

Cardio Good Fitness

Eugenie Seholm

Contents

- Business Problem Overview and Solution Approach
- Data Overview
- Exploratory Data Analysis (EDA)
- Business Insights and Recommendations

Business Problem Overview and Solution Approach

- Core business idea
 - Gross sales vs. profit margins vs. user advocacy
- Problem to tackle
 - Identify key characteristics of customers based on potential strategic initiatives
- Financial implications
 - Simple gross sales model: simplistic marketing
 - Net margin model: targeted high-profit
 - Product advocacy model: long-term growth

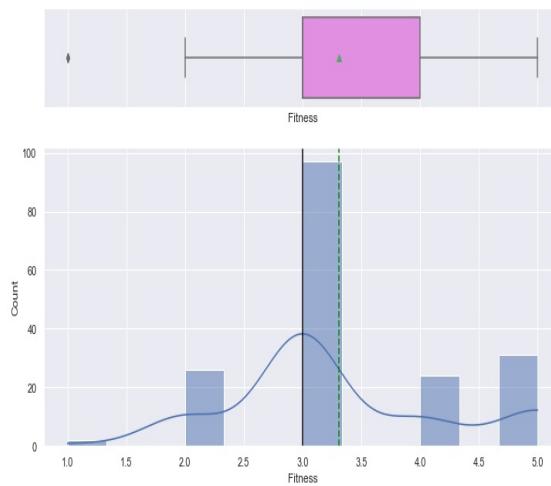
Data Overview

- 180 observations
- 6 numerical variables
- 3 categorical variables
- No missing values
- Age variable has been cut into ordinal data, i.e., teens, 20s, 30s, etc. for extra analysis

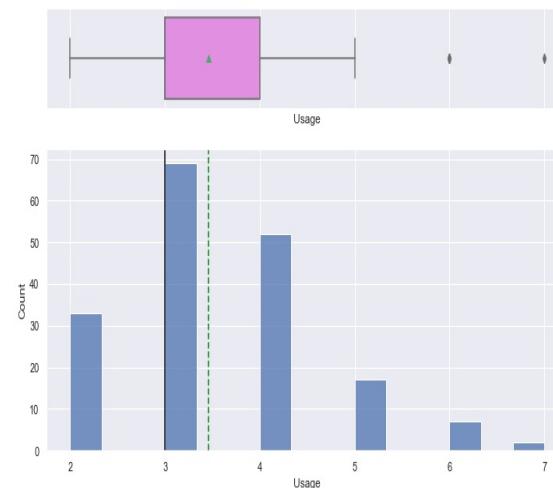
Variable	Description
Product	The model no. of the treadmill
Age	Age of the customer in no. of years
Gender	Gender of the customer
Education	Education of the customer in no. of years
MaritalStatus	Marital status of the customer
Usage	Average # times customer wants to use the treadmill weekly
Fitness	Self-rated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

EDA: General Customer

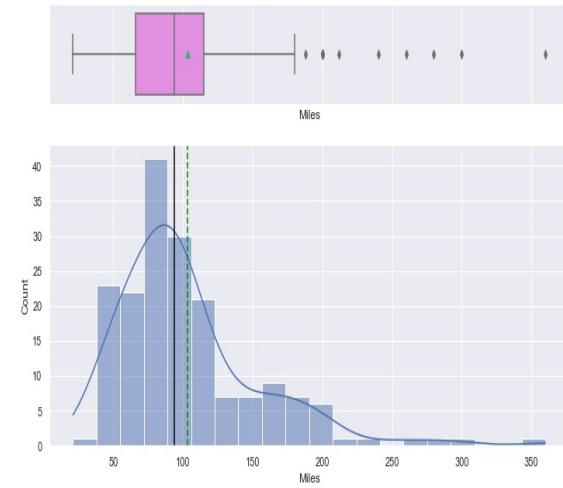
Fitness Score



Weekly Usage



Miles Per Week



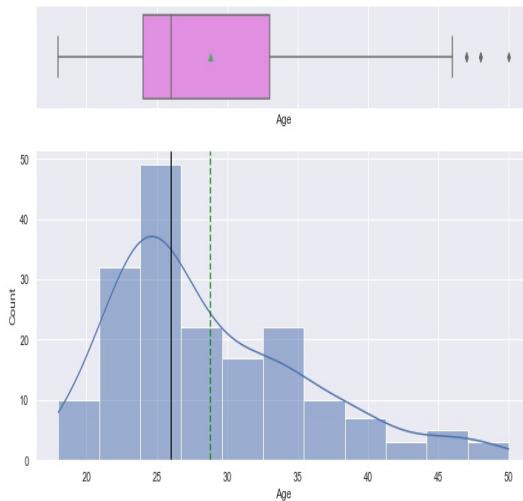
- Fitness score data is positively skewed, with outliers in the negative direction
- $Q1 = Q2$
- Mean is slightly higher than median
- More than half customers rate fitness as 3

- Usage data is positively skewed, with outliers in the positive direction
- $Q1 = Q2$
- Majority of customers planned to use treadmill 2 to 4 times per week
- Everyone intended to use treadmill at least twice weekly

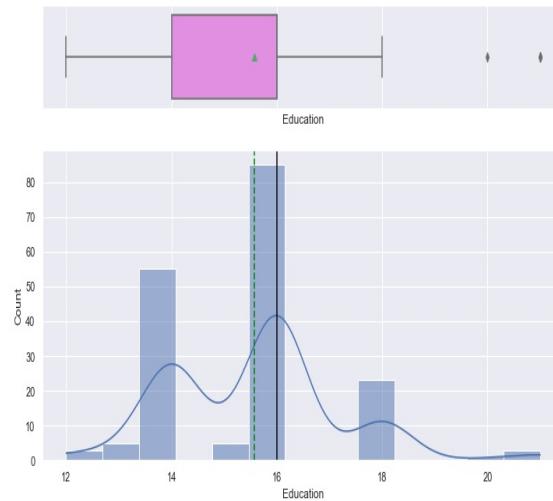
- Miles per week data is positively skewed
- Whisker ends at approximately 175, but there are outliers greater than 350

EDA: General Customer

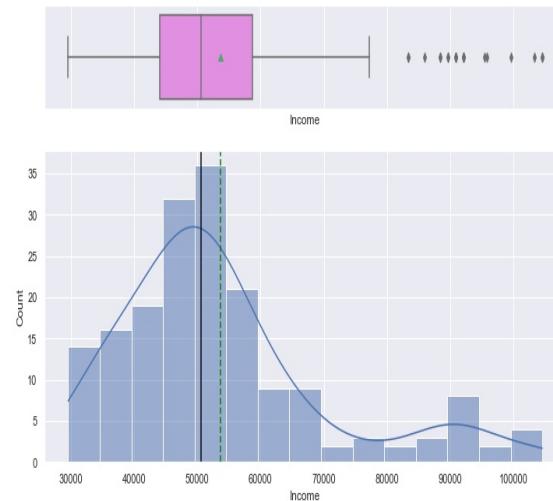
Age



Education



Income



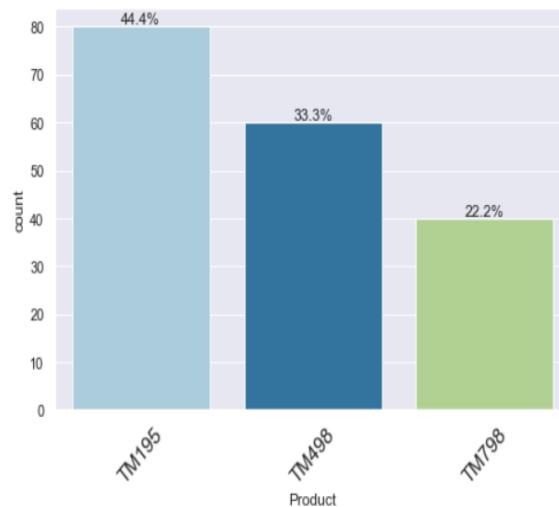
- Age is positively skewed with outliers in the positive direction
- Median age 26
- Majority of customers (>75%) are between 18 to 35, but at least one customer as old as 50

- Education is negatively skewed, with outliers in the positive direction
- All customers are at least high school graduates
- $Q2 = Q3$
- The majority of customers (at least 75%) college grads, but a few have 9 years education after HS

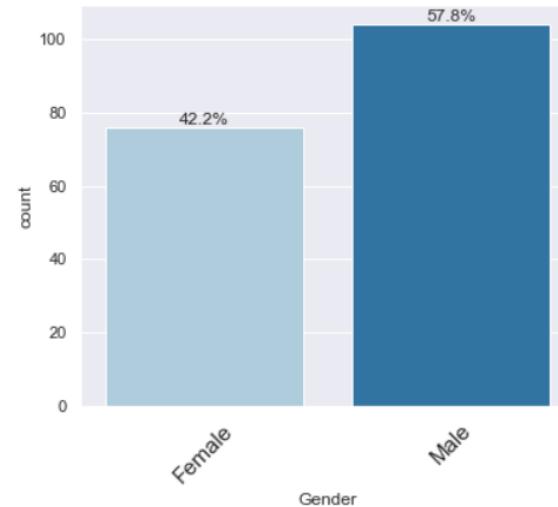
- Income is positively skewed with many outliers in the positive direction
- Large range:
- $Q3$ approximately 59K but max is approximately 105K

EDA: General Customer

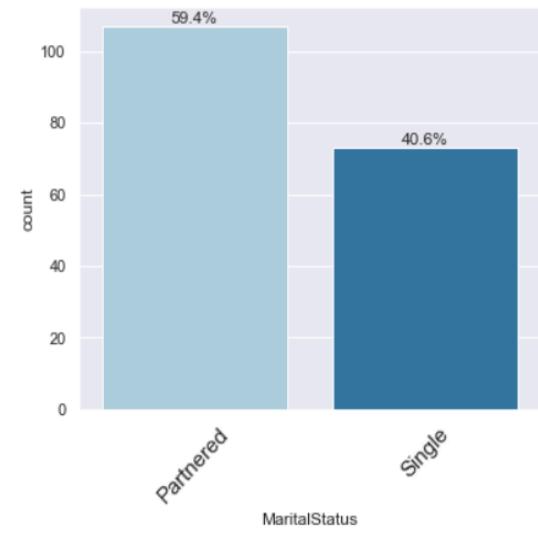
Product



Gender



Marital Status



- TM195 is the most popular product
- TM798 is the least popular product

- More men purchase these products than women
- It would be interesting to compare it to the overall clientele of the store to see if there is a difference

- More of the customers who have purchased these products are partnered than single
- It would be interesting to compare it to the overall clientele of the store

EDA: Correlation Matrix

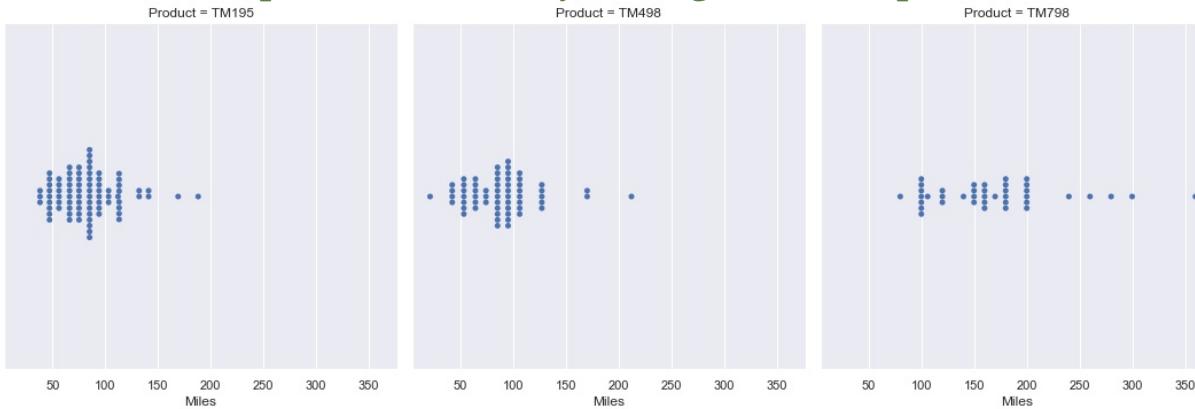


Observations:

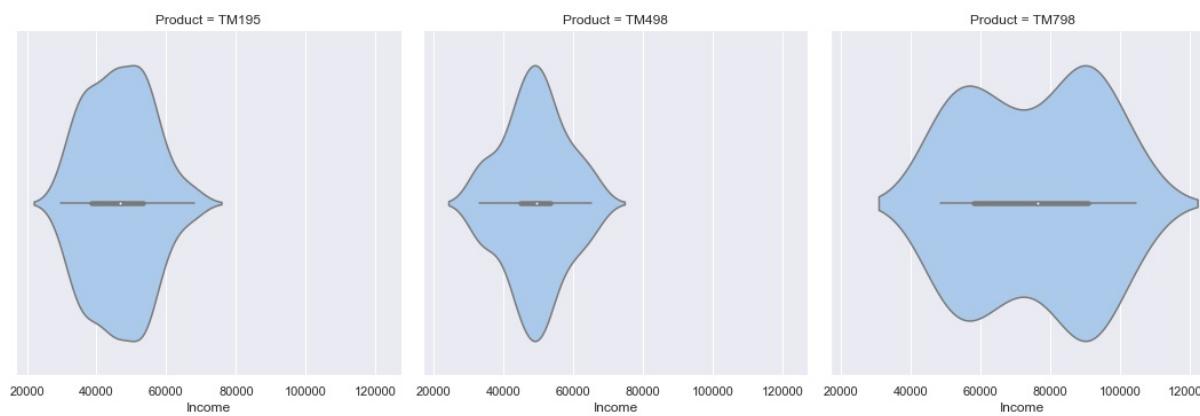
- Miles has a relatively high correlation with Usage and Fitness
- Use and Fitness are not as highly correlated as with miles, but still relatively strong.
- Income has relatively strong correlation with education
- Age has weakest correlations, followed by education. Both have moderate correlation with income
- No negative correlations

EDA: Product and Numerical Variables

Dispersion of weekly mileage for each product



Dispersion of income for each product

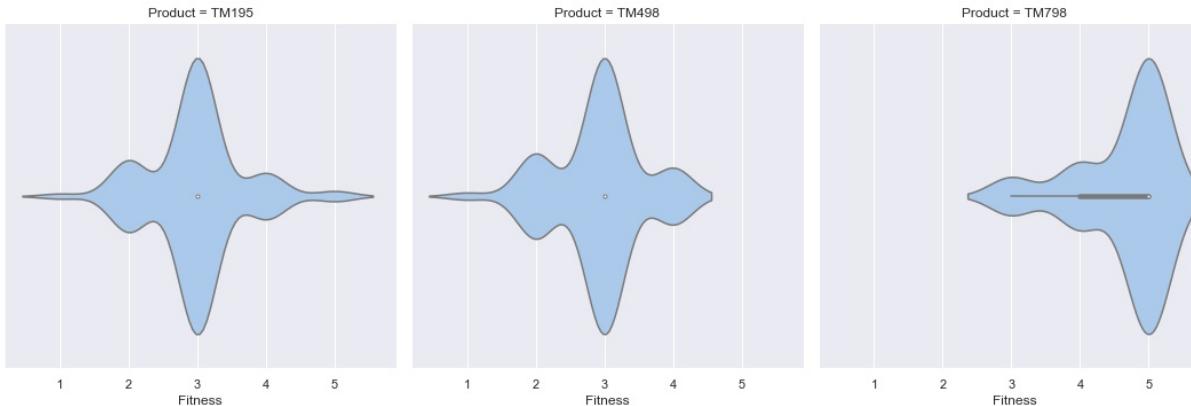


- All distributions are positively skewed.
- TM195 & TM498 have similar looking distributions
- TM798 seems to appeal to those who intend to use it for much more mileage.

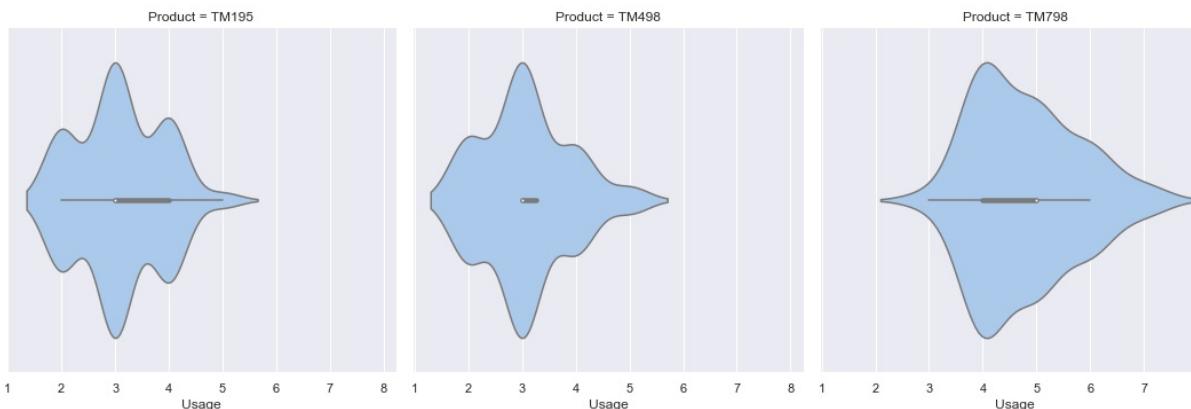
- TM195 and TM495 have somewhat normal distributions, with similar range.
- TM798 appears to be bimodal with much larger range than the other two

EDA: Product and Numerical Variables

Dispersion of fitness score for each product



Dispersion of usage for each product

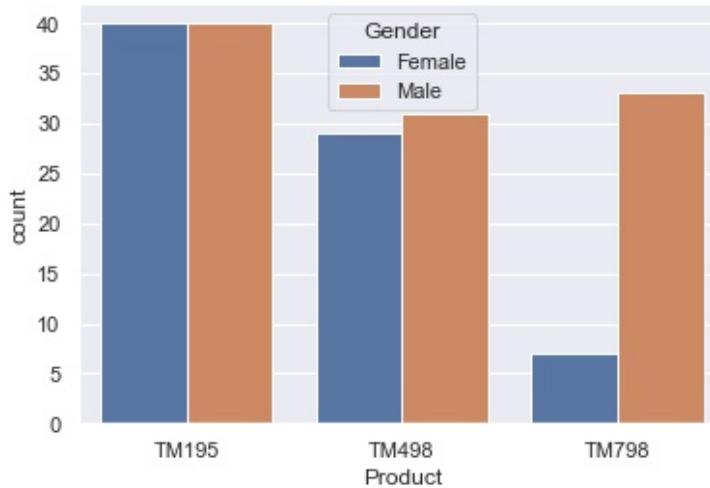


- TM195 and TM498 have similar shapes, aside from the right tail present in TM195 that is absent in TM498
- TM798 is negatively skewed and the majority of values are high

- Usage is similar for TM195 and TM498
- Usage for TM798 is generally higher overall and has a larger range

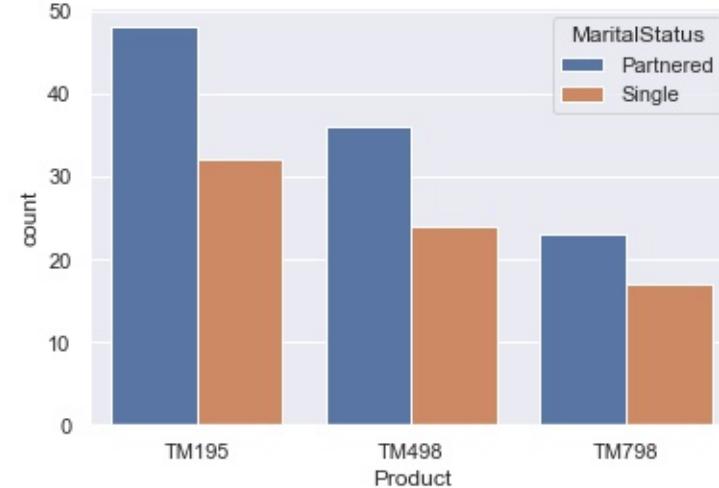
EDA: Product and Numerical Variables

Count of customers that purchased each product by gender



- TM195 most popular for both genders
- TM798 has significantly fewer females than males

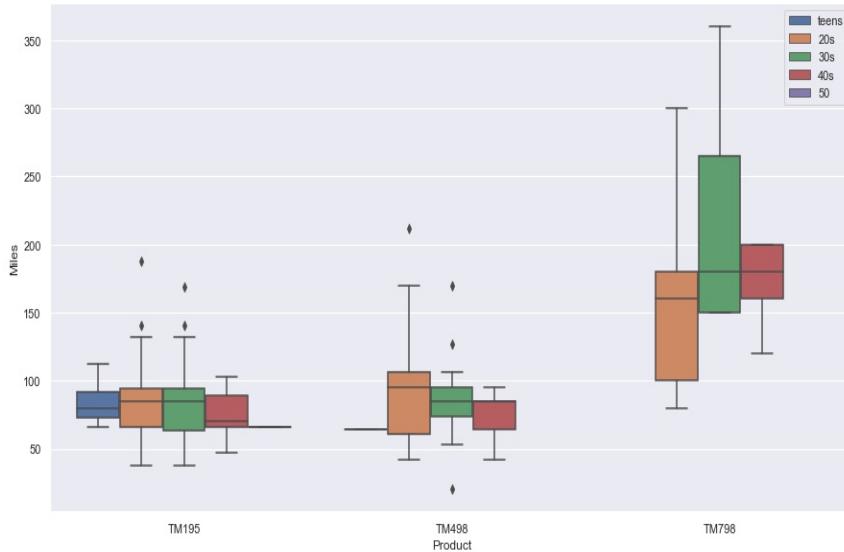
Count of customers that purchased each product by marital status



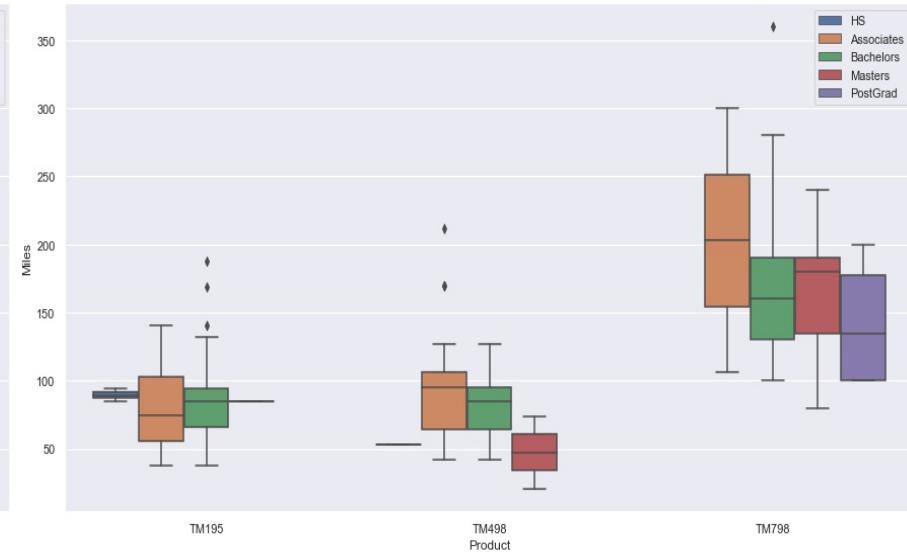
- Relative difference between partnered and single is consistent from product to product

EDA: Product and Numerical Variables

Miles per product, broken down by age groups



Miles per product, broken down by education groups

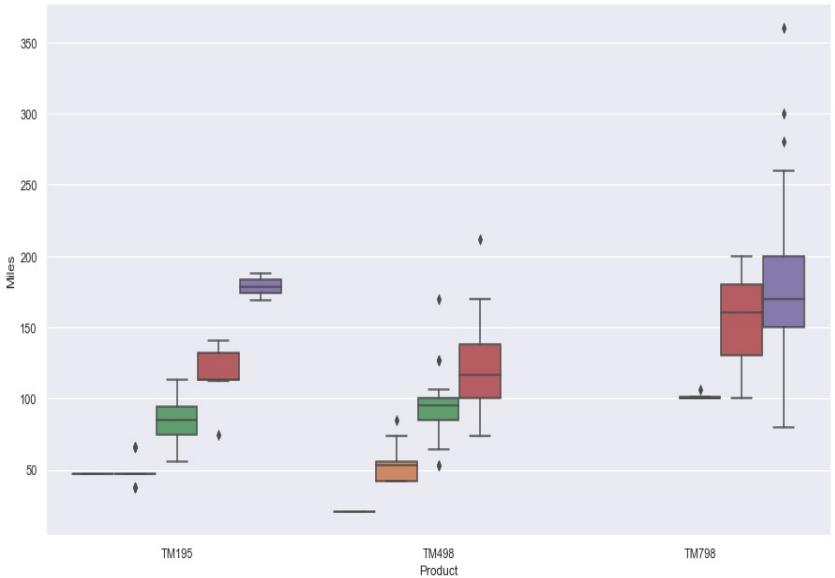


- TM498 and TM798 do not appeal to customers under 20 years old
- Customers who run more miles tend to purchase TM798, particularly customers in 30s

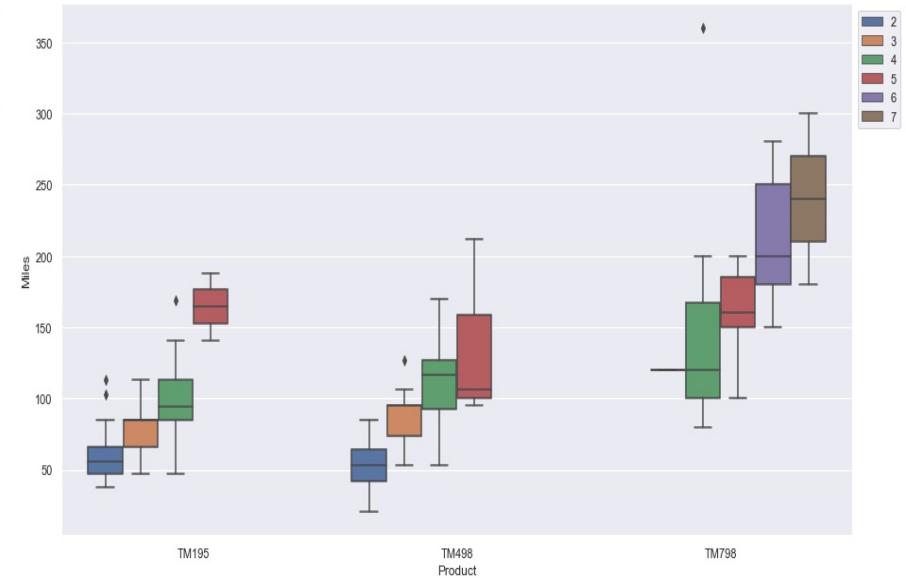
- Customers with education beyond bachelors do not tend to purchase TM195
- Highest educated customers tend to purchase TM798

EDA: Product and Numerical Variables

Miles per product, broken down by fitness score



Miles per product, broken down by usage



- As model number increases, spread increases for the miles expected to run for most fitness scores

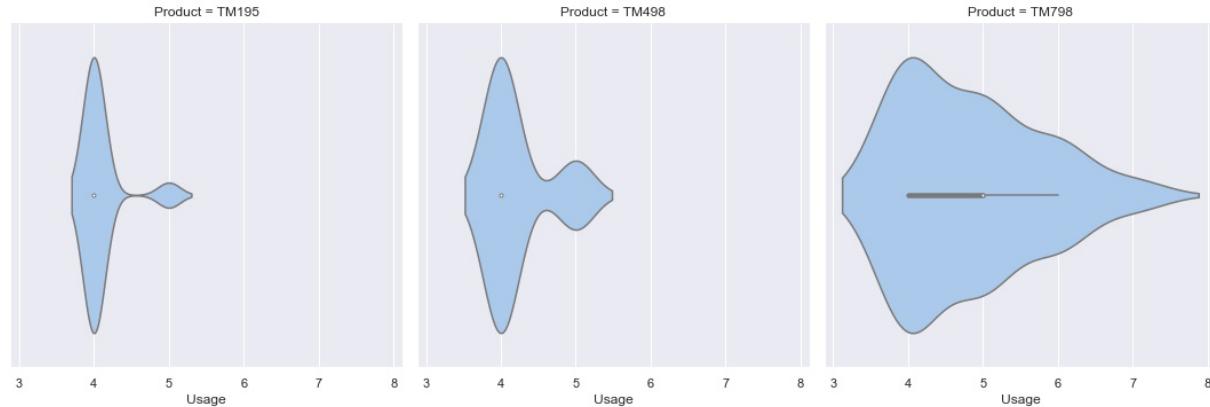
- Customers who plan to use the treadmill fewer times weekly purchase TM195 and TM498
- Daily users buy TM798

EDA: High Usage Customers

Usage of 4 or more times weekly

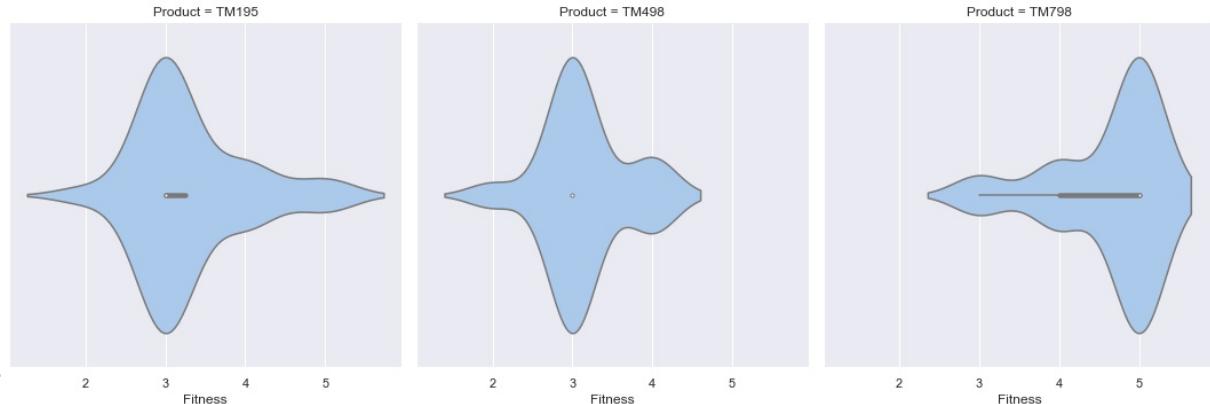
Dispersion of usage for high usage customers

- Higher model numbers have increasingly more high usage customers
- Notably, TM798 has far more high usage customers than both the other two

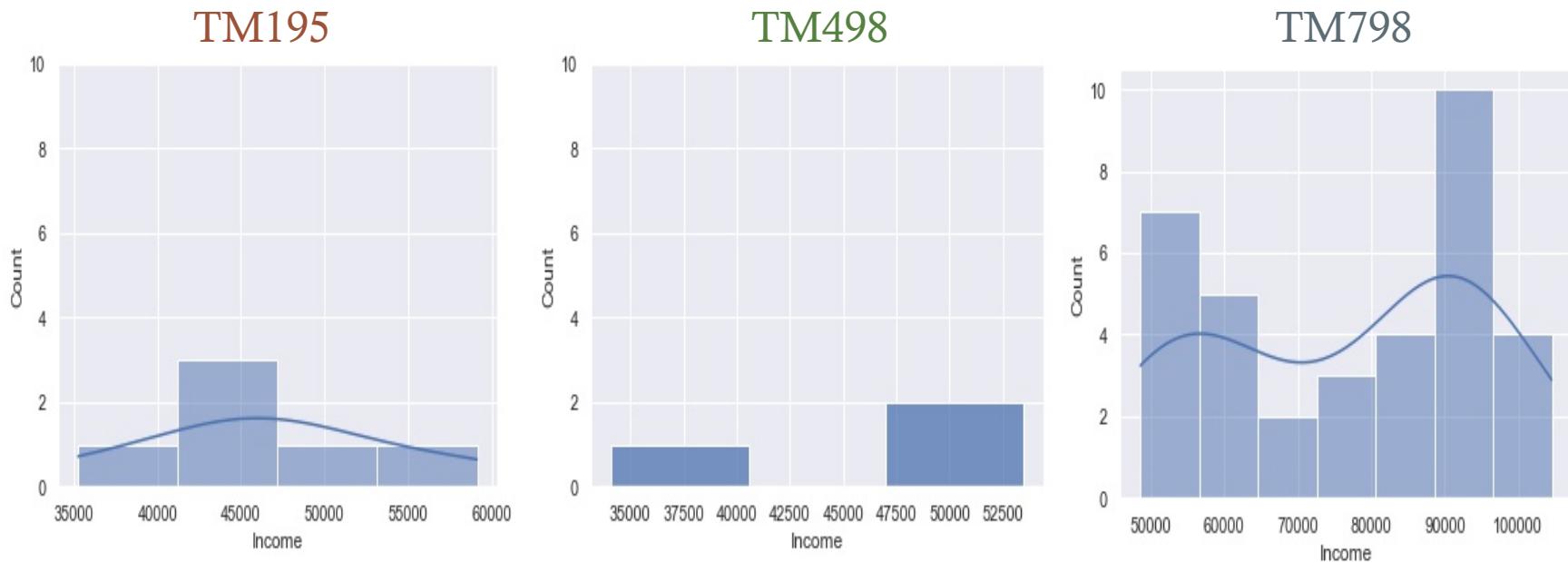


Dispersion of fitness for high usage customers

- TM195 positively skewed
- TM498 similar to TM195 but truncated in the positive direction
- TM798 negative skewed
- TM798 has notably higher fitness scores



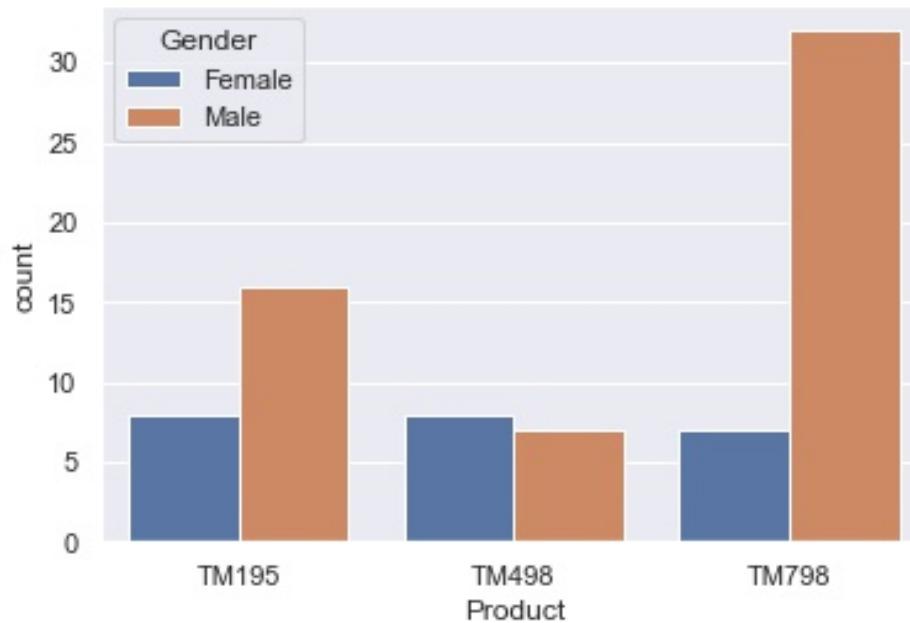
EDA: High Usage Customers



- Few high usage customers choose models besides TM798
- High usage customers who have lower income choose TM195 or TM498

EDA: High Usage Customers

Count plot of high usage customers across models by gender



- High usage females are not as affected by model number
- High usage males tend to purchase TM798 most, TM195 about half frequency of TM798
- TM495 is purchased least of all for high usage males

Business Insights and Recommendations

- Simple gross sales model:
 - Simplistic marketing
 - Target general customer base for rapid gross sales
 - 24-33 year old
 - Middle class
 - Male
 - Casual runner
 - College grad
 - Average fitness

Business Insights and Recommendations

- Net margin model
 - Targeted high-profit by targeting customer base for model with highest margin
 - TM195 & TM498
 - Lower-middle to middle class
 - College-educated
 - Male or female
 - Average fitness
 - TM798
 - Middle to upper-middle class
 - Highly-educated
 - Male
 - Daily runner
 - High fitness score

Business Insights and Recommendations

- Product advocacy model
 - Long-term growth
 - Target customers who would use the product the most and advocate for the product
 - Middle to upper-middle class
 - Highly-educated
 - Male
 - Daily runner
 - High fitness score

Further Recommendations

- More data—potentially over a longer period of time if possible—should be collected for better analysis
 - Data may differ significantly around the new year
 - Data may differ during years that broadcast Olympic events
- Compare data for these treadmills with that of the rest of the clientele for this store to find insights