

Data Intake Report

Name: Cross Selling recommendation – How to increase cross selling of Banking Products

Report date: 19/09/2023

Internship Batch: LISUM24

Version:1.0

Data intake by: DataFind

Data intake reviewer: Kumar Mitra

Data storage location: <https://github.com/kbmPractice/Week7>

Tabular data details:

File: Train.csv

Total number of observations	13647309
Total number of files	1
Total number of features	48
Base format of the file	.csv
Size of the data	2.13 GB

File: Test.csv:

Total number of observations	929615
Total number of files	1
Total number of features	24
Base format of the file	.csv
Size of the data	105 MB

Proposed Approach:

1. No duplicates were found
2. NaN values present, will need to explore the percentage of NaN values present if it will be suitable to drop these NaN values or else if any imputation method can be used for the missing values.