

Week 8 - Data Analyst: Cross-selling recommendation

Team Member Details: Individual project (no team)

Name: Asha K. C.

Email: aasha.kc@outlook.com

Country: India

Company: DataGlacier

Specialization: Data Analyst

Problem Description:

The XYZ bank is having difficulty Cross-selling its products to existing customers. Customers are not buying additional products sold by their bank. Hence, as data analysts, we need to provide the required information to enhance their cross-selling methods.

Data Understanding:

Dataset provided for this project consists of 48 customer features. The size of the whole dataset is 2.13MB. There are a total of 13647309 customer records prior to cleaning. Below steps were followed for initial inspection of the dataset.

1. Load the Data:

- The data provided for this project consists of a csv file (train.csv). Load the file into a pandas dataframe
- Record pandas load time
- Check for errors and warnings during load

2. Check dimensions:

- Identify the number of rows and columns in the dataset.
- Display the first few rows of the csv file. (df.head)

3. List Column Names:

- Identify all the column names and its datatype (df.info)

4. Check Data type for each column:

- Identify the columns that need to be validated
- Convert column names to meaningful English text for easy understanding

- Correct the data type of columns that showed warnings when loading the data

5. Check missing values:

- Identify rows with missing values (`isnull().sum()`)
- Identify duplicates
- Identify values that need to be replace/remove with valid values

6. Understand column values:

- Go through the data set values and understand the column values based on the details given in the project description.

Github Repo link:

https://github.com/Asha-KC-07/Data-Glacier-Internship-2025---LISUM43/blob/main/Week%208%20-%20Data%20Analyst_Cross-selling%20recomendation/LoadData.ipynb