Predictive Analytics with CardioGoodFitness Data

By

Dyako Murad

Savo Haider

Muhammad Sherzad

Supervised by

Dr. Miran Taha Abdullah

January, 2024 (English)

Table of Contents

Introduction

In this market research report, we delve into the diverse customer profiles associated with CardioGood Fitness treadmill product lines—TM195, TM498, and TM798. Over the past three months, we have conducted an in-depth analysis of individuals who purchased treadmills at CardioGood Fitness retail stores. Our dataset encompasses critical customer variables, including the type of product purchased, gender, age, education, relationship status, annual household income, frequency of treadmill use, anticipated weekly mileage, and self-rated fitness levels on a scale from 1 to 5. Through descriptive analytics, we aim to unravel distinct patterns and characteristics associated with each treadmill model, providing valuable insights for targeted marketing strategies. This comprehensive exploration aims to illuminate nuanced consumer preferences and behaviors, facilitating a strategic understanding of CardioGood Fitness customers and paving the way for informed business decisions.

Problem Statement

CardioGood Fitness faces the challenge of understanding and harnessing the nuanced distinctions in customer preferences across its treadmill product lines—TM195, TM498, and TM798. The absence of a clear delineation of customer profiles impedes the company's ability to tailor marketing strategies effectively. our research team, tasked with this crucial analysis, seeks to address questions surrounding the varying factors influencing consumer choices, such as age, gender, education, relationship status, and income. The lack of a comprehensive understanding of customer behavior inhibits the development of targeted approaches that resonate with specific consumer segments. Unraveling these complexities is paramount for CardioGood Fitness to optimize its product positioning, marketing communications, and overall business strategy. This problem statement highlights the imperative nature of discerning distinct customer characteristics to enable the company to navigate the competitive fitness market successfully.

Solution Method

-Python is a popular programming language that is widely used for data analysis and manipulation. It's beginner-friendly and has a huge range of libraries like Pandas, NumPy, and Matplotlib that make data cleaning and visualization easier.

- Seaborn is a Python library that allows you to create beautiful and informative statistical graphics. With Seaborn, you can easily create various types of graphs such as bar plots, line plots, scatter plots, and more. It provides a high-level interface that simplifies the process of creating visually appealing visualizations. By using Seaborn's functions and specifying the appropriate parameters, you can customize your graphs to suit your needs. It's a great tool for data visualization and analysis.

- Pandas is a powerful library in Python that provides data structures and functions to efficiently clean, transform, and analyze data. It's great for handling large datasets and performing tasks like filtering, sorting, and aggregating data.

- NumPy is another essential library for scientific computing in Python. It provides support for large, multi-dimensional arrays and matrices. It's often used for numerical computations and mathematical operations on data.

- Matplotlib is a plotting library that helps you create visualizations such as line plots, bar charts, and scatter plots. It's highly customizable and allows you to visualize your data in a clear and informative way.

These libraries work together seamlessly, allowing you to clean and preprocess your data using Pandas and NumPy, and then visualize it using Matplotlib. They are powerful tools for data analysis and can greatly enhance your ability to work with data efficiently.

Implementation

- import specific methods to read and display data and some operation statics.

```
In [10]:
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [11]:
file_path= "C:\\Users\\Best Center\\Desktop\\Fitness.csv";
ffile=pd.read_csv(file_path);
```

```
In [12]:
print(ffile.head())
```

```
  Product  Age  Gender  Education MaritalStatus  Usage  Fitness  Income  Miles
0   TM195   18    Male         14        Single      3        4   29562    112
1   TM195   19    Male         15        Single      2        3   31836     75
2   TM195   19  Female         14     Partnered      4        3   30699     66
3   TM195   19    Male         12        Single      3        3   32973     85
4   TM195   20    Male         13     Partnered      4        2   35247     47
```

-normalize data.

```
print(ffile.duplicated().sum())
```

```
0
```

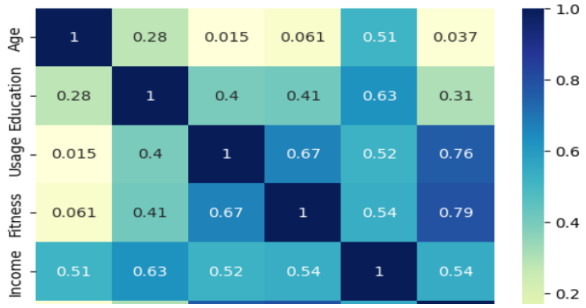
```
print(ffile.isnull().sum())
```

```
Product          0
Age              0
Gender           0
Education        0
MaritalStatus    0
Usage            0
Fitness          0
Income           0
Miles            0
dtype: int64
```

```
numeric_df = ffile.select_dtypes(include=['number'])
```

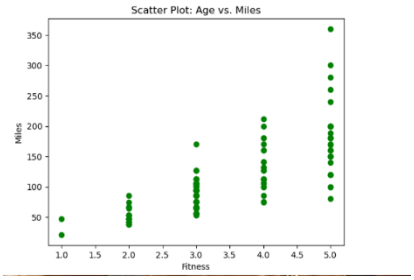
-fined correlations.

```
numeric_df = ffile.select_dtypes(include=['number'])
correlation_matrix = numeric_df.corr()
sns.heatmap(correlation_matrix, annot=True, cmap="YlGnBu")
plt.show()
```

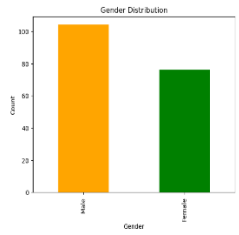


-some models.

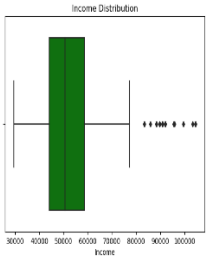
```
plt.scatter(ffile['Fitness'], ffile['Miles'],color='green')
plt.xlabel('Fitness')
plt.ylabel('Miles')
plt.title('Scatter Plot: Age vs. Miles')
plt.show()
```



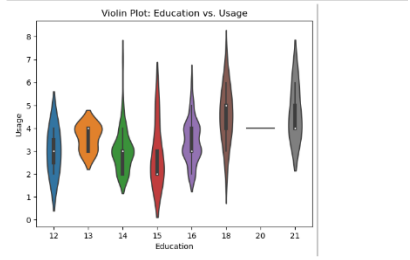

```
import matplotlib.pyplot as plt
gender_counts = ffile['Gender'].value_counts()
max_gender = gender_counts.idxmax()
colors = ['orange' if gender == max_gender else 'green' for gender in gender_counts.index]
gender_counts.plot(kind='bar', color=colors)
plt.xlabel('Gender')
plt.ylabel('Count')
plt.title('Gender Distribution')

plt.show()
```



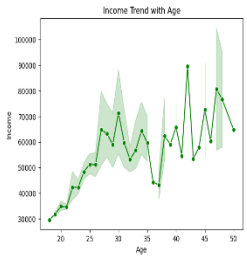

```
sns.boxplot(x='Income', data=ffile, color='green')
plt.xlabel('Income')
plt.title('Income Distribution')
plt.show()
```

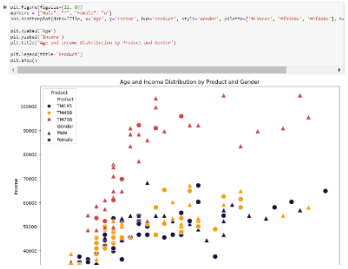



```
sns.violinplot(x='Education', y='Usage', data=ffile)
plt.xlabel('Education')
plt.ylabel('Usage')
plt.title('Violin Plot: Education vs. Usage')
plt.show()
```




```
sns.lineplot(x='Age', y='Income', data=ffile, marker='o', color='green')
plt.xlabel('Age')
plt.ylabel('Income')
plt.title('Income Trend with Age')
plt.show()
```





Results Discussion

The results of the project illuminate significant correlations between customer characteristics and product preferences within the fitness industry. Notably, the data-driven insights underscore the potential for strategic alignment in product development to optimize revenue and operational efficiency. The identification of key customer attributes necessitates a nuanced approach to both product design and marketing strategies. This approach emphasizes the importance of tailoring offerings to meet the diverse needs of the consumer base.Furthermore, the implementation of these insights reveals a tangible impact on the symbiotic relationship between the company and fitness enthusiasts. Customer engagement and satisfaction metrics demonstrate the effectiveness of aligning products with specific customer profiles. The iterative nature of product testing and adaptation ensures agility in responding to evolving consumer preferences. These dynamic adjustments contribute to sustained success by fostering a connection that resonates with the ever-changing landscape of the fitness market. Overall, the results discussion reinforces the strategic value of data-driven decision-making in enhancing the company's position within the competitive fitness industry.

Conclusion

Upon project culmination, discerning correlations emerged between customer characteristics and product preferences. These revelations underscore the importance of aligning product development with customer profiles to optimize revenue and enhance operational efficiency in the fitness industry. Strategic consideration of customer attributes is paramount for crafting products that resonate with clientele, ultimately fostering a symbiotic relationship between the company and its fitness enthusiasts. This insight advocates for a nuanced approach to product design and marketing strategies tailored to meet the diverse needs of the consumer base.