High Impact Skills Development Program
in Artificial Intelligence, Data Science, and Blockchain

Project Title: Online Retail Segmentation.

Learning Objectives:

1. **Understanding the fundamentals of Data Mining**
2. **Learn how to use SQL in data mining**
3. **Learn how to implement mining concepts in MySQL Workbench**

NOTE: (Dataset are available on LMS)

Overview:

Customer segmentation is a typical strategy used by organizations to categories clients based on their demographics, shopping patterns, or other traits.

Customer segmentation constitutes a widely adopted strategic approach within organizations, aimed at systematically categorizing clients by analyzing a spectrum of factors such as demographics, shopping patterns, and other discernible traits. By effectively segmenting the customer base, businesses gain the ability to tailor their offerings, communications, and marketing efforts with a higher degree of precision, ultimately leading to enhanced customer satisfaction, targeted engagement, and improved overall business performance.

MetaData:

Look for a dataset that contains information about customers such as demographic information, purchasing history, and customer interactions. The data set contains the following variables:

InvoiceNo: The invoice number for each transaction

StockCode: The unique code for each product sold

Description: The description of each product sold

Quantity: The quantity of each product sold in each transaction

InvoiceDate: The date and time of each transaction

UnitPrice: The price of each product sold

CustomerID: The unique identifier for each customer

Country: The country where each transaction occurred

SQL Project Idea: Use SQL queries to answer the following questions:

Beginner Queries.

- Define meta data in mysql workbench
- What is the distribution of order values across all customers in the dataset?
- How many unique products has each customer purchased?
- Which customers have only made a single purchase from the company?
- Which products are most commonly purchased together by customers in the dataset?

Advance Queries

1. Customer Segmentation by Purchase Frequency

Group customers into segments based on their purchase frequency, such as high, medium, and low frequency customers. This can help you identify your most loyal customers and those who need more attention.

2. Average Order Value by Country

Calculate the average order value for each country to identify where your most valuable customers are located.

3. Customer Churn Analysis

Identify customers who haven't made a purchase in a specific period (e.g., last 6 months) to assess churn.

4. Product Affinity Analysis

Determine which products are often purchased together by calculating the correlation between product purchases.

5. Time-based Analysis

Explore trends in customer behavior over time, such as monthly or quarterly sales patterns.

Requirements:

1. Create account on GitHub (Already created)
2. Share your profile link in the project report (to be submitted at the end of your project)
3. Prepared a presentation of the project.
4. Prepare a short 3–4-page report about your project that should contain.
   a. Capture screenshots of the executed queries and incorporate them into the report.
   b. Elaborate upon the procedural aspects of the queries.