Data Intake Report

Name: Cross Selling Recommendation
Report date: 8/21/2023
Internship Batch: LISUM23: 30
Version: 1.0
Data intake by: Erwin Linares
Data intake reviewer: n/a
Data storage location: https://github.com/elinares996/Case-study-Cross-Selling-Recommendation

Tabular data details:

Total number of observations	929616 rows
Total number of files	2
Total number of features	48
Base format of the file	.csv
Size of the data	2.23 GB

Proposed Approach:
- Mention your assumptions (if you assume any other thing for data quality analysis)
- My approach to this problem would be to get to know the data, look over it see what kind of columns I have; whether or not these columns are homonyms or synonyms of each other. Also would have to look at the values inside the data itself to find any human errors or values that were not correctly integrated. Once the cleaning is done I would look at some simple observations within the data begin my analysis. Finally visualize my findings in a way that stakeholders can quickly and effectively get what's being portrayed.
  - My assumptions from this data set are:
    - The customers from Spanish countries have different purchasing behaviors from those in more western countries.
    - There are more co-owner accounts then any other by a vast margin.
    - The co-owner accounts are tend to only purchase things like deposit accounts. Also it may be likely that some customers hold co-owner accounts but then go and get primary accounts for themselves
    - Gross income of primary accounts is higher than co-owner ones