Academic Task Number: 01

Course code: CSI323

Date of allotment: 04-Mar-2025

Course Title: Predictive Analysis

Date of submission: 04-Mar-2025

Maximum Marks: 20 Practical +10 theories

Sales Performance Analysis in IBM SPSS Modeler

A retail company is analyzing its sales performance across different store locations. The dataset consists of multiple files, including sales transactions, product details, store locations, and employee performance. The goal is to integrate and analyze these datasets to gain insights into revenue generation and store-wise performance.

Steps to Follow in IBM SPSS Modeler:

1. **Import the Data**: Load all five datasets into IBM SPSS Modeler.
2. **Append the Data**:
   - Append **Sales_Data_1** and **Sales_Data_2** as they have the same structure.
3. **Merge the Data**:
   - Merge the appended sales data with **Product_Details** using the Product_ID column.
   - Merge the resulting dataset with **Store_Locations** using the Store column.
4. **Apply a Set-to-Flag Transformation**:
   - Convert the **Store** column into a flag variable to categorize store locations.
5. **Calculate Revenue Earned by Each Store**:
   - Compute the total revenue (Revenue = Quantity × Price) for each store location.
6. **Create a New Stream**:
   - Design a new SPSS Modeler stream to process the cleaned and transformed data.
7. **Visualization**:
   - Generate a bar chart to visualize revenue distribution across different store locations.

Data set link -
https://drive.google.com/drive/folders/1zLdwP4HDxIIyQXqijh0G6PcvTQsFwaBV?usp=drive_link