**CUSTOMER SEGMENTATION USING DATA SCIENCE**

**Phase1:Problem definition and design thining**

**PROBLEM DEFINITION:**

The problem is to implement data science techniques to segment customers based on their behavior, preferences, and demographic attributes.

Customer segmentation is the practice of dividing a customer base into group of individuals that have similar characteristics relevant to marketing, such as age, gender, interests and spending habits.

The goal of customer segmentation is to reach out to customers more effectively, thereby leading to more sales or customer conversions. Companies also hope to gain a deeper understanding of their customers' preferences and needs by discovering what each segment finds most valuable and more accurately tailoring marketing materials toward that segment.

The goal is to enable businesses to personalize marketing strategies and enhance customer satisfaction. This project involves data collection, data preprocessing, feature engineering, clustering algorithms, visualization, and interpretation of results.

**DESIGN THINKING**

1. **Data Collection:** Collect customer data, including attributes like purchase history, demographic information, and interaction behavior.Gather from various sources of data to do the customer segments such as surveys, interviews, feedback forms, web analytics, CRM reports, social media, and email campaigns,kaggle.
2. **Data Preprocessing:** Clean and preprocess the data, handle missing values, and convert categorical features into numerical representations.

The raw data we downloaded is complex and in a format that cannot be easily ingested by customer segmentation models. We need to do some preliminary data preparation to make this data interpretable.

The informative features in this dataset that tell us about customer buying behavior include “Quantity”, “InvoiceDate” and “UnitPrice.” Using these variables, we are going to derive a customer’s RFM profile - Recency, Frequency, Monetary Value.

**RFM** is commonly used in marketing to evaluate a client’s value based on their:

**Recency:** How recently have they made a purchase?

**Frequency:** How often have they bought something?

**Monetary Value:** How much money do they spend on average when making purchases?

1. **Feature Engineering:** Create additional features that capture customer behavior Preferences, such as total spending, frequency of purchases.

Selecting the most important existing features.

Creating new features from existing features.

Aggregating features by customer.

1. **Clustering Algorithms:** Apply clustering algorithms like K-Means, DBSCAN, ,GMM algorithm,Mean shift or hierarchical clustering to segment customers.
2. **Visualization:** Visualize the customer segments using techniques like scatter plots, bar charts, graphs ,histograms, pie charts and heatmaps.
3. **Interpretation:** Analyze and interpret the characteristics of each customer segment to derive actionable insights for marketing strategies.Customer segmentation is the process of dividing customers into groups based on common characteristics, such as demographics or behaviors. The goal is to help marketing and sales teams reach out to customers more effectively