**Data Intake Report**

**Name:** Bank Marketing Campaign Project

**Report Date:** 19/09/2022

**Internship Batch:** LISUM12

**Version**: 1.0

**Data Intake:** Rohit Sunku

**Data Intake Reviewer**: Data Glacier

Data Storage Location: <https://archive.ics.uci.edu/dataset/222/bank+marketing>

|  |  |
| --- | --- |
| **Total Number of Observations** | 45211 (bank-full), 41188(bank-additional) |
| **Total Number of Files** | 4 |
| **Total Number of Features** | 17 |
| **Base Format of Files** | .csv |
| **Size of the Data** | bank-full(3,664), bank(451), bank-additional-full(5,699), bank-additional(571) |
| **Number of Null Values** | 0 |
| **Number of Duplicate Values** | 0 (bank-full), 12(bank-additional) |

**Proposed Approach:**

* There are no missing values in this dataset
* Remove duplicate values for one of the dataset
* The data was separated by semicolons in one column, so we split this using the Text-to-Column Excel option
* We will run ML models on bank-full and bank-additional-full data – we compare and analyse the differences between these two
* Lastly, we will use the bank and bank-additional datasets to test more computationally demanding ML algorithms