



ÉCOLE NATIONALE DE LA STATISTIQUE ET DE L'ANALYSE DE L'INFORMATION



RAPPORT DE STAGE

Deuxième Année

**LA BUSINESS INTELLIGENCE (BI) POUR LA CONCEPTION ET LA RÉALISATION DES
TABLEAUX DE BORD BDES (BASE DE DONNÉES ECONOMIQUE ET SOCIALE)**



Réalisé par :

Caleb Carloss AGUIDA

Encadrant :

Axelle DACQUIN

Coach :

Laurent BRAULT

Année académique 2019-2020

Remerciements

Tout d'abord, j'adresse mes remerciements à Mme Axelle Dacquin qui m'a aidé dans ma recherche de stage et m'a permis de postuler dans cette entreprise et de trouver ce stage. Je tiens aussi à la remercier vivement pour son accueil, le temps passé ensemble et le partage de son expertise au quotidien. Grâce aussi à sa confiance j'ai pu m'accomplir totalement dans mes missions. Elle fut d'une aide précieuse dans les moments les plus délicats.

Je ne pourrai oublier Mr Laurent BRAULT, directeur des systèmes d'information et Mme Estelle MOISAN, directrice des ressources humaines qui n'ont ménagé aucun effort pour la réussite de mon stage.

Je remercie également toute l'équipe de Centravet pour leur accueil, leur esprit d'équipe et les valeurs humaines qu'elles incarnent.

Table des matières

Remerciements	i
Liste des abréviations et des acronymes	iii
Liste des tableaux et graphiques	iv
1 Description du stage	1
1.1 Présentation de la structure d'accueil	1
1.2 Présentation du sujet de stage : Problématique	2
1.3 Planification, suivi et évaluation	3
1.4 Analyse et Critique de l'existant	4
1.5 Analyse des besoins	5
2 Description de la méthodologie	7
2.1 L'informatique décisionnelle	7
2.2 Etapes de la mise en oeuvre d'une solution BI	8
2.3 Outils de modélisation	9
2.4 Modélisation dimensionnelle et ses concepts	11
2.5 Différents modèles de la modélisation dimensionnelle	13
3 Conception et développement de l'ETL	14
3.1 Modélisation	14
3.1.1 Statistiques ressources humaines	14
3.1.2 Activités ressources humaines	14
3.1.3 Modèle en flocon	15
3.2 Développement et création de l'entrepôt de données	17
3.2.1 Définition des variables de contexte	17
3.2.2 Métadonnées	18
3.2.3 Description des jobs	19
3.3 Restitution	24
3.3.1 Création des vues et des index	25
3.3.2 Tableaux de bords	25
Conclusion	32
Annexes	viii

Liste des sigles et abréviations

BI : Business Intelligence
RH : Ressource Humaine
ETL : Extraction, Transformation, Loading

Liste des tableaux

1.1	Informations sur Centravet	2
2.1	Comparaison entre les tables de faits et les tables de dimensions	12
3.1	Dimension pour les Statistiques ressources humaines	14
3.2	Dimension pour les Activités ressources humaines	15

Table des figures

1.1	Planning Jira	4
1.2	Processus	5
2.1	Talend	10
2.2	Interaction entre les outils	11
3.1	Modèle	16
3.2	Contexte	17
3.3	Context Dev/Prod	18
3.4	Connexion	18
3.5	Requête	18
3.6	Fichier Excel	18
3.7	Job Type Tiers	19
3.8	Job Type Contrat	19
3.9	Job Type Contrat Plan	19
3.10	Job Type Activités	19
3.11	Job Société	20
3.12	Job Situation Matrimoniale	20
3.13	Job Genre	20
3.14	Job Contrat	21
3.15	Job Contrat, tmap 2	22
3.16	Job Contrat, tmap 3	22
3.17	Job Stat_Tiers	23
3.18	Formule filtre	23
3.19	tmap Stattiers	24
3.20	tmap Stattiers	24
3.21	Interface filtre dynamique	26
3.22	Interface Emploi filtre dynamique	27
3.23	Répartition des niveaux par poids	28
3.24	Répartition des niveaux par Etablissement	29
3.25	Répartiton géographique des Tiers autours des Etablissements	30
3.26	Ancienneté des tiers par genre	31
3.27	Job Menu	viii

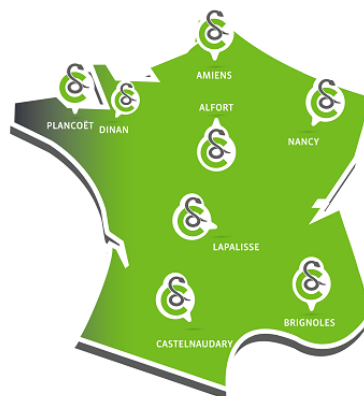
Description du stage

Dans cette partie, nous présenterons dans un premier temps la structure d'accueil i.e celle au sein de laquelle s'est déroulé le stage, dans un second temps le sujet traité et pour finir les résultats attendus et le planning du stage.

1.1 Présentation de la structure d'accueil

L'organisme qui nous a accueilli dans le cadre de ce stage est **CENTRAVET**. Centravet est une coopérative vétérinaire, c'est-à-dire que ses clients sont également ses actionnaires (voir Tableau 1.1). C'est la première centrale de distribution vétérinaire à avoir été créée en France. Il s'agit d'une entreprise de niveau national.

A ce titre, Centravet dispose de différents sites implantés en France, à savoir Dinan, Plancoët, Lapalisse, Nancy, Castel, Alfort, Brignoles et Amiens. Certains sont des sites de productions et d'autres sont des sites de logistiques. Le site de Dinan est le siège de l'entreprise. Cependant, l'ensemble des relations commerciales et logistiques dépend directement de chaque établissement.



Les fonctions centrales sont regroupées au Siège à Dinan. Elles sont limitées aux secteurs où il y a un intérêt de regrouper les moyens comme la comptabilité, les relations fournisseurs, l'informatique, le service technique matériel et la communication. La gestion de tout le personnel (toutes les établissements y compris) est confié à la Direction des Ressources Humaines du siège qui à ce titre fait de l'égalité femmes-hommes une valeur incarnée par l'entreprise. En effet, l'index égalité femmes-hommes est de 99/100 pour Centravet. Cet index est le résultat d'une politique RH clairement orientée équité salariale.

L'objectif et slogan de l'entreprise est de simplifier la vie des vétérinaires praticiens et leurs apporter des solutions pour développer leurs structures. Centravet se charge donc de la distribution en médicaments et produits vétérinaires. Nous pouvons d'ailleurs noter que la direction de chaque site dispose d'au moins un directeur/directrice vétérinaire. Proches des clients, les Directeurs d'établissements avec leurs équipes commerciales et l'ensemble du personnel des entrepôts peuvent ainsi répondre au plus vite aux besoins des vétérinaires.

Toutefois Centravet ne se contente pas faire de la distribution vétérinaire même si cela reste

La Business Intelligence pour la conception et la réalisation des tableaux de bord BDES leur principale activité. En effet, Centravet propose à ses clients de les aider à développer leur communication. Notamment en leur proposant de développer leur site internet, e-boutique ou encore des cartes de visite. De plus, la coopérative offre également ses conseils en ce qui concerne le matériel et la logistique des clients et leur propose des solutions informatiques afin de simplifier au maximum leurs démarches.

Afin d'assurer tous ces services l'entreprise dispose d'équipes aux métiers et compétences très variés. Elle compte aujourd'hui plus de 350 personnes. Parmi elles GMV, qui est une structure appartenant au groupe Centravet, une filiale. Il s'agit d'une structure regroupant des spécialistes en informatique qui ont pour mission de s'occuper des innovations technologiques de la coopérative.




 CENTRAVET	
 Zone Artisanale Alleux, 22100 Taden	
 02 96 85 80 40	
<i>Site web</i> : https://cooperative.centravet.net/default.aspx	
Type	Société coopérative
Histoire	La société a vu le jour en 1972, au sein d'un entrepôt en Bretagne à proximité de Plancoët. Aujourd'hui Centravet à une activité qui s'étend sur tout le territoire. La société est devenue une coopérative de vétérinaires praticiens en 1978. Elle compte aujourd'hui plus de 5000 vétérinaires adhérents.
Effectif	Plus de 350 personnes
Secteur d'activité	Distribution vétérinaire
Champ d'action	Toute la France
Situation géographique	Siège : Dinan Autres sites : Plancoët, Lapalisse, Nancy, Castel, Alfort, Brignoles et Amiens.

Tableau 1.1: Informations sur Centravet



1.2 Présentation du sujet de stage : Problématique

Il est d'une obligation légale que « l'employeur d'au moins 50 salariés doit mettre à disposition du comité économique et social (CSE) ou des représentants du personnel une base de données économiques et sociales (BDES). Elle rassemble les informations sur les grandes orientations économiques et sociales de l'entreprise. La base de données comprend des mentions obligatoires qui varient selon l'effectif de l'entreprise. » (Voir [Pre20]). C'est déférant à cette prescription que s'inscrit ce stage dont l'objectif est de réaliser les tableaux de bords BDES (Base de Données Economique et Sociale) au sein de **CENTRAVET**. Il s'agira plus précisément de mettre la lumière sur l'évolution des effectifs et des conditions de travail en fonction des catégories sociaux économiques depuis 3 ans.

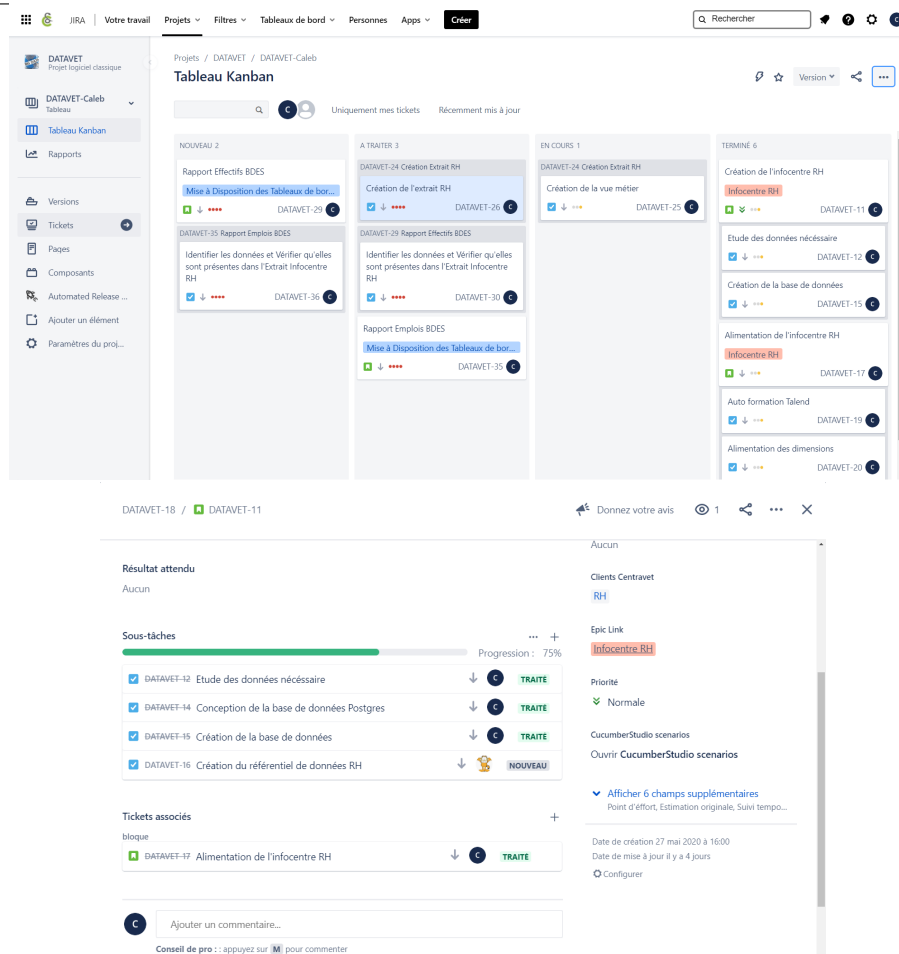
1.3 Planification, suivi et évaluation

Pour mener à bien ce projet, une phase de planification primordiale a été réalisée afin de lister les différentes tâches à réaliser et les ressources nécessaires. La méthode de planification utilisée est la **méthode agile** (adaptative) qui est une approche de planification par lequel on définit au début d'un projet les ressources et le délai visé et on adapte régulièrement en cours de projet le périmètre. Cette méthode s'oppose généralement à la méthodologie traditionnelle (on définit au début le planning, les ressources nécessaires, les délais pour réaliser le périmètre visé) en offrant une plus grande flexibilité et une meilleure visibilité dans la gestion du projet. Le choix porté sur cette méthode est également motivé par les possibles changements qui pourront être occasionnés par la non disponibilité de certaines informations dans les bases de données existantes et nécessaire pour les métiers, des changements des besoins en tableaux de bords et des imprévus mais aussi de favoriser la collaboration et la communication fréquente avec les métiers.

L'outil de gestion utilisé pour la planification est le duo **Confluence** et **Jira** de Atlassian. Jira est un excellent outil pour la gestion de projets. Il permet de définir le responsable d'une tâche, le workflow et la date de fin du projet mais fournir peu de détail sur le détail. D'où l'utilisation de confluence qui est un espace et une flexibilité illimités pour donner les détails nécessaires des projets à tous les membres de l'équipe. Spécifiquement à **CENTRAVET**,

- ① Confluence  Confluence : Outil de gestion documentaire de l'entreprise, elle regroupe la base de connaissance du support, et des services commerciaux Gmvet les spécifications fonctionnelles des différents projets développés par l'entreprise, ainsi que les documentations techniques liées à chaque pan fonctionnel des projets de l'entreprise ;
- ② Jira  Jira : Solution de Ticketing et suivi du workflow de développement, aussi bien utilisée par les responsables projets que par les développeurs et les membres du support.

La figure 1.1 réalisé sous Jira présente le planning des différentes tâches retenues, leur progression, les sous-tâches et le responsable.



Graphique 1.1: Planning Jira

1.4 Analyse et Critique de l'existant

L'ensemble des données utilisées par les RH sont issues principalement des bases de données Plannivet et CEGID. Ceci garantit la disponibilité de la donnée RH. Cependant, les différentes rencontres faites avec les responsables des services de ressources humaines nous ont permis de relever un certain nombre d'insuffisances ou problèmes qui ralentissent certaines activités au sein du service de la RH. L'un des principaux problèmes relevé est *la gouvernance de la donnée*¹. Ce dernier résulte des constats liés à des définitions incohérentes des données, la présence des doublons, des champs manquants, l'absence de règle de saisie claire et de maintenance des données (définitions reconnues, règles de style, etc). Il nous est permis par exemple de constater la non disponibilité d'un schéma de base de données pour PLANNIVET et CEGID, d'un dictionnaire des données (ce qui entraîne beaucoup de temps pour la recherche d'information), plusieurs règles de gestion au cours du temps sans source écrite (entraînant

1. C'est un ensemble de processus qui assurent que les données sont formellement gérées à travers l'entreprise. Elle garantit la fiabilité des informations utilisées pour les processus métier critiques, la prise de décision et la comptabilité.

La Business Intelligence pour la conception et la réalisation des tableaux de bord BDES
une incohérence dans les enrégistrement), multi-source de données, absence de norme claire
pour les noms de variables et tables de la base.

1.5 Analyse des besoins

Il est primordial de rendre possible la bonne gouvernance de la donnée RH afin de réduire les erreurs au sein des bases de données et surtout le gain du temps qui dans le cas contraire serait mobilisé pour corriger ces erreurs et rechercher une information (☺ le temps, c'est de l'argent).

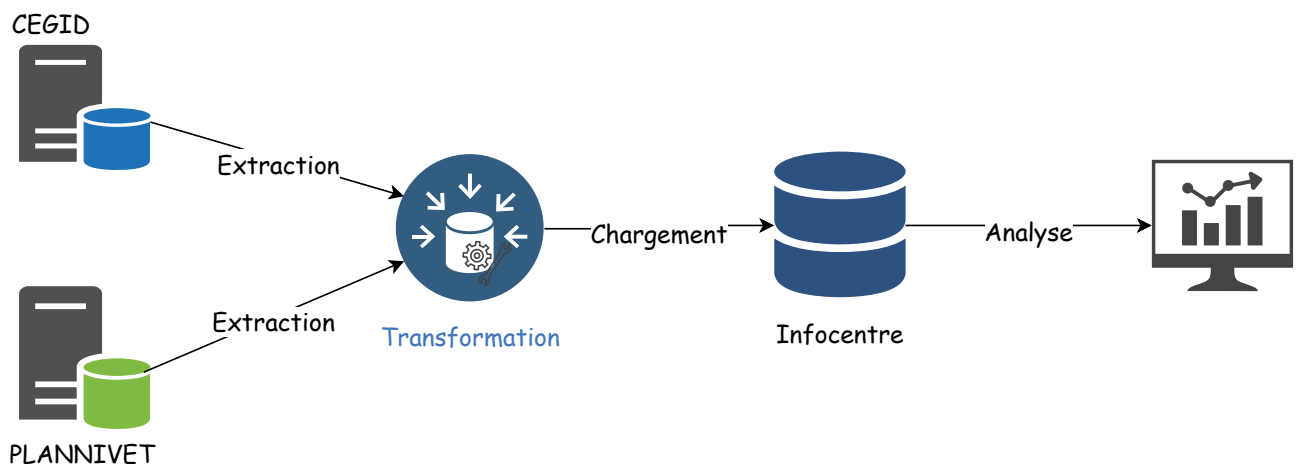
Pour cela, il s'avère indispensable de mettre en place un *entrepôt de données* (Datawarehouse) de telle sorte à avoir une source de données qui puisse non seulement être :

- ☞ unique ;
- ☞ adapté aux besoins spécifiques de la RH donc plus optimale et performante ;
- ☞ évolutive et adaptée aux besoins futurs ;
- ☞ mise à jour périodiquement eu égard à la granularité choisie,

mais aussi avoir :

- ☞ la source d'extraction de la donnée (base source, table source) ;
- ☞ un schéma de base de données pour l'infocentre (modélisation) ;
- ☞ un dictionnaire des données ;
- ☞ un système d'alerte en cas d'erreur d'alimentation de l'infocentre (gestion des erreurs) ;
- ☞ un processus pour la reprise des données.

Cet infocentre permettra de fournir à tout instant au métier des tableaux de bord explicites ; des indicateurs clairs ; des statistiques fiables et facilement analysables et constituera un véritable outil de décision quant au RH.



Graphique 1.2: Processus

En résumé, le travail pour la suite consistera à extraire les informations des principales

La Business Intelligence pour la conception et la réalisation des tableaux de bord BDES
sources de données (CEGID, Plannivet), à les agréger, transformer puis stocker dans un info-
centre qui servira à alimenter les tableaux de bords pour le métier. (Voir figure 1.2)

Description de la méthodologie

Cette partie donne une description de l'informatique décisionnelle, ses avantages, ses limites et ses étapes. Il sera aussi question de définir les notions d'ETL et de datawarehouse ainsi que de présenter les outils de modélisation.

2.1 L'informatique décisionnelle

Définition

L'**informatique décisionnelle** (en anglais **business intelligence (BI)** ou **decision support system (DSS)**) est l'informatique à l'usage des décideurs et des dirigeants d'entreprises. Elle désigne les moyens, les outils et les méthodes qui permettent de collecter, consolider, modéliser et restituer les données, matérielles ou immatérielles, d'une entreprise en vue d'offrir une aide à la décision et de permettre à un décideur d'avoir une vue d'ensemble de l'activité traitée. Ce type d'application utilise en règle générale un datawarehouse (ou entrepôt de données) pour stocker des données transverses provenant de plusieurs sources hétérogènes et fait appel à des traitements lourds de type batch pour la collecte de ces informations.

Avantages

D'un point de vue général, la business intelligence présente les avantages tels :

① **une vue globale en temps réel de l'activité et des ressources de l'entreprise**

Elle permet d'obtenir une vue en direct de l'ensemble des indicateurs-clés de la société, mise à jour en temps réel. Une solution de BI offre un tableau de bord complet et intégré, nécessaire au pilotage d'une entreprise et de ses filiales.

② **un outil de référence pour la gouvernance**

La cohérence d'une gouvernance tient en ce que ses différentes parties parlent le même langage. Avec un outil d'informatique décisionnelle, les indicateurs et le vocabulaire sont commun à l'ensemble des collaborateurs jouant un rôle dans la stratégie d'entreprise : PDG, directeurs de business units, directeurs commerciaux, DRH... Cette uniformisation permet une meilleure compréhension des sujets et des enjeux et une synergie renforcée. Elle accélère le processus de décision.

③ **un outil de décision accélère vos performances**

La vue à 360° des activités de l'entreprise permet d'identifier immédiatement les secteurs performants et ceux en perte de vitesse. Elle offre une visualisation directe de toutes les anomalies à corriger. Les tableaux et rapports fournis par les outils sont donc des lanceurs d'alertes qui accélèrent la prise de décision et la résolution des problèmes. Les erreurs vite corrigées entraînent une augmentation de la performance globale et du chiffre d'affaires.

④ **anticiper les événements et se projeter dans l'avenir**

Associée au machine learning et à la gestion des big data, la business intelligence dispose d'un pouvoir de prédiction efficace. Elle permet d'anticiper les orientations d'un marché, elle dévoile les tendances des demandes clients, ou encore dresse les perspectives de croissance et d'amélioration des différents secteurs d'activité de l'entreprise. Le concept d'informatique décisionnelle prend ici tout son sens : grâce à cette vision augmentée, les dirigeants sont plus à même de choisir les orientations et les développements les plus pertinents pour leur société.

En résumé, la BI représente un vecteur de croissance pour l'entreprise.

Limites

La business intelligence présente toutefois quelques limites :

- ① une mauvaise restitution des données entraînera des mauvaises prises de décision qui peuvent être préjudiciables à l'entreprise ;
- ② La mise en place d'une solution de BI est complexe et prend beaucoup du temps.

2.2 Etapes de la mise en oeuvre d'une solution BI

La mise en oeuvre d'une solution BI se fait grâce à un processus ETL (Extraction, Transformation, Loading) du chargement des données, depuis les différentes sources d'information, dans des entrepôts de données décisionnels.

L'extraction consiste principalement en la collecte de l'information de sources hétérogènes : fichiers plats, fichiers Excel, base de données (DB2, Oracle, SQL Server, etc.), service web, données massives.

La transformation implique le nettoyage, la traitement et la consolidation des données extraites et leur adaptation aux besoins décisionnels. Elle est nécessaire pour nettoyer les données, supprimer les redondances et vérifier la conformité des données chargées. [LAA]

Les données collectées et traitées, vient ensuite le stockage et la consolidation dans des bases de données spécifiques à l'usage décisionnel : **Entrepôt de données ou Datawarehouse**. Ces derniers sont spécialement conçus et développés pour stocker les données décisionnelles traitées et supporter l'usage décisionnel.

La Business Intelligence pour la conception et la réalisation des tableaux de bord BDES

Après ces étapes, on procède à la restitution (tableaux de bords) et l'exploitation des données (prise de décisions, métier).

2.3 Outils de modélisation

Afin d'atteindre nos objectifs, nous avons utilisé les outils ci-après : Talend Open Studio(TOS), pgModeler, pgAdmin, SQL Server Management Studio (SSMS) et Tableau Serveur. Chacun de ces outils est bien spécifique quant à la tâche à accomplir.

2.3.1 Talend Open Studio (TOS)

Talend est un ETL (Extract Transform and Load) qui permet d'extraire des données d'une source, de modifier ces données, puis de les recharger vers une destination. La source et la destination des données peuvent être une base de données, un service web, un fichier csv et bien d'autres... C'est un ETL de type « générateur de code ». Pour chaque traitement d'intégration de données, un code spécifique est généré, ce dernier pouvant être en Java ou en Perl. Talend Open Studio utilise une interface graphique, le « Job Designer » (basée sur Eclipse RCP) qui permet la création des processus de manipulation de données.

La définition de quelques concepts clés est nécessaire pour la compréhension du fonctionnement de Talend lors de la mise en oeuvre. Il s'agit de :

Job

Un job est un ensemble de tâches, considérées comme une unité. Il regroupe de nombreux composants liés entre eux. Il peut être exécuté en mode batch ou interactif, avec de multiples inputs/outputs autorisés au sein d'un même job.

Les jobs peuvent être synchronisés entre eux de diverses façons. La parallélisation et le multithreading sont envisageables. De plus les jobs peuvent être hiérarchisés : des jobs principaux peuvent lancer des sous-jobs

Composant

Un composant est un sous-ensemble d'un job offrant une panoplie de fonctions. Il peut être considéré comme une unité de traitement (exécution d'une fonction précise). Il en existe environ 600 sous TOS 5.x.

L'objectif des composants est d'éviter la rédaction de lignes de code simples les plus courantes; (exemple : mettre en majuscule tout une chaîne de caractères)

Ex : tFileInputDelimited (csv en entrée)-> tMap (composant de transformation) -> tLogRow (affichage sur la console des données en sortie)

Connecteurs

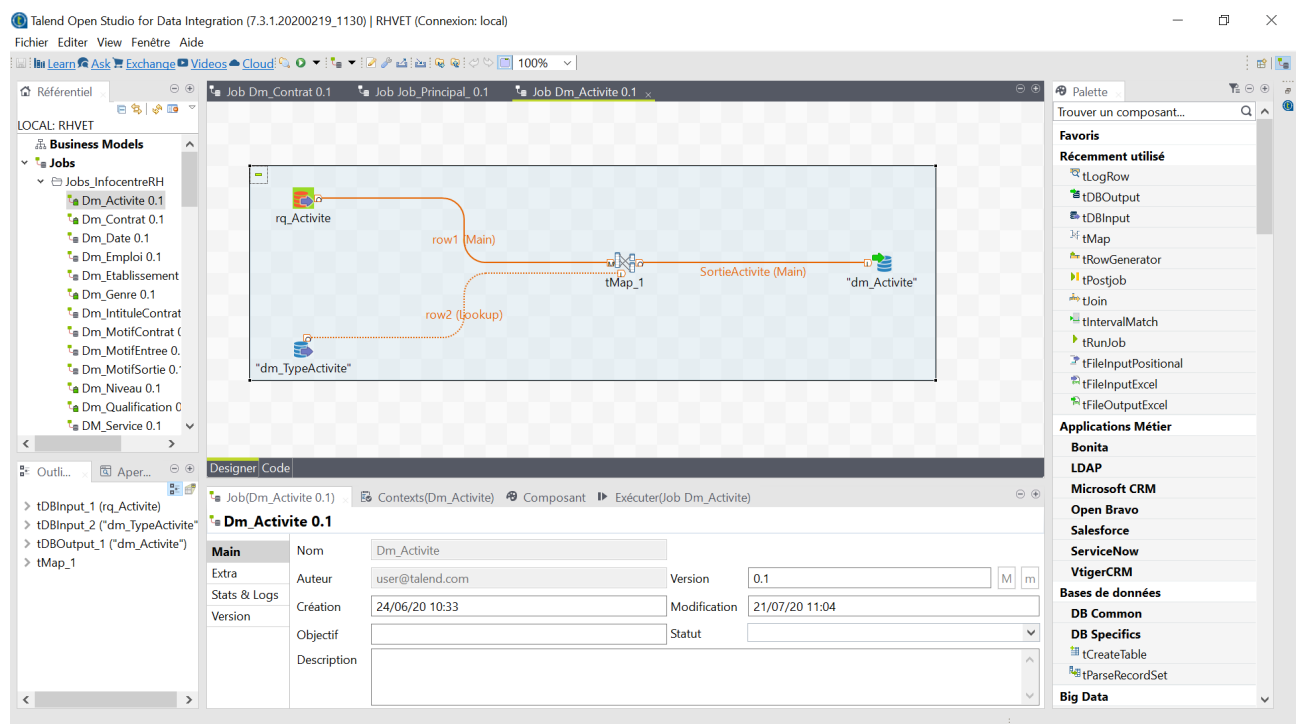
Plus de 450 connecteurs sont disponibles pour se connecter aux principaux SGBD (SQL-Server, Oracle, PostgreSQL, MySQL,...) ainsi que pour traiter tous les types de fichiers plats (CSV, Excel, XML), aussi bien en lecture qu'en écriture.

Métadonnées

Les métadonnées sont des données qui définissent les données traitées. Un référentiel permet de les stocker afin de pouvoir les exploiter dans différents jobs.

L'objectif des métadonnées est d'obtenir une application plus robuste, plus facile à maintenir et plus rapide.

La figure 2.1 donne un aperçu de l'interface graphique du TOS.



Graphique 2.1: Talend

2.3.2 pgAdmin

Le logiciel pgAdmin est une plate-forme d'administration et de développement pour la base de données PostgreSQL. Elle a été conçue pour répondre aux besoins de tous les utilisateurs, depuis l'écriture de requêtes simples jusqu'au développement de bases de données complexes. L'interface graphique donne accès aux fonctionnalités de PostgreSQL les plus récentes, faisant de l'administration un véritable jeu d'enfant. L'application comprend également un constructeur de requêtes, un éditeur SQL, un éditeur de code server-side et bien plus encore.

2.3.3 SQL Server Management Studio (SSMS)

SQL Server Management Studio (SSMS) est un logiciel lancé avec Microsoft SQL Server 2005. Il est utilisé pour la configuration, la gestion et l'administration de tous les composants avec Microsoft SQL Server.

2.3.4 Tableau Serveur

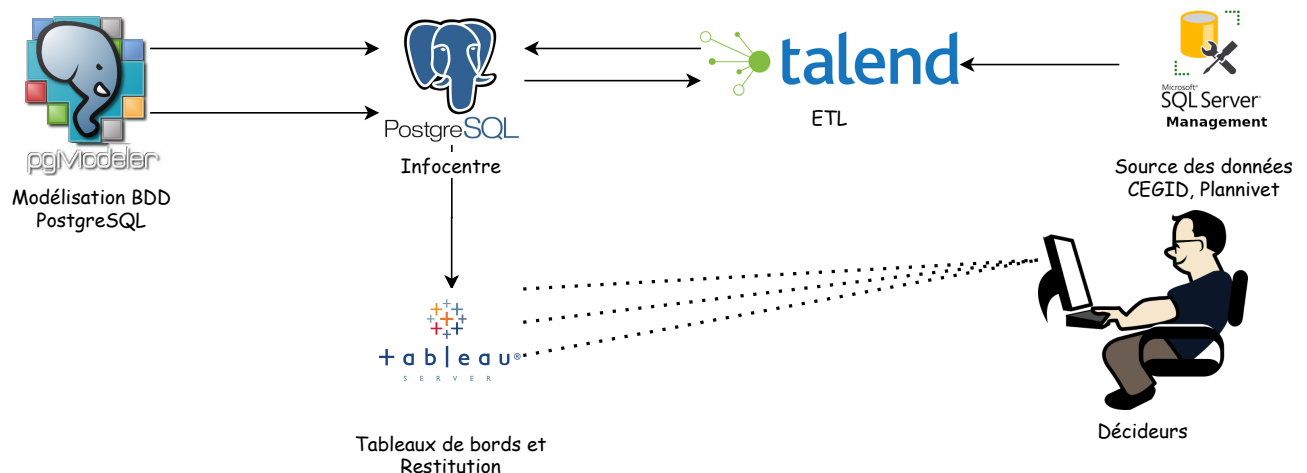
Tableau Server est l'un des produits principaux proposés par Tableau software dont l'activité consiste à interroger les bases de données relationnelles, cubes, les bases de données cloud et les feuilles de calcul et de générer des nombre de graphiques qui peuvent être combinés dans des tableaux de bord et partagés sur un réseau intranet ou sur internet.

Plus spécifiquement, Tableau Server est une plate-forme en ligne pour héberger et conserver tous les classeurs de tableaux, les sources de données et les données de tableaux connexes

2.3.5 pgModeler

pgModeler est un acronyme de « PostgreSQL Database Modeler » et fait référence à un outil de modélisation de base de données open source spécialement conçu pour le système de gestion de base de données relationnelle PostgreSQL.

Les interactions entre ces différents outils est représenté sur la figure 2.2



Graphique 2.2: Interaction entre les outils

2.4 Modélisation dimensionnelle et ses concepts

La différence entre l'entrepôt de données (la Datawarehouse) et une simple BD relationnelle (BDR) est sa modélisation. Alors que la modélisation classique se présente sous la forme

La Business Intelligence pour la conception et la réalisation des tableaux de bord BDES de modèle entité-relation, où l'on représente les données sous la forme d'entités (tables) et de relations (références ou tables) avec un schéma normalisé, la modélisation d'un entrepôt de donnée représente les données comme des faits et des dimensions non normalisées. On parle de la **modélisation dimensionnelle**.

La modélisation dimensionnelle est une technique de conception logique permettant de structurer les données de manière à les rendre intuitives aux utilisateur d'affaires et offrir une bonne performance aux requêtes. Elle divise les données en dimensions et en faits.

2.4.1 Concept de fait

Une **table de faits** est la table centrale d'un modèle dimensionnel, où les mesures de performances sont stockées. Une ligne d'une table de faits correspond à une mesure. Ces mesures sont généralement des valeurs numériques, additives; cependant des mesures textuelles peuvent exister mais sont rares. Une table de faits assure les liens plusieurs à plusieurs entre les dimensions. Elles comportent des clés étrangères, qui ne sont autres que les clés primaires des tables de dimension.

2.4.2 Concept de dimension

Les **tables de dimension** sont les tables qui raccompagnent une table de faits, elles contiennent les descriptions textuelles de l'activité. Une table de dimension est constituée de nombreuses colonnes qui décrivent une ligne. C'est grâce à cette table que l'entrepôt de données est compréhensible et utilisable; elles permettent des analyses en tranches et en dés. Une dimension est généralement constituée : d'une clé artificielle, une clé naturelle et des attributs.

Bref, les faits stockent les indicateurs analysés et les dimensions contiennent la description des données analysées.

2.4.3 Comparaison entre les tables de faits et les tables de dimensions

Les différences entre la table de fait et celle des dimensions est illustrée dans le tableau ci après :

	Tables de faits	Tables de dimensions
Structure	Peu de colonnes beaucoup de lignes	Peu de lignes beaucoup de colonnes
Données	Mesurable, généralement numérique	Descriptives généralement textuelles
Référentiel	Plusieurs clés étrangères	Une clé primaire
Valeur	Prend de nombreuses valeurs	Plus ou moins constantes
Manipulation	Participe à des calculs	Participe à des contraintes
Signification	Valeurs de mesure	Descriptive
Rôle	Assure les relations entre les dimensions	Assure l'interface homme / entrepôt

Tableau 2.1: Comparaison entre les tables de faits et les tables de dimensions

2.5 Différents modèles de la modélisation dimensionnelle

Il existe différents types de modèles de données qui sont mis en œuvre dans les entrepôts de données et qui utilisent bien évidemment les notions de table de faits et de table de dimensions. Parmi ces modèles, on peut citer :

Le modèle en étoile : dans ce schéma, la table de fait est liée à toutes les tables de dimension.

Modèle en constellation : Ce n'est rien d'autre que plusieurs modèles en étoile liés entre eux par des dimensions communes. **Le modèle en flocon** : dans le schéma en flocon, la table de faits est toujours au cœur du modèle et les dimensions sont autour de la table centrale. Le modèle en flocon évite les redondances d'information mais nécessite des jointures. **C'est cette approche qui sera utiliser dans ce document.**

Une autre notion importante à définir lors de la modélisation dimensionnelle est la **granularité**. Elle se définit comme niveau de détail attendu par l'utilisateur au niveau des tables de faits. Autrement dit, c'est le niveau de finesse de la donnée. Par exemple, dans la dimension temps, le jour est plus fin que le mois, et l'heure est plus fin que le jour. Le choix revient donc à l'utilisateur du niveau de détail escompté.

Conception et développement de l'ETL

Les besoins des utilisateurs étant collectés, la méthodologie et les outils à utiliser présentés, il sera question dans cette partie, de présenter les résultats de la modélisation dimensionnelle et l'alimentation de l'entrepôt de données.

3.1 Modélisation

De l'analyse des besoins, deux principaux plans ressortent. Il s'agit de : *Statistiques ressources humaines* et *Activités ressources humaines*.

3.1.1 Statistiques ressources humaines

« Les ressources humaines sont l'ensemble des salariés de tous statuts (ouvriers, employés, cadres) appartenant à une organisation mais aussi – et de plus en plus – liés à elle par des rapports de sujétion (ainsi, les prestataires extérieurs ou sous-traitants sont considérés comme faisant partie de fait du périmètre des ressources humaines de l'entreprise) ».

Ce plan permettra de fournir toutes les informations concernant les ressources humaines. La mesure ici est le nombre ou l'effectif des ressources humaines. Cette mesure pourra être effectuée suivant les différents axes ou dimension ci après :

Type de tiers (Salarié, Stagiaire, Interimare)	Qualification	Type de Contrat
Situation Matrimoniale	Etablissement	Société
Service	Niveau	Emploi
Genre	Intitulé du contrat	Motif du contrat
Motif d'entrée	Motif de sortie	

Tableau 3.1: Dimension pour les Statistiques ressources humaines

3.1.2 Activités ressources humaines

Elles retracent les absences, les CET, les congés maladies, les congés travaux des ressources humaines. Il s'agit de mesurer pour chaque tiers la durée liées à ses différentes activités. Ces analyses peuvent être réalisés selon :

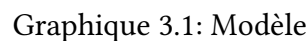
Type d'activité	Qualification	Type de Contrat
Situation Matrimoniale	Etablissement	Société
Service	Niveau	Emploi
Genre	Contrat	Motif du contrat
Motif d'entrée	Motif de sortie	Intitulé du contrat
Type de tiers (Salarié, Stagiaire, Interimaire)	Tiers	

Tableau 3.2: Dimension pour les Activités ressources humaines

Ainsi, nous avons deux faits nommés respectivement *fm_StatTiers* et *fm_CompteurTemps* pour les statistiques des ressources humaines et les Activités ressources humaines. La granularité adoptée est le **jour**.

3.1.3 Modèle en flocon

L'analyse sur ces deux plans nous a permis d'élaborer le modèle ci après :



Rapport de stage ★ ★

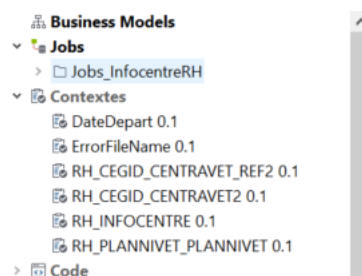
3.2 Développement et création de l'entrepôt de données

Comme expliqué dans la méthodologie, notre ETL de travail est Talend (TOS) qui fonctionne principalement avec des jobs. La création des jobs TOS suit généralement les étapes ci-après :

- Définition des variables de contexte
- Création des métadonnées
 - Création des connexions avec les sources (création des requêtes, récupération des schéma de table ...)
 - Création des fichiers (fichier délimité, fichier excel,...)
- Définition du job

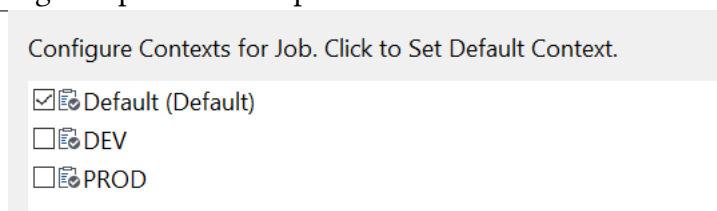
3.2.1 Définition des variables de contexte

La figure 3.2 illustre les différentes variables de contexte déclarées. Une variable de contexte est une variable définie par l'utilisateur pour un contexte particulier. Les contextes donnent la possibilité de gérer différemment vos Jobs selon les conditions d'utilisation.



Graphique 3.2: Contexte

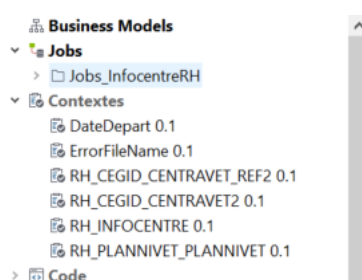
La variable de contexte *DateDepart* permet de définir la date à partir de laquelle les statistiques sont établies. La variable de contexte *ErrorFileName* permet de donner dynamiquement un nom au fichier d'erreur ou de succès renvoyé à l'administrateur chaque jour après l'exécution des jobs. Les variables de contexte *RH_CEGID_CENTRAVET_REF*, *RH_CEGID_CENTRAVET*, *RH_INFOCENTRE*, *RH_PLANNIVET_PLANNIVET* permettent de définir les paramètres de connexion aux bases de données lors des deux phases : Développement(DEV) et Production(PROD). Voir 3.3



Graphique 3.3: Context Dev/Prod

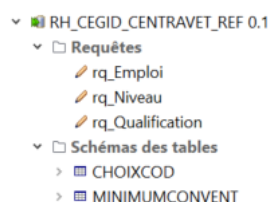
3.2.2 Métadonnées

La figure 3.4 présente les différentes connexions créées. Il s'agit de choisir le type de la BD et du contexte pour remplir les différentes informations de connection à la base.



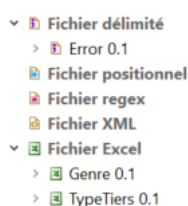
Graphique 3.4: Connexion

Pour chaque connexion, plusieurs requêtes ont été écrites pour récupérer exclusivement les informations utiles (attributs concernés) dans chaque source (Voir 3.5).



Graphique 3.5: Requête

Les informations non disponibles ou non explicites dans les sources et utiles pour l'infocentre ont été créées en dur (fichier délimité, fichier excel). Ce fut le cas pour les dimensions *Genre* et *Type tiers*.

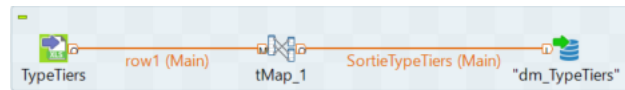


Graphique 3.6: Fichier Excel

3.2.3 Description des jobs

Dm_TypeTiers

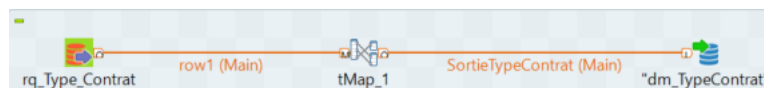
Le job **Dm_TypeTiers** est le job d'extraction et de chargement des différents type de tiers. Il a en entrée un fichier csv qui renseigne sur les types.



Graphique 3.7: Job Type Tiers

Dm_TypeContrat

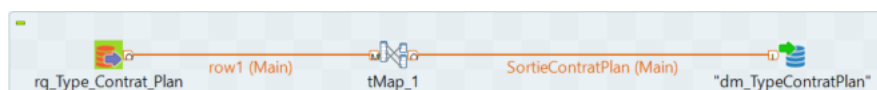
Ce job permet d'extraire les types de contrat utilisé dans la source CEGID et d'alimenter la dimension dm_TypeContrat dans l'infocentre. Il a en entrée la requête sur les types de contrats de CEGID.



Graphique 3.8: Job Type Contrat

Dm_TypeContrat_Plan

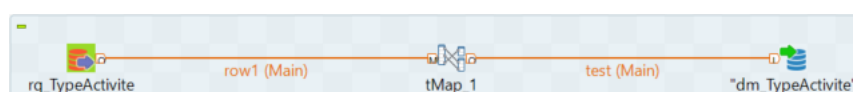
Ce job permet d'extraire les types de contrat utilisés dans PLANNIVET pour alimenter la dimension dm_TypeContratPlan dans l'infocentre.



Graphique 3.9: Job Type Contrat Plan

Dm_TypeActivité

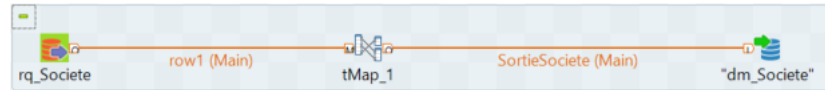
Le job **Dm_TypeActivité** permet d'alimenter la dimension dm_TypeActivité. Ses informations proviennent de PLANNIVET.



Graphique 3.10: Job Type Activités

Dm_Societe

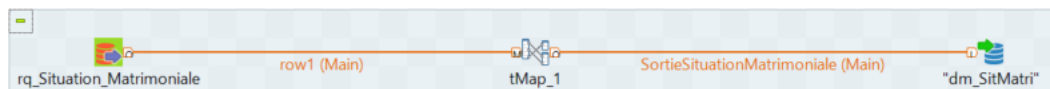
C'est le job qui permet d'alimenter la dimension dm_Societe . Il a entrée une requête sur la base PLANNIVET qui extrait les informations nécessaires.



Graphique 3.11: Job Société

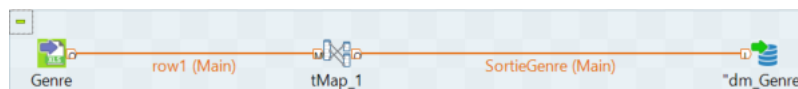
Dm_Situation_Matrimoniale

Ce job permet de renseigner les champs de la dimension dm_SituationMatrimoniale en puisant sa source par une requête sur PLANNIVET.



Graphique 3.12: Job Situation Matrimoniale

Dm_Genre



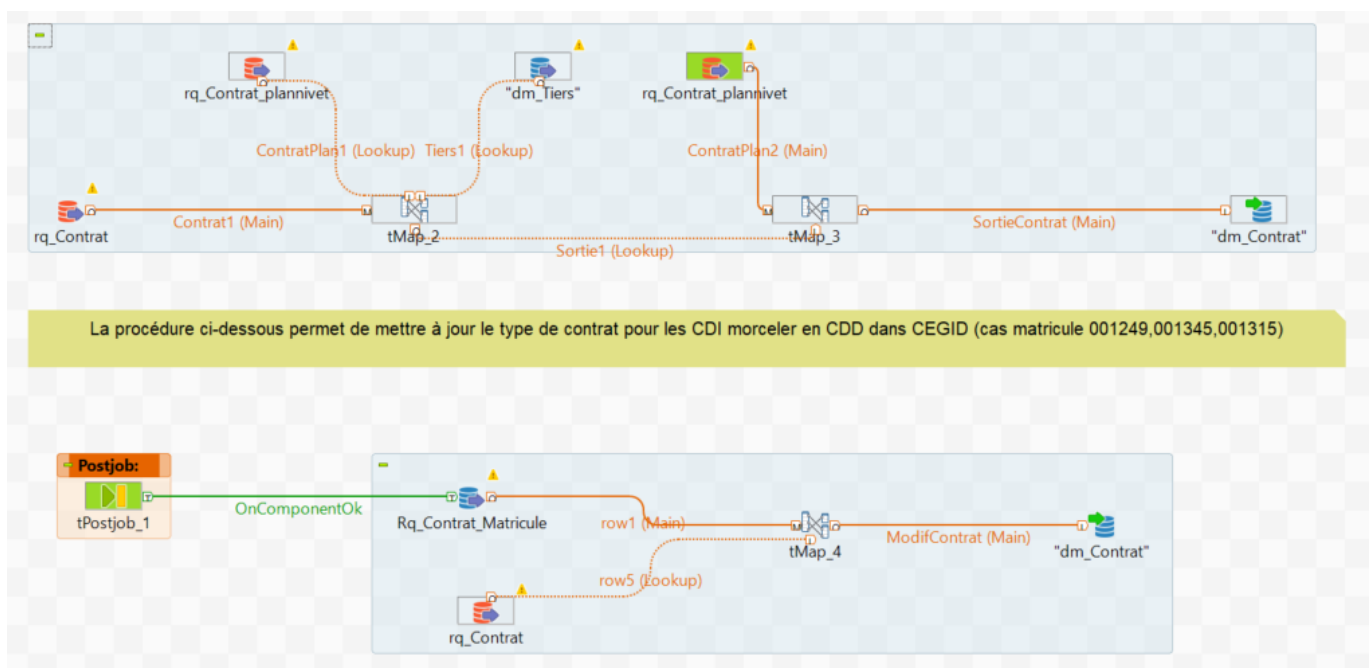
Graphique 3.13: Job Genre

Job_Principal

Ce job represente le menu . Il permet de synchroniser l'exécution des sous jobs et de procéder à la gestion des erreurs. Il est composé d'un tlogRow de lancement qui permet de signaler le début d'exécution et un tlogRow qui signale la fin de l'exécution. Ces deux tlogRow permettent d'avoir la durée de l'exécution. En plus de ces deux composant, on a un tlogRow associé à chaque sous jobs qui capte les erreurs d'exécutions. Une fois l'exécution terminé, le tLogCather récupère toutes les erreurs possibles s'ils existent, les stockent dans un fichier csv et les envoie par mail à l'administrateur ou sinon notifie par mail à l'administrateur la bonne marche de l'exécution.(Voir 3.27)

Dm_contrat

C' est la dimension qui permet de réunir les contrats de PLANNIVET et de CEGID dans une même table de l'infocentre : dm_Contrat. Dans une telle situation, on ne pourrait obtenir toutes ces informations par le biais d'une seule requête SQL car les sources sont distinctes. La première difficulté qui se présente est de trouver les éléments de liens entre les deux bases afin d'établir la correspondance entre un contrat de CEGID et un ou plusieurs autres de PLANNIVET. La date de début du contrat et le matricule semblent être une solution mais sont très tôt limités car il existe dans PLANNIVET des contrats groupés qui sont des contrats de CEGID morcelés en de petits CDD d'une semaine de durée dans PLANNIVET donc n'ayant pas les mêmes dates de début du contrat. Une autre solution est d'utiliser un between pour récupérer les contrats morcelés de PLANNIVET dont la date de début et fin sont contenues dans la date de début et de fin du contrat dans CEGID pour chaque matricule. Cette option n'existe malheureusement pas dans la version de Talend que nous utilisons.



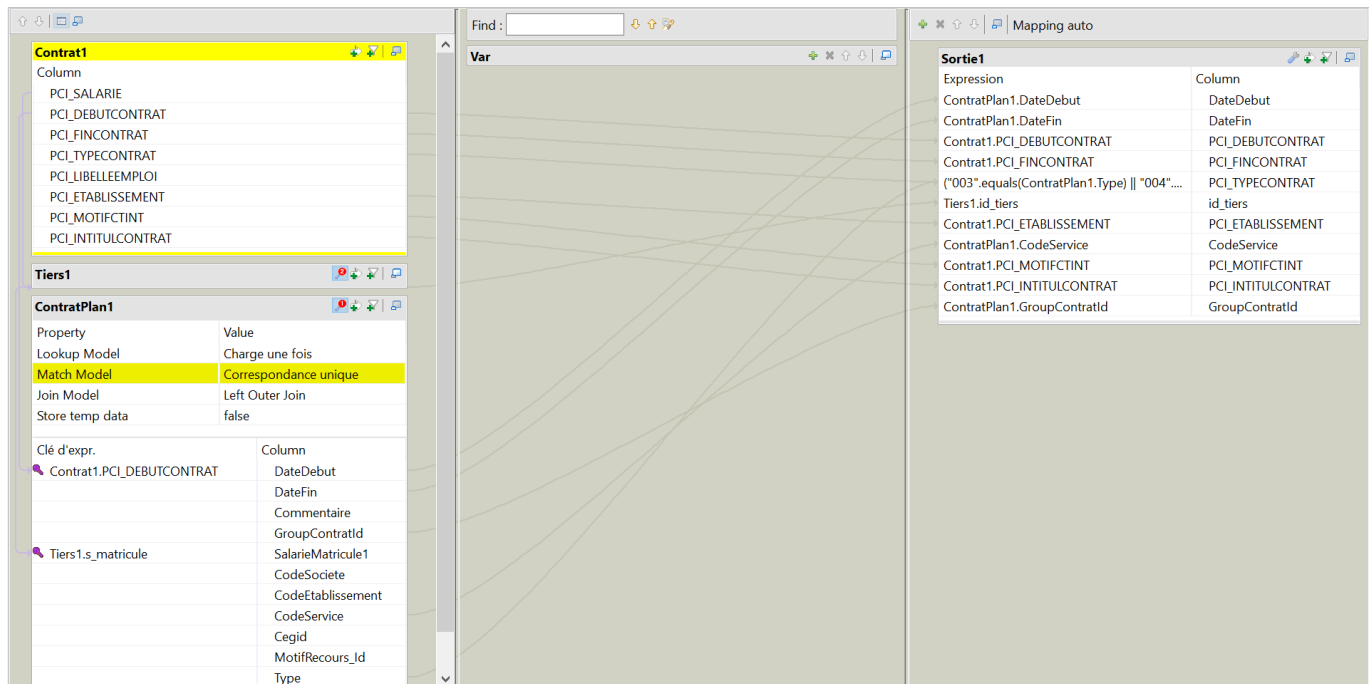
Graphique 3.14: Job Contrat

Pour contourner ce problème, nous avons procédé en deux étapes :

— 1^{re} étape tmap :

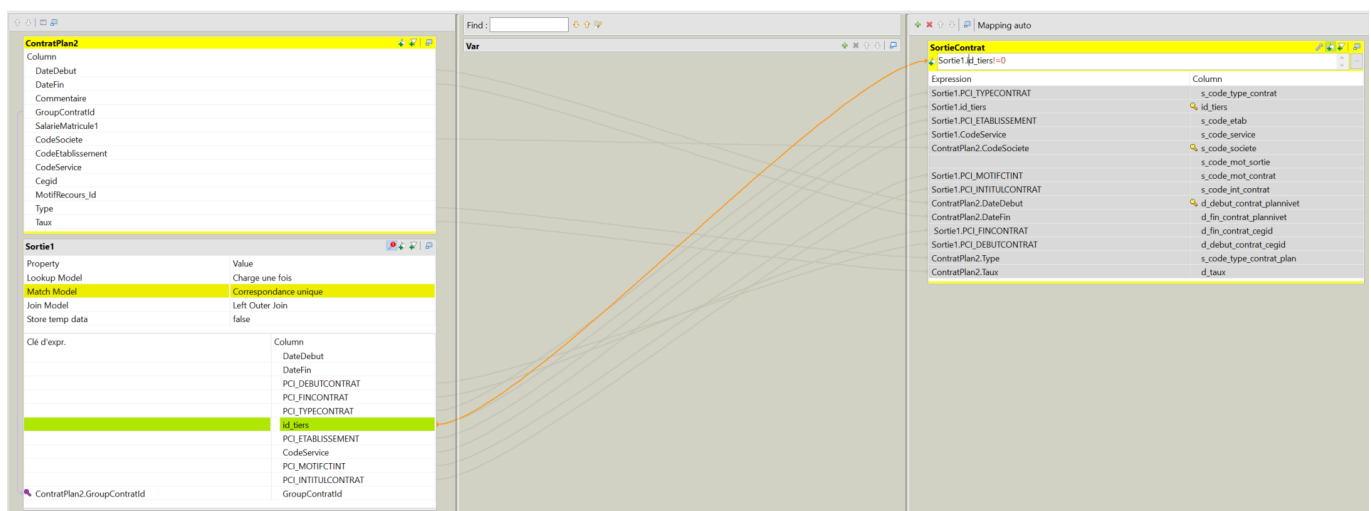
récupérer dans un premier temps pour chaque matricule la date du début du contrat dans CEGID dont nous cherchons la correspondance avec les contrats de PLANNIVET en faisant une jointure "Left Outer Join" sur la date de debut du contrat et du matricule, nous créons une première sortie (Sortie1) où nous récupérons pour le contrat PLANNIVET correspondant l'attribut "groupContratId" qui permet d'identifier les autres contrats de plannivet faisant partie du même groupe. l'id_tiers étant uniquement gérer dans l'in-

La Business Intelligence pour la conception et la réalisation des tableaux de bord BDES focentre, nous nous basons sur le matricule pour récupérer l'id_tiers correspondant.



Graphique 3.15: Job Contrat, tmap 2

- 2^e étape tmap : Dans la deuxième étape tmap, nous prenons tous les contrats de PLAN-NIVET auxquels nous rattachons le contrat CEGID correspondant en faisant un Left Outer Join en Correspondance unique entre la précédente sortie et la liste des contrats de PLANNIVET . Nous filtrons sur id_tiers!=0 pour ne récupérer que les tiers de Centravet car PLANNIVET enregistre également les contrats de GMV.

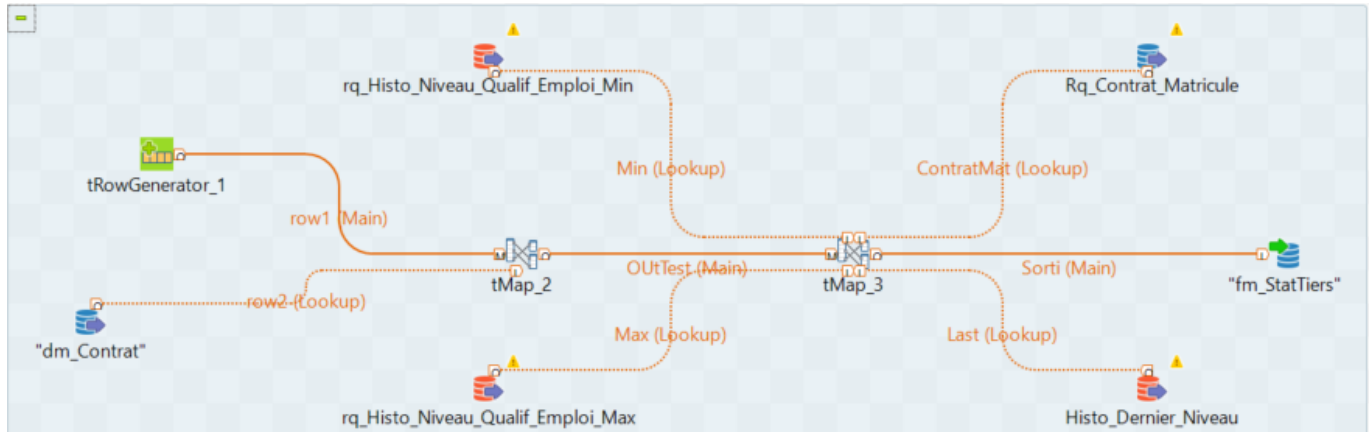


Graphique 3.16: Job Contrat, tmap 3

Une fois les données chargées dans l'infocentre, nous récupérerons tous les contrats chargés afin de procéder à une mise à jour du type de contrat pour certains contrats CDI de CEGID qui ont été morcelés dans CEGID . Il s'agit des contrats n'ayant pas les mêmes

dates de début de contrat en plannivet et en CEGID et qui ont soit les mêmes dates de fin dans les deux bases ou soit une date de fin null en PLANNIVET et égale à "01-01-1900" en CEGID.

fait_StatTiers



Graphique 3.17: Job Stat_Tiers

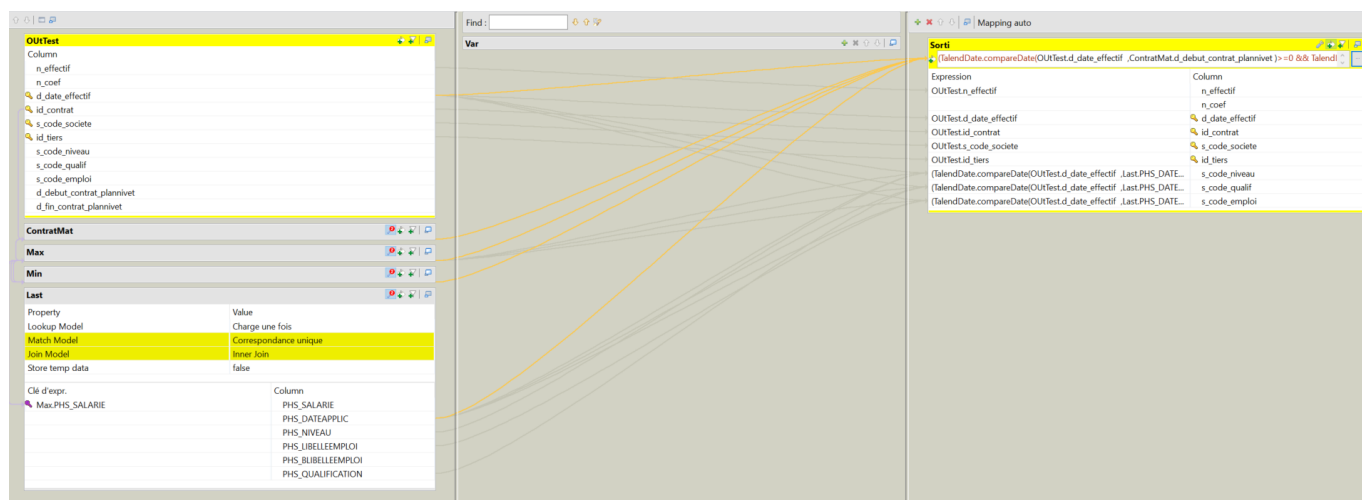
Le job Fait_StatTiers ci-dessus permet d'alimenter notre fait StatTiers. Son objectif principal est de donner pour chaque jour de la semaine depuis la date d'initialisation (ici 01-01-2019) tous les contrats en cours, les salariés présents, leur niveau, leur qualification, leur emploi et la société d'appartenance. Pour ce fait, nous avons utilisé un tRowgenerator qui génère tous les jours depuis la date d'initialisation et jusqu'au jour d'exécution du job. Avec un tmap, on réalise pour chaque jour un filtre pour générer autant de lignes qu'il y a de contrats dont la date de début et de fin contient le jour concerné. Les contrats utilisés dans ce cadre sont issus de l'infocentre (dm_contrat). On prend en compte dans le filtre les spécificités de chaque type de contrat (Voir 3.21).

```
((TalendDate.compareDate(row1.Date,row2.d_debut_contrat_plannivet )>=0 && TalendDate.compareDate
(row1.Date,TalendDate.getCurrentDate())<=0 && "CDI".equals(row2.s_code_type_contrat) &&
row2.d_fin_contrat_plannivet ==null)|| (TalendDate.compareDate(row1.Date,
row2.d_debut_contrat_plannivet )>=0 && TalendDate.compareDate
(row1.Date,row2.d_fin_contrat_plannivet ) <= 0 && "CCD".equals(row2.s_code_type_contrat)) ||
(TalendDate.compareDate(row1.Date, row2.d_debut_contrat_plannivet )>=0 && TalendDate.compareDate
(row1.Date,row2.d_fin_contrat_plannivet ) <= 0 && "CDI".equals(row2.s_code_type_contrat)
&&row2.d_fin_contrat_plannivet !=null))
```

Graphique 3.18: Formule filtre

Après cela nous avons procédé à la récupération de l'historique des contrats qui montre les différents changements de niveau ou de qualification d'un tiers au cours du temps. Ces données nous ont permis à chaque jour depuis le jour d'initialisation de récupérer le niveau et

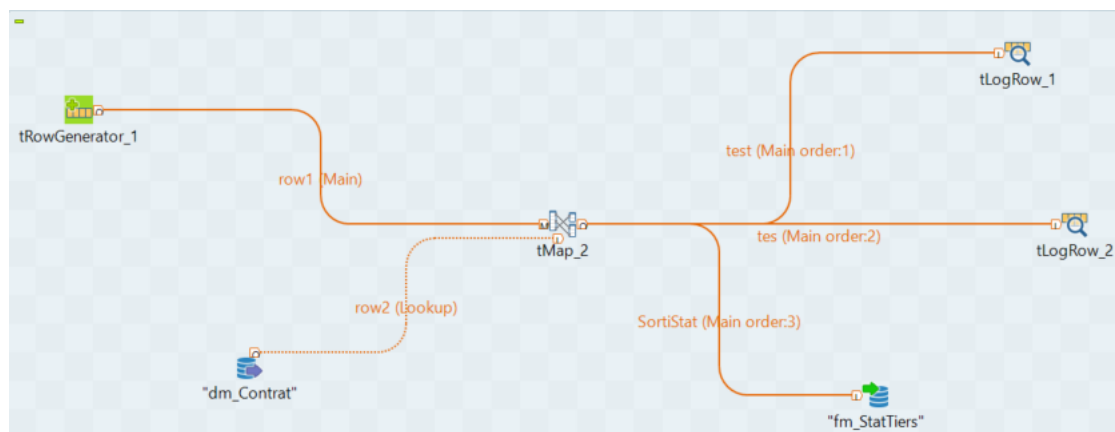
La Business Intelligence pour la conception et la réalisation des tableaux de bord BDES la qualification de chaque tiers. Pour cela, nous récupérerons pour chaque matricule dans l'historique les différents changement sur le niveau , la qualification et l'emploi en leur associant leur date de première application et de dernière application.



Graphique 3.19: tmap Stattiers

fait_CompteurTemps

Le job fait_CompteurTemps permet d'alimenter la table de fait fm_CompteurTemps. Dans ce job, on récupère toutes les informations liés aux activités des tiers par une requête sur PLANNIVET puis on associe à chaque activité le type d'activité, l'id_tiers ainsi que le contrat concerné (en se basant sur le matricule du tiers et la date de début du contrat). Un filtre est aussi réalisé pour exclure les tiers de GMV.



Graphique 3.20: tmap Stattiers

3.3 Restitution

Deux étapes ont suivies la mise en place de l'infocentre. Il s'agit de la création des vues et des index puis les tableaux de bords.

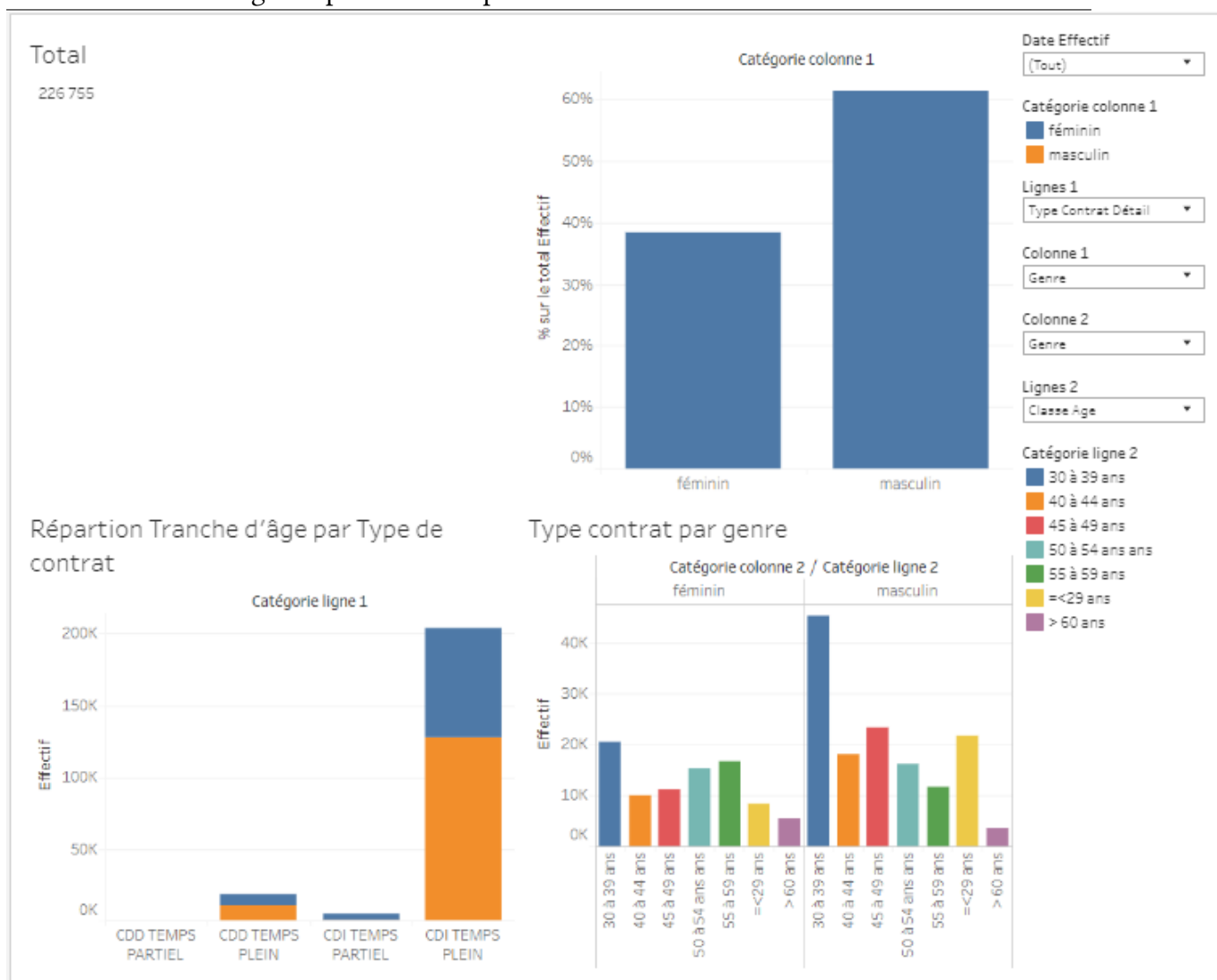
3.3.1 Création des vues et des index

Les vues représentent une sorte d'intermédiaire entre la base de données et l'utilisateur. Elles sont un grand avantage quand à la gestion des données, vis-à-vis de l'utilisateur final. En effet, elles permettent tout d'abord de simplifier la structure des tables, qui peuvent parfois comporter une multitude de colonnes. On pourra alors choisir, en fonction de l'utilisateur, les colonnes dont il aura besoin, et n'inclure que ces colonnes dans notre vue. L'atout majeur des vues : la sécurité d'accès aux données, très important pour les données RH. Il sera possible au travers de vues, de ne donner accès à un ou des utilisateurs particuliers, que les colonnes d'une table que nous voulons qu'il voit, et pas les autres. La requête de création de nos vues est présentée en annexe .

L'importance de la création des index est qu'elle nous permet d'accéder beaucoup plus rapidement aux enregistrements de la base de données pour les opérations de mise à jour (INSERT pour l'insertion de nouvelle(s) ligne(s) dans une table, DELETE pour la suppression et enfin UPDATE qui permet de modifier tout ou partie des valeurs des colonnes). Les index se présentent comme des catalyseurs de nos requêtes. Ils vont permettre sur le serveur, d'obtenir des temps de réponse de plus en plus courts, du fait que la requête n'aura pas à accéder et à parcourir la totalité des données dans les tables, mais accèdera à un nombre beaucoup moins conséquent de données dans notre index.

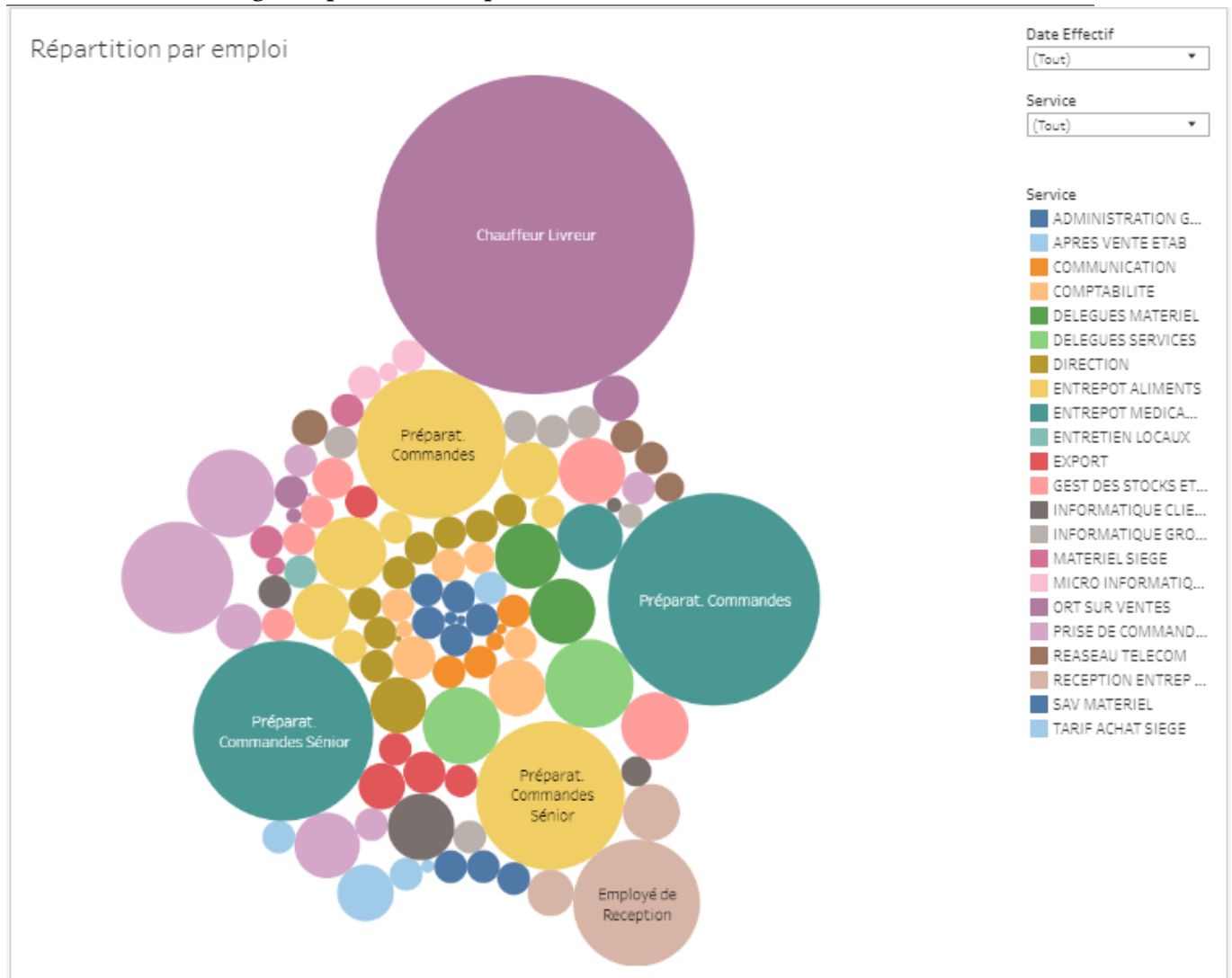
3.3.2 Tableaux de bords

C'est la partie de restitution des données aux métiers. Il s'agit de construire un ensemble d'indicateurs de pilotage, à l'intention des responsables, utile pour une prise de décision. Les figures ci dessous donnent quelques aperçus des indicateurs construits (les données présentées sont des données de tests).



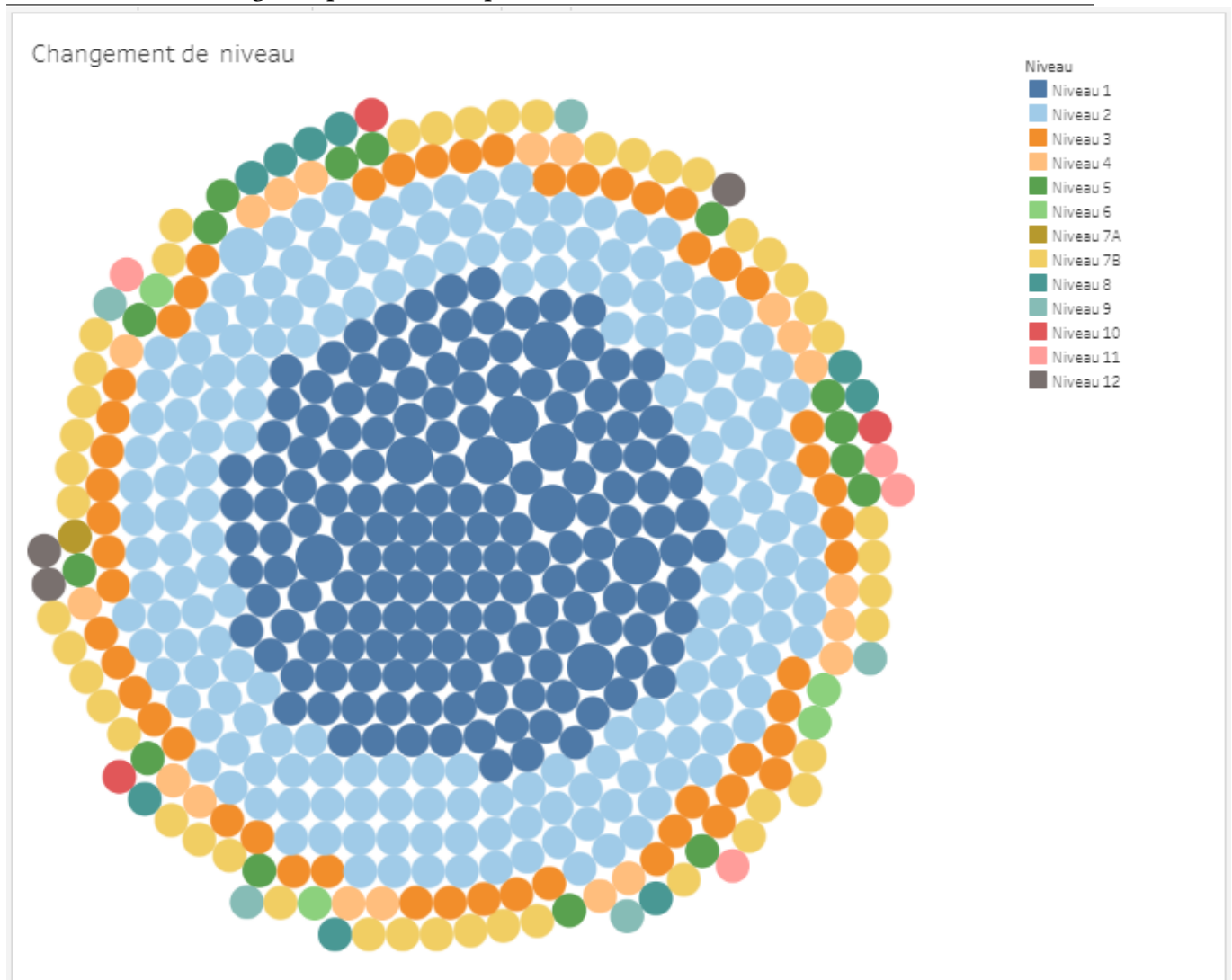
Graphique 3.21: Interface filtre dynamique

La figure 3.21 montre une interface de filtre dynamique. Elle permet à l'utilisateur d'inter-agir avec les graphiques pour calculer toutes sortes d'indicateurs .



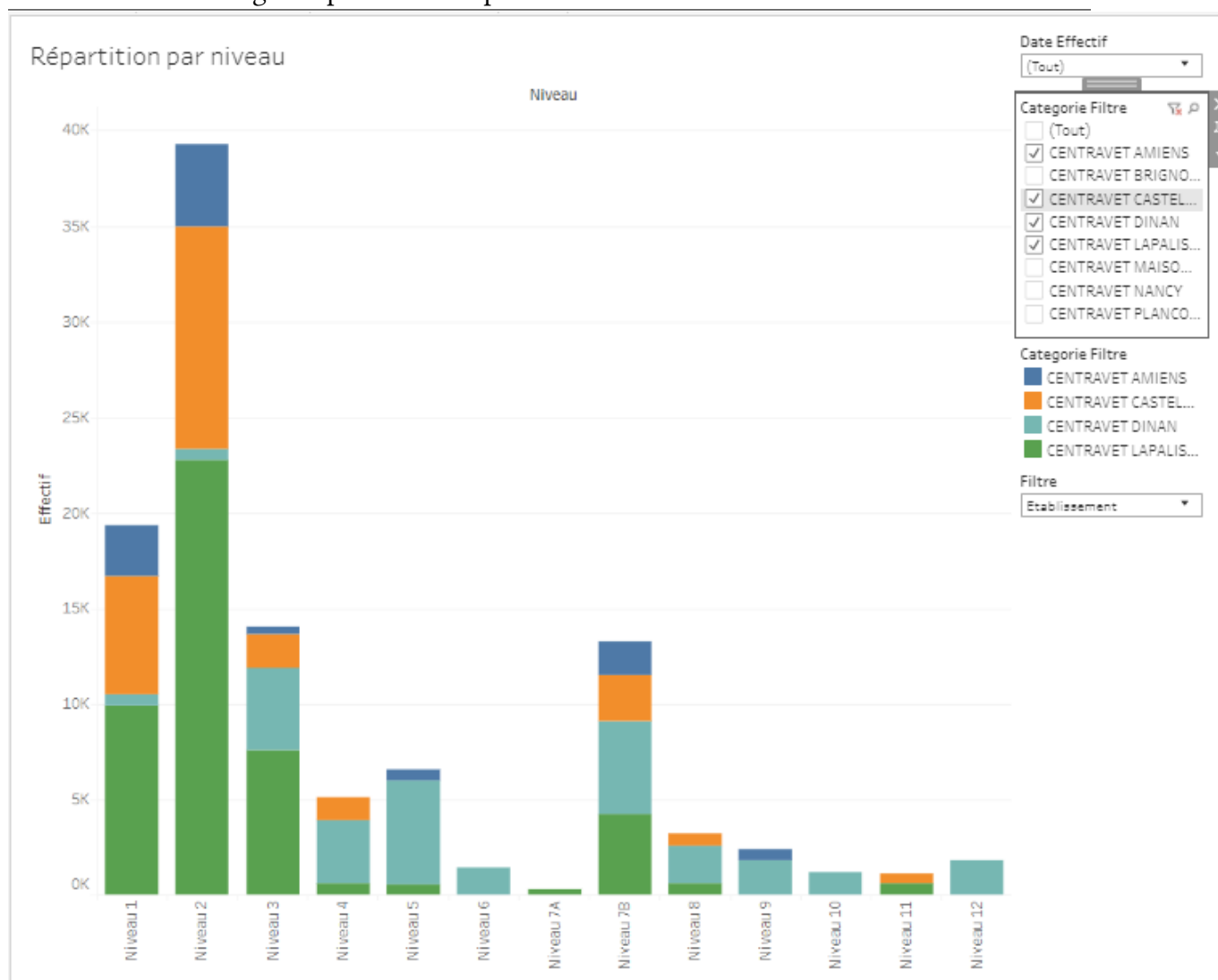
Graphique 3.22: Interface Emploi filtre dynamique

La figure 3.22 donne une interface pour présenter la répartition des emplois par service à une date donnée au sein de l'entreprise.



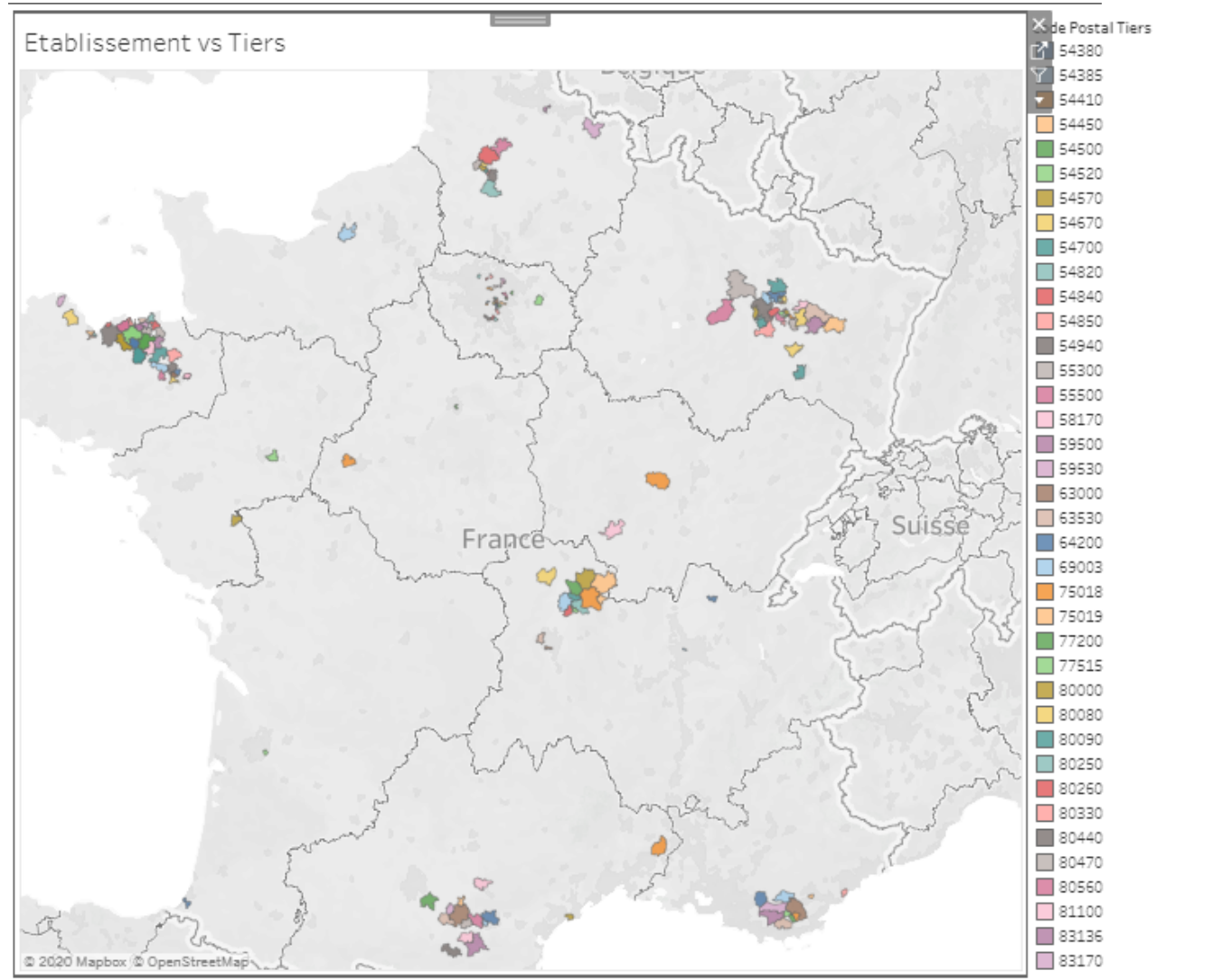
Graphique 3.23: Répartition des niveaux par poids

La figure 3.23 montre en proportion le nombre de différents niveaux par tiers .



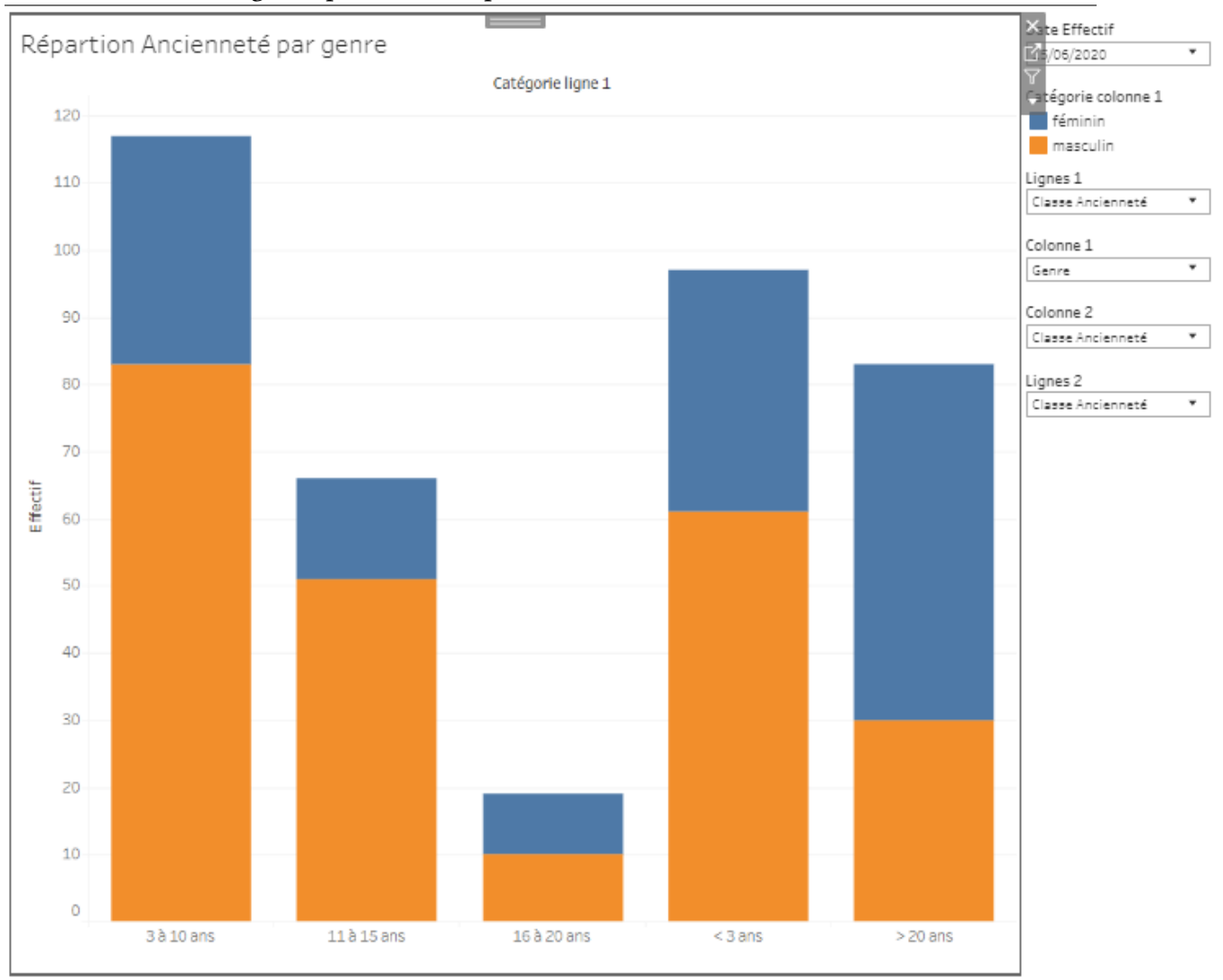
Graphique 3.24: Répartition des niveaux par Etablissement

La figure 3.24 présente la répartition des niveaux des tiers par établissement.



Graphique 3.25: Répartition géographique des Tiers autour des Etablissements

La figure 3.25 montre la répartition des salariés autour de leur établissement.



Graphique 3.26: Ancienneté des tiers par genre

La figure 3.26 donne un aperçu de la répartition de l'ancienneté des tiers par le genre.

Conclusion

Dans ce rapport, nous nous sommes intéressés à la mise en place d'une Base de Données Economiques et Sociales. Divers constats liés à l'existence et à la structure des données RH au sein de l'entreprise nous ont conduit à concevoir dans un premier temps un entrepôt de données selon l'approche de l'informatique décisionnelles puis à en extraire des données pour calculer des indicateurs utiles à la mise en place de Tableaux de Bords.

Au terme du stage, nous avons atteint l'objectif fixé dans le temps malgré les difficultés rencontrées notamment dans la compréhension des données disponibles, les règles de gestion associées et l'étape de la fouille de données. Il nous a fallu en conséquence un temps assez important pour comprendre la structure des données.

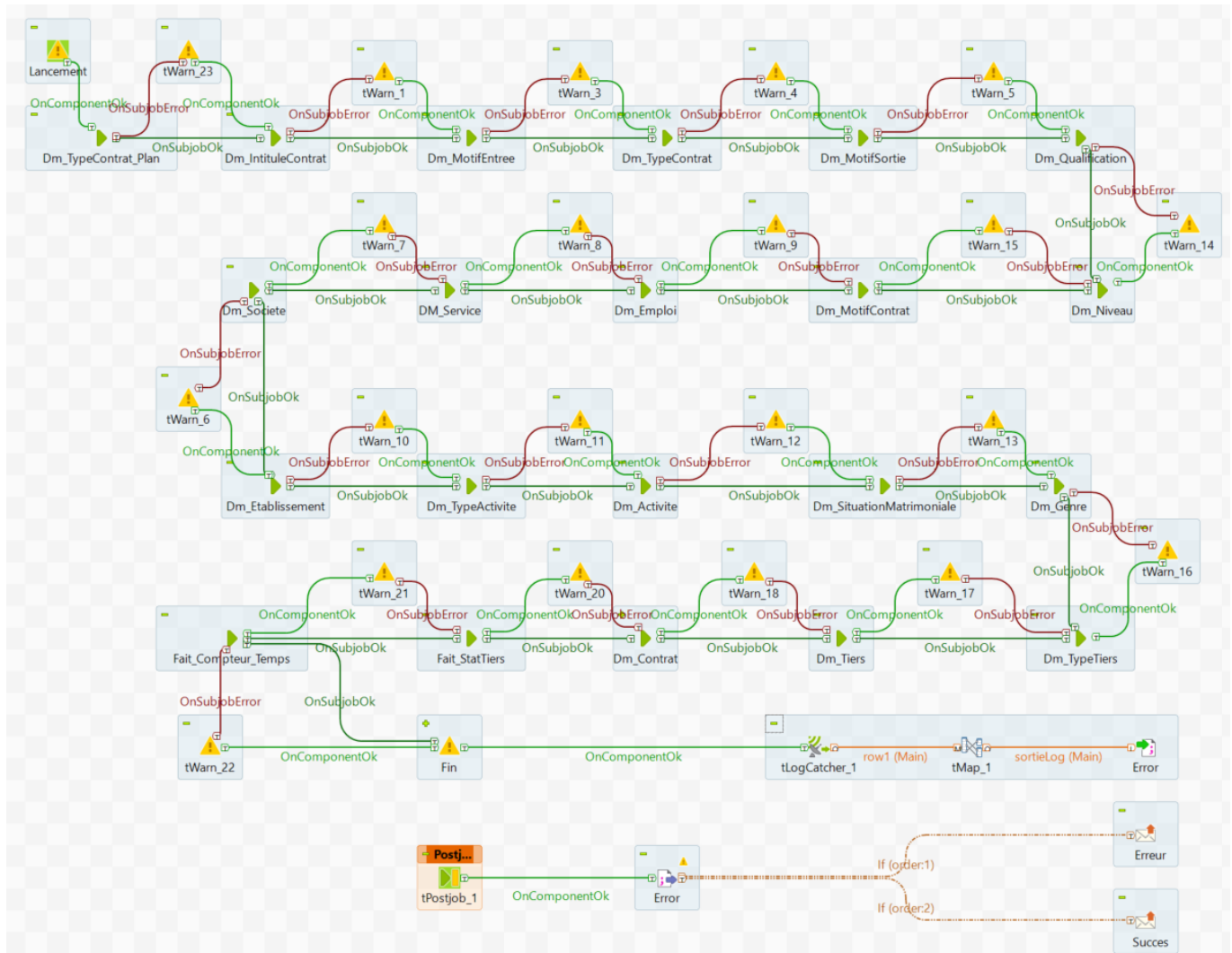
Le contenu du stage est toutefois assez intéressant car très recherché sur le marché de l'emploi et nous avons pu parcourir toutes les étapes d'un projet BI.

Bibliographie

- [MAR16] M KHALIS & A MARZAK. "Business Intelligence ETL (Extract, TransformLoad), Quelles limites?" In : *REVUE DE L'ENTREPRENEURIAT ET DE L'INNOVATION* Vol. 1.No. 1 (2016).
- [20a] "https://fr.wikipedia.org/wiki/Informatique_décisionnelle". Consulté le 12 Juin 2020.
- [20b] "<https://cooperative.centravet.net/Content.aspx?code=38775parent=38768> ". Consulté le 05 Juin 2020.
- [20c] "<https://www.supinfo.com/articles/single/4278-comprendre-l-informatique-decisionnelle-business-intelligence> ". Consulté le 21 Juillet 2020.
- [20d] "<https://stph.scenari-community.org/contribs/dwh/TalendOpenStudio/co/talend-open-studio9.html>". Consulté le 14 Juillet 2020.
- [20e] "<https://www.salesforce.com/fr/blog/2016/06/5-raisons-business-intelligence.html>". Consulté le 23 Juillet 2020.
- [20f] "<https://sites.google.com/site/lemondedecisionnel/home/sgbd/sql-serveur/les-vues-et-les-indexTOC-Pr-sentation-des-vues> ". Consulté le 23 Août 2020.
- [Pre20] Direction de l'information légale et administrative (PREMIER MINISTRE). "*Base de données économiques et sociales (BDES)*". <https://www.service-public.fr/professionnels-entreprises/vosdroits/F32193>. Vérifié le 11 février 2020, consulté le 2 Juin 2020.
- [KRA] Ines BEN KRAIEM. "Modélisation et intégration de données de capteurs/compteurs du SGE". In : ().
- [LAA] Brahim LAABOUBI. "Contribution à l'étude de la performance perçue de la Fonction Business Intelligence des grandes entreprises marocaines". In : ().

Annexes

Dictionnaire



Graphique 3.27: Job Menu

Dictionnaire des variables

Définition	Type objet T : Table V : Vue	Nom	Champs (PK) : PRIMARY KEY (FK) : FOREIGN KEYS	Type	Clé	Source	Bibliothèque	Table	Champ
Definition	T	dm_Qualification	s_code_qualif (PK)	varchar(17) NOT NULL	X	CEGID	CENTRAVET_REF	MINIMUMCO NVENT	[PMI_CODE] (où [PMI_NATURE] = 'QUA')
		dm_Qualification	s_libelle_categorie_sociale	varchar(35)					[PMI_LIBELLE],
Definition	T	dm_TypeContrat	s_code_type_contrat (PK)	varchar(3) NOT NULL	X	CEGID	CENTRAVET	COMMUN	CO_CODE (où [CO_TYPE] = 'PCT')
		dm_TypeContrat	s_libelle_type_contrat	varchar(105)					[CO_LIBELLE]
Definition	T	dm_SitMatri	code_sit_mat (PK)	varchar(3)	X	CEGID	CENTRAVET	COMMUN	Créer en dur, fichier excel
		dm_SitMatri	s_libelle_sit_mat	varchar(35)					
Definition	T	dm_Etablissement	s_code_etab (PK)	varchar(3) NOT NULL	X	CEGID	CENTRAVET	ETABLISS	[ET_ETABLISSEMENT]
		dm_Etablissement	s_libelle_etab	varchar(35)					[ET_LIBELLE]
		dm_Etablissement	s_adresse1	varchar(35)					[ET_ADRESSE1]
		dm_Etablissement	s_adresse2	varchar(35)					[ET_ADRESSE2]
		dm_Etablissement	s_adresse3	varchar(35)					[ET_ADRESSE3]
		dm_Etablissement	s_code_post	varchar(9)					[ET_CODEPOSTAL]
		dm_Etablissement	s_ville	varchar(35)					[ET_VILLE]
		dm_Etablissement	s_pays	varchar(3)					[ET_PAYS]
		dm_Etablissement	s_tel	varchar(25)					[ET_TELEPHONE]
		dm_Etablissement	s_code_societe (FK)	varchar(7)					
Definition	T	dm_Societe	s_code_societe (PK)	varchar(7) NOT NULL	X	PLANNIVET	PLANNIVET	Societe	[CodeSociete]
		dm_Societe	s_libelle_societe	varchar(15)					[LibelleSociete]
Definition	T	dm_Service	s_code_service (PK)	varchar(3) NOT NULL	X	CEGID	CENTRAVET	CHOIXCOD	[CC_CODE] où [CC_TYPE] = 'PUN'
		dm_Service	s_libelle_service	varchar(105)					[CC_LIBELLE]
Definition	T	dm_TypeTiers	s_code_type_tiers (PK)	varchar(7) NOT NULL	X				Créer en dur, fichier excel
		dm_TypeTiers	s_libelle_type_tiers	varchar(15)					
Definition	T	dm_Niveau	s_code_niveau (PK)	varchar(17) NOT NULL	X	CEGID	CENTRAVET_REF	MINIMUMCO NVENT	[CC_CODE]
		dm_Niveau	s_libelle_niveau	varchar(35)					[CC_LIBELLE]
Definition	T	dm_Emploi	s_code_emploi(PK)	varchar(3)	X	CEGID	CENTRAVET_REF	CHOIXCOD	[CC_CODE] (où [CC_TYPE] = 'PLE')
		dm_Emploi	s_libelle_emploi	varchar(105)					[CC_LIBELLE]
Definition	T	dm_Genre	s_code_genre (PK)	varchar(3) NOT NULL	X				Créer en dur, fichier excel
		dm_Genre	s_libelle_genre	varchar(10)					
Definition	T	dm_TypeActivite	s_code_type_act (PK)	varchar NOT NULL	X	PLANNIVET	PLANNIVET	TypeActivite	
		dm_TypeActivite	s_libelle_type_act	varchar(105)					distinct [Type] de [TypeActivite]
Definition	T	dm_Activite	s_code_activite (PK)	varchar(10) NOT NULL	X	PLANNIVET	PLANNIVET	TypeActivite	[Code]
		dm_Activite	s_libelle_activite	varchar(60)					[Libelle]

		dm_Activite	s_code_cegid	varchar(10)					[CodeCegid]
		dm_Activite	id_type_act (FK)	integer					
Definition	T	fm_CompteurTemps	id_fm_compteur (PK)	bigint NOT NULL GENERATED ALWAYS AS IDENTITY	X	PLANNIVET			
		fm_CompteurTemps	d_date_compteur	date					
		fm_CompteurTemps	n_heure_abs	decimal(10)					
		fm_CompteurTemps	n_heure_normale	decimal(10)					
		fm_CompteurTemps	n_jour_abs	decimal(10)					
		fm_CompteurTemps	id_contrat (FK)	bigint					
		fm_CompteurTemps	s_code_activite (FK)	varchar					
		fm_CompteurTemps	id_tiers (FK)	bigint					
Definition	T	fm_StatTiers	id_fm_stat_tiers (PK)	bigint NOT NULL GENERATED ALWAYS AS IDENTITY	X	CEGID			
		fm_StatTiers	n_effectif	integer					
		fm_StatTiers	d_date_effectif	date					
		fm_StatTiers	id_contrat (FK)	bigint					
		fm_StatTiers	s_code_societe (FK)	varchar(7)					
		fm_StatTiers	id_tiers (FK)	bigint					
		fm_StatTiers	s_code_niveau (FK)	varchar(17)					
		fm_StatTiers	s_code_emploi (FK)	varchar(3)					
		fm_StatTiers	s_code_qualif (FK)	varchar(17)					
Definition	T	dm_Tiers	id_tiers (PK)	bigint NOT NULL GENERATED ALWAYS AS IDENTITY	X	CEGID	CENTRAVET	SALARIES	
		dm_Tiers	s_matricule	varchar(17)					[PSA_SALARIE]
		dm_Tiers	n_num_carte_sejr	varchar(15)					[PSA_NUMERO SS]
		dm_Tiers	s_pays_naiss	varchar(3)					[PSA_PAYSSNAI SSANCE]
		dm_Tiers	s_civilite	varchar(3)					[PSA_CIVILITE]
		dm_Tiers	s_nomjf	varchar(35)					[PSA_NOMJF]
		dm_Tiers	s_pays	varchar(3)					[PSA_PAYS]
		dm_Tiers	s_nom_usage	varchar(35)					[PSA_LIBELLE]
		dm_Tiers	s_prenom	varchar(35)					[PSA_PRENOM]
		dm_Tiers	d_date_naissance	date					[PSA_DATENAI SSANCE]
		dm_Tiers	s_telephone_fixe	varchar(35)					[PSA_PAYS]
		dm_Tiers	s_telephone_mobile	varchar(35)					[PSA_TELEPHO NE]
		dm_Tiers	s_adresse1	varchar(35)					[PSA_ADRESS E1]
		dm_Tiers	s_adresse2	varchar(35)					[PSA_ADRESS E2]
		dm_Tiers	s_adresse3	varchar(35)					[PSA_ADRESS E3]
		dm_Tiers	s_code_post	varchar(9)					[PSA_CODEPO STAL]
		dm_Tiers	s_ville	varchar(35)					[PSA_VILLE]
		dm_Tiers	s_num_ss	varchar(20)					[PSA_NUMERO SS]
		dm_Tiers	d_date_exp_sejour	date					[PSA_DATEXPI RSEJOUR]
		dm_Tiers	d_date_entree	date					[PSA_DATEENT REE]

		dm_Tiers	s_motif_sortie	varchar(3)					[PSA_MOTIFSO RTIE]
		dm_Tiers	s_motif_sus_paie	varchar(3)					[PSA_MOTIFSU SPPAIE]
		dm_Tiers	s_delivre_par	varchar(35)					[PSA_DELIVPA R]
		dm_Tiers	s_code_sit_mat (FK)	varchar(3)					
		dm_Tiers	s_code_genre (FK)	varchar(3)					
		dm_Tiers	s_code_type_tiers (FK)	varchar					
		dm_Tiers	s_code_societe (FK)	varchar(7)					
		dm_Tiers	s_code_mot_entree (FK)	varchar(3)					
Definition	T	dm_Contrat	id_contrat (PK)	bigint NOT NULL GENERATED ALWAYS AS IDENTITY	X	CEGID & PLANNIVET	CENTRAVET	CONTRATT RAVAIL (CEGID) & contrat (PLANNIVET)	
		dm_Contrat	d_debut_contrat_planni vet	date		PLANNIVET			
		dm_Contrat	d_fin_contrat_plannivet	date		PLANNIVET			
		dm_Contrat	d_debut_contrat_cegid	date		CEGID			
		dm_Contrat	d_fin_contrat_cegid	date		CEGID			
		dm_Contrat	d_taux	decimal(3)					
		dm_Contrat	s_code_type_contrat (FK)	varchar(3)					
		dm_Contrat	id_tiers (FK)	bigint					
		dm_Contrat	s_code_etab (FK)	varchar(3)					
		dm_Contrat	s_code_service (FK)	varchar(3)					
		dm_Contrat	s_code_societe (FK)	varchar(7)					
		dm_Contrat	s_code_mot_sortie (FK)	varchar(3)					
		dm_Contrat	s_code_mot_contrat (FK)	varchar(7)					
		dm_Contrat	s_code_int_contrat (FK)	varchar(3)					
		dm_Contrat	s_code_type_contrat_pl an (FK)	varchar(4)					
Definition	T	dm_Motif_Contrat	s_code_mot_contrat (PK)	varchar(7) NOT NULL	X	CEGID	CENTRAVET	CHOIXCOD	
		dm_Motif_Contrat	s_libelle_mot_contrat	varchar(105)					
Definition	T	dm_Motif_Entree	s_code_mot_entree (PK)	varchar(3) NOT NULL	X	CEGID	CENTRAVET	MOTIFENTR EESAL	
		dm_Motif_Entree	s_libelle_mot_entree	varchar(75)					
Definition	T	dm_Motif_Sortie	s_code_mot_sortie (PK)	varchar(3) NOT NULL	X	CEGID	Non identifié	Non identifié	
		dm_Motif_Sortie	s_libelle_mot_sortie	varchar(105)					
Definition	T	dm_Intitule_Contrat	s_code_int_contrat (PK)	varchar(3)	X	CEGID	CENTRAVET	COMMUN	
		dm_Intitule_Contrat	s_libelle_int_contrat (PK)	varchar(105)					
Definition		dm_TypeContratPlan	s_code_type_contrat_pl an	varchar(4)	X	CEGID	CENTRAVET	CHOIXCOD	
		dm_TypeContratPlan	s_libelle_type_contrat_ plan	s_libelle_type_contra t_plan(105)					

