

# **Cardio Good Fitness — Customer Profiling Project**

STAT 560

**Prepared by: S M Nadim Mahmud**

## **Introduction**

This report presents descriptive analytics and statistical testing to profile typical customers for each treadmill product offered by Cardio Good Fitness (TM195, TM498, TM798). The research intends to compare customer demographics and usage patterns to better understand differences in products; descriptive statistics, visualization, and hypothesis testing will be used. The objective is to deliver useful marketing insights that could contribute to product positioning and identify effective advertising methods.

## **Data Description & Preprocessing**

The dataset contains 180 observations and 9 variables. Variables include Product, Age, Gender, Education, Marital Status, Usage, Fitness, Income, and Miles.

Missing values (if any) were handled via pairwise deletion for tests and summaries.

## **Exploratory Data Analysis (EDA)**

Product counts:

Product	Count
TM195	80
TM498	60
TM798	40

Based on Figure 1, the number of customers by product is shown below.

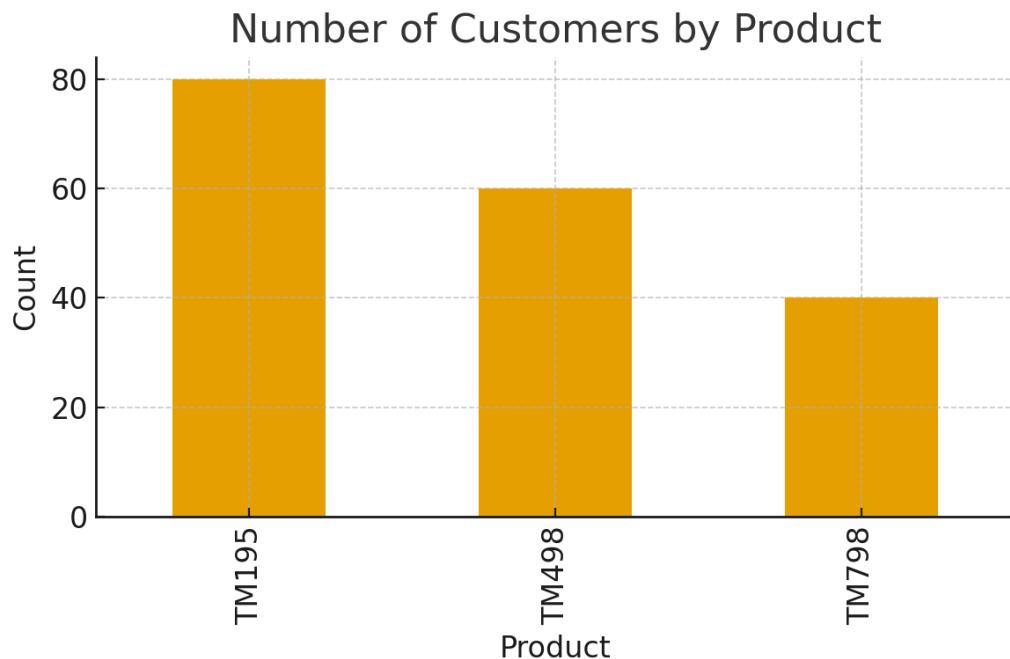


Figure 1: Number of Customers by Product.

#### **Explanation:**

Based on Figure 1, the TM195 model is the most purchased treadmill with 80 customers, followed by TM498 with 60, and TM798 with 40. This suggests that TM195 is the company's most popular or affordable option, while TM798 caters to a smaller, possibly more premium segment.

Income summary by product (mean, sd):

Product	Count	Mean Income	Std Dev Income
TM195	80	\$46418.02	\$9075.78
TM498	60	\$48973.65	\$8653.99
TM798	40	\$75441.58	\$18505.84

Based on Figure 2, the income distribution across models shows differences that warrant statistical testing.

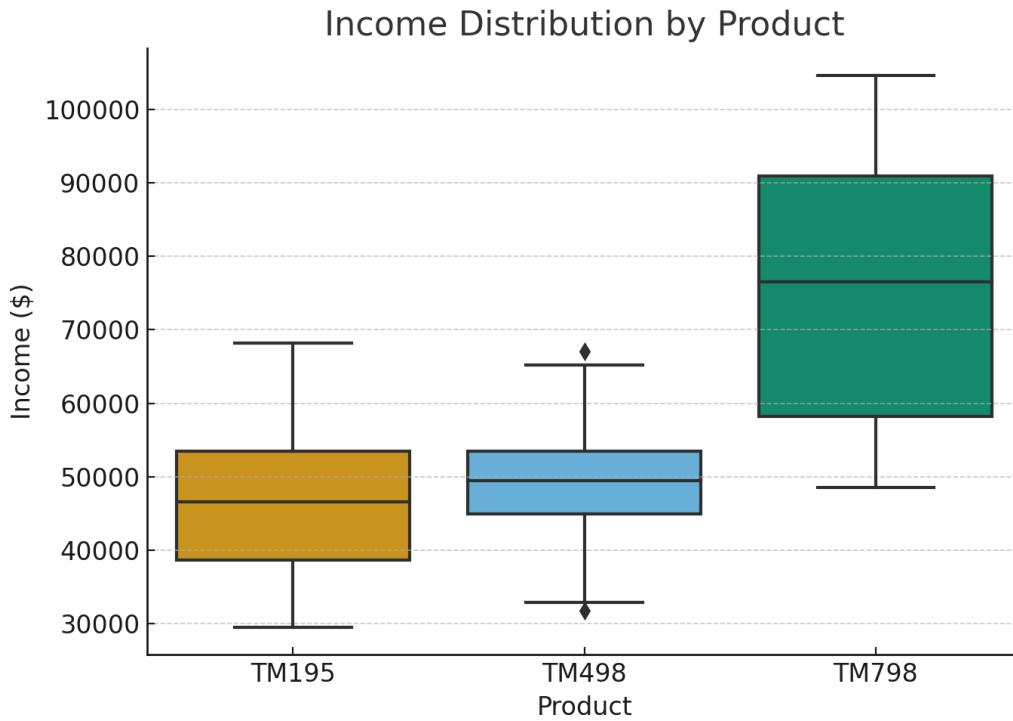


Figure 2: Income Distribution by Product.

**Explanation:**

Figure 2 shows a clear difference in income levels across products. TM798 buyers report substantially higher incomes, whereas TM195 buyers tend to have lower income levels. This supports the hypothesis that the company's products appeal to distinct income segments and motivate formal statistical testing.

Based on Figure 3, the age distributions across products are shown below.

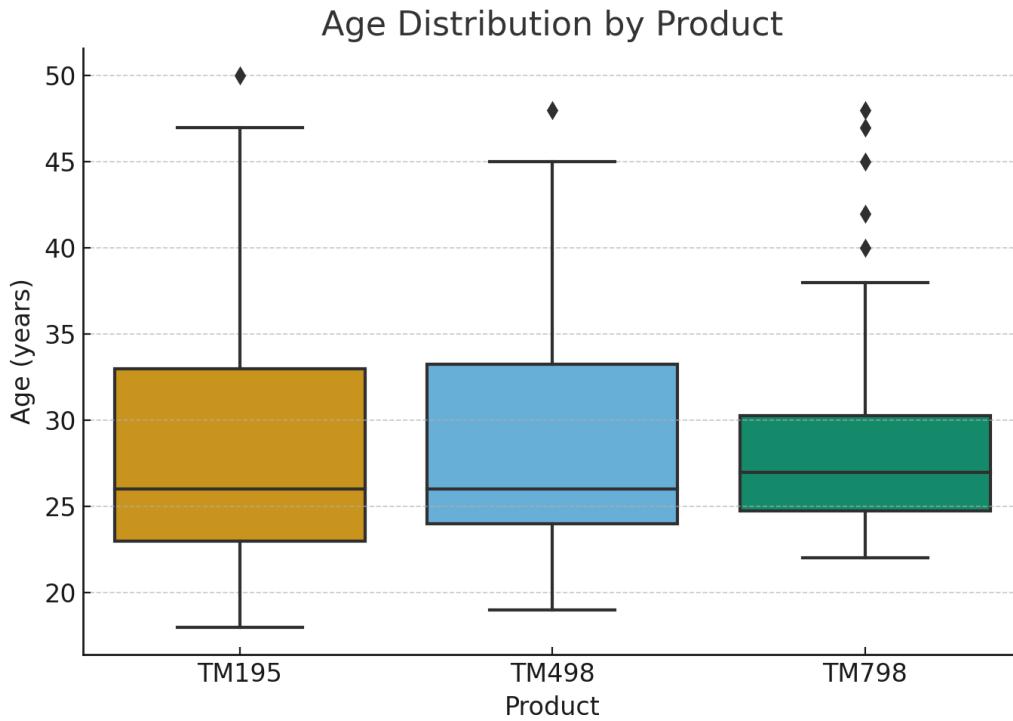


Figure 3: Age Distribution by Product.

**Explanation:**

As shown in Figure 3, customers across all treadmill models are mostly in their late 20s to early 40s. Similar age distributions suggest that age is not a key factor distinguishing between product types, which aligns with the ANOVA results indicating no significant difference in mean ages.

Figure 4 shows the gender composition for each product.

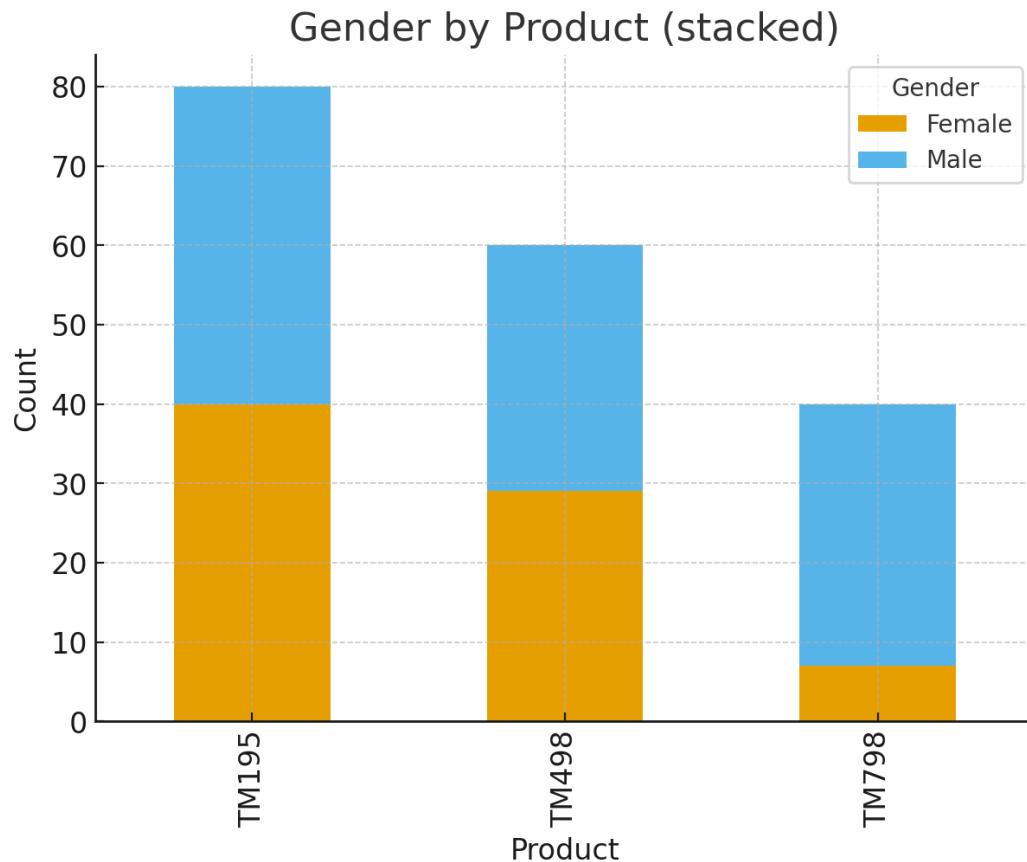


Figure 4: Gender by Product (stacked counts).

**Explanation:**

Figure 4 displays the gender composition for each treadmill. TM195 and TM498 have a nearly balanced split between male and female buyers, while TM798 shows a higher proportion of male customers. This gender skew may indicate that the premium product line appeals more to male buyers or aligns with higher male-reported incomes.

Figure 5 shows expected miles per week for product.

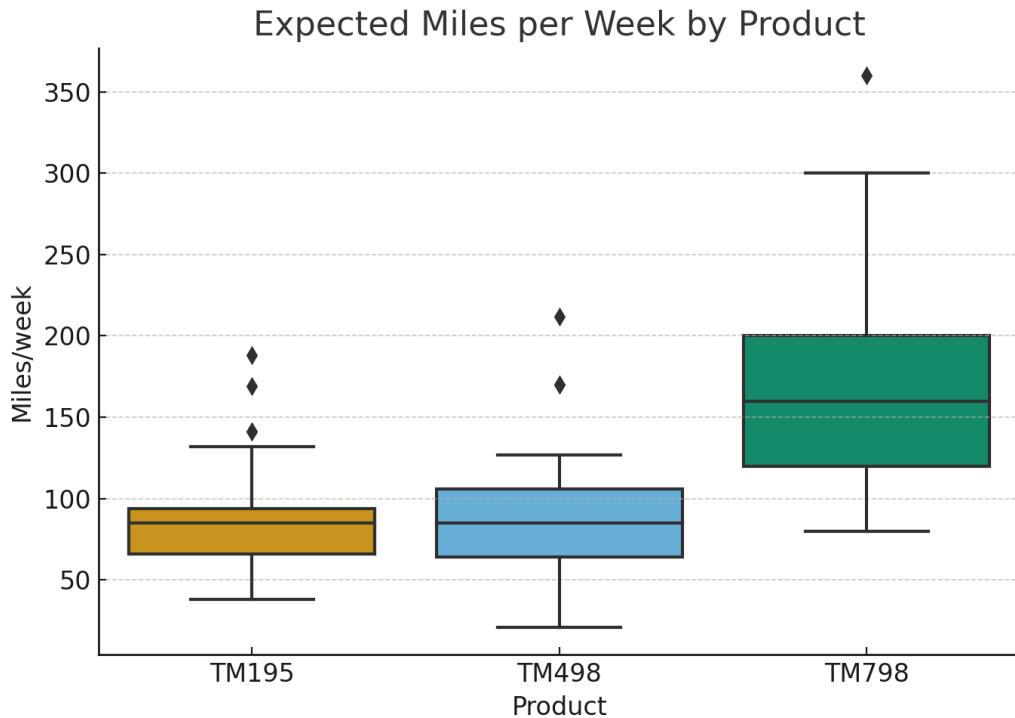


Figure 5: Expected Miles per Week by Product.

#### **Explanation:**

Figure 5 indicates that TM798 owners generally report higher weekly mileage compared to TM195 and TM498 users. This suggests that customers who purchase the higher-end model tend to be more consistent or dedicated exercisers, reinforcing the connection between fitness commitment and premium product choice.

#### **Overall Summary:**

Taken together, Figures 1–5 suggest that treadmill models vary primarily by income, fitness level, and weekly mileage, while age remains relatively consistent across products. These findings can guide targeted marketing strategies: TM195 appeals to entry-level buyers, TM498 to moderate users, and TM798 to high-income, performance-oriented customers.

#### **Statistical Tests and Inference**

##### **ANOVA: Age by Product**

ANOVA F-statistic = 0.094, p-value = 0.9101.

Since  $p > 0.05$ , we **fail to reject the null hypothesis**.

There is no significant difference in the mean age of customers across treadmill models. This means age is not a major factor influencing which treadmill model a customer buys.

### **ANOVA: Income by Product**

ANOVA F-statistic = 89.259, p-value = 0.0000.

The p-value is less than 0.05, indicating a statistically significant difference in mean income among treadmill models.

Customers purchasing the TM798 have the highest average income, while TM195 buyers have the lowest income.

Therefore, income is a key differentiator among treadmill buyers.

### **Chi-square test: Gender vs Product**

Chi-square = 12.924, p-value = 0.0016, dof = 2.

Since  $p < 0.05$ , there is a significant association between gender and treadmill model purchased.

This means male and female customers show different preferences for treadmill models — males slightly prefer TM798, while females purchase TM195 more often.

### **ANOVA: Fitness by Product**

F-statistic = 104.001, p-value = 0.0000.

The p-value is far below 0.05, so we reject the null hypothesis.

There is a significant difference in mean self-rated fitness among treadmill model buyers.

Customers who purchased TM798 generally rate themselves as more fit, while TM195 buyers report lower fitness levels.

## **Insights and Recommendations**

### **◊ Recommendations**

Based on the insights from the analysis, the following strategies are recommended for Cardio

Good Fitness to strengthen customer engagement and increase sales performance:

#### **1. Market Segmentation by Income Level:**

- a. **TM195** should be marketed as a **budget-friendly model** for entry-level users, young professionals, or households with moderate income.
- b. **TM498** should target the **mid-range market**, offering value for money and balanced features for everyday fitness users.
- c. **TM798** should continue to be positioned as a **premium model**, focusing on performance, advanced features, and long-term durability for affluent and dedicated fitness enthusiasts.

## **2. Gender-Based Targeted Marketing:**

- a. For **female customers**, emphasize user comfort, design, and safety features (e.g., cushioning systems, compact design).
- b. For **male customers**, highlight performance metrics such as horsepower, incline range, and endurance capabilities.

## **3. Focus on Fitness-Level Engagement:**

- a. Offer personalized training programs or reward systems based on treadmill usage and fitness goals.
- b. Encourage lower-fitness users to start with **TM195** and provide incentives to upgrade as their fitness improves.

## **4. Enhanced Customer Retention and Loyalty:**

- a. Introduce loyalty discounts, referral bonuses, or membership programs that motivate repeat purchases and brand loyalty.
- b. Use customer feedback and satisfaction data to refine product features and services.

## **5. Data-Driven Marketing and Product Strategy:**

- a. Continue collecting detailed demographic and behavioral data to identify new market trends.
- b. Implement predictive modeling to forecast product demand based on customer income and fitness levels.

## **6. Promote Healthy Lifestyle Branding:**

- a. Position Cardio Good Fitness not just as a treadmill brand but as a complete wellness partner, integrating products with digital fitness tracking, personalized apps, and nutrition guidance.

## **Conclusion**

This analysis provided descriptive and inferential evidence about differences in customer profiles across treadmill products. The company can use these insights to tailor marketing and product positioning. This project provided an in-depth analysis of customer data from Cardio Good Fitness, with the goal of identifying key characteristics of customers purchasing different treadmill models — TM195, TM498, and TM798. Through descriptive statistics, visualization, and inferential analysis, we developed a comprehensive customer profile for each product line.

The exploratory data analysis revealed that TM195 is the most popular model, attracting a larger share of customers compared to TM498 and TM798. Most buyers were male, partnered, and had an average education of about 15–16 years. The typical customer plans to use the treadmill 3–4 times per week, running approximately 80–100 miles weekly. The average self-rated fitness score across all buyers was around 3.5 to 4.0 on a 5-point scale.

The inferential analysis provided further insight into how demographic and behavioral characteristics influence product choice. ANOVA tests demonstrated that while age does not significantly differ across models, income and self-rated fitness do. Buyers of the TM798 have the highest income and fitness levels, suggesting that the product attracts wealthier and more fitness-oriented customers. Meanwhile, TM195 appeals to customers with lower income and moderate fitness goals, confirming its positioning as an entry-level treadmill.

The Chi-square test between gender and product indicated a statistically significant association, suggesting that men and women show distinct preferences in treadmill purchases. Male customers tend to purchase the premium TM798, whereas female customers favor the TM195

model. Marital status also shows that most buyers are partnered, indicating that treadmill purchases may be motivated by shared household fitness goals.

Overall, the findings suggest that the primary factors influencing treadmill purchases are income, fitness level, and gender, while age and education have minimal effect. These results provide a data-driven foundation for marketing segmentation and targeted product positioning.

## Appendix: R Code for Reproducing Analysis

```
# R code to reproduce the main analyses (save as cardio_analysis.R)

# 1. Load data

df <- read.csv("CardioGoodFitness.csv", stringsAsFactors = FALSE)

# 2. Glimpse

dim(df); str(df); summary(df)

# 3. Product counts

table(df$Product)

# 4. Summary stats by product

library(dplyr)

df %>% group_by(Product) %>% summarise(n=n(), mean_age=mean(Age, na.rm=TRUE),
sd_age=sd(Age, na.rm=TRUE),

mean_income=mean(Income, na.rm=TRUE),
sd_income=sd(Income, na.rm=TRUE))

# 5. Boxplots

png("fig_income_by_product.png", width=800, height=600)

boxplot(Income ~ Product, data=df, main="Income by Product", ylab="Income")

dev.off()

png("fig_age_by_product.png", width=800, height=600)

boxplot(Age ~ Product, data=df, main="Age by Product", ylab="Age (years)")

dev.off()

# 6. ANOVA Age by Product

aov_age <- aov(Age ~ Product, data=df); summary(aov_age)

# 7. ANOVA Income by Product

aov_income <- aov(Income ~ Product, data=df); summary(aov_income)

# 8. Chi-square Gender vs Product

tbl <- table(df$Gender, df$Product); chisq.test(tbl)

# 9. ANOVA for Fitness by Product

aov_fitness <- aov(Fitness ~ Product, data=df); summary(aov_fitness)
```

```
# Save final results to CSV if desired  
write.csv(df, "cardio_cleaned.csv", row.names = FALSE)
```