

RAPPORT DE STAGE DE FIN D'ETUDES

Master 2 **DATASCALE**

(Gestion des données et extraction de connaissance à large échelle)

Du 06 Septembre 2021 au 04 Mars 2022

STAGE DATA ANALYST ALM

Sujets : Reporting Grall (Accès à la donnée) | Tour de Contrôle OPAL Data Quality

Kossi Eloi Fabius DEFLY

Tuteur de Stage : **Sylvain OLIVA**

Responsable ALM Data-Management chez **Natixis**

Responsable Pédagogique : **Zoubida KEDAD**

Professeur, Maître de conférences à l'**UVSQ**

Université de Versailles Saint-Quentin en Yvelines

45 avenue des Etats Unis 78000 Versailles

Natixis

30 avenue Pierre Mendes 75013 Paris

Remerciements

A travers ces lignes j'exprime mes remerciements d'abord à toute l'équipe pédagogique du Master 2 DATASCALE (Gestion des données et extraction de connaissance à large échelle) de l'Université de Versailles Saint-Quentin en Yvelines du groupe Paris-Saclay pour la formation qu'elle m'a donnée et à l'établissement financier Natixis pour l'opportunité de stage qu'elle m'a offerte.

Mes remerciements vont ensuite aux personnes suivantes :

- A Mme. Zoubida KEDAD et M. Phillipe PUCHERAL, responsables du Master Gestion des données et extraction de connaissance à large échelle (DATASCALE).
- A M. Benoit GUILLLOT-LE GOFF, Manager du département FDI : Financial Data Integrity
- A mon tuteur de stage et Manager de l'équipe ALM Data-Management, M. Sylvain OLIVA qui s'est montré disponible tout au long de mon stage ainsi que pour son aide et sa confiance.
- A L'ensemble des personnes de l'équipe ALM Data-Management, du département ALM Analytics & Transverse et de la direction de la gestion financière de Natixis pour leur disponibilité et leur soutien au cours de ce stage.

Enfin j'adresse mes remerciements à toute ma famille : Mon père M. Kouami Dewonou DEFLY, Ma mère Mme. Yawovi Valentine AGOUZE, mes frères et ma sœur de m'avoir soutenu tout au long de mes études en particulier durant ce parcours de fin d'études et tous les amis.

Résumé

Le succès de tous type d'activité de gestion des risques, y compris la surveillance de la stabilité financière et la mise en œuvre de la politique des autorités, dépend de la qualité, de l'exhaustivité et de la disponibilité en temps voulu des données nécessaires à l'analyse.

La problématique de qualité de données prend une ampleur plus grande dans le contexte actuel de mondialisation des grands établissements bancaires.

Dans ce contexte d'augmentation des volumes de données et de leur hétérogénéité, les établissements bancaires doivent adopter une approche de "Master Data Management", s'appuyer sur des outils de contrôle automatisés et renforcer leurs filières de gouvernance de la donnée.

Une automatisation des contrôles de qualité sur les données permet de les exécuter plus fréquemment et de pouvoir anticiper pour agir (alerter, corriger) avant les périodes sensibles d'arrêt mensuel ou trimestriel.

L'équipe ALM Data-Management du département ALM Analytics & Transverse de la Direction Financière de Natixis est en charge de la mise en qualité des données ALM. C'est pour cela que j'ai été ravi d'intégrer cette équipe dans le cadre de mon stage de fin d'études dans le but de travailler sur des projets d'intégration, d'analyse et de mise en qualité des données ALM intitulé « Grall Reporting et Tour de contrôle OPAL Data Quality ».

Abstract

The success of any type of risk management activity, including financial stability monitoring and policy implementation, depends on the quality, completeness, and timely availability of the data needed for analysis.

The issue of data quality takes on greater importance in the current context of globalization of large banking institutions.

In this context of increasing data volumes and heterogeneity, banking institutions must adopt a "Master Data Management" approach, rely on automated control tools and strengthen their data governance processes.

Automating data quality controls allow them to be performed more frequently and to be able to anticipate and act (alert, correct) before the sensitive monthly or quarterly closing periods.

The ALM Data-Management team of the ALM Analytics & Transverse department of Natixis is in charge of ALM data quality. That's why I was delighted to join this team as part of my end-of-studies internship with the aim of working on ALM data access and quality projects entitled "Grall Reporting and OPAL Data Quality Control Tower".

Table des matières

Remerciements.....	2
Résumé.....	3
Abstract.....	4
Table des matières.....	5
Table des illustrations.....	7
I - Introduction.....	8
II - Natixis.....	9
II.1 – Présentation générale de Natixis.....	9
II.2 – Mission de Natixis.....	9
II.3 – Historique de Natixis.....	9
II.4 – Quelques chiffres de Natixis.....	10
II.5 – Département ALM Analytics & Transverse.....	11
II.6 - Equipe Data Management.....	11
II.7 - Observation.....	11
III – Contexte et Objectif du stage.....	12
IV – Approche.....	12
IV.1 – Acteurs du travail.....	12
IV.2 – Organisation du travail.....	12
V – Monter en compétence.....	14
V.1 – l'ALM (Gestion Actif-Passif).....	14
V.1.1 - Définition générale.....	14
V.1.2 – Objectif de l'ALM.....	14
V.1.3 – Risques traités par l'ALM.....	14
V.1.4 – Les indicateurs ALM.....	15
V.2 – Les produits financiers.....	15
V.2.1 – Définition.....	15
V.2.2 – Quelques produits financiers.....	15
V.3 – Notion de Deal et contrats RP.....	15
V.4 – Dbeaver.....	16
V.5 – Power Query.....	17
V.6 - OPAL.....	18
VI – Mission Reporting GRALL (Accès à la donnée).....	19
VI.1 – Contexte.....	19
VI.2 – Objectifs.....	19

VI.3 – Mes productions	19
VI.3.1 – Analyse de l'existant	19
VI.3.2 – L'architecture cible	19
VI.3.3 – Dictionnaire de données OPAL	20
VI.3.4 – Créations des vues OPAL.....	20
VI.4 – Préparation des données avec Power Query	21
VI.4.2 – Formatage et unification des tables de deals et contrat RP	22
VI.4.3 – Restitution des données OPAL.....	23
VII – Mission Data Quality Control Tower.....	24
VII.1 – Contexte.....	24
VII.2 – Objectifs	24
VII.3 – Analyse du besoin	24
VII.4 – Mes productions	25
VII.4.1 – Identifications des données essentielles	25
VII.4.2 – Contrôles et points d'alertes.....	25
VII.4.3 – Architecture de la tour de contrôle	26
VII.5 – Développement	26
VII.5.1 – Présentation de Power BI	26
VII.5.2 – Chargement des données dans Power BI	28
VII.5.	28
VII.5.4 – Les autres rapports de la tour de contrôle	29
VII.6 – Les formules DAX	30
VII.6.1 – Quelques formules DAX utiliser sur la Tour de Contrôle	30
VII.7 – Mise en production de la Tour de Contrôle.....	31
VIII – Organisation et évaluation.....	32
VIII.1 – Organisation de la mission.....	32
VIII.2 – Auto-évaluation et problèmes rencontrés	32
IX – Conclusion	33
X – Bibliographies.....	34

Table des illustrations

Figure 1 - Organigramme Natixis	9
Figure 2 - Chiffres Collaborateurs Natixis	10
Figure 3 - Résultats Natixis.....	10
Figure 4 - Interface Dbeaver	16
Figure 5 - Interface de préparation des données Power Query	17
Figure 6 - Architecture d'OPAL.....	18
Figure 7 - Tableau récapitulatif des Tables dans la base OPAL.....	18
Figure 8 - Architecture Existante.....	19
Figure 9 - Architecture cible de l'accès à la données	20
Figure 10 - Requête SQL pour créer une vue	20
Figure 11 - Requête SQL pour créer une vue	21
Figure 12 - Script de récupération dynamique des dates	21
Figure 13 - Requête d'extraction dynamique	22
Figure 14 - Union des tables de Deals.....	22
Figure 15 - Processus ETL de chargement des données dans le rapport.....	23
Figure 16 - Interface principale de Power BI.....	27
Figure 17 - La page principale de la tour de contrôle	28
Figure 18 - Architecture finale de la tour de contrôle	29
Figure 19 - Tour de contrôle déployer sous forme d'application web	31
Figure 20 - Mise en place d'un Gateway d'actualisation automatique de la tour de contrôle	31
Figure 20 - Mise en place d'un Gateway d'actualisation automatique de la tour de contrôle	31

I - Introduction

Dans le cadre de la dernière année de mon cycle de Master intitulé Master DATASCALE : Gestion des données et extraction de connaissance à large échelle de l'Université de Versailles Saint-Quentin en Yvelines (Université Paris Saclay), j'ai eu l'opportunité d'effectuer mon stage de fin d'étude au sein de la société Natixis (établissement financier français) et plus précisément dans l'équipe ALM Data-Management du département le département ALM Analytics & Transverse.

Cette équipe est chargée de l'analyse des données et de la production des rapports permettant le pilotage de la liquidité, et des reportings règlementaires.

Durant ma période d'apprentissage chez Natixis, j'ai été encadré par mon tuteur M. Sylvain OLIVA et j'ai eu l'occasion d'intervenir sur deux projets tous axés sur la Data. Le premier intitulé Reporting GRALL a été de développer des restitutions des données d'OPAL (Bases de données de marchés financiers) afin d'alimenter l'outil d'analyse de données Power BI.

Le deuxième projet qui est le sujet principal du stage consistait à créer une tour de contrôle Data-Quality sur les données de la base de données OPAL (Bases de données de marchés financiers). L'objectif de la tour de contrôle est de permettre au data sourcing de qualifier la qualité des données quotidiennement en se basant sur des KPI permettant de couvrir le périmètre de données consommées par la base de données OPAL sur différents axes d'analyses pertinentes.

Dans ce rapport je présenterais tout d'abord la société Natixis, puis les deux projets réalisés en mettant un focus particulier sur le projet Tour de contrôle OPAL - Data Quality. Et pour finir la conclusion de ce stage qui a été une suite logique à ma formation et un excellent tremplin afin de me permettre de lier mes apprentissages académiques et les pratiques en milieu professionnel.

II - Natixis

II.1 – Présentation générale de Natixis

Natixis est un établissement financier français de dimension internationale spécialisé dans la gestion d'actifs et de fortune, la banque de financement et d'investissement, l'assurance et les paiements. Filiale du groupe BPCE qui est issue elle-même d'une fusion de la Caisse nationale des caisses d'épargne et de la Banque fédérale des banques populaires.

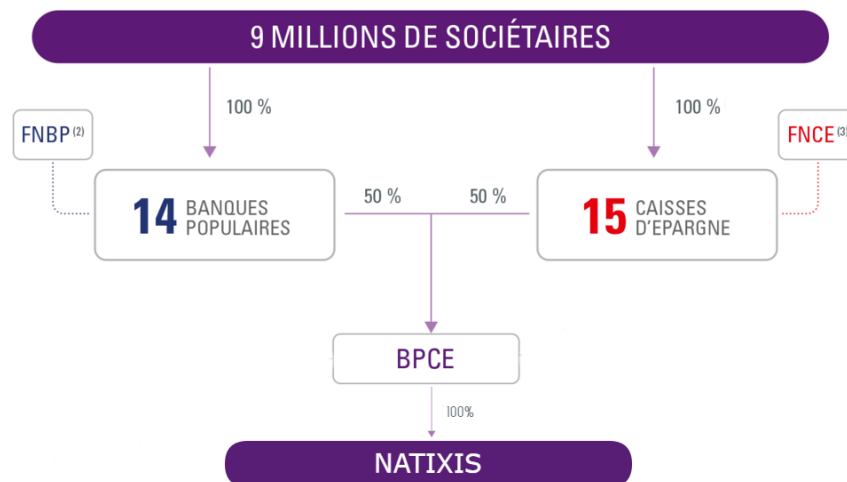


Figure 1 - Organigramme Natixis

II.2 – Mission de Natixis

Natixis a pour mission de concevoir des solutions financières sur mesure pour accompagner ses clients et leur permettre de réaliser leurs projets. L'objectif du groupe est de développer son expertise :

- Dans la gestion d'actifs et de fortune.
- Dans l'investissement, le financement.
- Dans les assurances pour proposer des solutions toujours plus innovantes.

II.3 – Historique de Natixis

Natexis SA a été créée en 1996 suite à la fusion de la Banque Française du Commerce Extérieur (BFCE) et du Crédit national.

En 1998, Natexis SA est rachetée par la Banque populaire. Cette dernière rattache à Natexis SA les activités opérationnelles de la Caisse Centrale des Banques Populaires (CCBP). En 1999, Natexis SA devient Natexis Banques Populaires. Parallèlement, La CCBP devient la Banque Fédérale des Banques Populaires (BFBP).

En 2004, la Caisse d'Épargne acquiert CDC IXIS et IXIS CIB et IXIS AM sont créés.

En 2006, la Caisse d'Épargne et la Banque Populaire annoncent le projet de rapprochement entre Natexis Banque et Ixis CIB. La nouvelle entité créée est dénommée Natixis. La Caisse d'Épargne et la Banque Populaire détenaient alors chacune 35% des actions du groupe, le reste du capital est en bourse. En 2009, le groupe BPCE est constitué suite à la fusion de la CNCE et de la BFBP. Le 4 juin 2021, BPCE, actionnaire majoritaire, lance une OPA pour racheter le capital de Natixis coté en bourse, atteindre plus de 90 % et achever par un retrait du titre.

II.4 – Quelques chiffres de Natixis

Aujourd'hui, Natixis compte plus de 16 000 collaborateurs et présent dans plus de 36 pays répartis sur cinq grandes zones géographiques : les Amériques, l'Asie-Pacifique, l'Europe et le Moyen-Orient avec un chiffre d'affaires de 9 616 milliards d'euros en 2018.

Natixis dispose de ressources globales et d'une expérience permettant de mobiliser des équipes à fortes compétences pour répondre aux défis de ses clients.



Figure 2 - Chiffres Collaborateurs Natixis

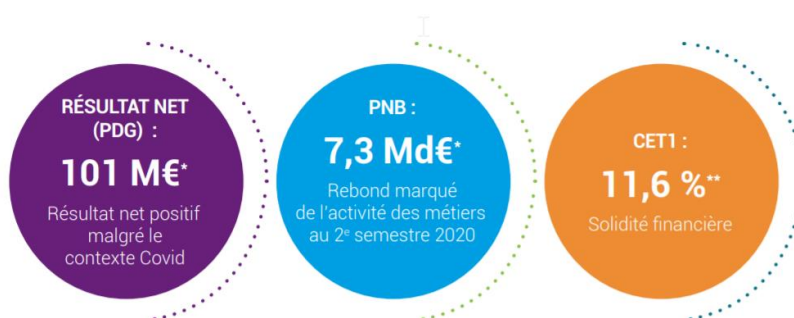


Figure 3 - Résultats Natixis

II.5 – Département ALM Analytics & Transverse

Le département ALM Analytics & Transverse au sein de la direction de la gestion financière est en charge de l'analyse et de la production des rapports permettant le pilotage de la liquidité, et de certains reportings réglementaires ainsi que du pilotage du système d'information de la direction. Il est composé de 3 équipes :

1. ALM Analytics : conduit des analyses hebdomadaires, mensuelles et trimestrielles et produit les indicateurs que ce soit le gap de liquidité, le gap de taux ou la marge nette d'intérêt ainsi que certains ratios associés.
2. Transverse : assure le pilotage du SI et le support fonctionnel.
3. Data management : est en charge de l'administration du SI, de la qualité de données et de la préparation des rapports.

II.6 - Equipe Data Management

L'équipe ALM Data Management est chargée de la gestion et de la mise en qualité des données qui sont destinées à la production des indicateurs ALM : GAP, LCR, MNI, NSFR, elle est aussi en charge de l'administration du SI et de la préparation des rapports du pilotage de la liquidité de la banque.

II.7 - Observation

L'accueil des nouveaux arrivants chez Natixis se fait durant un Onboarding de présentation de l'entreprise suivi d'une présentation des locaux, différents services et des collaborateurs sur le plateau. Chez Natixis les collaborateurs travaillent en « Open Space » ce qui a pour but de favoriser la collaboration, le partage et l'échange d'idées. De plus on se tutoie tous mutuellement, ce qui met tout le monde sur le même piédestal.

On peut ainsi remarquer une intégration rapide des nouveaux arrivants grâce à la disponibilité des collaborateurs présents sur les différents plateaux.

Une des observations qui selon moi est très appréciée et encourageante est le sourire qu'on retrouve sur le visage de tous même avec les masques durant la journée et en particulier la mise en place d'un open space virtuel qui est entièrement une initiative de mon tuteur Mr. Sylvain OLIVA pour permettre aux membres de l'équipe, surtout ceux qui sont en télétravail de partager un espace d'échange convivial avec les collaborateurs qui sont sur le site afin de ne pas se sentir isolé.

Tout ceci a facilité mon intégration au quotidien et aujourd'hui je suis fier de travailler dans un cadre comme celui-ci où les relations humaines sont un facteur mis au premier rang.

III – Contexte et Objectif du stage

Afin d'accompagner leur projet de transformation qui consiste en la migration des outils d'analyse et processus interne vers d'autres technologies plus performantes et la mise en place d'une tour de contrôles « data quality » sur les données ALM de la banque, Natixis a émis le besoin de recruter un stagiaire. Au sein du département ALM Analytics & Transverse et plus précisément de l'équipe ALM Data-Management de la direction de la gestion financière, la mission du stagiaire serait d'accompagner la réalisation de certaines missions sous la responsabilité de l'expert Data et responsable de l'équipe ALM Data-Management M. Sylvain OLIVA.

IV – Approche

IV.1 – Acteurs du travail

Comme acteurs sur ce projet on avait une équipe composée de développeurs en la personne de Diaby DOUCOURE (Data Analyst ALM) et moi Kossi Eloi Fabius DEFLY (stagiaire Data Engineer/Analyst), Heila TEBIB (Data Analyst ALM), Nabil EL KAISOUNI (Data Analyst ALM) pour les tests des solutions et nous étions encadrés par M. Sylvain OLIVA Manager de l'équipe ALM Data-Management.

IV.2 – Organisation du travail

Au sein de l'équipe ALM Data-Management à laquelle j'ai été intégrée la méthodologie de travail adoptée est l'approche agile. Au sein de cette équipe est utilisé le Framework scrum.

Le Framework scrum consiste à traiter des problèmes complexes, en les subdivisant en des tâches individuelles, en priorisant ces tâches et en attribuant chacune d'entre elles à la personne la plus compétente. Scrum offre un cadre de gestion de projet. Un projet est constitué de plusieurs demandes et exigences par les utilisateurs. Pour gérer une telle demande de manière agile et efficace sans se laisser noyer, la méthode agile scrum s'est révélée la meilleure des solutions possibles. Les rôles principaux de la méthode sont :

- Le Product Owner : Il représente le client. Il définit les spécifications fonctionnelles et établit la liste des priorités de ce qu'il faut développer. Il valide également les fonctionnalités. Dans ce projet ce rôle est assuré par les membres de l'équipe ALM Data-Management.
- Le Tech Lead : C'est la personne qui a une bonne vision technique, qui maîtrise le sujet et qui guide l'équipe. Il prend le rôle de responsable de la qualité du rendu. Ce rôle est occupé par mon tuteur de stage M. Sylvain OLIVA.
- Le Scrum Master : Il est garant du respect des processus scrum. Il assure la bonne communication entre les membres de l'équipe. C'est la personne qui s'assure que tout se passe bien et est chargé des différentes tâches administratives. Cette partie est gérée par Heila TEBIB.
- L'équipe : Composée de développeurs, testeurs, analystes et autres techniciens liés au besoin du projet. Ils mettent en œuvre les solutions techniques et réalisent les développements. C'est le cœur de l'équipe. Cette partie est assurée par Diaby DOUCOURE et moi-même.

Processus de la méthodologie scrum :

- La User Story : C'est la définition de l'expérience utilisateur dans un langage, un vocabulaire et une terminologie du client. Chacun d'elle contient, un identifiant, un nom (description du produit), l'importance (priorité représentée par des numéros), l'estimation du travail nécessaire, une Démo et les informations nécessaires à la réalisation de la story.
- Product Back log : représente le carnet de commande pour le produit. Il va constamment évoluer pour refléter les nouveaux besoins.

Une fois l'User Story et le Product Back log validé, commence la réalisation du projet qui est subdivisé en plusieurs itérations :

- Sprint : Correspond à la réalisation d'une itération. Il commence par une réunion de planification (sprint planning meeting). Lors de cette séance, il on décide des éléments prioritaires du Product Back log qui seront développés. Chaque sprint dure de 2 à 4 semaines. Il contiendra du développement, des contrôles de qualité et une livraison.
- Daily Scrum : C'est une réunion quotidienne de 20 à 30 minutes. Le but n'est pas de résoudre les problèmes, mais de faire remonter toutes les informations d'avancement du sprint en cours et d'en valider la qualité. Elle permet également de vérifier le respect des délais. Lors de cette réunion chaque membre à la parole et doit être capable de dire ce qu'il a fait la veille, ce qu'il va faire pendant la journée et les problèmes qui le bloquent.
- Burn Down Chart : C'est un graphique qui décrit l'avancement du projet tenu par le Scrum master. Il est mis à jour à la fin de chaque Daily.
- Sprint Meeting Review : C'est une présentation de la solution au client sous forme de démonstration et la réception des retours de ce dernier. Les améliorations suggérées et les problèmes rencontrés seront écrits dans le Product Back Log.
- Review : Généralement lieu après chaque revue de sprint elle permet à l'équipe de réfléchir sur ce qui a marché ou pas au cours du sprint et parler des correctifs et améliorations.
- Sprint Back log : C'est l'ensemble des livraisons des sprints cumulés.

V – Monter en compétence

Dès le début de mon stage, pour mieux entamer ma mission dans l'équipe ALM Data-Management j'ai été amené à suivre des formations afin de monter en compétence sur des notions de la finance de marché ainsi que sur certains outils et technologies.

V.1 – l'ALM (Gestion Actif-Passif)

V.1.1 - Définition générale

L'ALM est l'abréviation de l'expression anglo-saxonne **Asset and Liability Management**, qui se réfère à la notion de gestion actif-passif en milieu bancaire.

L'ALM est une démarche globale au sein de la banque visant à optimiser la rentabilité des fonds propres tout en préservant un niveau acceptable des risques financiers (risque de taux, de change et de liquidités) et en assurant une allocation des fonds propres de manière à adapter le volume de l'activité à l'évolution du marché et à l'environnement financier et réglementaire, notamment aux ratios prudentiels. On parle alors d'optimisation du couple risque/rentabilité.

V.1.2 – Objectif de l'ALM

La gestion ALM consiste à veiller aux équilibres des ressources (dépôts) et des emplois (crédits) selon la devise, le type de taux et les échéances (maturités) pour la stabilité du PNB (produit net bancaire) afin d'obtenir des marges positives stables.

Elle assure, également, la pérennité de la banque via une gestion contrôlée des risques (GRM : Global Risk management) conformément à la réglementation afin d'éviter une prise de risque excessive menant à la faillite (optimisation du couple rentabilité-risque).

Enfin, elle consiste aussi à prévoir et anticiper les évolutions futures, prendre des décisions stratégiques (stratégie financière).

V.1.3 – Risques traités par l'ALM

1. **Le risque de liquidités** : se matérialise par l'incapacité d'investir ou de respecter ses engagements. Pour une banque, ils correspondent donc soit à une perte d'opportunité, soit à l'occurrence d'une crise de liquidités lorsque celle-ci n'est plus en mesure d'honorer les sorties d'argent (retraits sur dépôts à vue, déblocage de nouveau prêt, remboursement d'emprunt) en dépit de ses liquidités : rentrées d'argent (remboursement de prêt, émission de nouveaux emprunts) et disponibilités (compte à vue, actifs liquides, fonds propres).
2. **Le risque de taux** : traduit l'incertitude sur les résultats financiers induits par l'exposition aux variations de la courbe des taux. Ce risque est analysé en intérêts courus (impact sur la marge d'intérêt) ainsi qu'en valeur de marché (pour les éléments comptabilisés en juste valeur, et les indicateurs réglementaires en valeur actuelle nette).
3. **Le risque de change** : il est lié aux activités en devises, qui se reflètent par les positions au bilan (prêteuse ou emprunteuse nette contre la devise de référence), et le niveau de couverture retenue.

V.1.4 – Les indicateurs ALM

LCR : (ratio de liquidité à court terme) est un indicateur qui permet d'assurer que la banque dispose d'un niveau adéquat d'actifs liquides de haute qualité non grevés pouvant être convertis en liquidité pour couvrir ses besoins sur une période de 30 jours calendaires en cas de graves difficultés de financement.

MNI : La marge nette d'intérêt est un indicateur de la rentabilité de la banque. Il fait référence à la différence entre les intérêts reçus et les intérêts payés.

NSFR : (Net Stable Funding Ratio) est un indicateur qui permet d'assurer que la banque dispose suffisamment de ressources stables (c'est-à-dire des ressources de maturité initiale supérieure à un an).

V.2 – Les produits financiers

V.2.1 – Définition

Un produit financier est un placement qui génère des intérêts. Il peut ainsi s'agir d'un placement ou d'un investissement, qu'une banque effectuera avec sa trésorerie lors de l'exercice comptable. Il existe une variété de produits financiers, comme des obligations, actions, titres de créance.

V.2.2 – Quelques produits financiers

Parmi tous les produits financiers qui existent on peut citer :

- Les Titres : Un titre représente une part d'une créance à court ou moyen ou long terme, ou une part du capital d'une entreprise. Pour l'émetteur du titre il s'agit d'un instrument de financement, pour l'acheteur c'est un instrument de placement.
- Les Produits Dérivés : regroupent toutes les opérations qu'on appelle également "hors bilan" puisqu'elles ne font pas l'objet d'un enregistrement au bilan de l'établissement financier.
- Les Actions : Une action représente une part du capital de la société émettrice. A ce titre, elle confère des droits. La valeur des actions varie en fonction de nombreux facteurs.

V.3 – Notion de Deal et contrats RP

Un deal est une opération, un événement contractuel d'achat ou de vente pour échanger un actif contre un paiement. Le marché financier permet de faire des deals (opérations) de produits financiers.

Un contrat RP est un deal transformé et préparé pour le calculateur ALM en d'autres termes les Contrats RP sont dérivés des deals.

V.4 – Dbeaver

DBeaver est un logiciel permettant de gérer plusieurs bases de données différentes sur une seule et même application, et donc de manipuler des données de différentes sources. Ci-dessous je vous présente l'interface principale de Dbeaver.

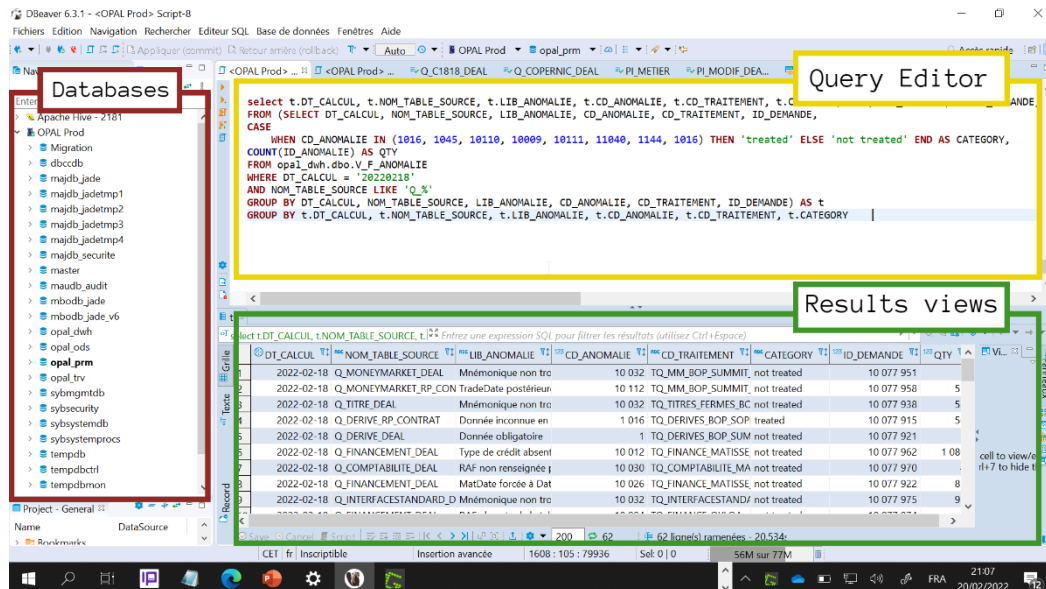


Figure 4 - Interface Dbeaver

A gauche on a la partie Data bases où sont présentes les différentes connexions et bases de données. Au centre on a la partie Query editor où sont éditées les requêtes SQL et en bas la partie results views où on peut voir les résultats des requêtes exécutées.

Dbeaver permet de faire l'administration et le requêtage des bases de données. Il fournit un éditeur qui prend en charge la complétion de code et la coloration syntaxique, ce qui a pour but d'accélérer la productivité du développeur.

V.5 – Power Query

Power Query est un moteur de transformation des données et de préparation des données. Power Query est fourni avec une interface graphique pour obtenir les données à partir de sources et un éditeur de Power Query pour appliquer les transformations.

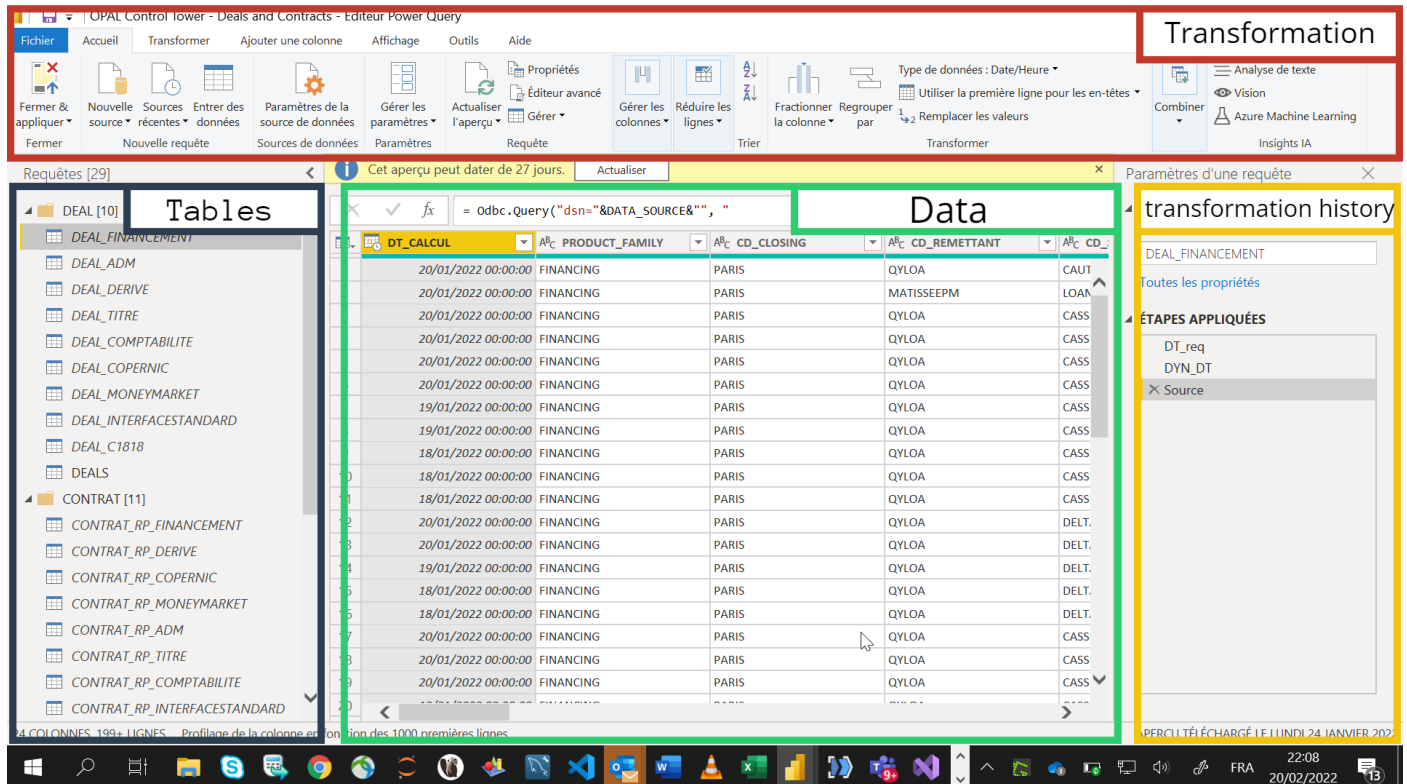


Figure 5 - Interface de préparation des données Power Query

L'interface de Power Query est composée de plusieurs parties avec des rôles différents, on a :

- La partie Tables : où on peut retrouver les tables qui ont été requêtées.
- La partie transformation : nous permet de faire la préparation des données, elle nous permet notamment de supprimer ou renommer certaines colonnes, filtrer certaines valeurs, faire du formatage, faire des fusions de plusieurs requêtes.
- La partie Data visualisation où on peut avoir un aperçu sur les données et visualiser l'effet des transformations sur les données.
- La partie Transformation History : où on retrouve l'historique des transformations qui ont été appliquées sur les données, cette partie nous permet de modifier ou de supprimer une étape de transformation.

V.6 - OPAL

OPAL est un serveur de bases données SAP Sybase qui abrite les données financières de la banque. Les données contenues dans les bases de données d'OPAL sont de types divers et variées on peut donc avoir des données sur les opérations (Transactions, Acquisitions, Comptabilité...), des taux de change, les anomalies.

Les données d'OPAL proviennent de plusieurs sources appelées systèmes remettant, un remettant une source de données qui peut se présenter sous plusieurs formes : application de marché, bases de données, etc.

Parmi les systèmes remettant d'OPAL on peut citer : Sophis, Summit, Qyloa, Loan IQ, Murex, Matissegl pour n'en citer que ceux-là.

V.6.1 – Architecture d'OPAL

OPAL est composée de plusieurs bases de données mais nous allons nous intéresser aux bases suivantes : OPAL_dwh, OPAL_prm. Ces dernières contiennent à leur tour plusieurs tables subdivisées par périmètres de données, par système remettant, par famille de produits financiers et par type de traitement (Daily, Monthly) :

- OPAL_dwh : rassemble les tables de traitements et d'opérations (les tables de deals, contrats RP et les flux) par famille de produit et aussi, des tables d'anomalies et des tables référentielles.
- OPAL_prm : rassemble les tables de paramètres fonctionnelles.

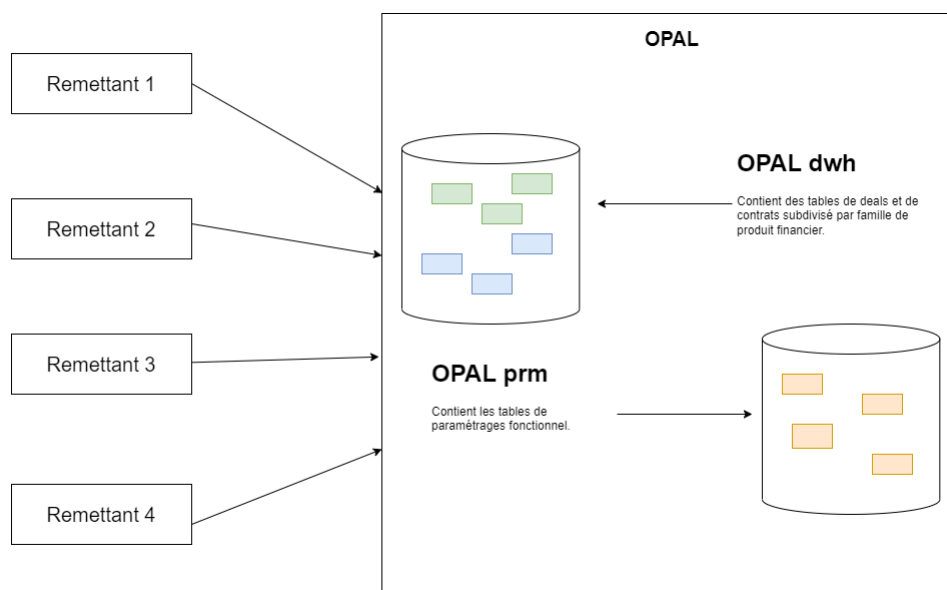


Figure 6 - Architecture d'OPAL

	X = M or Q (Monthly or daily)			
Product family	DEAL Table	Cashflow Table	Contract Table	Flow Table
Financement	X_FINANCEMENT_DEAL	X_FINANCEMENT_CASHFLOW	X_FINANCEMENT_RP_CONTRAT	X_FINANCEMENT_RP_FLUX
Titre	X_TITRE_DEAL	X_TITRE_CASHFLOW	X_TITRE_RP_CONTRAT	X_TITRE_RP_FLUX
Copernic	X_COPERNIC_DEAL	no cashflow table	X_COPERNIC_RP_CONTRAT	no cashflow table
ADM	X_ADM_DEAL	X_ADM_CASHFLOW	X_ADM_RP_CONTRAT	X_ADM_RP_FLUX
Dérivé	X_DERIVE_DEAL	X_DERIVE_CASHFLOW	X_DERIVE_RP_CONTRAT	X_DERIVE_RP_FLUX
AjustGap	No T1	No T1	X_AJUSTEMENT_RP_CONTRAT	X_AJUSTEMENT_RP_FLUX
Money Market	X_MONEYMARKET_DEAL	X_MONEYMARKET_CASHFLOW	X_MONEYMARKET_RP_CONTRAT	X_MONEYMARKET_RP_FLUX
Comptabilité	X_COMPTABILITE_DEAL	no cashflow table	X_COMPTABILITE_RP_CONTRAT	no cashflow table
Interface Standard	X_INTERFACESTANDARD_DEAL	X_INTERFACESTANDARD_CASHFLOW	X_INTERFACESTANDARD_RP_CONTRAT	X_INTERFACESTANDARD_RP_FLUX

Figure 7 - Tableau récapitulatif des Tables dans la base OPAL

VI – Mission Reporting GRALL (Accès à la donnée)

VI.1 – Contexte

Dans le cadre de la migration de la solution SAP Business Objects vers Microsoft Power BI, démarré en septembre 2021, l'équipe ALM Data-Management qui est en charge de la base de données OPAL a choisi de restituer les données dans l'outil de reporting Power BI et via une couche d'extraction et de préparation de données assurée par Power Query.

La stratégie adoptée par l'équipe ALM Data-Management pour migrer intégralement les données OPAL vers cette nouvelle solution, consiste à créer des vues sur les tables des bases de données OPAL afin de limiter les risques de crash liés à la forte volumétrie des bases de données OPAL.

VI.2 – Objectifs

La mission Reporting Grall a pour objectifs de rendre accessible les données des bases de données OPAL via l'outil de reporting Power BI, de rendre facile le requêtage des données grâce à l'unification et l'homogénéisation des tables et du décommissionnement de la solution SAP Business Object au profit de Power BI.

VI.3 – Mes productions

VI.3.1 – Analyse de l'existant

Dans l'architecture existante illustrée par la figure ci-dessous, on a un serveur de bases de données SAP Sybase qui contient plusieurs bases de données et tables et sur lequel est connecté un serveur d'applications hébergeant le logiciel d'informatique décisionnel SAP Business Objects.

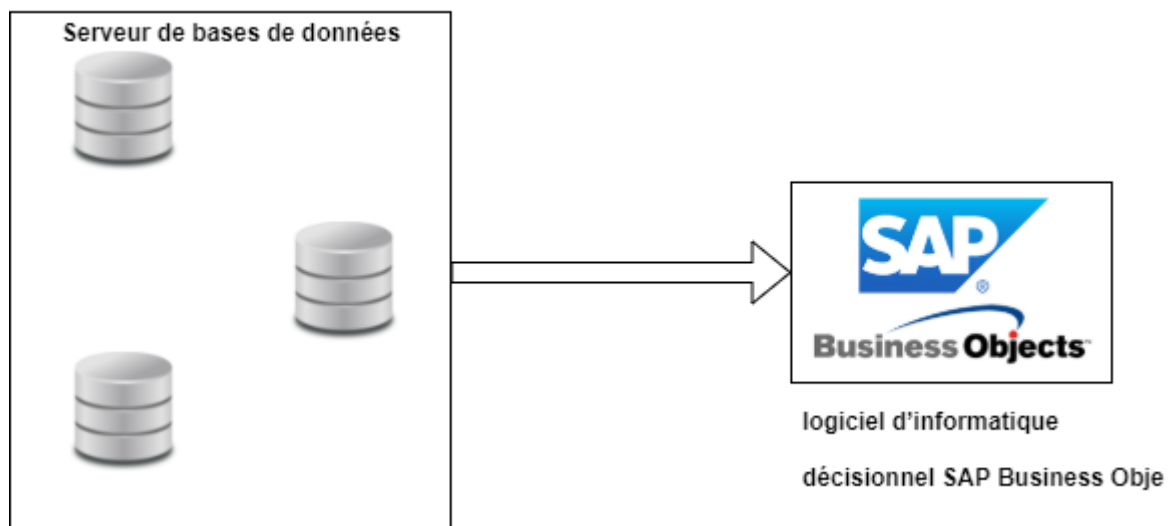


Figure 8 - Architecture Existante

VI.3.2 – L'architecture cible

Dans l'architecture cible illustrée sur l'image ci-dessous, on voit l'implémentation d'une couche de préparation des données entre les bases de données et Power BI par l'outil Power Query. Cette architecture a pour but de mettre au meilleur format pour faciliter l'exploration et l'analyse des données aux équipes métier.

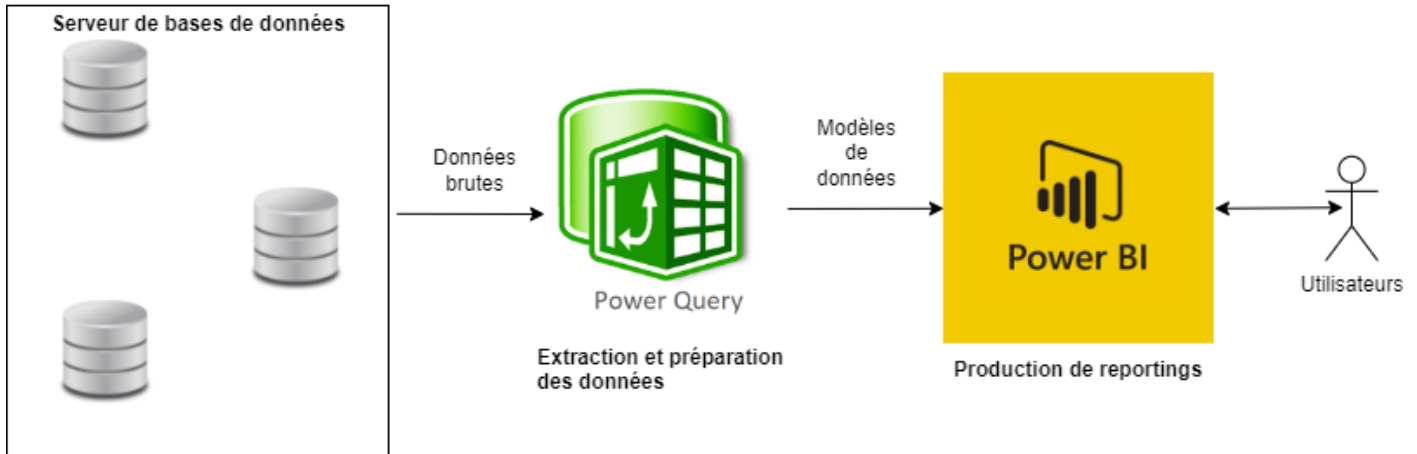


Figure 9 - Architecture cible de l'accès à la donnée

VI.3.3 – Dictionnaire de données OPAL

Après l'analyse de l'existant et compréhension du besoin a commencé la première phase de la mission qui consiste à réaliser un dictionnaire des données de la base de données OPAL. L'objectif de ce dictionnaire de données est de faire un inventaire des actifs immatériels que sont les données dans le but de permettre aux utilisateurs de découvrir et d'explorer tous les jeux de données disponibles, d'améliorer leur compréhension de ces données, de faciliter la collaboration avec les autres utilisateurs afin d'enrichir la qualité de ces actifs, et de créer plus de valeur à partir de ces données.

VI.3.4 – Créations des vues OPAL

Par mesure de sécurité et de restriction de certains utilisateurs, nous avons décidé de créer des vues sur les tables afin de masquer certains champs à certains utilisateurs. J'ai donc après la création du dictionnaire des données identifié et créer les vues sur les tables de la base de données OPAL.

Sur les illustrations suivantes quelques scripts SQL des vues créées :

```

/* Adaptive Server has expanded all '*' elements in the following statement */
create view V_Q_COPERNIC_DEAL
AS SELECT
  opal_dwh..Q_COPERNIC_DEAL.ID_SUIVI_TRAITEMENT, opal_dwh..Q_COPERNIC_DEAL.ID_Q_COPERNIC_DEAL,
  opal_dwh..Q_COPERNIC_DEAL.DT_CALCUL, opal_dwh..Q_COPERNIC_DEAL.DT_CRISTALLISATION,
  opal_dwh..Q_COPERNIC_DEAL.DH_FICHER, opal_dwh..Q_COPERNIC_DEAL.CD_FAMILLE_PRODUIT,
  opal_dwh..Q_COPERNIC_DEAL.CD_CLOSING, opal_dwh..Q_COPERNIC_DEAL.ID_A_COPERNIC_POST_ACQ,
  opal_dwh..Q_COPERNIC_DEAL.TYPE_BILAN, opal_RTICIPATION, opal_dwh..Q_COPERNIC_DEAL.CD_DEVISE_PARTICIPATION,
  opal_dwh..Q_COPERNIC_DEAL.MNT_DEVISE_PARTICIPATION, opal_dwh..Q_COPERNIC_DEAL.MNT_INIT_DEVISE_PARTICIPATION,
  opal_dwh..Q_COPERNIC_DEAL.SCHED_VALEUR_DEFAUT, opal_dwh..Q_COPERNIC_DEAL.CD_TIERS_PARAMETRAGE,
  opal_dwh..Q_COPERNIC_DEAL.CD_DEVISE_RISKPRO_PARTICIP,
  opal_dwh..Q_COPERNIC_DEAL.TOP_DELTA, opal_dwh..Q_COPERNIC_DEAL.VERSION_DEAL, opal_dwh..Q_COPERNIC_DEAL.DH_MAJ,
  opal_dwh..Q_COPERNIC_DEAL.CD_REMETTANT, opal_dwh..Q_COPERNIC_DEAL.CD_SOUS_REMETTANT
FROM opal_dwh..Q_COPERNIC_DEAL
  
```

Figure 10 - Requête SQL pour créer une vue

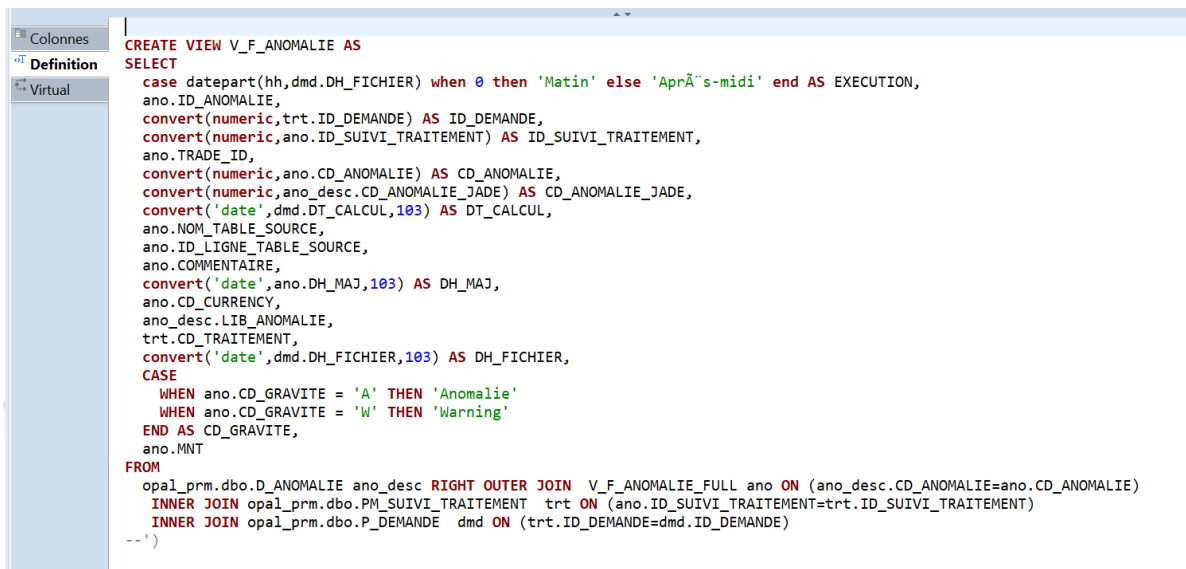


Figure 11 - Requête SQL pour créer une vue

VI.4 – Préparation des données avec Power Query

VI.4.1 – Requête des données

Après la phase création des vues on a mis en place une connexion sur le serveur et les bases de données OPAL via le module ODBC.

OPAL étant un serveur de bases de données contenant des bases de données qui reçoivent une volumétrie importante de données par jour (environ 3 000 000 de lignes) les risques de crash qui peuvent être liées à une mauvaise requête des données sont importantes. Nous avons donc décidé de mettre à disposition un accès limiter aux données en créant une connexion en mode export et en mettant à disposition les données des trois dernières dates de calcul c'est-à-dire les données du J-1, du J-2 et du J-3.

Les données sont rechargées dynamiquement par rapport à ces trois dernières dates de calcul grâce à un script de paramétrage des requêtes, écrit en langage M et qui permet la récupération et le formatage automatique des trois dernières dates de calcul depuis les données, ce qui permet entre autres d'éviter le requête sur une date fériée ou un jour de weekend où on n'a pas de données disponibles. Ci-dessous les unes image du script M et une requête paramétré par ce script.

DYNAMIC_DATES

Options d'affichage ▾

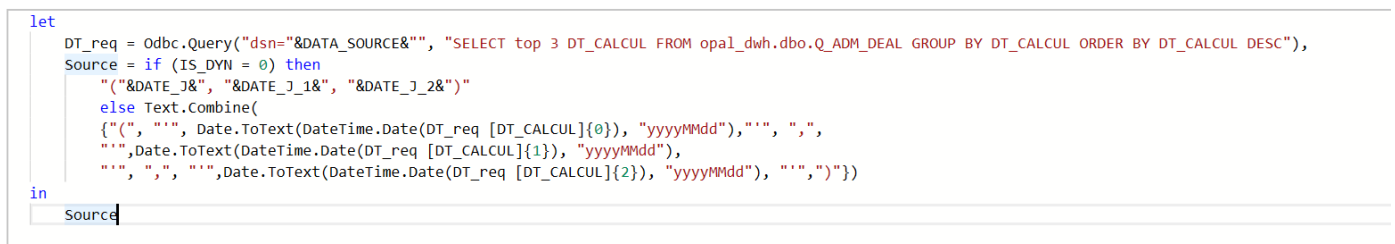


Figure 12 - Script de récupération dynamique des dates



Figure 13 - Requête d'extraction dynamique

VI.4.2 – Formatage et unification des tables de deals et contrat RP

Après la phase d'extraction des données, j'ai ensuite réalisé les tâches de formatage suivantes sur les données :

- Choix des champs
- Renommage de certains champs
- Conversion des champs « datetime » en date
- Conversion des valeurs des champs montants en euros

Le formatage et l'homogénéisation des tables étaient la phase préparatoire de l'unification des tables de Deals et Contrats RP. L'unification des tables a pour objectif de faciliter l'accès aux données d'une manière à pouvoir accéder à tous deals et contrat RP peu importe la famille de produit dans une seule table. L'unification des données a été faite par le module « fusion » de Power Query. L'illustration ci-dessous montre un aperçu de la fusion des tables deal.

On peut observer dans la case de droite (dans la partie Tables à ajouter) la sélection de toutes les tables de Deal à fusionner.

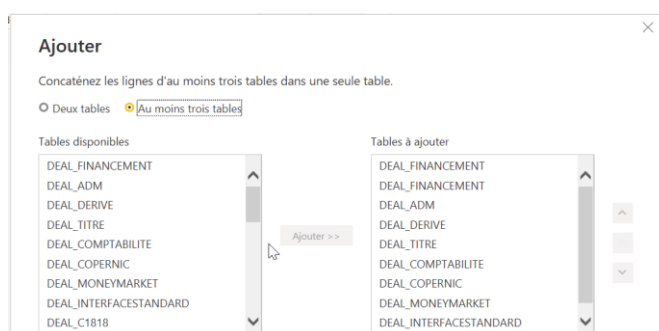


Figure 14 - Union des tables de Deals

VI.4.3 – Restitution des données OPAL

Après toutes les phases précédentes, vient la phase de chargement des données dans Power BI qui est décomposée en plusieurs étapes :

- Extraction : durant cette étape, Power BI se connecte aux sources de données pour collecter les données.
- Transformation : au cours de cette étape s'applique les différentes transformations et formatages définis au préalable dans Power Query.
- Chargement : représente l'étape finale où les données transformées sont chargées dans Power BI et prêtes à l'utilisation.

La figure ci-dessous donne une illustration de tout ce processus.

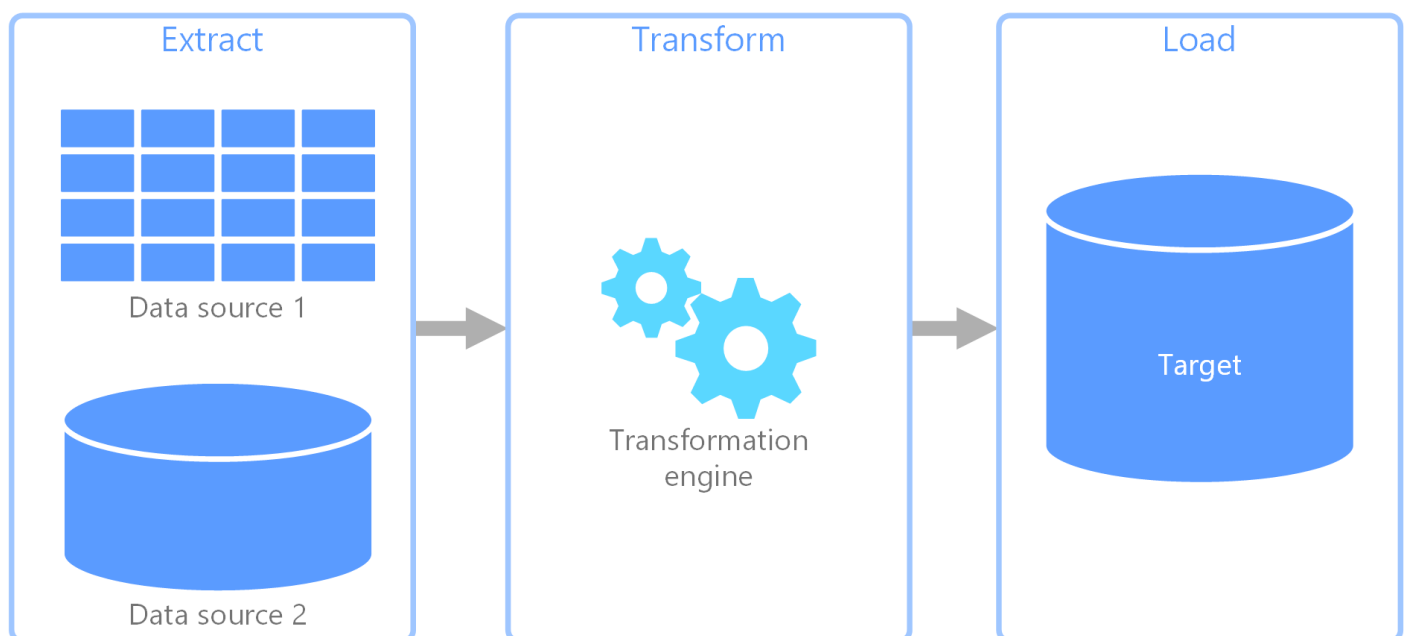


Figure 15 - Processus ETL de chargement des données dans le rapport

VII – Mission Data Quality Control Tower

VII.1 – Contexte

Le principal objectif de l'équipe ALM Data-Management est de toujours fournir toutes les données attendues par les différents consommateurs avec la meilleure qualité possible. Parce que les données sont au cœur du métier de la finance, la tour de contrôle de la qualité des données est un état de suivi quotidien pour surveiller et communiquer sur les principales données disponibles dans les bases de données OPAL.

VII.2 – Objectifs

La tour de contrôle doit permettre de qualifier la qualité du run daily en se basant sur des KPIs permettant de couvrir le périmètre de données consommées par OPAL sur différents axes d'analyse pertinentes. La Tour de contrôle doit satisfaire les points suivants :

- Qualifier, décrire et donner du sens aux données.
- Analyser l'étendue des données requises et définir un schéma pertinent pour afficher l'information attendue comparaison basée sur le benchmark quotidien.
- Alerter sur le manque de données.
- Informer sur le statut ou la variation observée sur les Deals et les Contrats Risk Pro d'un jour à l'autre.
- Alerter sur l'augmentation anormale des éléments et montants erronés.
- Sécuriser par l'application de toutes les actions nécessaires au maintien de la qualité et de la pérennité des données.

VII.3 – Analyse du besoin

Pour la bonne résolution de ce projet l'une des étapes après l'identification des besoins est l'analyse de ceux-ci et leur transcription en exigences du futur système. Ces exigences constituent les propriétés que le système doit remplir. Une exigence doit être :

- Mesurable : expression quantitative et non qualitative.
- Utile : éviter les redondances ou informations non pertinentes.
- Simple : facile à comprendre et non soumise à l'interprétation.
- Vérifiable et confirmable : prouver que l'exigence est effectuée dans le système et qu'elle répond parfaitement au besoin.

VII.4 – Mes productions

VII.4.1 – Identifications des données essentielles

La première brique de la conception de la tour de contrôle de la qualité des données OPAL a été d'identifier les données à observer. En effet l'identification du périmètre des données à observer est l'étape cruciale de cette mission.

Afin d'identifier les données critiques qui doivent être observées dans la tour de contrôle, une étude approfondie sur chaque périmètre de données des bases de données OPAL a été réalisé. Le résultat de cette étude nous a permis d'identifier les périmètres de données suivants :

- Les deals
- Les contrats
- Les devises
- Les taux de change
- Les anomalies
- Les paramétrages fonctionnels

VII.4.2 – Contrôles et points d'alertes

Après l'étape d'identification des données à observer, nous avons ensuite identifié les contrôles pertinents et les points d'alertes à mettre en place sur chaque périmètre de données identifier. Le tableau ci-dessous récapitule l'ensemble des contrôles et points d'alertes par périmètre de données identifier.

Périmètre de données	Type de contrôle	Points d'alertes
Les données des Deal	Contrôle sur les variations des montants et les quantités de deals par famille de produit et par remettant et en prenant les données du run J-1 comme benchmark.	Remontée d'éléments anormaux : montants aberrants, forte variation de l'état d'une opération par rapport à la date du j-1
Les données des Contrat RP	Contrôle sur les variations des montants et les quantités de contrats RP par famille de produit et par remettant et en prenant les données du run J-1 comme benchmark.	Remontée d'éléments anormaux : montants aberrants, forte variation de l'état d'une opération par rapport à la date du j-1
Les données de devises et de change	<ul style="list-style-type: none"> - Contrôle sur les variations des taux de change par parité de devises. - Contrôle sur le nombre de parités de l'euro avec d'autres devises. 	<ul style="list-style-type: none"> - Alerte sur le manque d'une parité de devises (par exemple manque de la parité EUR/USD) - Alerte sur la forte variation du taux de change d'une parité de devises par rapport au J-1
Les données d'anomalies	<ul style="list-style-type: none"> - Contrôle sur la quantité d'anomalies par type d'anomalies et par remettant. - Détection de nouvelles anomalies 	<ul style="list-style-type: none"> - Alerte sur les anomalies non traitées. - Alerte sur l'explosion du nombre d'anomalies sur un périmètre de données
Les données de paramétrages fonctionnels	Contrôle sur la quantité de paramètres actifs.	Contrôle sur les variations du nombre de paramétrages valides et activés

VII.4.3 – Architecture de la tour de contrôle

L'architecture de la tour de contrôle est illustrée sur la figure ci-dessous et est composée des étapes suivantes :

1. Extraction des données avec le langage SQL et import des données dans Power Query
2. Préparation et traitements des données avec Power Query
3. Chargement des données dans Power BI
4. Production d'indicateurs de qualité de données

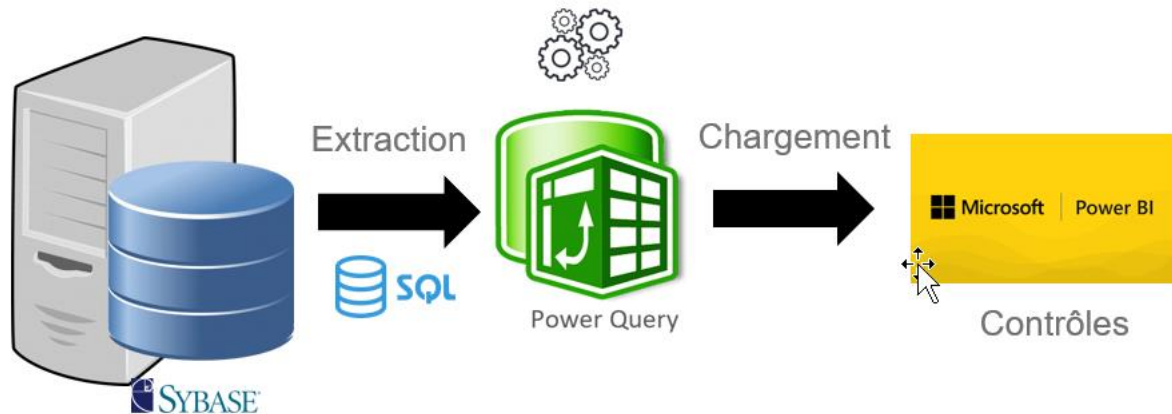


Figure 16 - Architecture de la Tour de contrôle

VII.5 – Développement

Après la phase de cadrage du projet, vient la phase du choix de la forme de la tour de contrôle. Les outils et technologies candidats étudiés sont :

1. Power BI
2. Programme python avec la bibliothèque Pandas

Après études des solutions précédemment citées Power BI répondait aux besoins. La tour de contrôle va donc se présenter sous forme de Dashboard regroupant les contrôles qui se présenteront sous forme de visuels et graphiques Power BI.

VII.5.1 – Présentation de Power BI

Power BI est une solution d'analyse de données créée par Microsoft. Il permet de créer à partir des données des visualisations, des rapports et tableaux de bord personnalisés et interactifs.

Power BI est un ensemble de services logiciels, d'applications et de connecteurs qui fonctionnent ensemble pour transformer différentes sources de données en informations visuelles, immersives et interactives. Ci-dessous une image de l'interface principale de Power BI.

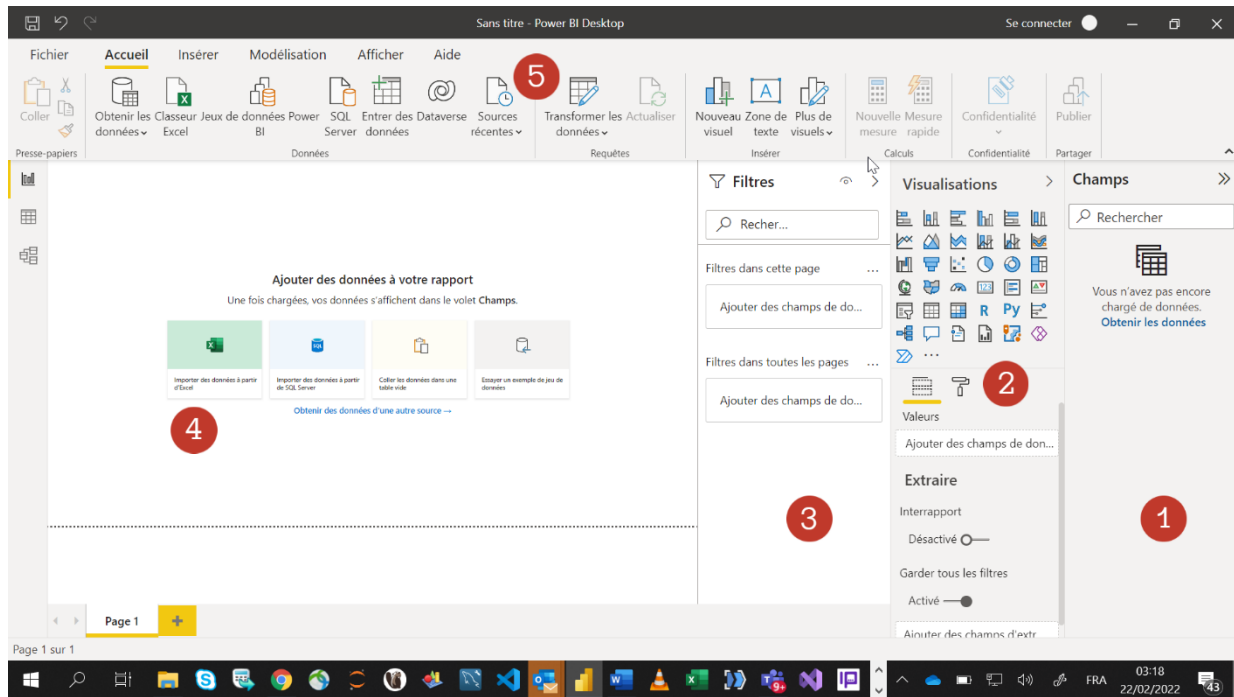


Figure 17 - Interface principale de Power BI

Description de l'interface utilisateur de Power BI :

1. Champs : représente le bloc où on trouve l'ensemble des tables chargées dans le rapport Power BI
2. Visualisations : représente le bloc où on peut choisir le type de visuels et les champs à visualiser
3. Filtres : ce bloc nous permet d'appliquer des filtres avancés sur les visuels du rapport
4. Rapport : ce bloc représente le rapport et permet d'afficher les visuels
5. Menu : représente le bloc menu de Power BI

VII.5.2 – Chargement des données dans Power BI

La réalisation de la mission reporting Grall (accès à la donnée) à constituer un socle pour la partie, préparation des données de la tour de contrôle Data Quality. Les données étant déjà préparé au préalable et disponible dans Power BI je suis passé directement au développement des différents éléments de la tour de contrôle Data quality.

VII.5.3 – Présentation de la tour de Contrôle Data Quality

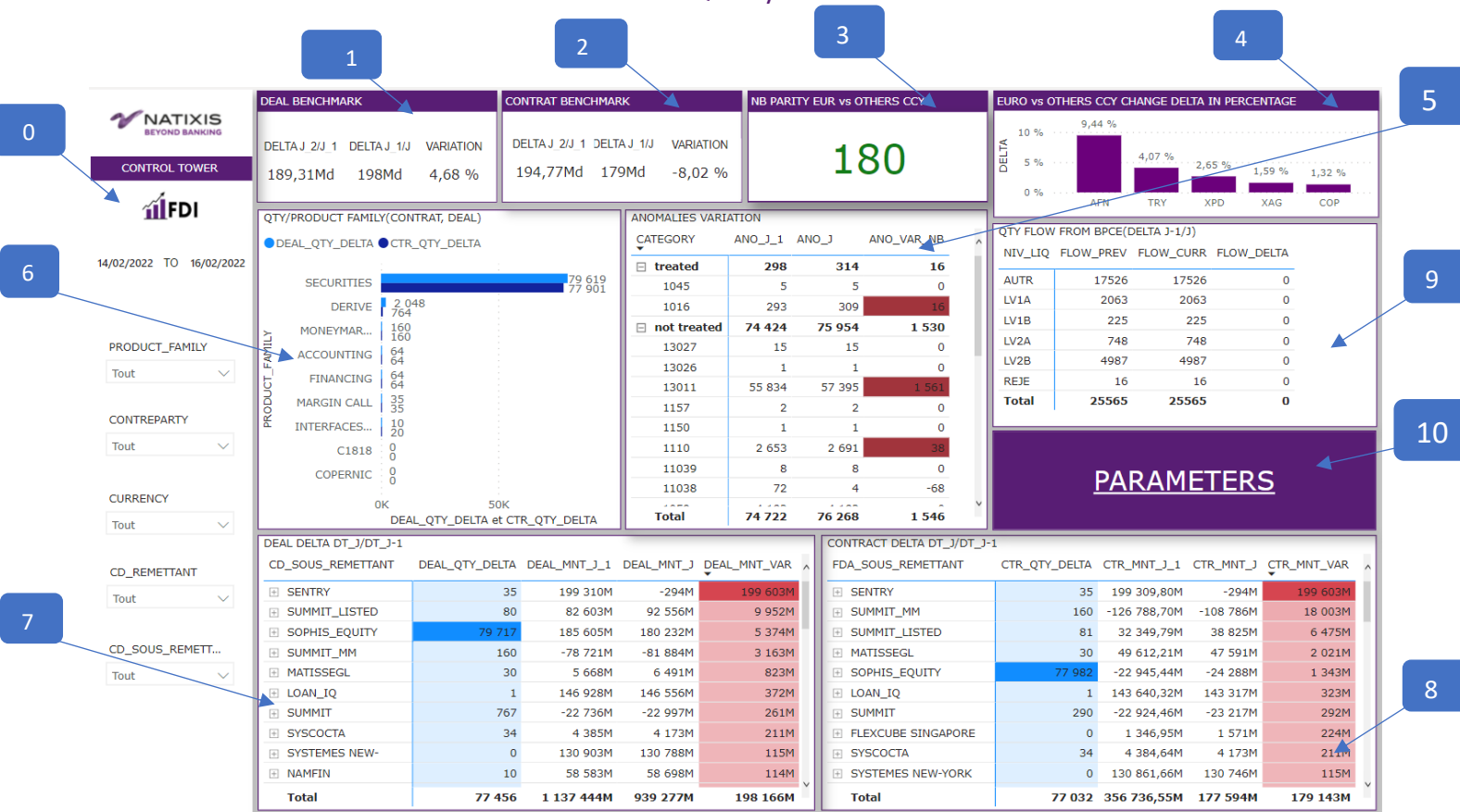


Figure 18 - La page principale de la tour de contrôle

La tour de contrôle Data quality est composée de plusieurs rapports Power BI, L'image ci-dessus représente la page principale de la tour de contrôle Data quality. C'est un Dashboard rassemblant plusieurs contrôles et alertes présentées sous forme de visuels graphiques ou tableaux.

Ci-dessous je vais présenter le détail des contrôles et visuels numérotés sur l'image ci-dessus :

- Deal Benchmark** : Présente la variation sur les deltas des montants cumulés de tous les deals réalisés au(J-1/J) par rapport au (J-2/J-3) qui nous sert de benchmark. Une variation de plus de 30% déclenche automatiquement une alerte.
- Contrat RP Benchmark** : Présente la variation sur les deltas des montants cumulés de tous les **contrats RP** réalisés au(J-1/J) par rapport au (J-2/J-3) qui nous sert de benchmark. Une variation de plus de 30% déclenche automatiquement une alerte.

3. **NB Parités EUR vs autre CCY** : Nombre de parités de devises en prenant l'EURO comme monnaie de base. La valeur attendue est de 180. Exemple : EUR GBP ; EUR USD
4. **Euro vs autres CCY change delta in percentage** : Les cinq plus grandes variations du taux de change de l'euro vers les autres devises d'un jour à l'autre. Une variation de plus de 20% déclenche une alerte.
5. **Variation Nb anomalies** : Variation de la volumétrie des anomalies OPAL d'un jour à l'autre. La détection de la même anomalie sur la même table d'un jour à l'autre déclenche une alerte
6. **Nbr / Famille de produit / contrat deal** : Variation de la quantité de deals et de contrats RP par famille de produits d'un jour à l'autre
7. **Deal delta DT_J / DT_J-1** : Tableau comparatif des données des deals du J par rapport au J-1. Permet la détection et l'alerte des éléments anormaux et montant aberrant
8. **Contrat RP delta DT_J / DT_J-1** : Tableau comparatif de des données des contrats RP du J par rapport au J-1. Permet la détection et l'alerte des éléments anormaux et montant aberrant
9. **Nb Flux retour Delta DT_J / DT_J-1** : Tableau comparatif du nombre de Flux de liquidité valide reçus de BPCE. Il s'agit d'évaluer le retour d'information de BPCE sur l'enrichissement de la liquidité.
10. Lien vers le rapport des contrôles sur les données de paramétrages.

VII.5.4 – Les autres rapports de la tour de contrôle

Dans un souci de performance et afin de permettre le chargement rapide de la tour de contrôle Data Quality le rapport principal (le Dashboard) de la tour de contrôle a été créé avec des données agrégées et limitées à un certain nombre de champs.

Pour aller plus dans les détails des contrôles, la tour de contrôle permet de faire des extractions vers d'autres rapports afin de mettre un focus plus détaillé sur une opération (deal, contrat RP), ou une anomalie particulière, ce procédé s'appelle « Drillthrough ». Ci-dessous l'illustration de ce procédé.

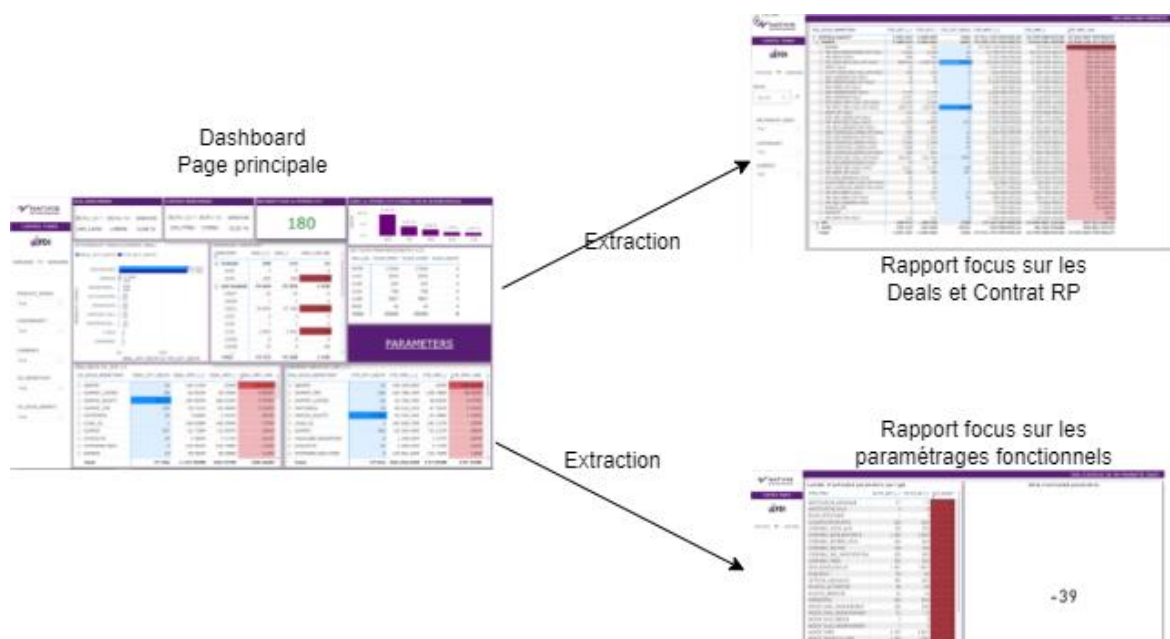


Figure 19 - Architecture finale de la tour de contrôle

VII.6 – Les formules DAX

Le langage DAX (Data Analysis Expressions) est utilisé par Power BI. Il s'agit d'un langage de requête conçu pour résoudre des problèmes d'analyse de données et de calcul de base. Il est inspiré en partie des fonctions Excel mais adapté à la réalité d'analyse de données. Ce langage permet de créer des formules efficaces qui permettent de tirer le meilleur parti possible des données et de résoudre des problèmes concrets. Il permet la création d'indicateurs sur mesure répondant aux problématiques des utilisateurs.

VII.6.1 – Quelques formules DAX utiliser sur la Tour de Contrôle

- La formule suivante permet de faire la somme des montants des deals du J-1

```
DEAL_MNT_J_1 =
    VAR min_date = MIN(dt_mid[DATES])
    RETURN SUMX(
        FILTER(
            DEALS, DEALS[DT_CALCUL] = min_date
        ),
        DEALS[AMOUNT]
```

- La formule suivante compte le nombre de parités de devises

```
1 PARITY/J =
2 VAR DT_MAX = MAX(F_COURS[DT_CALCUL])
3 RETURN COUNTROWS(
4     FILTER(F_COURS, F_COURS[DT_CALCUL] = DT_MAX))
```

- La formule suivante calcul la variation des montants des deals d'un jour à l'autre (du J/J-1)

```
1 DEAL_MNT_VAR = abs(DEALS[DEAL_MNT_J] - DEALS[DEAL_MNT_J_1])
```

- La formule suivante calcul la variation en pourcentage des montants des deals d'un jour à l'autre

```
1 DEAL_DELTA_BENCHMARK = ((DEALS[DEAL_MNT_VAR] - DEALS[DEAL_MNT_VAR_prev]) / DEALS[DEAL_MNT_VAR_prev])
```

- La formule suivante calcul la quantité des deals de la date du Jour

```
1 DEAL_QTY_J =
2     VAR max_date = MAX(DEALS[DT_CALCUL])
3     RETURN SUMX(
4         FILTER(
5             DEALS, DEALS[DT_CALCUL] = max_date
6         ),
7         DEALS[QUANTITY]
8     )
```

VII.7 – Mise en production de la Tour de Contrôle

Après la phase de développement l'ensemble des rapports de la tour de contrôle ont été publiés dans un workspace dédié à cet effet, La tour de contrôle a donc été déployée et mise en production sous d'application web dans le service Power BI.

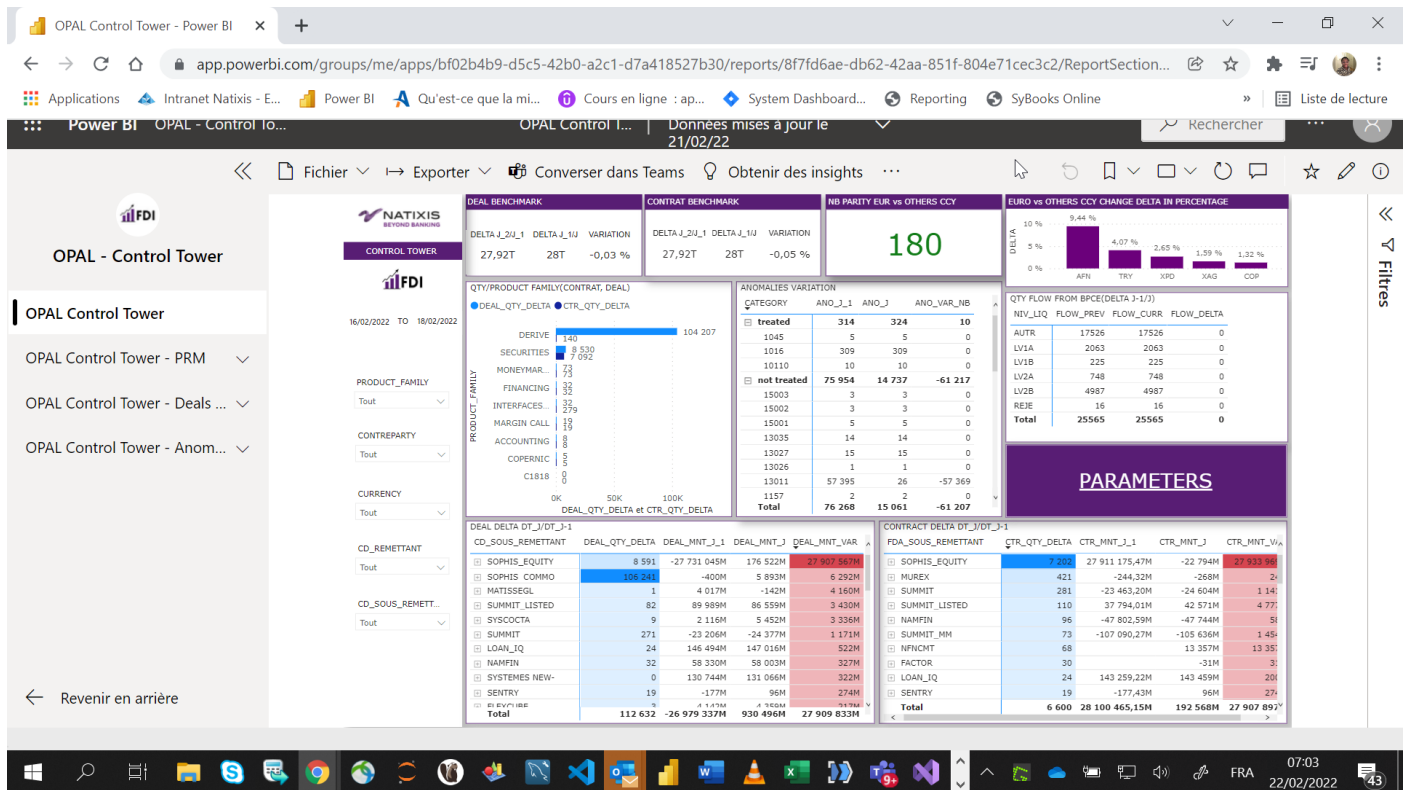


Figure 20 - Tour de contrôle déployer sous forme d'application web

Sur la bannière gauche sur l'image ci-dessus on peut constater un menu regroupant l'ensemble des rapports de la tour de contrôle ce qui permet à l'utilisateur d'explorer facilement les différents rapports de la tour de contrôle et d'y accéder directement via un lien https.

La mise en place d'un Gateway planifier tous les jours à 11H permet la mise à jour et le rafraichissement automatique des données et des visuels de la tour de contrôle.

NATIXIS - NXDEV - FDI - Data sourcing - ALM-Data-Management - Control Tower

+ Nouveau Créer un pipeline

Tout Contenu Jeux de données + flux de données

Nom	Type	Propriétaire	Actualisé	Prochaine actualisation	Approbation
OPAL Control Tower	Jeu de données	NATIXIS - NXDEV - F...	21/02/22, 11:38:34	22/02/22, 11:00:00	—
OPAL Control Tower - Anomalies	Jeu de données	NATIXIS - NXDEV - F...	21/02/22, 11:17:06	22/02/22, 11:00:00	—
OPAL Control Tower - Deals and Contracts	Jeu de données	NATIXIS - NXDEV - F...	21/02/22, 11:44:16	22/02/22, 11:00:00	—

Figure 21 - Mise en place d'un Gateway d'actualisation automatique de la tour de contrôle

VIII – Organisation et évaluation

VIII.1 – Organisation de la mission

Pendant mon stage chez Natixis, j'ai travaillé sur deux projets. Le premier a été la phase préparatoire du deuxième projet. Après une brève formation de quelques semaines sur Dbeaver et Power Query, j'ai débuté le premier projet qui s'intitule « Reporting Grall », un projet qui a duré un mois. Puis j'ai débuté le projet « Control Tower OPAL Data Quality », j'avais des checkpoints tous les jours avec mon tuteur M. Sylvain OLIVA pour le suivi de mon stage et ma montée en compétences, des conseils sur les axes réguliers d'évolutions à suivre pour améliorer le rendu de la migration. J'avais des objectifs de réalisations tous les deux jours.

Ma capacité d'adaptation sur les projets a été un grand plus pour moi. Ainsi j'ai pu dans un premier temps découvrir un nouvel environnement et des projets très bien structurés.

Je prenais au fur et à mesure des notes pour me faciliter la rédaction de mon rapport de stage.

VIII.2 – Auto-évaluation et problèmes rencontrés

J'ai connu une montée en compétence rapide sur l'outil Power BI, Power Query et Dbeaver. Au bout de quelques semaines j'étais presque opérationnel. D'un point de vue formation aux outils et prise de connaissance des projets auxquelles j'ai pu participer et j'en suis vraiment satisfait. Grâce aux différents points avec mon tuteur j'ai pu acquérir rapidement des compétences et travailler avec des objectifs suivant les tâches. Pour des besoins sur le plan technique, au début comme durant mes projets, les collègues aînés de l'équipe ALM Data-Management étaient à disposition pour répondre à mes questions. Mon tuteur de stage s'arrêtait très souvent pour m'expliquer les points durs ou pas clairement exprimés. Mon tuteur de stage s'est montré très souvent disponible à ma formation et mon intégration ce qui m'a permis de travailler sans aucun blocage. Souvent amené à travailler tout seul, apprendre et résoudre certains problèmes rencontrés, je pense avoir fait preuve d'autonomie.

J'ai parfois rencontré quelques difficultés comme l'accès aux ressources par exemple au tout début, l'installation de certains outils ont tardé à être disponibles pour nous, ce qui a rendu le démarrage de la formation un peu moins efficient. La confrontation à ces difficultés m'a appris à être plus patient et beaucoup plus organisé. D'un point de vue humain, j'ai davantage développé mes qualités relationnelles grâce aux activités au sein de l'équipe et grâce à la disponibilité des uns et des autres. Je pense avoir été à la hauteur et relevé quelques défis avec NATIXIS.

IX – Conclusion

Ces six mois de stage chez Natixis m'ont permis de me positionner dans un contexte professionnel et de travailler en équipe sur des projets de grande envergure. Heureux de conclure ce parcours avec l'université de Versailles Saint-Quentin en Yvelines et d'intégrer le monde professionnel en entreprise.

Mon stage s'est avéré particulièrement formateur du point de vue technique car j'ai amélioré mes connaissances sur les bases de données Sybase, l'extraction et la préparation de données et l'analyse de données. Plusieurs workshops techniques ont été menés sur Microsoft Power Query et Power BI. J'y ai trouvé un grand intérêt d'autant plus que ces connaissances qui ne nous ont pas été toutes enseignées à l'université et viennent compléter ma formation.

Mon travail sur la tour de contrôle Data Quality m'a permis de prendre davantage l'assurance sur des projets de migration, d'intégration, d'extraction et d'analyse de données.

D'un point de vue global, l'ambiance de travail était très bonne le contact avec le personnel fut très rapide et convivial, cela est peut-être dû à la culture d'entreprise assez chaleureuse qui règne chez Natixis et plus précisément au sein de l'équipe ALM Data-Management. Ce stage est totalement dans la continuité de ma formation et m'a permis d'enrichir mes connaissances dans l'extraction, la préparation et l'analyse de données et à l'utilisation d'outils ETL et d'analyse. J'ai compris l'importance de l'analyse de données en milieu bancaire.

X – Bibliographies

<https://docs.microsoft.com/fr-fr/power-bi/>

<https://docs.microsoft.com/fr-fr/power-query/>

<https://docs.microsoft.com/fr-fr/powerquery-m/m-spec-introduction>

<https://docs.microsoft.com/fr-fr/dax/dax-function-reference>

<https://dbeaver.com/docs/wiki/>

https://natixis.groupebpce.com/natixis/fr/accueil-j_6.html

Documentation interne de Natixis.