**AML – KYC**

**Outils de pilotage du Flux et du Stock**

|  |  |
| --- | --- |
| Date : 21/02/2018 | **AML - KYC**  **Outils de pilotage du Flux et du Stock**  **Documentation technique et fonctionnelle** |
| Version : 1.0 |
| Rédacteur : Alain Le Gall  (SAGALINK Consulting) |

Table des matières

[1 Macro description du besoin 3](#_Toc507141130)

[1.1 Contexte 3](#_Toc507141131)

[1.2 Besoin 3](#_Toc507141132)

[1.3 Types de clients 3](#_Toc507141133)

[1.4 Outils 3](#_Toc507141134)

[2 Expression des Besoins 3](#_Toc507141135)

[3 Implémentation technique 3](#_Toc507141136)

[3.1 Architecture 3](#_Toc507141137)

[3.2 Bases de données 4](#_Toc507141138)

[3.2.1 Données BCT / Sidéral 4](#_Toc507141139)

[3.2.2 Données CRM 5](#_Toc507141140)

[3.2.3 Chaînes de connexion 6](#_Toc507141141)

[3.3 Outil de pilotage Excel 6](#_Toc507141142)

[3.3.1 Logique générale 6](#_Toc507141143)

[3.3.2 Récupération des dates d'extraction disponibles en base 7](#_Toc507141144)

[3.3.3 Mise à jour des données source et des indicateurs 8](#_Toc507141145)

[3.3.4 Transcodifications 10](#_Toc507141146)

[3.3.5 Onglets des indicateurs 10](#_Toc507141147)

[3.4 Log , mode DEBUG & langue 13](#_Toc507141148)

[4 Mode d’emploi utilisateur 14](#_Toc507141149)

[4.1 Logique générale 14](#_Toc507141150)

[4.2 Mise à jour des dates d'extraction 15](#_Toc507141151)

[4.3 Mise à jour des indicateurs 16](#_Toc507141152)

[4.4 Consultation des indicateurs 16](#_Toc507141153)

[4.4.1 Emplacement 16](#_Toc507141154)

[4.4.2 Personnalisation 16](#_Toc507141155)

[4.4.3 Regroupements de données 16](#_Toc507141156)

[4.4.4 Chronologies 17](#_Toc507141157)

[4.5 Historisation des indicateurs 18](#_Toc507141158)

[5 Divers 18](#_Toc507141159)

[5.1 Comparaison de fichiers de grande taille 18](#_Toc507141160)

[5.2 Utilisation par d'autres entités du groupe CACEIS 19](#_Toc507141161)

[5.2.1 Utilisation des bases de l'IP Luxembourg 19](#_Toc507141162)

[5.2.2 Duplication dans les IP locales 20](#_Toc507141163)

[5.3 Indicateurs dépriorisés 20](#_Toc507141164)

[5.3.1 Modifications réalisées par une autre entité sur des Tiers partagés 20](#_Toc507141165)

[5.3.2 Cohérence des Tiers "communs" à l'activité TA et Bank 20](#_Toc507141166)

# Macro description du besoin

## Contexte

Les équipes AML KYC de CACEIS Luxembourg s’assurent, via la collecte et l’analyse de données et de documents, de la conformité des clients face aux législations visant à contrer le blanchiment et le financement du terrorisme.

Les informations collectées sont saisies dans la base de données BCT (Base Clients Tiers), utilisée par l’ensemble des entités du groupe CACEIS. Les documents collectés sont « encodés » (saisie en base d’informations), scannés (et alors stockés dans une GED) et stockés physiquement lorsque la réglementation l’impose.

## Besoin

Afin de piloter l’activité de l’équipe AML KYC, les managers de l’équipe sont demandeurs d’outils produisant des indicateurs portant tant sur la qualité des informations saisies, leur cohérence, leur complétude, que sur la charge de travail associée à l’alimentation et à la maintenance de la base de données.

## Types de clients

Les clients de CACEIS Luxembourg sont :

* Soit des « Investisseurs », clients du Transfer Agent (TA) ;
* Soit clients des autres activités (tenue de compte, valorisation, etc.) et alors considérés comme clients « Banque ».

## Outils

Les indicateurs de pilotage sont basés sur les données de deux outils principaux :

* BCT dont des extractions offrent une photo du « stock » des documents collectés auprès des clients. BCT couvre le périmètre Banque et Investisseurs.
* CRM, outil de « flux » utilisé pour gérer la relation avec les clients investisseurs (échanges d’e-mails, workflow de collecte, …). Actuellement, seule la relation avec les clients « Investisseurs » est gérée via la CRM, les échanges avec les clients « Banque » étant réalisés directement dans Outlook.

# Expression des Besoins

Les besoins exprimés lors des différents entretiens menés tant sur le stock que sur le flux ont été regroupés dans la présentation PowerPoint ci-dessous :



# Implémentation technique

## Architecture

L’IP (Informatique de proximité) a défini des bonnes pratiques et mis en place l’architecture associée. La production des indicateurs demandés par le métier se fait en conformité avec cette architecture (Cf. schéma ci-dessous) :

➊ : Exports de BCT (mainframe) vers l’entrepôt de données Sidéral (inchangé). En effet, BCT est une base Mainframe qui n’est pas requêtée mais via une base intermédiaire nommée Sidéral.

➋ : Procédures stockées SQL et planification des exports à destination d'une base de données de l'IP. Les données y sont "timestampées" (via le champ "date\_extract") et historisées. (NB : en pratique, il y a 2 bases IP : une pour les données BCT et une pour les données du CRM).

➌ : Les données sont alors récupérées dans des fichiers Excel à l'aide de requêtes enregistrées dans des fichiers ODC (Office Data Connections). Ces fichiers ODC sont centralisés dans SharePoint (versioning, gestion des accès).

➍&➎ : Les outils de pilotage au format Excel sont hébergés dans SharePoint : le programme rafraîchit des tableaux de données et réalise les filtres, compilations, transformations, ..., nécessaires. Le résultat est présenté dans Excel.



## Bases de données

Deux bases de données sont utilisées pour stocker les données nécessaires au calcul des indicateurs :

* Une base pour les données issues du CRM (BDP\_CRM\_2011);
* Une base pour les données issues de BCT / Sidéral (BDP\_FDS).

### Données BCT / Sidéral

La base de données contient les tables suivantes, correspondant aux extractions demandées dans le fichier d'expression de besoin (Cf. § 2 Expression des Besoins) :

FDS\_tab\_qry1\_STA\_KYC

FDS\_tab\_qry2\_STA\_OPE

FDS\_tab\_qry5\_vol\_cre

FDS\_tab\_qry6\_vol\_clo

FDS\_tab\_qry\_7\_8

FDS\_tab\_qry10\_DELAI\_CRE

FDS\_tab\_qry\_11b\_STA\_R

FDS\_tab\_qry\_9\_11a\_13\_AF

FDS\_tab\_qry\_9\_11a\_13\_compte

FDS\_tab\_qry\_9\_11a\_13\_tiers\_role

FDS\_tab\_qry12\_docs\_scan

FDS\_tab\_qry\_14\_15\_derog

A titre d'exemple, la table FDS\_tab\_qry1\_STA\_KYC est constituée des champs suivants :

C\_GROUPE\_RATTACH

C\_ETABLISSEMENT

C\_SOUS\_TYPO\_SYNT

C\_TYPE\_SYNTHESE

D\_DEB\_VALIDITE

D\_FIN\_VALIDITE

C\_TIERS\_BCT

C\_ETAT\_SYNTHESE

C\_TYP\_TIERS

C\_NATURE\_ECO\_BCTGR

period

date\_extract

Le champ period est présent dans chacune des tables. C'est un champ permettant de préciser si la donnée consultée résulte d'un export quotidien, hebdomadaire ou mensuel. Cela permet en particulier de gérer des exports multiples un même jour (le jour où l'export hebdo est aussi un jour d'export mensuel).

Le champ date\_extract est présent dans chacune des tables. C'est un champ datetime permettant de préciser la date et le jour où l'extraction a été réalisée. Les extractions étant au mieux quotidiennes, l'heure est nécessaire pour gérer les cas de reprises produisant plusieurs exports au cours d'une même journée.

Sauf exception, les noms des autres champs sont repris à l'identique de ceux présents dans Sidéral.

### Données CRM

Les données issues du CRM sont stockées dans la table Cases qui contient les colonnes suivantes :

|  |  |  |
| --- | --- | --- |
| caseid | new\_fundsid | new\_statusdate |
| ticketnumber | new\_fundsidname | new\_dateallocated |
| statecode | new\_investorid | new\_dateworkstarted |
| statecodename | new\_investoridname | new\_datepending |
| statuscode | ca\_promoternames | new\_datewaitingfeedback |
| statuscodename | new\_platinumflag | new\_datereturnfeedback |
| new\_sequestrestatus | new\_platinumflagname | new\_sequestredaterelance1 |
| new\_sequestrestatusname | new\_MessageQueue | new\_sequestredaterelance2 |
| ownerid | new\_overdue | new\_sequestredaterelance3 |
| owneridname | new\_overduename | new\_plannedcompletiondate |
| ca\_teamname | createdon | new\_closedate |
| owneridtype | modifiedon | new\_effectivecompletiondate |
| title | new\_initialdate | new\_reopendate |
| casetypecode | new\_calculateddate | new\_queuename |
| casetypecodename | new\_dateplanification | db\_extractdate |

Le champ db\_extractdate un champ datetime permettant de préciser la date et le jour où l'extraction a été réalisée. Les extractions étant quotidiennes, l'heure est nécessaire pour gérer les cas de reprises produisant plusieurs exports au cours d'une même journée.

### Chaînes de connexion

La liaison entre les tableaux Excel et les bases de données sont réalisées via des chaînes de connexion telle que :

ODBC;DSN=SQLS\_IPQ\_WRITE\_LUX;UID=BDP\_FDS\_READ;PWD=fi9EJ;APP=Microsoft Office 2010;WSID=FWPC-85232;DATABASE=BDP\_FDS;

ODBC;DSN=SQLS\_IPQ\_WRITE\_LUX;UID=BDP\_FDS\_READ;PWD=fi9EJ;APP=Microsoft Office 2010;WSID=FWPC-85232;DATABASE=BDP\_CRM\_2011;

Deux outils permettent de gérer ces chaînes :

|  |  |
| --- | --- |
| Nom | ListAllOdbcConnString |
| Paramètres | NA |
| Action | * Ajoute un nouvel onglet au fichier Excel * Boucle sur toutes les chaînes de connexion ODBC existantes dans le fichier Excel et identifie la chaîne de connexion et la requête SQL associée * Ecrit ces informations dans l'onglet préalablement ajouté |

|  |  |
| --- | --- |
| Nom | SetAllOdbcConnectionStrings |
| Paramètres | Chaîne de connexion souhaitée |
| Action | * Ajoute un nouvel onglet au fichier Excel * Boucle sur toutes les chaînes de connexion ODBC existantes dans le fichier Excel et identifie la chaîne de connexion et la requête SQL associée * Remplace la chaîne de connexion existante par celle fournie en paramètre |

NB : A ce jour, toutes les requêtes utilisent la même chaîne de connexion à l'exception des deux requêtes qui pointent vers le CRM. Lorsque la chaîne de connexion définitive est connue, il convient donc d'utiliser SetAllOdbcConnectionStrings avec la chaîne pointant vers les données BCT puis de modifier manuellement la chaîne pointant vers le CRM.

## Outil de pilotage Excel

### Logique générale

L’outil Excel fonctionne selon la logique générale suivante :

* Les onglets dont le nom est préfixé de "qry\_" sont dédiés à la récupération, via des requêtes SQL, des données stockées dans la base de l'IP. Les requêtes SQL sont du type SELECT \* FROM table\_name WHERE date\_extract = ?, la date d'extraction voulue étant passée en paramètre.
* Les indicateurs sont présentés dans des onglets nommés en lien avec l'indicateur contenu (ex : "1-NbParStatutKYC"). Les données présentes dans ces onglets sont essentiellement calculées par des Tableaux Croisés Dynamiques (TCD) basés sur les tableaux contenus dans les onglets *qry\_\**.
* Un onglet "Parametres" permet à l'utilisateur :
  + De mettre à jour la liste des dates d'extraction possibles ;
  + De choisir les indicateurs à mettre à jour et la date d'extraction voulue ;
  + De lancer la mise à jour des onglets d'indicateurs.

Le recours au code VBA est restreint au nécessaire et utilisé pour :

* Mettre à jour les données externes à l'outil et nécessaires aux indicateurs demandés par l'utilisateur ;
* Mettre à jour les TCD associés aux indicateurs demandés par l'utilisateur ;
* Mettre à jour les "chronologies" permettant de filtrer les TCD (Cf. 3.3.5.2).

Le fichier Excel et le code VBA utilisent fréquemment des "noms" pour identifier des sources de données (Cf. <https://support.office.com/fr-fr/article/D%c3%a9finir-et-utiliser-des-noms-dans-les-formules-4d0f13ac-53b7-422e-afd2-abd7ff379c64?NS=EXCEL&Version=16&SysLcid=1036&UiLcid=1036&AppVer=ZXL160&HelpId=25450&ui=fr-FR&rs=fr-FR&ad=FR>). L'usage de ces noms sera détaillé dans chaque paragraphe concerné.

### Récupération des dates d'extraction disponibles en base

Un onglet Excel (qry\_date\_extract) est dédié à la récupération des dates d'extraction disponibles. Pour chaque table source, une requête de type (SELECT DISTINCT date\_extract, period, 'table\_name' AS SourceTable FROM table\_name) alimente un tableau tel que celui-ci :

|  |  |  |
| --- | --- | --- |
| **date\_extract** | **SourceTable** | **period** |
| 05/01/2018 17:04 | FDS\_tab\_qry5\_vol\_cre | M |
| 09/01/2018 10:04 | FDS\_tab\_qry5\_vol\_cre | M |

Ces tableaux sont nommés Dates\_qry\_*xxx* en fonction de la table requêtée, par exemple :

Dates\_qry\_1

Dates\_qry\_2

Dates\_qry\_5

Dates\_qry\_6

Dates\_qry\_7\_8

Etc.

La mise à jour VBA des données est automatisée par UpdateExtractionDates.

|  |  |
| --- | --- |
| Nom | UpdateExtractionDates |
| Paramètres | NA |
| Constantes utilisées | * WS\_EXTRACTION\_DATES As String = "qry\_date\_extract", nom de l'onglet Excel contenant les dates d'extraction disponibles. * WS\_PARAMETERS As String = "Parametres", nom de l'onglet Excel contenant les paramètres. * RNG\_PARAMETERS As String = "Parametres", nom Excel de la zone contenant les paramètres. * COL\_DATES As String = "Date à utiliser", nom de la colonne contenant les dates d'extraction voulues par l'utilisateur. * COL\_MAX\_DATES As String = "MaxDatesDispo", nom de la colonne contenant les dates d'extraction les plus récentes possibles pour chaque table. |
| Action | * Mise à jour des dates dans WS\_EXTRACTION\_DATES * Mise à jour de la colonne COL\_DATES avec les dates les plus récentes possibles |

La mise à jour des dates dans WS\_EXTRACTION\_DATES se fait par simple "Actualisation du tableau", la définition de la requête SQL étant intégrée dans l'objet tableau.

La mise à jour de la colonne COL\_DATES avec les dates les plus récentes possibles s'appuie sur du paramétrage Excel :

* Des noms DATES\_*xxx* sont définis comme étant =Dates\_qry\_*xxx*[date\_extract]

DATES\_1 = Dates\_qry\_1[date\_extract]

DATES\_2 = Dates\_qry\_2[date\_extract]

DATES\_5 = Dates\_qry\_5[date\_extract]

DATES\_6 = Dates\_qry\_6[date\_extract]

DATES\_7\_8 = Dates\_qry\_7\_8[date\_extract]

Etc.

* Dans l'onglet Parametres, pour chaque indicateur, le nom DATES\_*xxx* associé est indiqué dans la colonne "Zones des dates".
* Dans l'onglet Parametres, pour chaque indicateur, la colonne COL\_MAX\_DATES contient la formule Excel =MAX(INDIRECT(DATES\_*xxx*)).
* Post mise à jour des dates dans WS\_EXTRACTION\_DATES, la macro copie les dates de la colonne COL\_MAX\_DATES et les colle en valeurs dans la colonne COL\_DATES.

Le champ "period" indique si l'extraction concernée est une extraction mensuelle (M), hebdomadaire (H) ou quotidienne (Q). En effet, certains indicateurs sont demandés à la fois en hebdomadaire et en mensuel et ce champ permet de les différencier dans le cas où les deux extractions auraient lieu le même jour. Cette donnée est affichée dans le tableau des paramètres utilisateurs (via un RECHERCHEV).

### Mise à jour des données source et des indicateurs

La mise à jour VBA des données est automatisée par UpdateDataFromSqlServer.

|  |  |
| --- | --- |
| Nom | UpdateDataFromSqlServer |
| Paramètres | NA |
| Constantes utilisées | * WS\_PARAMETERS As String = "Parametres", nom de l'onglet Excel contenant les paramètres. * RNG\_PARAMETERS As String = "Parametres", nom Excel de la zone contenant les paramètres. * COL\_BOOL\_UPDATE As String = "MAJ", nom de la colonne contenant les VRAI/FAUX saisis par l'utilisateur pour sélectionner les indicateurs à mettre à jour. * COL\_QUERIES As String = "Onglets des requêtes", nom de la colonne contenant les onglets à mettre à jour pour requêter les données. * COL\_PIVOT\_TABLES As String = "Onglets des TCD", nom de la colonne contenant les onglets dont les TCD sont à mettre à jour. |
| Action | Pour les indicateurs dont COL\_BOOL\_UPDATE = VRAI :   * Mise à jour des tableaux de données dans les onglets listés dans COL\_QUERIES * Mise à jour des TCD des onglets listés dans COL\_PIVOT\_TABLES |

#### Mise à jour des données

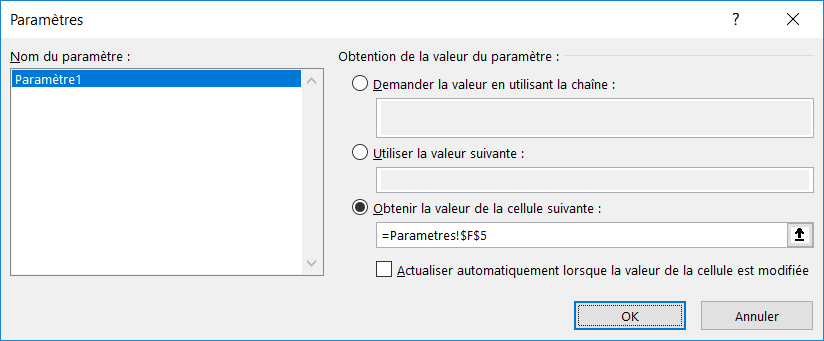
La mise à jour des tableaux de données fait appel à RefreshTables :

|  |  |
| --- | --- |
| Nom | RefreshTables |
| Paramètres | colSheetsToUpdate As Collection |
| Constantes utilisées | NA |
| Action | Pour chacun des onglets listés dans la collection, mise à jour de tous les tableaux de l'onglet via ListObject.QueryTable.Refresh BackgroundQuery:=False |

On voit ici que la date d'extraction choisie par l'utilisateur n'apparaît ni via une constante, ni via un paramètre. En effet, l'application du choix de l'utilisateur se fait en pratique directement dans la définition des requêtes qui sont formulées ainsi :

SELECT \* FROM FDS\_tab\_qry5\_vol\_cre WHERE date\_extract = ?

La présence du ? indique que la date est passée comme paramètre dont la source est définie dans l'écran dédié :



Chaque requête pointe ainsi vers la ligne adaptée de la colonne Date SQL dans le tableau RNG\_PARAMETERS, colonne renseignée par la formule =FormatDateSQL([@[Date à utiliser]]) permettant de passer la date choisie par l'utilisateur au format attendu par SQL. Cette formule fait appel à une fonction VBA afin de contourner les problèmes de langues Excel. La fonction est définie ainsi :

FormatDateSQL = Format(dateToFormat, "yyyy-mm-dd hh:mm:ss")

#### Mise à jour des TCD

La mise à jour des TCD fait appel à RefreshPivotTables :

|  |  |
| --- | --- |
| Nom | RefreshTables |
| Paramètres | colSheetsToUpdate As Collection |
| Constantes utilisées | NA |
| Action | Pour chacun des onglets listés dans la collection, mise à jour de tous les TCD de l'onglet via PivotTable.PivotCache.Refresh |

#### Cas particulier de l'onglet qry\_2\_t

Du fait de la modélisation des données dans BCT, la production de l'indicateur 2 sur les Tiers nécessite de requêter deux fois les données :

* Une première fois de manière classique avec un SELECT \*
* Une seconde fois de façon à récupérer 4 champs uniquement sur les Tiers qui ont un indicateur d'activité saisi :

SELECT C\_ID\_OBJ\_GENBCT, C\_IDENTIFIANT\_LUX, C\_VAL\_CODIF\_ADMIN, date\_extract

FROM FDS\_tab\_qry\_2\_TIERS\_ROLE

WHERE (C\_IDENTIFIANT\_LUX = 'I-ACT-SAI ' AND date\_extract = ?)

Cela permet, via un RECHERCHEV, de récupérer sur une même ligne les statuts déduits et saisis et de les comparer.

### Transcodifications

Les données récupérées comportent parfois des codes nécessitant des transcodifications afin de rendre les indicateurs plus lisibles. Cela est réalisé via des RECHERCHEV définis sur la droite des tableaux de données et utilisant des noms définis dans l'onglet correspondances tels que Statuts, Typologie, TypeDerog, etc. Ainsi, Statuts est utilisé dans des formules telles que = RECHERCHEV([@[C\_ETAT\_SYNTHESE]]; Statuts; 2; 0) et correspond au tableau suivant :

|  |  |
| --- | --- |
| **C\_ETAT\_SYNTHESE** | **Statut** |
| C | Complet |
| D | Complet avec dérogation |
| I | Incomplet |
| R | A renouveler |
| M | Original manquant |
| J | Incomplet, en constitution |

Cet onglet contient également un tableau Equipes permettant d'associer facilement un agent à une équipe et offrant toute la souplesse nécessaire en cas de modification des équipes, d'arrivée ou de départ d'agents, de constitution de nouvelles équipes, etc.

### Onglets des indicateurs

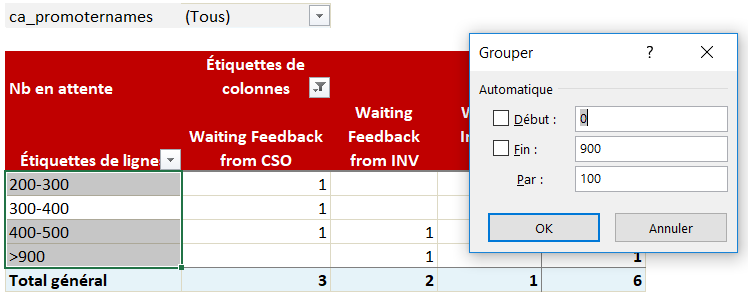
Les indicateurs sont calculés par des TCD qui ont pour source les tableaux de données. Ces derniers sont nommés tab\_qry\_*xxx* sans que ce nom ait une importance particulière pour le programme (il sert juste à la lisibilité).

Il y a une correspondance 1 pour 1 entre les onglets de données et les onglets d'indicateurs.

En plus du paramétrage classique des TCD (sélection des lignes, colonnes, données, filtres), les TDC utilisés dans les indicateurs font appel à des groupes et à des chronologies.

#### Groupes

Les groupes permettent de définir des regroupements de données tel que ci-dessous :

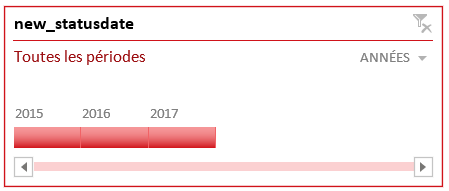


Les bornes des segments sont paramétrables (mais avec un pas fixe).

#### Chronologies (timelines)

*NB : Les chronologies présentées ci-dessous sont apparues avec Office 2013 alors que les utilisateurs sont actuellement sous Office 2010. Dans l'attente d'une migration sous Office 2016, l'utilisation des chronologies est donc impossible et peut être remplacée par l'utilisation des groupes mentionnés ci-dessus. Excel propose naturellement des groupes Années/Trimestres/Mois sur les champs Date. Les TCD initialement basés sur des chronologies sont donc livrés en double (chronologie et groupes).*

Lorsque l'on souhaite filtrer les données d'un TCD en fonction d'une date présente dans les données source, il est possible d'utiliser une "Chronologie" faisant apparaître une fenêtre telle que celle-ci :



Lorsque l'étendue du champ date choisi (ici new\_statusdate) est grande, la définition d'un filtre sur une période très courte peut être assez fastidieuse. Pour contourner ce problème, la sélection des dates à appliquer à la chronologie peut être définie par VBA :

|  |  |
| --- | --- |
| Nom | UpdateTimeline |
| Paramètres | sXlNameFrom As String  sXlNameTo As String  sSlicer As String |
| Constantes utilisées | NA |
| Action | Mise à jour de la chronologie nommée sSlicer afin de définir un filtre entre la valeur de la cellule Excel dont le nom est sXlNameFrom et celle dont le nom est sXlNameTo .  En fonction de la taille de la fenêtre de dates choisie, la précision de l'affichage de la chronologie évolue (en jours, mois, trimestres ou années). |

La fonction UpdateTimeline est appelée par autant de procédures VBA qu'il y a de chronologies à mettre à jour (chaque chronologie a ses propres dates de début et de fin). Ainsi, à ce jour, 4 procédures ont été créées pour autant de chronologies :

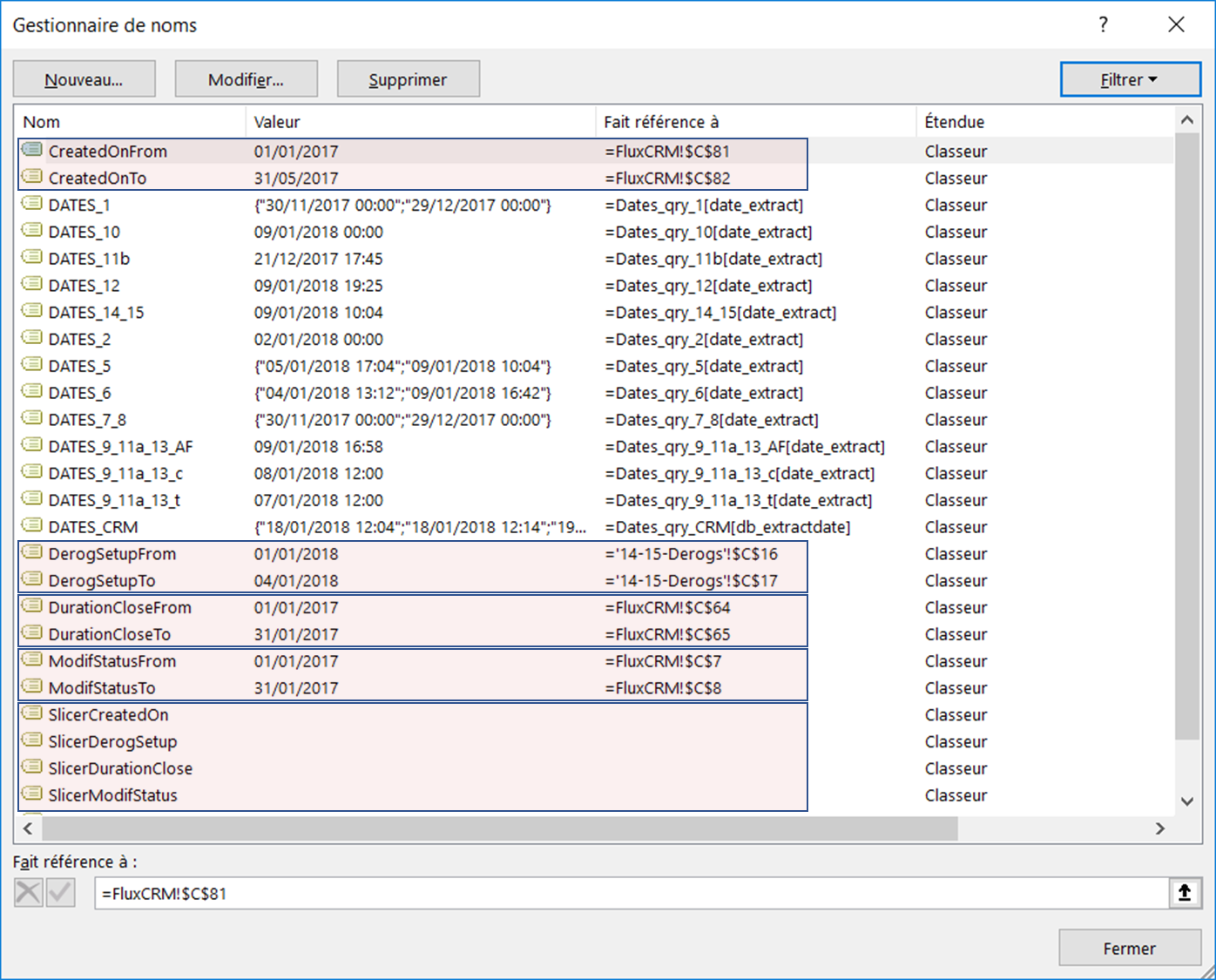
|  |  |
| --- | --- |
| Nom | UpdateTimelineModifStatus |
| Paramètres | NA |
| Noms Excel utilisés | ModifStatusFrom : nom de la cellule contenant la date de départ  ModifStatusTo : nom de la cellule contenant la date de départ  SlicerModifStatus : nom de la chronologie |
| Action | Appel de la fonction UpdateTimeline pour mettre à jour la chronologie. |

|  |  |
| --- | --- |
| Nom | UpdateTimelineDurationClose |
| Paramètres | NA |
| Noms Excel utilisés | DurationCloseFrom : nom de la cellule contenant la date de départ  DurationCloseTo : nom de la cellule contenant la date de départ  SlicerDurationClose : nom de la chronologie |
| Action | Appel de la fonction UpdateTimeline pour mettre à jour la chronologie. |

|  |  |
| --- | --- |
| Nom | UpdateTimelineCreatedOn |
| Paramètres | NA |
| Noms Excel utilisés | CreatedOnFrom : nom de la cellule contenant la date de départ  CreatedOnTo : nom de la cellule contenant la date de départ  SlicerCreatedOn : nom de la chronologie |
| Action | Appel de la fonction UpdateTimeline pour mettre à jour la chronologie. |

|  |  |
| --- | --- |
| Nom | UpdateTimelineDerogSetup |
| Paramètres | NA |
| Noms Excel utilisés | DerogSetupFrom : nom de la cellule contenant la date de départ  DerogSetupTo : nom de la cellule contenant la date de départ  SlicerDerogSetup : nom de la chronologie |
| Action | Appel de la fonction UpdateTimeline pour mettre à jour la chronologie. |

Les adresses des dates concernées sont visibles dans le gestionnaire de noms :



NB : Lors de la création d'une nouvelle chronologie, Excel lui attribue un nom par défaut tel que ChronologieNative\_*nomduchamp* qu'il est possible de modifier dans le gestionnaire de noms ci-dessus.

## Log , mode DEBUG & langue

Les réalisations successives du programme et les éventuelles erreurs sont listées dans une feuille Excel :



Par défaut, le programme écrit ces informations dans la feuille "log" du programme lui-même. Cependant, particulièrement en cas de plantage d'Excel, il peut être intéressant d'écrire ces informations dans un fichier Excel différent sauvegardé au fur et à mesure. Pour cela, il faut passer la constante DEBUG\_MODE à True et choisir le chemin de sauvegarde voulu avec la constante DEBUG\_PATH.

La langue des informations enregistrées peut être choisie en modifiant la constante PROG\_LANGUAGE (FR ou EN).

NB : Les noms des onglets Excel ou les intitulés des colonnes peuvent être modifiés (par exemple pour être en anglais) à condition de modifier en conséquence les constantes suivantes du programme :

Noms des onglets :

Public Const WS\_EXTRACTION\_DATES As String = "qry\_date\_extract"

Public Const WS\_PARAMETERS As String = "Parametres"

Public Const WS\_CRM As String = "Flux"

Noms des colonnes du tableau de paramètres :

Public Const COL\_BOOL\_UPDATE As String = "MAJ"

Public Const COL\_DATES As String = "Date à utiliser"

Public Const COL\_MAX\_DATES As String = "MaxDatesDispo"

Public Const COL\_QUERIES As String = "Onglets des requêtes"

Public Const COL\_PIVOT\_TABLES As String = "Onglets des TCD"

# Mode d’emploi utilisateur

## Logique générale

La production des indicateurs demandés par le métier se fait en conformité avec cette architecture (Cf. schéma ci-dessous) :

➊ : Exports de BCT (mainframe) vers l’entrepôt de données Sidéral (inchangé). En effet, BCT est une base Mainframe qui n’est pas requêtée mais via une base intermédiaire nommée Sidéral.

➋ : Procédures stockées SQL et planification des exports à destination d'une base de données de l'IP. Les données y sont "timestampées" (via le champ "date\_extract") et historisées. (NB : en pratique, il y a 2 bases IP : une pour les données BCT et une pour les données du CRM).

➌ : Les données sont alors récupérées dans des fichiers Excel à l'aide de requêtes enregistrées dans des fichiers ODC (Office Data Connections). Ces fichiers ODC sont centralisés dans SharePoint (versioning, gestion des accès).

➍&➎ : Les outils de pilotage au format Excel sont hébergés dans SharePoint : le programme rafraîchit des tableaux de données et réalise les filtres, compilations, transformations, ..., nécessaires. Le résultat est présenté dans Excel.



L’outil Excel fonctionne selon la logique générale suivante :

* Les onglets dont le nom est préfixé de "qry\_" sont dédiés à la récupération des données stockées dans les bases de données de l'IP.
* Les indicateurs sont présentés dans des onglets nommés en lien avec l'indicateur contenu (ex : "1-NbParStatutKYC"). Les données présentes dans ces onglets sont essentiellement calculées par des Tableaux Croisés Dynamiques (TCD) basés sur les tableaux de données contenus dans les onglets *qry\_\**.
* Un onglet "Parametres" permet à l'utilisateur :
  + De mettre à jour la liste des dates d'extraction possibles ;
  + De choisir les indicateurs à mettre à jour et la date d'extraction voulue ;
  + De lancer la mise à jour des onglets d'indicateurs.

## Mise à jour des dates d'extraction

Les extractions automatiques mises en place par l'IP permettent d'historiser les données dans des bases dédiées. Ainsi, les extractions pour lesquelles les extractions sont mensuelles seront disponibles :

* En date du dernier jour du mois précédent (End Of Last Month : EOLM)
* En date EOLM - 1M (End Of Last Month moins 1 mois)
* En date EOLM - 2M
* En date EOLM - 3M
* …
* En date EOLM - 12M

A ce jour, il est demandé que les historiques sont conservés sur 13 mois glissants.

Pour chacun des indicateurs, l'utilisateur a donc la possibilité de choisir la date des données à récupérer mais cela nécessite avant tout de mettre à jour la liste des dates disponibles dans le fichier Excel.

A cet effet, l'utilisateur doit :

* se rendre dans l'onglet "Parametres" ;
* cliquer sur le bouton "Maj des dates d'extractions disponibles" ;
* attendre que le programme lui indique que cette mise à jour est terminée.

## Mise à jour des indicateurs

Une fois mises à jour les dates d'extraction disponibles, l'utilisateur doit :

* se rendre dans l'onglet "Parametres" ;
* choisir les indicateurs qu'il souhaite mettre à jour en renseignant "VRAI" dans la colonne "MAJ" ;
* choisir la date d'extraction des données à utiliser pour chacun des indicateurs dans la colonne "Date à utiliser" ;
* cliquer sur le bouton " Maj des indicateurs" ;
* attendre que le programme lui indique que cette mise à jour est terminée.

La colonne "Fréquence" située à côté de la colonne "Date à utiliser" indique si l'extraction concernée est une extraction mensuelle (M), hebdomadaire (H) ou quotidienne (Q). En effet, certains indicateurs sont demandés à la fois en hebdomadaire et en mensuel et ce champ permet de les différencier dans le cas où les deux extractions auraient lieu le même jour.

NB : par défaut, la colonne "Date à utiliser" est toujours renseignée par la date d'extraction la plus récente disponible.

## Consultation des indicateurs

### Emplacement

Les différents indicateurs peuvent être consultés dans chacun des onglets en bleu. La date d'extraction utilisée est indiquée dans le titre de l'onglet (ligne 2).

### Personnalisation

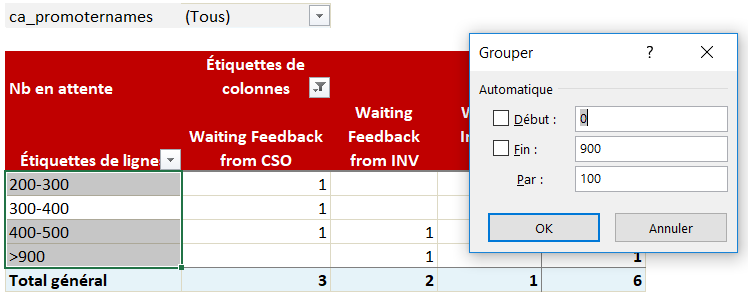
Les indicateurs sont produits en utilisant des Tableaux Croisés Dynamiques (TCD) basés sur les données extraites. L'utilisateur a donc la possibilité de modifier ces TCD (par exemple d'échanger les lignes et les colonnes, ou de définir des critères différents, etc.).

Attention : Si les modifications sont enregistrées, elles seront alors conservées par la suite. En effet, il n'y a pas de "remise à zéro" des TCD par le programme. En cas d'incertitude, il peut être bénéfique de commencer par copier le TCD existant avant de le modifier.

Par ailleurs, les noms des onglets Excel ou les intitulés des colonnes peuvent être modifiés (par exemple pour être en anglais) **à condition de modifier en conséquence les constantes VBA du programme (à réaliser en relation avec l'IP uniquement)**.

### Regroupements de données

Les groupes permettent de définir des regroupements de données tel que ci-dessous :

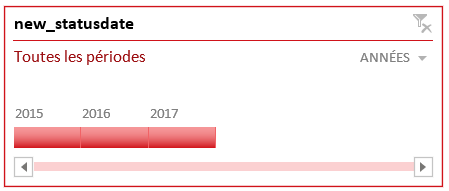


Les bornes des segments sont paramétrables (mais avec un pas fixe). Pour modifier les critères de regroupement, faire un clic droit sur les étiquettes de ligne et cliquer sur "Grouper…".

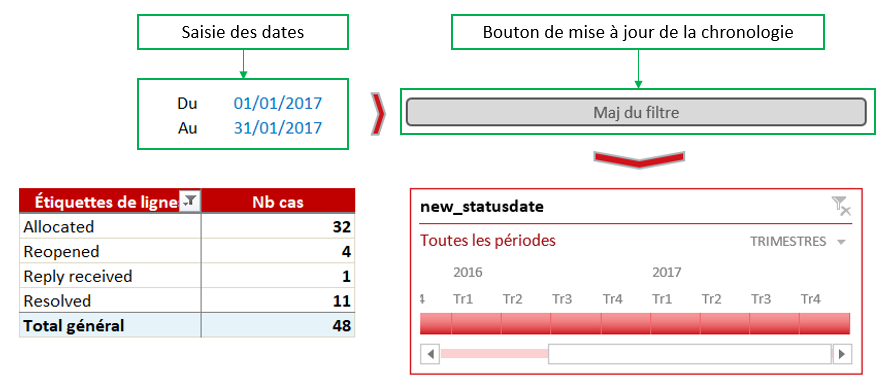
### Chronologies

*NB : Les chronologies présentées ci-dessous sont apparues avec Office 2013 alors que les utilisateurs sont actuellement sous Office 2010. Dans l'attente d'une migration sous Office 2016, l'utilisation des chronologies est donc impossible et peut être remplacée par l'utilisation des groupes mentionnés ci-dessus. Excel propose naturellement des groupes Années/Trimestres/Mois sur les champs Date. Les TCD initialement basés sur des chronologies sont donc livrés en double (chronologie et groupes).*

Lorsque l'on souhaite filtrer les données d'un TCD en fonction d'une date présente dans les données source, il est possible d'utiliser une "Chronologie" faisant apparaître une fenêtre telle que celle-ci :



Lorsque l'étendue du champ date choisi (ici new\_statusdate) est grande, la définition d'un filtre sur une période très courte peut être assez fastidieuse. Pour contourner ce problème, l'utilisateur a la possibilité de saisir les dates de filtre qu'il souhaite appliquer dans 2 cellules Excel avant de cliquer sur le bouton " Maj du filtre".



L'utilisateur a toujours la possibilité de modifier la frise manuellement en utilisant sa souris. Il peut réinitialiser la chronologie en cliquant sur le bouton  en haut à droite de la chronologie.

## Historisation des indicateurs

Comme indiqué ci-avant, les données servant au calcul des indicateurs sont historisées et conservées sur 13 mois glissants. Il n'est donc pas nécessaire de conserver une copie du fichier Excel pour accéder à des données historiques.

Cependant, si l'historisation des fichiers "en l'état" doit être faite, deux options existent :

* Enregistrer une copie du fichier sur un répertoire adapté ;
* Ou enregistrer une "version majeure" directement dans SharePoint.

# Divers

## Comparaison de fichiers de grande taille

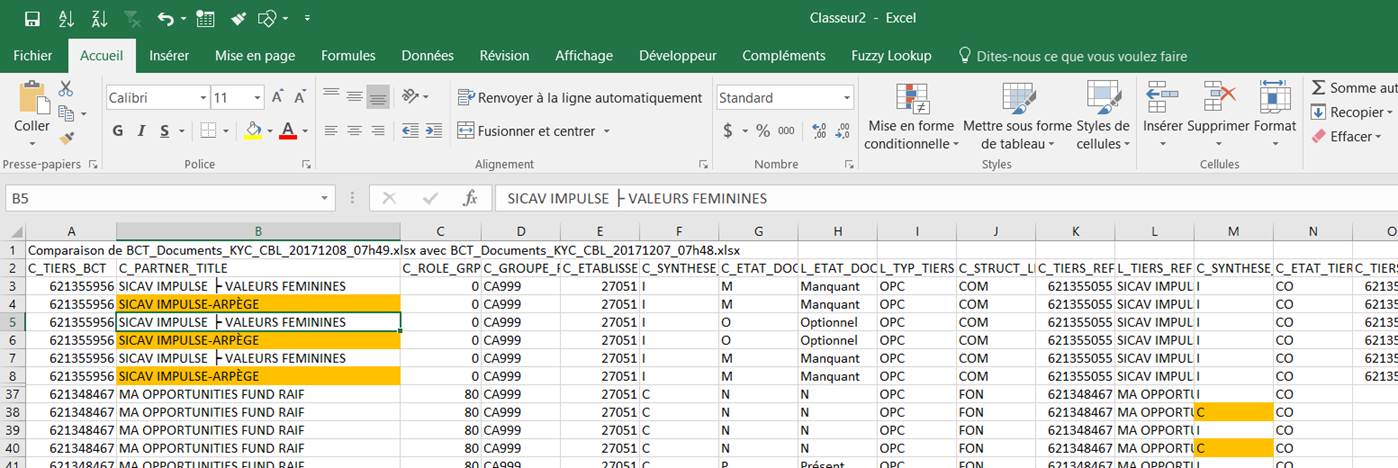
Un des besoins identifiés lors de l'atelier sur le "flux" de l'activité "Bank" consiste à récupérer deux extractions BCT, actuellement diffusées par des e-mails issus de l'équipe BMG (IP Paris), et à comparer les photos de la base à deux dates différentes.

*Extrait d'un tel fichier :*



Ces fichiers contiennent une 50aine de colonnes et plus de 80 000 lignes. Afin de remplacer des comparaisons par RECHERCHEV, un outil Excel est développé et fonctionne ainsi :

* Sélectionner les colonnes devant servir de clé d'identification dans la comparaison des deux fichiers (en indiquant le numéro de colonne du fichier source dans la colonne "Clé" du tableau de paramètres).
* Cliquer sur le bouton orange et choisir les deux fichiers à comparer (le plus ancien puis le plus récent).
* Attendre (environ 45 secondes avec des fichiers en local sur mon poste) et 4 fichiers Excel sont alors ouverts :
  + Celui du programme (CompareLargeDatasets.xlsm)
  + Les deux fichiers à comparer (ouverts en lecture seule)
* Le fichier résultat qui contient les noms des deux fichiers comparés, les entêtes de colonnes et un nombre pair de lignes de données :
  + - Les lignes numérotées impaires dans Excel viennent du premier fichier choisi (données les plus anciennes).
    - Les lignes numérotées paires dans Excel viennent du second fichier choisi (données les plus récentes).
    - Les données qui ont été modifiées d'un fichier à l'autre sont présentées sur fond jaune (Cf. ci-dessous).



*Programme de comparaison des fichiers :*



*Exemple de résultat produit :*



## Utilisation par d'autres entités du groupe CACEIS

Sous réserve d'une compréhension et d'une utilisation similaires des aspects KYC de la base BCT, la partie "Stock" des indicateurs développés peut être réutilisée par d'autres entités du groupe CACEIS.

La mise en œuvre de cette adaptation dépend essentiellement des choix techniques de l'informatique de proximité.

### Utilisation des bases de l'IP Luxembourg

Une première possibilité consiste à ajouter un paramètre "entité" dans les tables de l'IP :

* Nécessite de modifier tables de l'IP et les requêtes de chargement des données (SQL dans SSIS) ;
* Nécessite de modifier les requêtes de lecture des données (SQL dans Excel) pour ajouter un filtre sur l'entité ;
* Se traduit par l'existence d'un fichier Excel différent par entité ;
* N'offre pas toutes les garanties en matière de sécurité (un utilisateur pourrait modifier la requête de lecture pour récupérer des données d'une entité qu'il n'est pas supposé voir).

Une bonne gestion de la sécurité imposerait la mise en place d'une restriction des périmètres visibles user par user (user au sens de celui mentionné dans la chaîne de connexion) :

* Nécessite d'étudier puis de mettre en place des solutions techniques côté SQL Server ;
* Se traduit par l'existence d'un fichier Excel différent par entité ;
* Nécessite de modifier les chaînes de connexion des fichiers des nouvelles entités adopter le user adapté.

### Duplication dans les IP locales

Techniquement la plus simple, cette solution consiste à dupliquer les bases et requêtes mises en place au Luxembourg dans les différentes IP locales :

* Implique de modifier la table de paramétrage des entités ;
* Se traduit par l'existence d'un fichier Excel différent par entité ;
* Les chaînes de connexion du fichier Excel sont modifiées de manière à pointer sur les bases nouvellement créées avec le User adapté.

## Indicateurs dépriorisés

### Modifications réalisées par une autre entité sur des Tiers partagés

La production de cet indicateur a été dépriorisée car il n'a pas été trouvé de méthode permettant d'identifier précisément la personne à l'origine d'une modification sur l'un ou l'autre d'un sous-ensemble de champs sans utiliser la piste d'audit (à laquelle nous ne souhaitons pas recourir pour des raisons de performance).

### Cohérence des Tiers "communs" à l'activité TA et Bank

La production de cet indicateur peut être réalisée en quatre étapes :

* Identification de similitudes textuelles : sur base des noms et adresses des Tiers, calcul d'un "ratio de similitude" (NB : ne permettra pas d'identifier que NAM est similaire à Natixis Asset Management). Ce travail peut a priori être réalisé de façon efficace par le complément "Fuzzy Lookup Add-In For Excel" disponible sur le site de [Microsoft](https://blogs.msdn.microsoft.com/business_intelligence_labs/2011/04/27/fuzzy-lookup-add-in-for-excel/).
* Etude du résultat de l'étape 1 et saisie manuelle d'une table de mapping des tiers en doublon. Un "Tiers principal" est choisi. A cette occasion, la mise en place d'une gestion "humaine" des acronymes de type NAM évoqués ci-dessus est envisageable.
* Extraction SQL des données sous-jacentes au niveau de vigilance et à la matrice KYC.
* Pour chaque ensemble de doublons, comparaison des données du "Tiers principal" avec les données du doublon.