**Project 8: Customer Segmentation using data science**

**Problem Definition**

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

* Data Collection
* Data Preprocessing
* Feature Engineering
* Clustering Algorithms
* Visualization
* Interpretation

**Data Collection**

Collect customer data, including attributes like purchase history, demographic information, and interaction behavior.

The dataset we have for this project that contains the following data

* Customer ID
* Genre
* Age
* Annual Income
* Spending score

Dataset link: [**https://www.kaggle.com/datasets/akram24/mall-customers**](https://www.kaggle.com/datasets/akram24/mall-customers) **.**

**Data Preprocessing**

Clean and preprocess the data, handle missing values, and convert categorical features into numerical representations.

Missing values can be handled by deleting the rows or columns having null values.

This process of converting categorical data into numeric representation is known as encoding. Qualitative and Quantitative Data - Image Source. There are two types of categorical data: nominal and ordinal.

**Feature Engineering**

Create additional features that capture customer behavior and preferences, such as total spending, frequency of purchases, etc.

It is the process of transforming raw data into features that are suitable for machine learning models.

**Clustering Algorithms**

Apply clustering algorithms like K-Means, DBSCAN, or hierarchical clustering to segment customers.

Using a clustering algorithm means you're going to give the algorithm a lot of input data with no labels and let it find any groupings in the data it can.

**Visualization**

Visualize the customer segments using techniques like scatter plots, bar charts, and heatmaps.

Data visualization tools provide an accessible way to see and understand trends, patterns in data, and outliers.

**Interpretation**

Analyze and interpret the characteristics of each customer segment to derive actionable insights for marketing strategies.

It is the process of reviewing data and arriving at relevant conclusions using various analytical research methods.