Aerofit Treadmill Customer Segmentation

About Aerofit

Aerofit is a leading brand in fitness equipment, offering a wide range of products like treadmills, exercise bikes, and gym equipment catering to diverse customer needs.

Business Problem

The Aerofit market research team seeks to understand their treadmill customer base better to recommend the most suitable product for each individual. They aim to identify potential relationships between customer characteristics and their treadmill choices (KP281, KP481, or KP781).

Dataset

The provided dataset (`Aerofit_treadmill.csv`) contains information on customers who purchased treadmills from Aerofit stores in the past three months. It includes the following features:

- **Product Purchased:** KP281, KP481, or KP781 (treadmill model)
- **Age** (in years)
- **Gender** (Male/Female)
- **Education** (in years)
- **Marital Status** (Single or Partnered)
- **Usage:** Average planned weekly treadmill usage
- **Income** (Annual income in $)
- **Fitness:** Self-rated fitness level (1-5 scale, 1 = poor, 5 = excellent)
- **Miles:** Average expected weekly walking/running distance

Product Portfolio:

- **KP281:** Entry-level treadmill ($1,500)
- **KP481:** Mid-level treadmill for runners ($1,750)
- **KP781:** Advanced features treadmill ($2,500)

Analysis Steps:

1. **Data Import and Cleaning:**
    - Import the `Aerofit_treadmill.csv` dataset.
    - Analyze data structure and characteristics.
    - Identify and handle missing values (if any) using appropriate techniques.

    - Detect and address outliers using methods like boxplots or IQR analysis.
2. **Exploratory Data Analysis:**
    - Analyze the distribution of customer characteristics using visualizations (histograms, boxplots)
    - Investigate potential relationships between features (e.g., age vs. income, usage vs. fitness level) using scatter plots or correlation matrices.
3. **Customer Profiling by Product:**
    - Use crosstabulation to calculate the marginal probability of purchasing each treadmill model (KP281, KP481, KP781).
    - Analyze customer characteristics (age, income, usage, etc.) for each treadmill type using descriptive statistics and visualizations.
4. **Conditional Probability Analysis:**
    - Construct two-way contingency tables to analyse the relationship between specific variables like gender and product purchased.
    - Calculate conditional probabilities to understand, for example, the probability of a male customer buying a KP781 treadmill.
5. **Customer Segmentation:**
    - Based on the analysis, segment customers into meaningful groups with similar characteristics and buying behaviour.

Expected Outcomes:

- Identify key customer segments for each Aerofit treadmill model.
- Understand the relationship between customer characteristics and their treadmill preference.
- Develop data-driven recommendations to improve customer targeting and product recommendations.

Actionable Insights:

- **Product Recommendations:** Based on customer profiles, refine online product recommendations to suggest the most suitable treadmill for each user.
- **Targeted Marketing:** Develop targeted marketing campaigns based on customer segments to attract and convert specific customer groups.

- **Product Development:** Use customer insights to guide future product development efforts and cater to diverse customer needs.