**France Data Profiling Strategy**

**Ver 1.0**

***Version History***

|  |  |  |  |
| --- | --- | --- | --- |
| **Version #** | **Created By** | **Date** | **Remarks** |
| 0.1 | Prashant Maroli | 18-Feb-2019 | Draft version |
| 0.2 | Prashant Maroli | 20-Mar-2019 | Post initial IBG review |
| 0.3 | Ravi Chandran Bale | 21-Mar-2019 | Post Umair Merchant review |
| 0.4 | Ravi Chandran Bale | 24-Mar-2019 | Post Roland Michel Doummar review |
| 1.0 | Ravi Chandran Bale | 25-Mar-2019 | Post walkthrough with Roland and Team |

***Document Approval***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Final Version #** | **Submitted by** | **Date** | **Approver** | **Date** | **Comments** |
|  |  |  |  |  |  |

***Signoff Matrix***

|  |  |  |
| --- | --- | --- |
| **Approver Name** | **Designation** | **Date** |
| Virginie Seror | COO – France |  |
| David Dauer | Finance – France |  |
| Karima Abdesslami | Executive Director – France |  |
| Viktoria Szmidt | CRO – France |  |
| Samir Medjana | Head of Compliance |  |
| Nitin C/Marwa | HO FIN |  |
| Samrina/Soniya | HO CAD |  |
| Wael Qahoush | Compliance |  |

Table of Contents

[1 About the Document 5](#_Toc4437801)

[2 Data Profiling Approach 5](#_Toc4437802)

[2.1 Challenges 5](#_Toc4437803)

[2.2 Data requirement 5](#_Toc4437804)

[2.3 Data Profiling Dimensions 6](#_Toc4437805)

[2.4 Important terms 7](#_Toc4437806)

[2.4.1 Data Entity – 7](#_Toc4437807)

[2.4.2 Data Attributes 7](#_Toc4437808)

[2.4.3 Observation 8](#_Toc4437809)

[2.4.4 Data Profiling Environment 8](#_Toc4437810)

[2.5 Data Profiling Process Flow: 9](#_Toc4437811)

[2.6 Responsibilities 10](#_Toc4437812)

[2.7 Data Mapping 10](#_Toc4437813)

[3 Data Profiling Report 12](#_Toc4437814)

[3.1 Customer 12](#_Toc4437815)

[3.1.1 Overview 12](#_Toc4437816)

[3.1.2 Customer Name 12](#_Toc4437817)

[3.1.3 Corporate Customer (Non-Individual) Attributes 13](#_Toc4437818)

[3.1.4 Individual Customer Attributes 15](#_Toc4437819)

[3.1.5 CASA Account Attributes 17](#_Toc4437820)

[3.1.1 Account Address 19](#_Toc4437821)

[3.1.2 Account Statement 19](#_Toc4437822)

[3.1.3 Account Interest 20](#_Toc4437823)

[3.1.4 Accruals 21](#_Toc4437824)

[3.1.5 Account Block 21](#_Toc4437825)

[3.1.6 Account Balance 21](#_Toc4437826)

[3.1.7 Account Inactivity 22](#_Toc4437827)

[3.1.8 Account Future Balance 22](#_Toc4437828)

[3.1.9 Account Covering Instruction 23](#_Toc4437829)

[3.1.10 Demand Draft & Manager’s Cheques 23](#_Toc4437830)

[3.1.11 Cheques Issue Register 24](#_Toc4437831)

[3.1.12 Stop Cheques 24](#_Toc4437832)

[3.1.13 Cheque Returns for a Account Counter 25](#_Toc4437833)

[3.1.14 Deposit Attributes 25](#_Toc4437834)

[3.1.15 Loan Attributes 27](#_Toc4437835)

[3.1.16 Loan Repayment schedule check 29](#_Toc4437836)

[3.1.17 Trade Finance: 29](#_Toc4437837)

[3.1.18 Credit, Limits and Collateral: 29](#_Toc4437838)

[3.1.19 Global Markets (FX, Swap): 30](#_Toc4437839)

[3.1.20 Finance: 30](#_Toc4437840)

[3.1.21 Product Interest Rate (Fixed and Floating): 30](#_Toc4437841)

[3.2 Uniqueness 31](#_Toc4437842)

[4 Data Cleansing Approach 32](#_Toc4437843)

# About the Document

Data profiling is a precursor to the data migration activity to analyse the quality of current data of France. The analysis of the data will be on various different dimensions to investigate the quality of the data. Based on the data quality, data cleansing exercise needs to be conducted to ensure, we enhance the data quality for fitness for data migration.

In order to achieve the data quality goals, this document will detail the following –

1. Data profiling approach
2. Data cleansing approach
3. Roles and responsibilities for the various data profiling activities
4. Format and content of the data profiling report

# Data Profiling Approach

Data profiling entails defining the data requirement for profiling, preparing infrastructure and build scripts to extract the required data, analysing and measuring data migration impacts of the profiled data, presenting the profiled data to the key stakeholders, identifying the data attributes that will require cleansing and identifying data attributes that will not be considered for cleansing and management of cleansing items through JIRA.

## Challenges

There are various challenges that are foreseen for Data profiling and this needs to be addressed sufficiently in our approach to the data profiling. The challenges are listed below –

1. Access to source data
2. Understanding of the data dictionary to extract the data
3. Regulatory restrictions on accessing data in Project office (UAE)
4. Understanding of product catalogue and system modelling
5. Existing processes and the account mapping within the department
6. Understanding existing workflows within the products and user profile
7. Signature and archived documents are stored in shared folders (not in system)

## Data requirement

The source of all core banking data for profiling will be vBank core banking system , as for the Trade, it will be Meta Data (Excel Files / documents) will be validated against vBank GL balance.

As part of the data mapping, additional sources may be identified.

Since VBank Platform is not supported by the vendor and due to GDPR and security policy, Syncordis, vendor having expertise in vBank will be the custodian of the delivery of:

* Initial data profiling and assessment on overall data quality
* Extracting the required data efforts from VBank platform.
* Accountable for the data mapping/profiling deliverables

A fixed and agreed list of data attributes along with the templates will be listed as a requirement for Syncordis to provide the profiled data.

A data profiling environment consisting of only the vBank database will be provided to Syncordis for executing the data profiling scripts. This vBank database will be refreshed with the latest production database prior to data cleansing on the production environment.

## Data Profiling Dimensions

During the data quality assessment exercise, data quality has been defined and measured across four data quality dimensions as mentioned below.

| **Dimension** | **Significance** | **Description** |
| --- | --- | --- |
| **Completeness** | Is all data present and intact? | This check identifies the information which is missing in the data attributes  This check involves Identification of data fields with no content. |
| **Conformity** | Does the data conform to the metadata values for each attribute? | This check verifies if the data conforms to the allowed metadata format or data type for specific attributes. This check involves:  Determining the conformity of the data type format (date, string, currency, number etc.)  Determining the conformity of the data pattern (e.g. email, phone, Identifiers etc.) |
| **Accuracy** | Is the data correct & valid? | This check verifies if the information in the data attributes is accurate and maintains the referential integrity with other related attributes.  This check involves:   * Identification of suspect field values (e.g. containing special character, only one character etc.) * Ascertaining the data integrity within specific data attributes (Gender & Salutation correlation, Resident Flag and Residence) * Ascertaining if junk data has been provided (e.g. XX people with same phone number/email address etc.) * Determining the range of value for specific attributes (e.g. date, amount etc.) * Date Interdependency Checks between related date attributes (e.g. between Issue Date and Expiry Date etc.) * Analysing the valid data values based on reference data set. * Provision for international address and mobile number within EU. |
| **Uniqueness** | Is each record unique? | This check identifies duplicate records based on business defined rules. |

## Important terms

In order to ensure clear communication, we need to define the terms used in Data Profiling. This section may contain apparent and well understood terms yet it is documented for the purpose of unambiguity.

### Data Entity –

Data Entity groups the data elements for a high level function. A functional approach to grouping helps us to assign responsibility and accountability of the data quality. Since France has mainly data for Private banking and corporate customers, data profiling can be presented with the further breakdown for each of the entities.

Following are the Data Entities identified for Data profiling –

* 1. Customer
  2. Current Accounts and Savings Account (CASA) – Will be referred as CASA and will be part of liability track.
  3. Deposits – Will be part of the liability track
  4. Loans – Will be part of the asset track
  5. Trade Finance
  6. Global Markets
  7. Finance

### Data Attributes

Data attributes are the elements of a data entity. For e.g. Customer First Name is a data attribute of Customer Entity. Loan opening date is a data attribute of Loan entity. Data attributes also defines the significance of the data element for the overall data quality. For e.g. Date of birth is a significant data attribute of a customer and if it is missing it leads to a higher severity observation as compared to a lesser significant data element like date of graduation of the customer. Data attributes attached in the current document is base set of attributes. From completion perspective, all business validation reports and regulatory reports data elements will cross checked during execution.

### Observation

Deviation identified for any of the dimensions (Completeness, Conformity, Accuracy and Uniqueness) for any data attribute will be recorded as an observation in Jira. This observation will then be marked with severity based on the impact on data migration and BAU.

### Data Profiling Environment

Data profiling environment consists of vBank database and SQL query access to run scripts to retrieve profiling data. The database must be secured based on FAB security policies.

## Data Profiling Process Flow:

Once the data profiling environment is ready, the data profiling process is kicked off. This process will run prior to every migration execution.



## Responsibilities

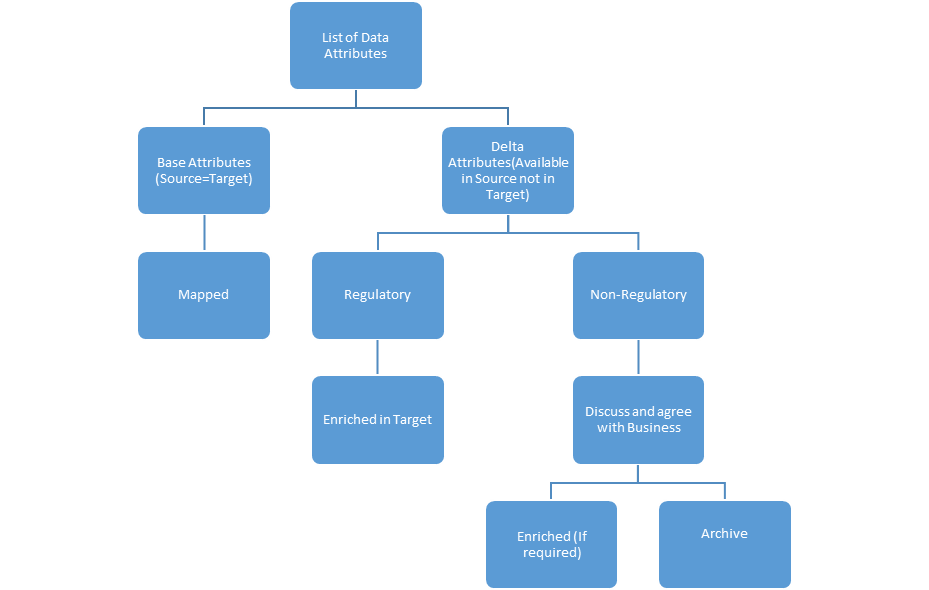
Following is the responsibility Matrix –

|  |  |  |
| --- | --- | --- |
| Activity | Responsibility | Location |
| Setup Data Profiling Environment | Group IT | France |
| Restore vBank Production Database | Group IT | France |
| Provide Data profiling Requirement | Evolvus (IBG) | UAE |
| Build Data Profile Scripts | Syncordis | France |
| Execute Data Profile Scripts | Syncordis | France |
| Provide Profiled Reports | Syncordis | France |
| Analyze Profiling Reports | Evolvus (IBG) | UAE |
| Identify and Raise Issues | Evolvus (IBG) | UAE |
| Prepare and Present Report to Stakeholder | Evolvus (IBG) | UAE |
| Create Data Cleansing Action Plan | IBG/France Business | UAE/France |
| Manual Data Cleansing | France Business | France |
| Track cleansing updates | Evolvus (IBG) | UAE |

## Data Mapping

An important prerequisite for data profiling will be the data mapping exercise. Data mapping activity will be executed in following steps by Syncordis along with IT DM.

1. Base target mapping will be used to map the source data to the target data.
2. Identify source data attributes that was not mapped in the base target. This will be called delta source attributes.
3. The delta source attributes will consists of
   1. Regulatory data attributes
   2. Non-regulatory data attributes
4. Regulatory attributes will be verified with software enablement for completeness
5. Non-regulatory data attributes will be analysed and business confirmation will be taken for
   1. Either enrichment is required in the target
   2. Or attributes need to be archived



# Data Profiling Report

The section provides the data profile requirement. The report consists of summary and details for all the data attributes of each entity.

## Customer

### Overview

|  |  |  |  |
| --- | --- | --- | --- |
| Customer | Total | Private Banking | CIB |
| Total Customer | 0 | 0 | 0 |
| Total Inactive / Dormant Customer  (based on flag in vBank) | 0 | 0 | 0 |
| Total Active Customer | 0 | 0 | 0 |
| Total Active Customer with all historical accounts | 0 | 0 | 0 |

### Customer Name

Corporate Customer (Non-Individual)

|  |  |  |  |
| --- | --- | --- | --- |
| Customer | Total | Private BankingNote3 | CIB |
| Total Corporate Customer | 0 | 0 | 0 |
| Total Active Corporate Customer | 0 | 0 | 0 |
| Full name will be available as part of the proposed solution that will be implemented in US (Having an additional custom data element as Customer full name) and same will be implemented for France. | 0 | 0 | 0 |

Individual Customer

|  |  |  |  |
| --- | --- | --- | --- |
| Customer | Total | Private Banking | CIB Note4 |
| Total Individual Customer | 0 | 0 | 0 |
| Total Active Individual Customer | 0 | 0 | 0 |
| Full name will be available as part of the proposed solution (Having an additional custom data element as Customer full name) that will be implemented in US and same will be implemented for France. | 0 | 0 | 0 |

### Corporate Customer (Non-Individual) Attributes

In addition to below, regulatory and payments fields should also be factored in for data profiling once discovery is completed by Synchordis.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **TOTAL CUSTOMER** | | | | **TOTAL ACTIVE CUSTOMERS** | | | |
| **Total** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** | **Total** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Salutation |  |  |  |  |  |  |  |  |
| Line of Business |  |  |  |  |  |  |  |  |
| Customer Type |  |  |  |  |  |  |  |  |
| Segment |  |  |  |  |  |  |  |  |
| Sub Segment |  |  |  |  |  |  |  |  |
| Business code |  |  |  |  |  |  |  |  |
| POS Code |  |  |  |  |  |  |  |  |
| Citizen |  |  |  |  |  |  |  |  |
| Nationality |  |  |  |  |  |  |  |  |
| Current delinquency status |  |  |  |  |  |  |  |  |
| Date of Birth |  |  |  |  |  |  |  |  |
| Open Date |  |  |  |  |  |  |  |  |
| Residence Country |  |  |  |  |  |  |  |  |
| Legacy Country code |  |  |  |  |  |  |  |  |
| Office country code |  |  |  |  |  |  |  |  |
| Visa Issue date |  |  |  |  |  |  |  |  |
| Visa Expiry Date |  |  |  |  |  |  |  |  |
| Company Registration |  |  |  |  |  |  |  |  |
| SWIFT CODE |  |  |  |  |  |  |  |  |
| Residence Address1 |  |  |  |  |  |  |  |  |
| Residence Address2 |  |  |  |  |  |  |  |  |
| Residence Address3 |  |  |  |  |  |  |  |  |
| Residence Address4 |  |  |  |  |  |  |  |  |
| Residence country code |  |  |  |  |  |  |  |  |
| Residence pin code |  |  |  |  |  |  |  |  |
| Residence City Code |  |  |  |  |  |  |  |  |
| Residence Phone Number |  |  |  |  |  |  |  |  |
| Fax number |  |  |  |  |  |  |  |  |
| Customer Risk Rating |  |  |  |  |  |  |  |  |
| Relationship Manager |  |  |  |  |  |  |  |  |
| Email Id1 |  |  |  |  |  |  |  |  |
| Email Id2 |  |  |  |  |  |  |  |  |
| FATCA Expiry Date |  |  |  |  |  |  |  |  |
| Security Number |  |  |  |  |  |  |  |  |
| TAX ID |  |  |  |  |  |  |  |  |
| Signature Id |  |  |  |  |  |  |  |  |
| POA First Name |  |  |  |  |  |  |  |  |
| POA Middle Name |  |  |  |  |  |  |  |  |
| POA Last name |  |  |  |  |  |  |  |  |
| GIIN |  |  |  |  |  |  |  |  |
| FAX Indemnity |  |  |  |  |  |  |  |  |
| POA Relationship |  |  |  |  |  |  |  |  |
| POA Customer Number |  |  |  |  |  |  |  |  |
| Corresponding Bank Info and mapping |  |  |  |  |  |  |  |  |

### Individual Customer Attributes

Below list should be aligned with France related requirement.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **TOTAL CUSTOMER** | | | | **TOTAL ACTIVE CUSTOMERS** | | | |
| **Total** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** | **Total** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Salutation |  |  |  |  |  |  |  |  |
| Profession |  |  |  |  |  |  |  |  |
| Customer Type |  |  |  |  |  |  |  |  |
| Segment |  |  |  |  |  |  |  |  |
| Sub Segment |  |  |  |  |  |  |  |  |
| POS Code |  |  |  |  |  |  |  |  |
| Citizen |  |  |  |  |  |  |  |  |
| Nationality |  |  |  |  |  |  |  |  |
| Current delinquency status |  |  |  |  |  |  |  |  |
| Date of Birth |  |  |  |  |  |  |  |  |
| Open Date |  |  |  |  |  |  |  |  |
| Residence Country |  |  |  |  |  |  |  |  |
| Legacy Country code |  |  |  |  |  |  |  |  |
| Office country code |  |  |  |  |  |  |  |  |
| Passport Number |  |  |  |  |  |  |  |  |
| Passport Issue Date |  |  |  |  |  |  |  |  |
| Passport Expiry Date |  |  |  |  |  |  |  |  |
| Visa Issue date |  |  |  |  |  |  |  |  |
| Visa Expiry Date |  |  |  |  |  |  |  |  |
| National ID |  |  |  |  |  |  |  |  |
| SWIFT CODE |  |  |  |  |  |  |  |  |
| Residence Address1 |  |  |  |  |  |  |  |  |
| Residence Address2 |  |  |  |  |  |  |  |  |
| Residence Address3 |  |  |  |  |  |  |  |  |
| Residence Address4 |  |  |  |  |  |  |  |  |
| Residence country code |  |  |  |  |  |  |  |  |
| Residence pin code |  |  |  |  |  |  |  |  |
| Residence City Code |  |  |  |  |  |  |  |  |
| Residence Phone Number |  |  |  |  |  |  |  |  |
| Fax number |  |  |  |  |  |  |  |  |
| Customer Risk Rating |  |  |  |  |  |  |  |  |
| Relationship Manager |  |  |  |  |  |  |  |  |
| Gender |  |  |  |  |  |  |  |  |
| Email Id1 |  |  |  |  |  |  |  |  |
| Email Id2 |  |  |  |  |  |  |  |  |
| FATCA Expiry Date |  |  |  |  |  |  |  |  |
| Security Number |  |  |  |  |  |  |  |  |
| TAX ID |  |  |  |  |  |  |  |  |
| Signature Id |  |  |  |  |  |  |  |  |
| POA First Name |  |  |  |  |  |  |  |  |
| POA Middle Name |  |  |  |  |  |  |  |  |
| POA Last name |  |  |  |  |  |  |  |  |
| GIIN |  |  |  |  |  |  |  |  |
| FAX Indemnity |  |  |  |  |  |  |  |  |
| POA Relationship |  |  |  |  |  |  |  |  |
| POA Customer Number |  |  |  |  |  |  |  |  |

### CASA Account Attributes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Product Code |  |  |  |  |  |
| POS Code |  |  |  |  |  |
| Customer number for account |  |  |  |  |  |
| Account name |  |  |  |  |  |
| Swift code |  |  |  |  |  |
| Financial Status |  |  |  |  |  |
| Inactive status |  |  |  |  |  |
| Dormant status |  |  |  |  |  |
| Facility Reference Number |  |  |  |  |  |
| Service Tier |  |  |  |  |  |
| Charge Type |  |  |  |  |  |
| Charge Amount |  |  |  |  |  |
| Charges Frequency |  |  |  |  |  |
| Account Open Date |  |  |  |  |  |
| Joint holder Number |  |  |  |  |  |
| POA Relationship |  |  |  |  |  |
| Alert Description |  |  |  |  |  |
| Alert Expiry Date |  |  |  |  |  |
| IBAN account Number |  |  |  |  |  |
| Statement required |  |  |  |  |  |
| Charge account number |  |  |  |  |  |
| Legacy account Number |  |  |  |  |  |
| Current delinquency status |  |  |  |  |  |
| IIS tracking flag |  |  |  |  |  |
| Interest in suspense amount |  |  |  |  |  |
| Recovery account number |  |  |  |  |  |
| Recovery full amount |  |  |  |  |  |
| Account Status |  |  |  |  |  |
| Account Opening POS code |  |  |  |  |  |
| Last interest application amount |  |  |  |  |  |
| Last accrual date |  |  |  |  |  |
| Interest scheme |  |  |  |  |  |
| Account for debit interest |  |  |  |  |  |
| Account for credit interest |  |  |  |  |  |
| Swift advice required or not |  |  |  |  |  |
| Credit / Debit Interest flag |  |  |  |  |  |
| Interest effective date |  |  |  |  |  |
| Interest expiry date |  |  |  |  |  |
| Interest rate |  |  |  |  |  |
| Amortization of Expenses and Revenue(To be checked whether upfront charges / Interest charged to the customer and how this is handled in T24) |  |  |  |  |  |

### Account Address

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Account Name |  |  |  |  |  |
| Pin code |  |  |  |  |  |
| Address 1 |  |  |  |  |  |
| Address 2 |  |  |  |  |  |
| Address 3 |  |  |  |  |  |
| Address 4 |  |  |  |  |  |
| City |  |  |  |  |  |
| Country code |  |  |  |  |  |

### Account Statement

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Account Number |  |  |  |  |  |
| Statement Frequency |  |  |  |  |  |
| Statement cycle |  |  |  |  |  |
| Next statement date |  |  |  |  |  |
| Statement delivery Mode |  |  |  |  |  |
| Last statement date |  |  |  |  |  |

### Account Interest

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Denominator |  |  |  |  |  |
| Debit interest rate structure |  |  |  |  |  |
| credit interest rate structure |  |  |  |  |  |
| Offset or Actual Flag |  |  |  |  |  |
| Debit minimum interest rate |  |  |  |  |  |
| Credit Maximum interest rate |  |  |  |  |  |
| Interest rate |  |  |  |  |  |
| Debit/Credit interest Flag |  |  |  |  |  |
| Offset rate |  |  |  |  |  |
| Amount slabs to be mapped for tiered interest schemes |  |  |  |  |  |

### Accruals

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Account No. |  |  |  |  |  |
| Debit Interest Accrual |  |  |  |  |  |
| Credit Interest Accrual |  |  |  |  |  |

### Account Block

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Account Number |  |  |  |  |  |
| Reason for the block |  |  |  |  |  |
| Block start date |  |  |  |  |  |
| Block End date |  |  |  |  |  |
| Block amount |  |  |  |  |  |
| Reference Number |  |  |  |  |  |

### Account Balance

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Account Number |  |  |  |  |  |
| Current Balance |  |  |  |  |  |
| Currency Code |  |  |  |  |  |
| Legacy Account |  |  |  |  |  |
| value date |  |  |  |  |  |
| Customer Number |  |  |  |  |  |

### Account Inactivity

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Legacy Account no |  |  |  |  |  |
| Last Credit Date |  |  |  |  |  |
| Last Credit amount |  |  |  |  |  |
| Last Debit Date |  |  |  |  |  |
| Last Debit amount |  |  |  |  |  |

### Account Future Balance

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Transaction Account Number |  |  |  |  |  |
| Transaction Amount |  |  |  |  |  |
| Transaction value Date |  |  |  |  |  |
| Customer Remarks |  |  |  |  |  |

### Account Covering Instruction

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Hunting Account Number |  |  |  |  |  |
| Full/Partial Flag |  |  |  |  |  |
| Status |  |  |  |  |  |
| Maturity Date |  |  |  |  |  |
| Effective date |  |  |  |  |  |
| Hunted account Number |  |  |  |  |  |
| Priority Account Number |  |  |  |  |  |

### Demand Draft & Manager’s Cheques

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Instrument Number |  |  |  |  |  |
| Branch code |  |  |  |  |  |
| Currency |  |  |  |  |  |
| Amount |  |  |  |  |  |
| Bank account |  |  |  |  |  |
| Purchaser Details |  |  |  |  |  |
| Issue Account |  |  |  |  |  |
| Payee Name |  |  |  |  |  |
| Issue Date |  |  |  |  |  |

### Cheques Issue Register

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Account Number |  |  |  |  |  |
| Currency |  |  |  |  |  |
| Cheques start number |  |  |  |  |  |
| Cheques end number |  |  |  |  |  |
| Cheques issue branch |  |  |  |  |  |
| Number of Cheques issued |  |  |  |  |  |
| Number of Cheques unused for the account |  |  |  |  |  |

### Stop Cheques

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Account Number |  |  |  |  |  |
| Currency |  |  |  |  |  |
| Customer No |  |  |  |  |  |
| Cheques number start for stop payment |  |  |  |  |  |
| Cheques number last for stop payment |  |  |  |  |  |
| No of leaves Low Cheques no |  |  |  |  |  |
| No of leaves up Cheques no |  |  |  |  |  |
| Stop date |  |  |  |  |  |
| Stop payment amount from |  |  |  |  |  |
| Stop payment amount To |  |  |  |  |  |
| Stop Active |  |  |  |  |  |

### Cheque Returns for a Account Counter

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Customer Number |  |  |  |  |  |
| Currency |  |  |  |  |  |
| Return Cheques Number |  |  |  |  |  |
| Return Cheques date |  |  |  |  |  |
| Transaction Reference |  |  |  |  |  |
| Return Cheques date |  |  |  |  |  |
| Transaction Reference |  |  |  |  |  |

### Deposit Attributes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Currency Code |  |  |  |  |  |
| Current Principal Amount |  |  |  |  |  |
| Value date of Deposit |  |  |  |  |  |
| Last interest Applied date |  |  |  |  |  |
| Maturity Date |  |  |  |  |  |
| Product Code |  |  |  |  |  |
| Interest rate type |  |  |  |  |  |
| Number of days deposit |  |  |  |  |  |
| Base Rate |  |  |  |  |  |
| Basis Point |  |  |  |  |  |
| Source Account Number |  |  |  |  |  |
| TIDE Account number |  |  |  |  |  |
| Target Account Number |  |  |  |  |  |
| Auto renewal condition |  |  |  |  |  |
| Auto renewal tenor |  |  |  |  |  |
| Auto renewal tenor |  |  |  |  |  |
| Auto renewal amount |  |  |  |  |  |
| Block reason code |  |  |  |  |  |
| Block reason text |  |  |  |  |  |
| Block till date |  |  |  |  |  |
| Legacy reference number |  |  |  |  |  |
| Original deposit Book date |  |  |  |  |  |
| Original Deposit amount |  |  |  |  |  |
| placement flag |  |  |  |  |  |
| Interest paid So far |  |  |  |  |  |
| Interest Accrued so for |  |  |  |  |  |
| last accrual date |  |  |  |  |  |

### Loan Attributes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Attribute Name** | **Number of records** | **Count for Data profiling** | **Completeness (%)** | **Conformity (%)** | **Accuracy (%)** |
| Current Balance |  |  |  |  |  |
| Product code |  |  |  |  |  |
| POS code |  |  |  |  |  |
| Loan Open Date |  |  |  |  |  |
| Currency |  |  |  |  |  |
| Disbursement date of the loan |  |  |  |  |  |
| Expiry Date of the loan |  |  |  |  |  |
| Category |  |  |  |  |  |
| Interest Rate type |  |  |  |  |  |
| Interest Rate type |  |  |  |  |  |
| Interest Rate type |  |  |  |  |  |
| Basis for interest calculation |  |  |  |  |  |
| Interest payment method |  |  |  |  |  |
| Interest rate of the loan |  |  |  |  |  |
| Interest calculation |  |  |  |  |  |
| Loan sub-product |  |  |  |  |  |
| Legacy system id |  |  |  |  |  |
| GLCMS Facility number |  |  |  |  |  |
| Accounts Officer |  |  |  |  |  |
| Original start date of the loan |  |  |  |  |  |
| Original Loan Amount |  |  |  |  |  |
| Salary Transfer |  |  |  |  |  |
| Base date definition |  |  |  |  |  |
| Maximum rate of the loan |  |  |  |  |  |
| Minimum rate of the loan |  |  |  |  |  |
| DDA reference number |  |  |  |  |  |
| DDA Expiry date |  |  |  |  |  |
| DDA Paying Bank |  |  |  |  |  |
| Indicator for Direct Debit setup |  |  |  |  |  |
| DDA Amount |  |  |  |  |  |
| New interest rate |  |  |  |  |  |
| Effective date of interest rate change |  |  |  |  |  |
| Deferment flag to indicate whether defer. is opted for by the customer or not |  |  |  |  |  |
| Algo Reference Number from NBAD |  |  |  |  |  |
| Remarks |  |  |  |  |  |
| Last Payment Amount |  |  |  |  |  |
| Total No of Payments Made |  |  |  |  |  |
| No Of Installments Paid |  |  |  |  |  |
| Rollover term |  |  |  |  |  |
| Type of schedule |  |  |  |  |  |
| Event date |  |  |  |  |  |
| Schedule amount |  |  |  |  |  |
| Rate Revision |  |  |  |  |  |
| Capit. decision flag for interest schedules |  |  |  |  |  |
| Description of the interest key |  |  |  |  |  |
| Actual interest rate |  |  |  |  |  |

### Loan Repayment schedule check

Below checks will be done in the Loan repayment schedule.

1. Change in interest rate / principal amount but rescheduling of the account is not taken place
2. Inconsistency in EMI amount
3. Outstanding interest accruals should match with the principal and Interest rate computation

### Trade Finance:

Currently Trade Finance fund based facilities (LC and LG ) are maintained in excel and Trade loans fund based facilities are in vBank. All GL balances are maintained in vBank. Data profiling perspective Trade Finance excel data will be validated against GL balance of vBank.

Also, by keeping in view of target system i.e TI+, TI+ data structure will be shared with France business to maintain the Trade Finance data.

### Credit, Limits and Collateral:

As per the current BAU, all borrowing customers along with limits and collateral are on- boarded in target stack i.e GLCMS.

No data profiling action to be expected. If any discrepancy is identified between country credit and HO credit, it will be taken up with Credit Integration team for the corrective action.

### Global Markets (FX, Swap):

Currently Global Markets (FX, Swap) are maintained in Murex XL and it is a standalone system. These transaction will be manually migrated to T24.

Following data attributes will be profiled for FX/SWAP transaction in vBank:

1. Amount bought
2. Amount sold
3. Exchange rate
4. Currency bought
5. Currency sold
6. Interest rate for SWAP
7. Counterparty details

### Finance:

Below is the list of mapping documents will be taken care by HO and Country finance as per FAB standards. Profiling is not required for the below.

* + 1. Internal Account
    2. PL Code
    3. Product Categories
    4. LOB POS
    5. Sector Code customer level
    6. Industry Code customer level
    7. Target Code customer level
    8. OneSumX hierarchy

### Product Interest Rate (Fixed and Floating):

All CASA and Lending products interest rate keys mappings, preferential interest rate from source to target and cleansing should be executed in coordination with HO Global Markets and Credit Risk and Operations. Attached base working file for easy reference.



### Uniqueness

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Business Key/Rule** | **Table Name** | **Total Count** | **Count for Data  profiling** | **Unique** | **Duplicate** | **Uniqueness %** |
| CIF Number |  |  |  |  |  |  |
| Passport Number |  |  |  |  |  |  |
| Email Id1 |  |  |  |  |  |  |
| Signature Id |  |  |  |  |  |  |
| Account Number |  |  |  |  |  |  |
| Account Name |  |  |  |  |  |  |
| Legacy Account number(CASA) |  |  |  |  |  |  |
| Debit Application Account number |  |  |  |  |  |  |
| Credit Application Account number |  |  |  |  |  |  |
| Signature Name |  |  |  |  |  |  |
| Legacy Account number(Deposit) |  |  |  |  |  |  |
| Legacy Account number(Loan) |  |  |  |  |  |  |

# Data Cleansing Approach

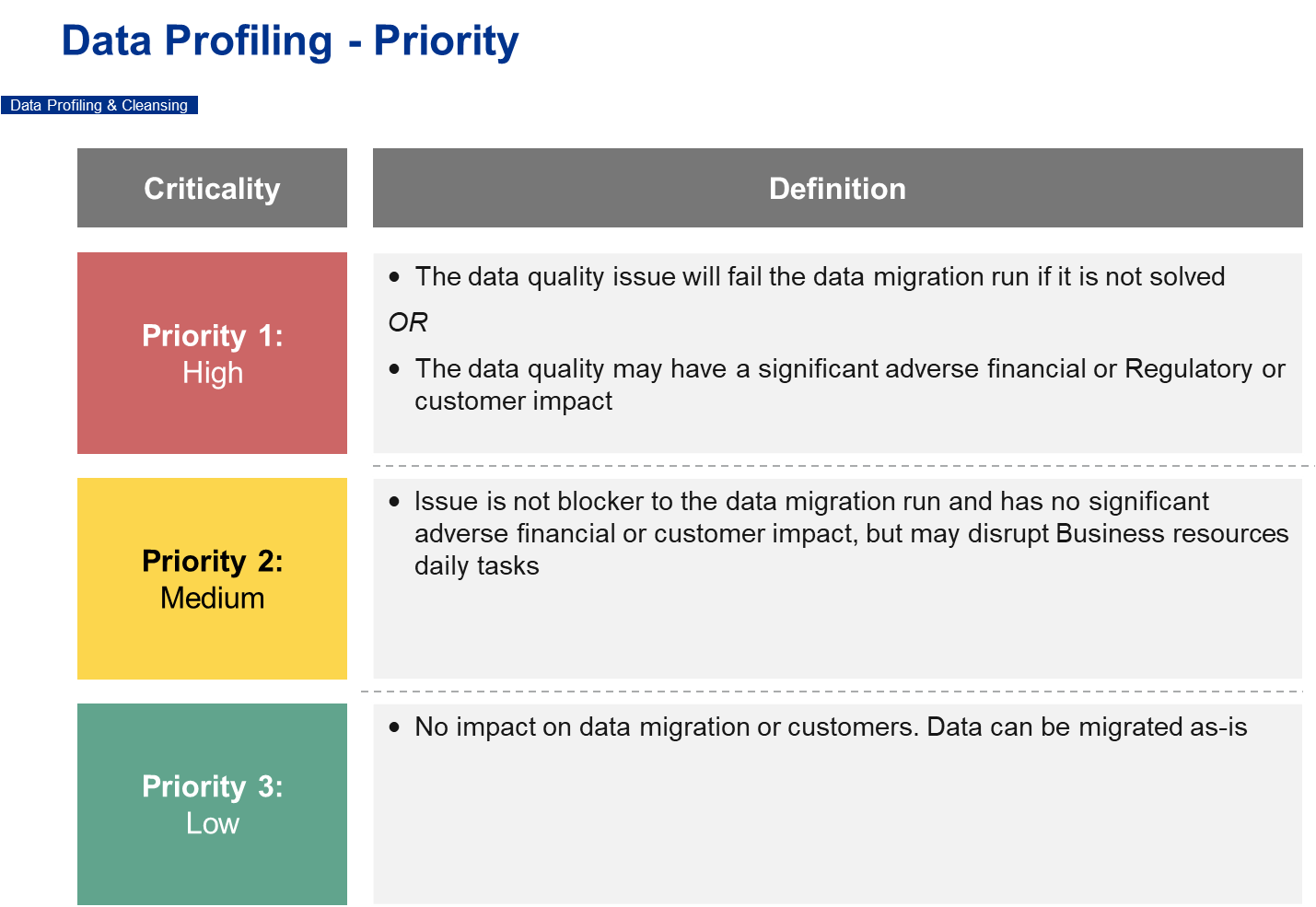
Once the data profiling exercise is completed, the output of the data profiling will result in Data Cleansing action report.

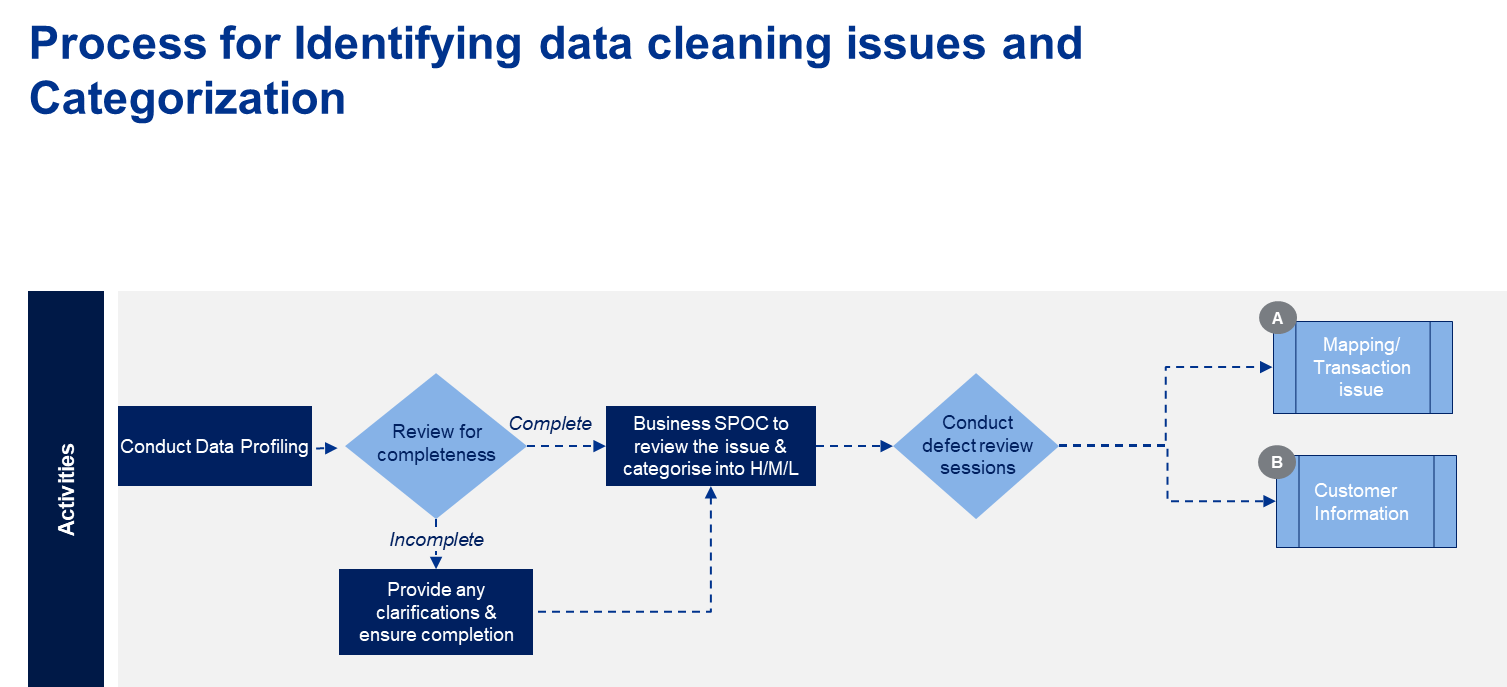
On each iteration of data profiling activity below Data cleansing action report with the all observations and actions will be published to all stakeholders and will be agreed with France business for manual correction of the data.

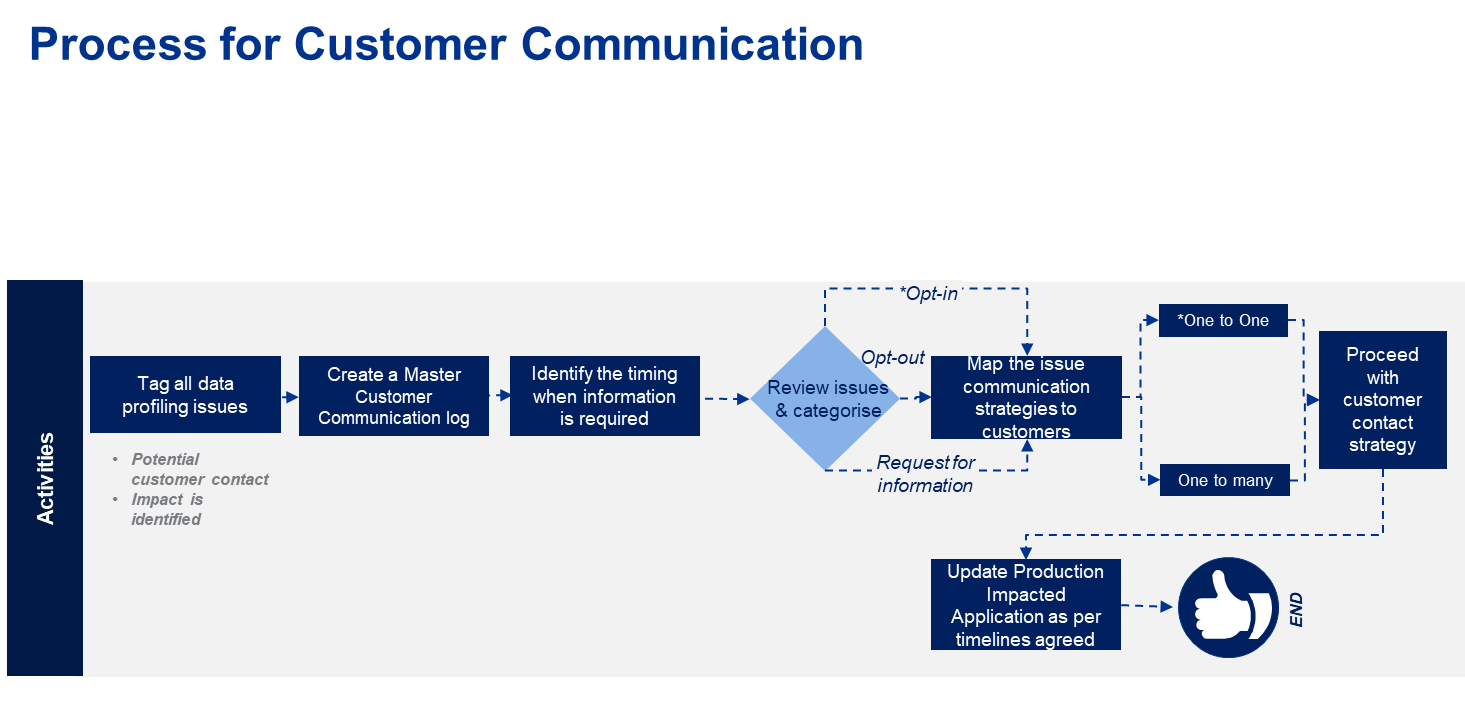
Below Data cleansing action report will be used for management of cleansing items through JIRA.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Jira No** | **Entity** | **Attribute** | **Issue Description** | **Action** | **Count of records** | **Status** | **Segment (CIB / PB)** | **Source System** | **Owner** | **Priority (H/M/L)** | **Target Date** |
| XXX | Customer | Passport No. | Passport No. is blank | Manually enter Passport no | **xx** | **Manual Correction** |  | **vBank or Excel** |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

Priority and categorization of each Jira is classified as per the below mentioned definition:







Below is the details process flow for data cleansing:

