# **Project Status Report**

## Name: Andrea Plunkett Jackson

## Date: 02/23/2024

## Status Report

| **Status** | Progress |
| --- | --- |
| **Supporting Documentation** | * The approach for modification of codebook was finalized based on data available through CSV file. Upon review of variables and identifiable datasets it contained, I eliminated the redundant factors prior to analysis. |
| **Preliminary Data** | * Variables of Interest per Original Codebook * Type: Characteristics of transactions to establish link between money movement type and fraud. * Amount: The monetary value of the transaction to be used to evaluate possible link between fraud occurrence and transactional charges. * NameDest: Receiver of transaction variable used to pinpoint possible commonality. * IsFraud: Variable determining current fraud behavior in order to establish connections if applicable. * IsFlaggedFraud: Review of account indicators on attempts on single transactions greater than $200,000 to establish controls in regard to Bank Secrecy Act. * OldBalanceOrg: Based on the research question on internal possibility of fraud incidents review of balances prior to fraud activity to see if pattern can be determined. * Updated codebook provided below * Process for Gathering Data * Fraud Data.csv mined and downloaded into Excel * Review of available information based and original codebook * Data sets pertaining to research question copied into new Excel file named La Banco * Data cleaned using revised codebook ((see updated codebook provided below) * All empty rows replaced with 0 for continuity. * Numeric datasets formatted to currency 2 decimal points * Text changed to numeric and scientific formatting removed * Removal of headers from Excel spreadsheet to prepare for importation. * Created schema “LaBanco”. * Created table using aforementioned variable names and data types (see updated codebook provided below) * Imported data from Excel into SQL table using Data Table Import Wizard. |
| **Data Analysis Requirements** | * While the available data allows for initial analysis of transactions considered fraudulent, more data should be gathered. In order to meet success criteria additional data in the form of data timestamp in place of step variable to be used to look for outline of fraudulent behavior as it is insufficient would be required This would allow for a clearer analysis statistically on establishing time of transactions meeting the standardized bank reporting and transparency enabling future systematical protocols to be established. Based on current information it restricts the full answer to the research questions established in order to fully protect the client from future fraud incidents as it is only a partial picture. |
| **Next Steps and Rationale** | * The provided dataset offers foundational statistics for the initial evaluation to address the client’s business need. It is noteworthy to state the inclusion of variables within the dataset aligns with aspects of the research question. While it suffices on base level enriching the endeavor with supplementary information would be required to yield valuable insights. Gaining a broader perspective of fraudulent indicators would shed light on correlations within fraud patterns that is limited by current identifiable records. |

## Supporting Documentation

***Updated Codebook***

|  |  |  |
| --- | --- | --- |
| **COLUMN** | **DESCRIPTION** | **DATA TYPE** |
| Type | CASH IN, CASH OUT, DEBIT, PAYMENT, and TRANSFER | Varchar |
| Amount | Amount of the transaction in local currency | Float |
| RecpName | Customer who is the recipient of the transaction | Varchar |
| IsFraud | Identifies a transaction as fraudulent (1) or non-fraudulent (0). | TinyInt |
| Flagged\_Fraud | Flags illegal attempts to transfer more than $200,000 in a single transaction. The business model aims to control massive transfers from one account to another and flags illegal attempts. An illegal attempt in this dataset is an attempt to transfer more than $200,000 in a single transaction. | TinyInt |
| OrgBalance | Initial balance before the transaction | Float |
| Count | Sequential numbering for analysis | INT |

## Preliminary Data

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated