CUSTOMER SEGMENTATION ANALYSIS– PROJECT

PROBLEM STATEMENT :

The purpose of this project is to perform customer segmentation analysis to gain insights into customer behavior, preferences, and characteristics. The goal is to identify distinct customer segments that can inform targeted marketing strategies, improve customer satisfaction, and optimize business operations.

PROJECT OBJECTIVE :

The primary objectives of this project are:

- Understand the diversity of customer behaviours and preferences.
- Segment customers based on common attributes and characteristics.
- Provide actionable insights to enhance marketing strategies.
- Improve customer engagement and satisfaction.

DATA DESCRIPTION:

The Data set used in this project was obtained from the Walmart.it contains

Dataset Information:
The online_retail.csv contains 387961 rows and 8 columns.

Feature Name	Description
Invoice	Invoice number
StockCode	Product ID
Description	Product Description
Quantity	Quantity of the product
InvoiceDate	Date of the invoice
Price	Price of the product per unit
CustomerID	Customer ID
Country	Region of Purchase

DATA PRE PROCESSING STEPS:

Data pre-processing involved the following steps:

- Handling missing values and outliers.
- Normalizing or scaling numerical features.
- Encoding categorical variables.
- Exploratory Data Analysis (EDA) to identify patterns and trends.

The inspiration for these steps came from the need to ensure the quality and reliability of the data for accurate segmentation analysis.

CHOOSING THE ALGORITHM FOR PROJECT:

The chosen algorithm for customer segmentation is K-means clustering. The algorithm was selected based on its ability to partition customers into distinct groups and its simplicity of implementation.

MOTIVATION AND REASON FOR THE ALGORITHM:

The motivation behind selecting [K-means clustering] includes:

- Efficient in handling large datasets.
- Ability to uncover hidden patterns and structures in the data.
- Proven effectiveness in customer segmentation tasks.

ASSUMPTIONS :

For the success of this project, the following assumptions were made:

- The dataset is representative of the overall customer population.
- The chosen algorithm is appropriate for the characteristics of the dataset.
- The features used for segmentation are relevant and accurately reflect customer behavior.

MODEL EVALUATION & TECHNIQUES :

Model evaluation was performed using [describe the evaluation metrics, e.g., silhouette score, within-cluster sum of squares]. Techniques such as cross-validation were employed to ensure the robustness of the results.

FUTURE POSSIBILITIES OF THIS PROJECT:

Potential future developments and enhancements for the project include:

- Incorporating additional data sources for more comprehensive segmentation.
- Implementing machine learning techniques for personalized marketing.
- Continuous monitoring and updating of customer segments based on evolving trends.

THANK YOU