Data Intake Report

Name: Project Week 7 (Bank Marketing Campaign)

Report date: 6 july 2025
Internship Batch: LISUM45
Version:<1.0>
Data intake by: Priyanjali Patel
Data intake reviewer: Data Glacier
Data storage location: https://archive.ics.uci.edu/dataset/222/bank+marketing

Tabular data details: bank-additional-full.csv

Total number of observations	41188
Total number of files	5
Total number of features	21
Base format of the file	.csv
Size of the data	5.8 MB

Proposed Approach:
- Mention approach of dedup validation (identification)
  : To identify duplicate records in the dataset, I will check for exact duplicates across **all columns** in the data. This means that if two or more rows have identical values in every column, they will be considered duplicates and flagged for removal or further investigation.