

Data Intake Report

Name: Week 7: Cross Selling Recommendation – Group Project

Report date: November 16th, 2024

Internship Batch: LISUM38

Version:1.0

Data intake by: Bilgan Kiris

Data intake reviewer: Data Glacier

Data storage location:

Tabular data details:

Total number of observations	929616
Total number of files	3
Total number of features	3
Base format of the file	.csv
Size of the data	2.23 GB

Proposed Approach:

Deduplication Validation: Use ncodpers as the unique identifier and check for duplicate rows based on this key.

Assumptions: Missing values in non-critical fields not significantly impact the analysis.