Reporting No:1   Week No:1   From :15/05/2025 To :22/05/2025
College ID**: 23AIML014**
Project Title: Exploratory Data Analysis and Customer Segmentation using Python

WEEKLY REPORT

Work done in last week (Attach supporting Documents):

1**. Acquired a comprehensive understanding of the fundamental principles of Exploratory Data Analysis (EDA)** and **Inferential Statistics**, with a specific emphasis on their practical application in addressing real-world challenges, including financial market analysis and consumer behaviour modelling.

2. **I have watched curated educational videos on sentiment analysis and customer data analytics, including:**

- What is EDA? | Exploratory Data Analysis with Python

- Inferential Statistics: Introduction and Concepts

- Customer Segmentation in Python using KMeans

- Sentiment Analysis using Python | NLP Tutorial

3. Proficient in employing industry-standard tools and techniques for data exploration and segmentation.

- Utilised pandas and matplotlib for data cleaning and visualisation.
- Executed statistical analysis, including calculating mean, variance, correlation, and identifying outliers.
- Implemented K-Means Clustering from scikit-learn for unsupervised segmentation of customer data.
- Generated visual representations using seaborn and plotly to effectively illustrate clusters and derive insights.

Reason for incomplete work: **N/A**   [All planned tasks completed for Week 1]

Plans for next week:

1. Project: Predicting House Prices (Linear Regression)

- **Goal: Predict house prices using multiple numerical features.**

- **Key Learnings: Multiple Linear Regression, model evaluation.**

- **Workflow:** Gather and clean data, select important features, train a model with Scikit-Learn, assess its performance using MSE, RMSE, and R², and visualise the actual and predicted prices.

2. Project: Wine Quality Prediction

- **Goal: Predict wine quality based on chemical properties.**

- **Dataset: Includes features like acidity, alcohol, and density.**

- **Approach:**

    - **Preprocess and explore data**

    - **Train 3 classifiers (e.g., Logistic Regression, Random Forest, SVM)**

    - **Evaluate using accuracy, F1-score, and confusion matrix**

    - **Visualize feature impact and model performance**

References:

**Oasis Infobyte : https://oasisinfobyte.com/
Clustering in Machine learning : https://www.geeksforgeeks.org/
clustering-in-machine-learning/**

Signature of External Guide   **Signature of Internal Guide**

Ayush Chaudhary
Program Coordinator

Student Id: 23AIML014

Student Name: Devang Dhandhukiya