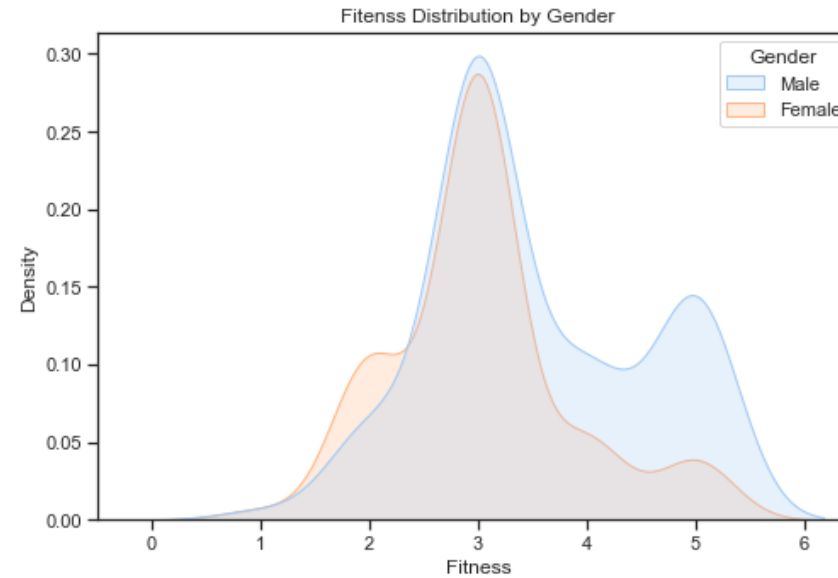
- More customers are partnered across all 3 products by about the same amount

Multi-Variate Analysis

Observations

- More males indicate a desire for higher usage
- More males report a higher fitness
- RECOMMEND: Increase marketing to female fitness enthusiasts

Fitness & Usage Distribution by Gender



Multi-Variate Analysis

Fitness & Usage Distribution by Product

Observations

- Higher reported fitness for the 798
- Higher desire for usage for the 798
- Fitness and usage are primarily the same with the TM195 having a slightly higher fitness range

