Project Title: Customer Segmentation Analysis in Online Retail

Module Code and Title: 7COM1039-0206-2024 - Advanced Computer Science Masters Project

Your Name and Student ID: Mihir Sanjaybhai Kalariya - 23029707

Aim of the Project

This project aims to analyze customer purchasing behavior in an online retail environment using **data-driven segmentation techniques**. The goal is to identify distinct customer groups based on their purchasing patterns, which can help businesses optimize marketing strategies, enhance customer retention, and improve revenue management.

Research Question/Hypothesis

- How can customer segmentation techniques be used to analyze purchasing behavior in an online retail setting?
- What insights can be gained from segmenting customers based on their transaction history?
- How can businesses leverage these insights for strategic decision-making?

Objectives

1. **Data Collection & Cleaning**: Acquire and preprocess transactional data, handling missing values and anomalies.
2. **Exploratory Data Analysis (EDA)**: Identify patterns and trends in customer behavior.
3. **Segmentation Techniques**: Apply methods such as **RFM (Recency, Frequency, Monetary) analysis, clustering, and statistical models** to segment customers.
4. **Evaluation & Business Insights**: Interpret segmentation results and propose actionable business strategies.
5. **Visualization & Reporting**: Present findings through dashboards or reports to facilitate stakeholder decision-making.

Short Description of Project Idea

Customer segmentation is a critical component of business strategy, enabling retailers to understand their customer base and tailor marketing efforts accordingly. This project will leverage **data analytics and machine learning techniques** to segment customers based on purchasing patterns. By analyzing transactional data, the project aims to provide insights that can improve **customer relationship management (CRM), personalized marketing, and business growth**.

Research Methodology

1. **Dataset Selection**: The project will use the "Online Retail" dataset from the UCI Machine Learning Repository (UCI Dataset).
2. **Data Preprocessing**:
    - Handle missing values and incorrect transactions.
    - Compute key metrics such as total purchase value and purchase recency.
3. **Segmentation Techniques**:
    - Utilize **RFM analysis** to categorize customers based on their purchasing behavior.
    - Explore **clustering techniques** (e.g., K-Means, DBSCAN, Hierarchical Clustering) to identify distinct customer groups.
4. **Analysis & Interpretation**:
    - Compare different segmentation models and evaluate effectiveness.
    - Extract business insights and suggest strategic improvements for customer engagement.
5. **Visualization & Reporting**:
    - Use charts and dashboards to present findings in an intuitive manner.

Citations

1. Hughes, A. (1994). **Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program**. McGraw-Hill.
2. Fader, P. S., Hardie, B. G., & Lee, K. L. (2005). **RFM and CLV: Using Iso-value Curves for Customer Base Analysis**. Journal of Marketing Research, 42(4), 415-430.
3. Han, J., Kamber, M., & Pei, J. (2011). **Data Mining: Concepts and Techniques**. Morgan Kaufmann.
4. UCI Machine Learning Repository. (2010). **Online Retail Dataset**. Retrieved from https://archive.ics.uci.edu/dataset/352/online+retail

---

Considerations & Feasibility

Data Availability: Open-source dataset from UCI ML Repository ensures easy access.
Realistic Timeline: The project scope is manageable within the given timeframe.

Business Impact: Findings can improve **marketing campaigns, customer loyalty strategies, and revenue forecasting**. **Practical Implementation**: The results can be used in real-world CRM applications.

Conclusion

This project provides a structured framework for **customer segmentation in online retail**, integrating **data science techniques and business strategy**. The insights derived from segmentation models will help businesses optimize marketing efforts and enhance customer relationships, ultimately driving growth and profitability.