

Team members' details

Group Name: Cross-Selling Recommendations

Name: Sri Saikamal Priyank Samireddi

Email: kp.samireddi@gmail.com

Country: United Kingdom

College: Bayes Business School

Specialisation: Data Analyst

Problem description

XYZ Credit Union has good performance in selling banking products to new customers but struggles with cross-selling additional products to existing customers. They want to increase cross-selling to boost revenue.

Data cleansing and transformation were done on the data

Importing CSV Data:

- The raw CSV data import resulted in mixed dtype warnings. Columns like Age, and Income had mixed types.
- This was fixed by specifying dtypes for the affected columns on import.

Handling Missing Values:

- Key columns like Employee Status had 27734 missing values. Dropped those rows.
- Numeric columns like Income had a large number of missing values. Filled with mean.
- Categorical like Spouse Index filled with a default 0 value.
- Dropped Payroll, Pensions columns due to high missing rate.

Encoding Categorical Columns:

- Sex column encoded to 1 for Male rows and 0 for Female rows.

- Employee Status mapped to numeric codes 1-5 based on categories.
- Customer Relationship Type mapped to 1-4 based on categories.
- Applied similar encoding for Residence Index, Foreigner Index etc.

Converting Date Columns:

- Date, Primary C Date, First Contract Date columns converted to datetime datatype.

Creating Aggregated Columns:

- New columns created to sum related columns - Total Deposits, Total Accounts etc.
- Provides an aggregated view of related variables.

Binning Numerical Columns:

- Binned Age into buckets like 18-25, 35-45 etc. to simplify analysis.
- Binned Income into brackets like Low, Medium etc.

Exporting Cleaned CSV:

- Final cleaned data exported to new CSV file after all transformations.
- Removed irrelevant columns like Customer Code, Payroll etc.
- Made data standardized and model ready.

The key focus was handling missing data, encoding categoricals, aggregating related columns, and simplifying continuous variables. This preprocessing ensures clean, consistent data for analysis

GitHub Repo link

<https://github.com/kamalpriyank/Cross-Selling-Reccommendations>