

## **Dédicace**

De plus profond de mon cœur, je dédie ce projet de fin d'études

### **Mes très chers parents Anoir & rim :**

Tous les mots du monde ne sauraient exprimer l'immense amour que je vous porte, ni la profonde gratitude que je vous témoigne pour votre affection, votre soutien et les valeurs que vous m'inculquez toujours. Merci, du plus profond de mon cœur, pour tout ce que vous avez fait pour moi. Sans vous, je ne serais pas la femme que je suis aujourd'hui.

### **Mes chers Frères Amin & Melek**

Pour leur amour, leur soutien, et leur écoute en cas de besoins. Je vous souhaite un avenir radieux plein de succès et de bonheur.

### **Mon cher fiancé Omar**

Mon amour mon meilleur ami et mon partenaire de vie je te remercie du fond du cœur pour ton amour et ton soutien inébranlables. Ta présence illumine ma vie et me donne la force de poursuivre mes rêves. Je suis reconnaissante de t'avoir à mes côtés. Merci mon cœur.

### **Ma précieuse binôme de travail Lina BEN HAMADOU**

Je tiens à exprimer ma sincère gratitude pour ta collaboration tout au long de notre projet je suis reconnaissante d'avoir pu partager cette expérience avec toi.

### **A mes amis**

Vous étiez à mes côtés tout au long de ce chemin, je suis très reconnaissante de faire votre connaissance, c'est un vrai honneur pour moi. Merci pour vos aides et vos conseils aux moments de faiblesse. Que dieu vous garde pour moi. Je vous adore.

## **Dédicace**

De plus profond de mon cœur, je dédie ce projet de fin d'études

### **Mes très chers parents Meher & Hayet**

Autant de phrases et d'expressions aussi éloquentes soient-elles ne sauraient exprimer ma gratitude et ma reconnaissance. Vous avez su m'inculquer le sens de la responsabilité, de l'optimisme et de la confiance en soi face aux difficultés de la vie. Je vous dois ce que je suis aujourd'hui et ce que je serai demain et je ferai toujours de mon mieux pour rester votre fierté et ne jamais vous décevoir.

### **Ma adorable sœur Sahar**

En signe de l'affection et du grand amour que je porte à toi, les mots sont insuffisants pour exprimer ma profonde estime.

### **Ma précieuse binôme de travail Eya LOUKIL**

Je tiens à te remercier du fond du cœur pour tout ce que tu as apporté à notre duo. Tu es bien tu apportes non seulement tes compétences exceptionnelles, mais aussi ton énergie positive et ta volonté de repousser les limites. Tu es plus qu'un simple binôme de travail pour moi.

### **A mes amis :**

En témoignage de ma profonde tendresse et reconnaissance, je vous souhaite une vie pleine de bonheur et de succès. A tous mes professeurs qui m'ont enseigné et contribué à ma formation pour monter l'escalier de la réussite, tout au long de mon parcours primaire, secondaire et universitaire.

## Remerciement

Avant de présenter notre travail, nous réservent ces quelques lignes pour exprimer mes sincères remerciements et nos gratitudes aux personnes qui ont apporté leur support et leur aide durant notre stage et qui ont contribué au succès de ce travail.

Nous tenons, tout d'abord, à présenter mes sincères remerciements à Monsieur **Anis JEDIDI**, notre encadrant à l'Institut Supérieur d'Informatique et de Multimédia de Sfax, pour sa disponibilité et ses remarques constructives et précieuses qui ont donné une valeur ajoutée à notre travail.

Nous tenons, aussi à manifester l'expression de notre respect le plus distingué et notre reconnaissance à Madame **Saida EDHIB** et monsieur **Aymen ZOUARI** nos encadrants à CliniSys, pour leur disponibilité, leurs conseils et l'aide qu'ils nous ont accordé et surtout pour leur patience interminable, qui a été déterminante dans l'élaboration de ce projet.

Nous considérons aussi remercier les membres du jury Madame **Raida KTARI** et Madame **Wafa NEIFAR** qui nous ont fait l'honneur d'examiner et d'évaluer notre modeste travail tout en espérant qu'elles trouveront dans ce rapport les qualités de motivation et de clarté qu'elles espèrent.

Nous exprimons également notre profonde gratitude et notre reconnaissance au corps professoral de l'Institut Supérieur d'Informatique et de Multimédia de Sfax pour la qualité de formation qu'ils nous ont offerte. En effet, nous avons entamé notre première expérience professionnelle avec des bases solides grâce à leurs contributions et leurs conseils.

# Sommaire

<b>Introduction générale</b> .....	1
<b>Glossaire</b> .....	3
<b>Chapitre 1 : Etude Préalable :</b> .....	4
1. Introduction : .....	5
2. Présentation de la société d'accueil : .....	5
3. Définition de la mission : .....	6
3.1 Présentation du projet : .....	6
3.2 Objectifs à atteindre : .....	6
4. Etude de l'existant et problématique : .....	7
4.1 Étude de l'existant : .....	7
4.2 Analyse de l'existant : .....	7
4.3 Critique de l'existant : .....	9
5. Solution proposée : .....	9
6. Définition des besoins fonctionnels et non fonctionnels : .....	10
6.1 besoins fonctionnels : .....	10
6.2 besoins non fonctionnels : .....	10
7. Planning Prévisionnel : .....	11
8. Conclusion : .....	11
<b>Chapitre 2 : Informatique décisionnelle :</b> .....	12
1. Introduction : .....	13
2. Architecture d'un système décisionnel : .....	13
3. Entrepôt de données et magasins de données : .....	14
3.1 Définition d'un entrepôt de données : .....	14
3.2 Les caractéristiques d'un entrepôt de données : .....	14
3.3 Magasins de données ou Data Marts : .....	15
4. Modélisation des entrepôts de données : .....	16
4.1 Définition de la modélisation multidimensionnelle : .....	16
4.2 Concepts de base de la modélisation multidimensionnelle : .....	16
4.3 Les schémas multidimensionnels : .....	18
5. L'alimentation de l'entrepôt de données : .....	20
6. Modélisation logique : .....	21
7. Gestion de projet décisionnel : .....	22
8. Langage DAX : .....	22
9. Conclusion : .....	23
<b>Chapitre 3 : Modélisation conceptuelle :</b> .....	24

1. Introduction :	25
2. Processus de modélisation d'un ED :	25
3. Méthodes de conception d'un ED :	25
4. Modélisation conceptuelle :	26
4.1 Méthode descendante :	26
4.1.1 Collecte des données :	26
4.1.2 Spécification des besoins :	27
4.1.3 Formalisation des besoins :	29
4.2 Méthode ascendante :	33
4.2.1 Détermination des faits :	33
4.2.2 Détermination des dimensions :	35
4.2.3 Hiérarchisation des dimensions :	37
4.2.4 Schéma multidimensionnel en constellation :	38
4.2.5 Méthode mixte :	39
5. Conclusion :	41
<b>Chapitre 4 : Construction de l'entrepôt de données :</b>	<b>42</b>
1. Introduction :	43
2. Outils utilisés :	43
2.1 Microsoft SQL Server Intégration Services :	43
2.2 Adobe Illustrator :	43
2.3 Microsoft SQL Server Management Studio 19 :	44
2.4 Visual Studio 2019:	44
2.5 Microsoft Excel:	44
2.6 Microsoft Power BI :	44
2.7 Entreprise Architect :	45
2.8 Redmine :	45
3. Conception détaillé de l'ETL :	45
4. Développement d'ETL :	47
Conclusion :	54
<b>Chapitre 5 : Restitution des données de l'ED :</b>	<b>55</b>
1. Introduction :	56
2. Modélisation de l'application BI :	56
2.1 Type d'application BI :	56
2.2 Type d'analyse BI :	56
3. Développement de l'application :	57
3.1 Projet Power BI :	57
3.2 Présentation de l'application :	59

<b>Conclusion générale :</b> .....	66
<b>Bibliographie :</b> .....	67
<b>Netographie :</b> .....	68

## Liste des figures

Figure 1: Organigramme général de la société CLINISYS. ....	6
Figure 2 : Tableaux de bord support.....	8
Figure 3 : Tableaux de bord Redmine. ....	8
Figure 4 : Planning prévisionnel.....	11
Figure 5 : Architecture des systèmes décisionnels. ....	13
Figure 6 : Données orientées sujet de l'entrepôt de données .....	14
Figure 7 : Données intégrées de l'entrepôt de données .....	15
Figure 8 : Données non volatiles et historisées de l'entrepôt de données .....	15
Figure 9 : Entrepôt de données et magasins de données .....	16
Figure 10 : Exemple d'un fait.....	17
Figure 11 : Exemples de dimensions .....	17
Figure 12 : Exemple d'un schéma en étoile.....	18
Figure 13 : Exemple d'un schéma en flocon de neige.....	19
Figure 14 : Exemple d'un schéma en constellation [Golfarelli 2009] .....	19
Figure 15 : Approche relationnelle (R-OLAP) .....	21
Figure 16 : Approche multidimensionnelle (M-OLAP) .....	21
Figure 17 : Approche hybride (H-OLAP) .....	22
Figure 18 : Schéma de cycle de vie dimensionnel [Kimball 2008] .....	22
Figure 19 : Processus de modélisation d'un entrepôt de données .....	25
Figure 20 : Les étapes de la modélisation descendante .....	26
Figure 21 : Schéma multidimensionnel en étoile (démarche descendante).....	32
Figure 22 : Les étapes de la modélisation ascendante .....	33
Figure 23 : Diagramme de classe de la source de données .....	34
Figure 24 : Graphes de dépendances entre classes .....	36
Figure 25 : Schéma multidimensionnel en constellation (démarche ascendante) .....	39
Figure 26 : Schéma multidimensionnel en constellation (démarche Mixte) .....	40
Figure 27 : Logo Microsoft SQL Server Intégration Services.....	43
Figure 28 : Logo Adobe Illustrator.....	43
Figure 29 : Logo Microsoft SQL Server Management Studio 19 .....	44
Figure 30 :Logo Visual Studio 2019 .....	44
Figure 31 :Logo Microsoft Excel .....	44
Figure 32 :Logo Microsoft Power BI .....	44

Figure 33 :Logo Entreprise Architect .....	45
Figure 34 :Logo Redmine.....	45
Figure 35 : Diagramme d'activités d'alimentation des tables de dimensions et de faits ....	46
Figure 36 : Tables de l'entrepôt de données.....	47
Figure 37 : Le projet SSIS. ....	47
Figure 38 :1er Étape de configuration de la connexion PostgreSQL. ....	48
Figure 39 :2eme Étape de configuration de la connexion PostgreSQL.....	48
Figure 40 :2eme Étape de configuration de la connexion PostgreSQL.....	49
Figure 41 : Étapes de configuration de la connexion PostgreSQL.....	49
Figure 42 : Gestionnaires de connexions.....	50
Figure 43 : Flux de contrôle du paquet « dimensions ».....	50
Figure 44 : Flux de contrôle du paquet « dimensions ».....	50
Figure 45 : Flux de données relative à la dimension « Tache » .....	51
Figure 46 : Chargement de données de la source Support. ....	52
Figure 47 : Destination de la dimension Tache. ....	53
Figure 48 : Flux de données « Fait_Suivi_Emploi » .....	53
Figure 49 : Chargement de données de la source Redmine.....	54
Figure 50 : Le projet Power BI. ....	57
Figure 51 : La requête chargement de la dimension date. ....	58
Figure 52 : Requête de la mesure Nbj_Emp_saisie avec DAX. ....	58
Figure 53 : Tableau de bord « Suivi Interne ». ....	59
Figure 54 : Tableau de bord « Vue d'ensemble des ressources et des performances ».....	60
Figure 55 : Tableau de bord « Gestion des clients Clinisys». ....	61
Figure 56 : Tableau de bord « Gestion de projet Clinisys». ....	62
Figure 57 : Tableau de bord « Gestion des jours ouvrables Clinisys ». ....	63
Figure 58 : Tableau de bord « Gestion de cout Clinisys ». ....	64
Figure 59 : Tableau de bord « suivi_congés Clinisys ».....	65



## Liste des tableaux

Tableau 1: La matrice des besoins.....	28
Tableau 2 : Simplification de la matrice des besoins .....	28
Tableau 3: Hiérarchies des dimensions (démarche descendante) .....	31
Tableau 4: Hiérarchies des dimensions (démarche ascendante) .....	38

# Introduction générale

L'informatisation de l'entreprise a permis à celle-ci de subir une véritable métamorphose tout en profitant de multiples avantages. Néanmoins, confrontés à une masse énorme et diversifiée de données, les systèmes d'informations de type relationnel, qui représentent depuis de nombreuses années la référence en matière de stockage de l'information, ne répondent plus aux besoins des décideurs.

La qualité de l'information et la rapidité de son acquisition augmente l'efficacité des prises de décisions et par la suite le développement rapide de l'entreprise. L'adoption de nouvelles technologies sur lesquelles reposent les systèmes d'information est désormais inévitable surtout avec la concurrence ascendante sur le marché.

Dans ce sens, l'informatique décisionnelle permet aux entreprises une meilleure homogénéité des données et y ajoute de l'intelligence afin de rendre les décisions plus adéquates. Ceci permet à l'entreprise d'être plus réactive et évolutive pour mieux répondre aux attentes des clients. Les divers avantages de cette technologie poussent un grand nombre d'entreprises à l'adopter et à bâtir dessus leurs systèmes d'information.

La sté CLINSYS, pionnière dans le domaine des applications informatiques dédiées à la santé, possède un système de gestion interne des employés qui fonctionnent sur différentes applications. Elle souhaite intégrer l'informatique décisionnelle afin de mieux suivre la gestion interne et garantir une meilleure exploitation des ressources. Dans ce cadre, s'inscrit notre projet de fin d'étude qui a pour but la conception et la réalisation d'un outil d'analyse et de reporting pour suivre la gestion interne en intégrant tous les systèmes déjà existant et y ajoutant les dimensions d'analyses nécessaires.

Le présent document décrit le travail réalisé pendant notre stage effectué au sein de l'entreprise CLINISYS. Ce rapport se compose de cinq chapitres décrits comme suit :

- Le premier chapitre présente l'organisme d'accueil. Il introduit le cadre général de notre projet pour expliquer le contexte et les objectifs à atteindre.
- Dans le second chapitre, nous introduisons les notions de base de l'informatique décisionnelle.

- Ensuite, dans le troisième chapitre nous présentons la modélisation conceptuelle de l'entrepôt de données selon la démarche mixte.

- Le quatrième chapitre décrit le développement du processus ETL (Extract-TransformLoad) des données. Une description des sources de données et des méthodes d'extraction / chargement de données seront introduites à ce niveau.

- Le dernier chapitre présente la restitution des données de notre entrepôt de données (ED). Notamment, il permet d'explicitier les principales fonctionnalités des outils de reporting que nous avons développés, ainsi que l'architecture technique utilisée.

# Glossaire

**Nbj\_Emp\_reel** : nombre de jour d'emploi réel

**Nbj\_Emp\_Est** : nombre de jour d'emploi estimé

**Ecart\_jr\_Emp** : écart jour d'emploi

**Indice\_PI** : Indice de performance individuel (focalisation)

**Cout\_Emp\_reel**: cout employer reel

**Cout\_Emp\_Est** : cout employer estimer

**Ecart\_cout** : ecart cout

**Nbj\_Congé** : nombre de jour de congé

**ERP**: Entrprise Resource Planing.

**ETL** : Extract Transform Load.

**ED** : Entrepôt de données.

# Chapitre 1 :

---

Etude Préalable :

## 1. Introduction :

Dans ce chapitre, nous abordons l'étude préalable de notre projet. Nous commençons par une présentation de l'organisme d'accueil, la société CLINISYS. Ensuite, nous présentons les grandes fonctionnalités de notre application et nous fixons les principaux objectifs à atteindre. La troisième partie de ce chapitre traite de l'étude de l'existant qui consiste en deux parties : la description de l'existant et sa critique. Enfin, La quatrième partie décrit la solution proposée.

## 2. Présentation de la société d'accueil :

CLINSYS est une société de services et d'ingénierie en informatique SSII. Elle a été créée en Tunisie depuis 1994. Sa filiale à Sfax est située à route manzel chaker km 4,5. CliniSys Group aide les professionnels de santé à tester, diagnostiquer et traiter chaque jour des millions de patients afin de préserver ou d'améliorer leur qualité de vie.

CLINISYS ERP est une solution de gestion hospitalière totalement intégrée installée dans 90% des cliniques en Tunisie. Elle est spécialement conçue afin de couvrir l'intégralité des fonctions des établissements de santé. Cet ERP gravite autour du dossier médical informatisé qui est conçu de façon à centraliser, en temps réel, l'ensemble des informations relatives à la prise en charge des patients, ce qui permet d'améliorer la qualité des soins.

Par ailleurs, CLINISYS ERP traite toutes les fonctions administratives et financières en appliquant les règles de la bonne gouvernance, en respectant les normes de qualité et d'accréditation HAS et JCI.

Depuis plus de 25 ans, elle a réussi à moderniser la gestion des hôpitaux dans des environnements difficiles et très hostiles. Elle a réussi à intégrer dans plus de 110 hôpitaux en Tunisie, en Egypte, en Libye et au Maroc tous les aspects administratifs, financiers, organisationnels, médicaux et décisionnels avec un seul produit appelé CLINISYS.

Elle migre vers des architectures modernes et vers le cloud. Elle innove avec l'analytique et l'intelligence artificielle.

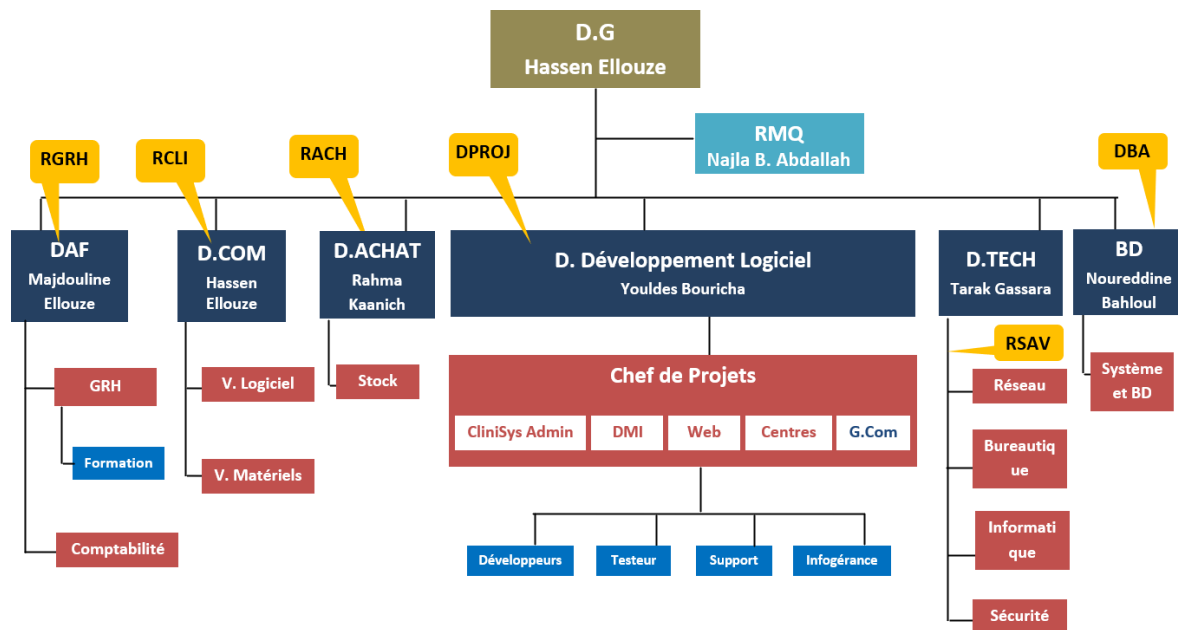


Figure 1: Organigramme général de la société CLINISYS.

### 3. Définition de la mission :

#### 3.1 Présentation du projet :

Dans le cadre de notre projet de fin d'étude, nous avons été appelés à concevoir et à réaliser un tableau de bord d'équipe support. Ce tableau de bord permettra d'assurer le suivi et l'analyse des domaines suivants :

- Gestion de ressource humain.
- Gestion de projet.

#### 3.2 Objectifs à atteindre :

Les principaux objectifs que nous souhaitons atteindre par notre application sont les suivants :

- Éliminer le travail ennuyant de collecte des données éparpillées dans les différentes bases de données sources.
- Constituer une source de données flexible et adaptable aux changements des données.
- Faciliter la lecture de grands volumes de données et améliorer l'accès aux informations.
- Avoir une source de données unique et non volatile, regroupant les données de 3

logiciels :

- Redmine
- support Clinisys
- Fiche d'intervention Clinisys
- Offrir une bonne visibilité des données par :
  - Élimination des champs inutilisables dans les sources de données.
  - Élimination des champs qui ne paraissent pas analytiques.
  - Organisation des données selon les besoins d'analyse des décideurs.
- Obtenir un seul tableau de bord contenant plusieurs interfaces visuelles.

#### 4. Etude de l'existant et problématique :

##### 4.1 Étude de l'existant :

L'étude de l'existant est la première étape du processus de la conception. Elle vise essentiellement à étudier l'existant et le critiquer afin de dégager les orientations et comprendre les fonctionnalités du futur système.

##### 4.2 Analyse de l'existant :

Nous décrivons dans ce paragraphe l'application Power BI utilisée dans la société Clinisys. Nous nous intéressons au tableau de bord analysant les données de la base Support et à celui analysant les données de la base Redmine.

##### ➤ Tableau de bord de la base support :

La figure 2 montre une interface de visualisation des informations générales fournies par le tableau de bord Support. Parmi ces informations, on note l'analyse des tâches accordées aux employés.



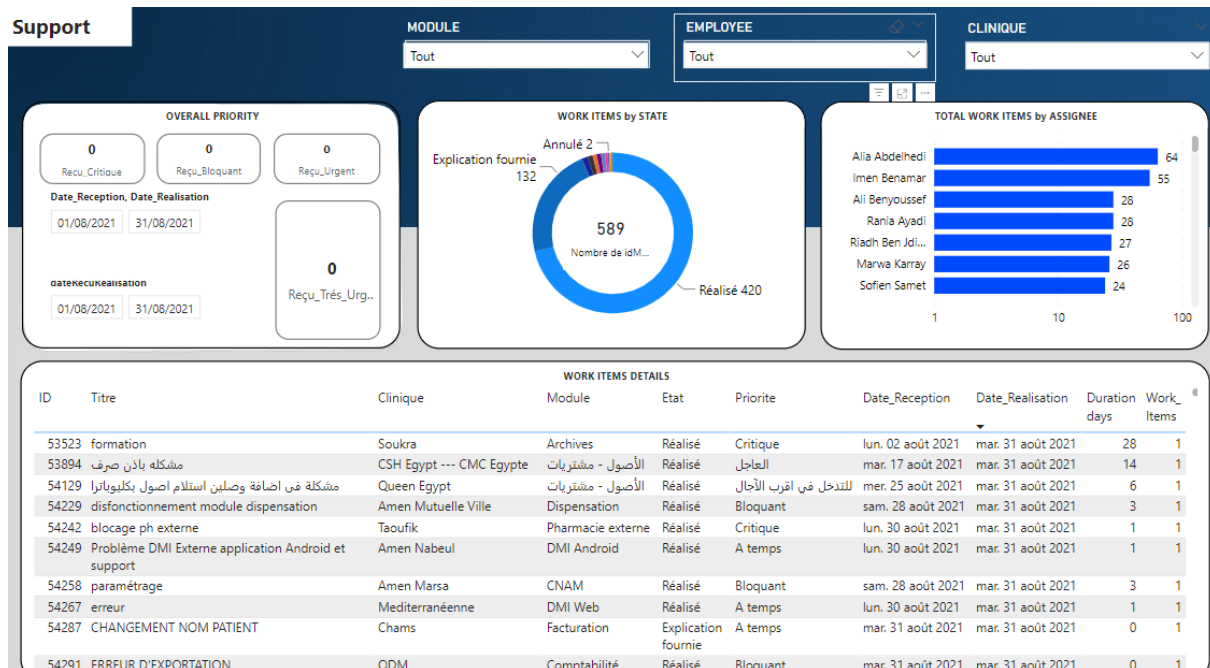


Figure 2 : Tableaux de bord support.

➤ Tableau de bord de la base Redmine :

La figure 3 montre une interface de visualisation des informations générales fournies par le tableau de bord Redmine. Parmi ces informations on note l'analyse des interventions en fonctions des régions.

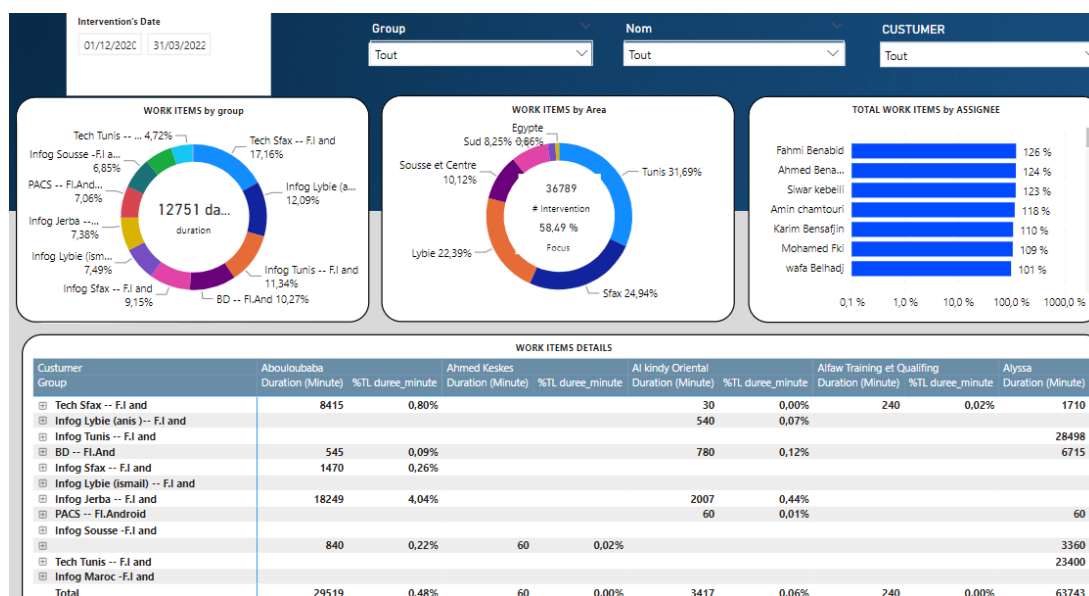


Figure 3 : Tableaux de bord Redmine.

#### 4.3 Critique de l'existant :

Même si CLINISYS utilise deux applications de gestion interne (REDMINE et SUPPORT), elle n'arrive pas à bien contrôler les activités internes des différents employés dans la réalisation de leurs tâches. Le conflit entre les deux applications peut engendrer les défaillances suivantes :

- ❖ Possibilité de doublons et de redondance : Si les tableaux de bord ne sont pas synchronisés de manière adéquate, il peut y avoir des doublons ou des redondances dans les données affichées.
- ❖ Complexité de la gestion : Avec des tableaux de bord séparés, il peut devenir complexe de gérer et de suivre les indicateurs clés de performance de manière globale. Les utilisateurs doivent naviguer entre les interfaces des différents logiciels pour obtenir une vue complète de l'entreprise, ce qui peut entraîner une perte de temps et une fragmentation des informations.
- ❖ Difficulté de comparaison : L'absence d'une vision consolidée rend difficile la comparaison directe des données et des performances entre les différents aspects de l'entreprise.
- ❖ Manque de cohérence : En maintenant des interfaces distinctes, il peut y avoir un manque de cohérence dans la présentation des informations et dans l'expérience utilisateur. Les utilisateurs peuvent se retrouver confrontés à des interfaces différentes en termes de fonctionnalités, de navigation et de présentation des données, ce qui peut rendre l'utilisation globale plus complexe et moins intuitive.
- ❖ Complexité des mises à jour et des intégrations : Maintenir et mettre à jour des tableaux de bord séparés pour chaque logiciel peut demander davantage de travail en termes de gestion des mises à jour logicielles et des intégrations avec d'autres systèmes. Cela peut entraîner des problèmes de compatibilité et des défis techniques supplémentaires.

#### 5. Solution proposée :

À la suite de notre critique de l'existant et compte tenu des objectifs fixés auparavant, nous avons constaté l'importance de construire un entrepôt de données permettant de :

- Extraire les données nécessaires à partir de 3 sources.
- Nettoyer les données avant de les charger afin de ne conserver que les données pertinentes pour l'analyse.

- Collecter les données dans une seule base de données bien structurée et facile à utiliser.
- Organiser les données dans un entrepôt de données pour faciliter leurs activités d'analyse.
- Mettre en place un processus de rafraîchissement périodique des données de l'entrepôt de données.
- Construire des tableaux de bord incluant différents visuels (Graphique en anneaux, Histogramme, Segment, ...) et des tableaux croisés pour rendre plus facile l'analyse des données et prendre ainsi une meilleure décision.
- Unifier les tableaux de bord des logiciels Redmine et Support.
- Viser à créer une interface centralisée offrant une vue d'ensemble complète des activités et des performances de l'entreprise. Cela facilitera l'analyse, la prise de décisions et la gestion globale, tout en améliorant l'efficacité et l'expérience utilisateur.

## 6. Définition des besoins fonctionnels et non fonctionnels :

### 6.1 besoins fonctionnels :

La construction d'un entrepôt de données pour analyser gestion du projet et la gestion des ressources humaines :

- Construire une source de données unique et non volatile.
- Rendre la base plus évolutive et plus adaptée à nos futurs besoins.
- Effectuer une modélisation multidimensionnelle des données et une hiérarchisation des données d'analyse allant d'un niveau moins fin à un niveau plus détaillé.
- Rendre la base de données plus optimale.
- Assurer la restitution des données de l'ED à travers la construction de tableaux de bord pour la société.
- Organiser et représenter les données d'une façon structurée.

### 6.2 besoins non fonctionnels :

Les principaux besoins non fonctionnels sont :

- Fournir des rapports lisibles.
- Fournir des tableaux de bord clairs et faciles à analyser.
- La solution doit être modulaire pour garantir la souplesse et l'évolutivité.

## 7. Planning Prévisionnel :

La planification est parmi les phases d'avant-projet les plus importantes. Elle consiste à indiquer et à planifier les tâches du projet et à estimer leurs charges respectives. Parmi les outils de planification de projet les plus connus, nous choisissons le diagramme de GANTT, qui permet de planifier le projet et de faciliter le suivi de son avancement. Ce diagramme permet aussi de visualiser l'enchaînement et la durée des différentes tâches durant le stage comme il est représenté dans la figure 4.







Taches	Janvier	Février	Mars	Avril	Mai
<b>Etude préalable</b>					
<b>Spécification des besoins</b>					
<b>Modélisation conceptuelle</b>					
<b>Conception et développement de l'ETL</b>					
<b>Power BI</b>					
<b>Rédaction du rapport</b>					

Figure 4 : Planning prévisionnel.

## 8. Conclusion :

Ce chapitre constitue une partie introductive dans laquelle une présentation de l'organisme d'accueil a été élaborée en premier lieu. En second lieu, nous avons présenté notre projet en définissant notre mission et en fixant les objectifs à atteindre. En troisième lieu, nous avons mené une étude de l'existant suivie de la proposition de la solution qui est basée sur la construction de l'entrepôt de données CLINISYS.

# Chapitre 2 :

---

Informatique décisionnelle :

## 1. Introduction :

Nous présentons dans ce chapitre les principes de base de l'informatique décisionnelle :

- L'entrepôt de données et ses différentes caractéristiques.
- Les concepts de la modélisation multidimensionnelle.
- Les outils ETL (Extract-Transform-Load).

## 2. Architecture d'un système décisionnel :

L'informatique décisionnelle ou "Business Intelligence" désigne l'ensemble des technologies offrant l'accès aux données, permettant d'analyser les informations pour améliorer et optimiser les décisions et les performances d'une entreprise [1]. Elle se base sur une architecture multi-tiers formée par les quatre composants essentiels de la figure 5 présentée ci-dessous.

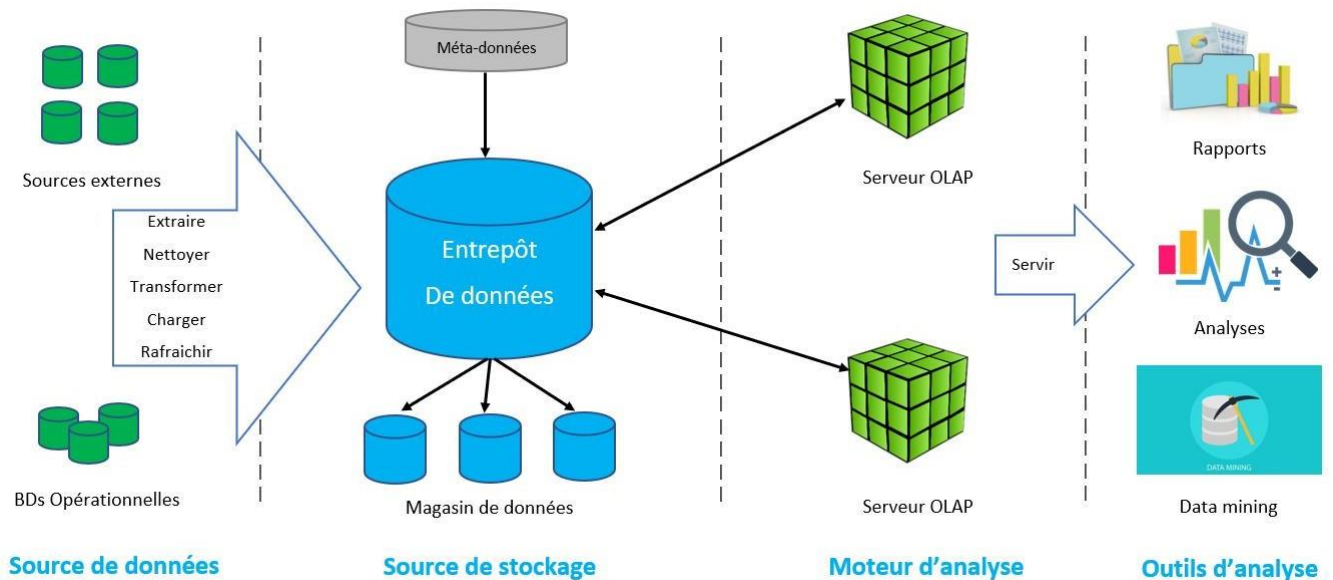


Figure 5 : Architecture des systèmes décisionnels.

- Les sources de données : sont nombreuses, variées, distribuées et autonomes. Elles peuvent être internes (bases de production) ou externes (Internet, bases des partenaires) à l'entreprise.
- ETL : permettant d'alimenter l'entrepôt de données à partir des données existantes.
- L'entrepôt de données : c'est le lieu de stockage centralisé des informations utiles pour les décideurs. Il met en commun les données provenant des différentes sources et conserve leurs évolutions.

- Les magasins de données : sont des extraits de l'entrepôt. Elles sont orientées sujet. Leurs données sont organisées de manière adéquate pour permettre des analyses rapides à des fins de prise de décision.
- Les outils d'analyse : permettent de manipuler les données selon des axes d'analyse. L'information est visualisée au travers d'interfaces interactives et fonctionnelles dédiées à des décideurs souvent non informaticiens (directeurs, chefs de services, ...).

### 3. Entrepôt de données et magasins de données :

#### 3.1 Définition d'un entrepôt de données :

D'après BILL Inmon : "Un entrepôt de données (ED) ou Data Warehouse (DW) est une collection de données orientées sujets, intégrées, non volatiles et historisées, organisées pour le support d'un processus d'aide à la décision" .

#### 3.2 Les caractéristiques d'un entrepôt de données :

Données orientées sujet : Les informations sont assemblées par thème. Grâce à cette orientation sujet, l'entreprise pourra développer son système décisionnel d'une manière incrémentale.

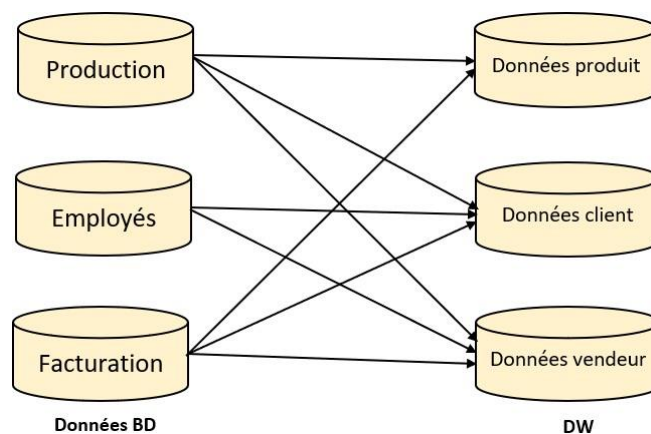


Figure 6 : Données orientées sujet de l'entrepôt de données

Données intégrées : L'entrepôt de données est composé de données intégrées, elles proviennent de systèmes sources hétérogènes donc un nettoyage préalable des données est nécessaire dans le but d'assurer la cohérence, la normalisation et la prise en compte des contraintes référentielles et des règles de gestion.

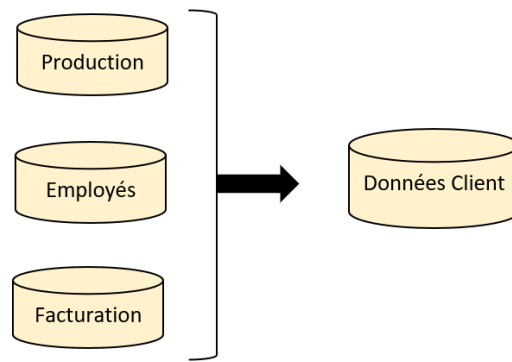


Figure 7 : Données intégrées de l'entrepôt de données

**Données non volatiles :** Afin de conserver la traçabilité des informations et des décisions prises, les informations stockées au sein de l'entrepôt de données ne peuvent être supprimées.

**Données Historisées :** L'historisation est nécessaire pour suivre dans le temps l'évolution des différentes valeurs des indicateurs à analyser. Chaque nouvelle insertion de données provenant du système de production ne détruit pas les anciennes valeurs, mais crée une nouvelle occurrence de la donnée.

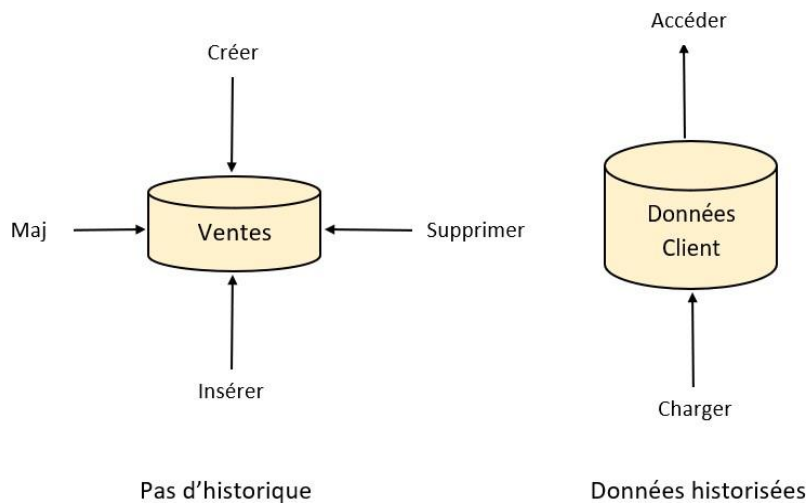


Figure 8 : Données non volatiles et historisées de l'entrepôt de données

### 3.3 Magasins de données ou Data Marts :

Le magasin de données est un sous-ensemble complet et naturel de l'entrepôt de données. Un entrepôt de données et un magasin de données se distinguent par le spectre qu'il recouvre :

- L'entrepôt recouvre l'ensemble des données et problématiques d'analyse visées par l'entreprise.



- Le magasin de données recouvre une partie des données et problématiques liées à un métier ou un sujet d'analyse en particulier.

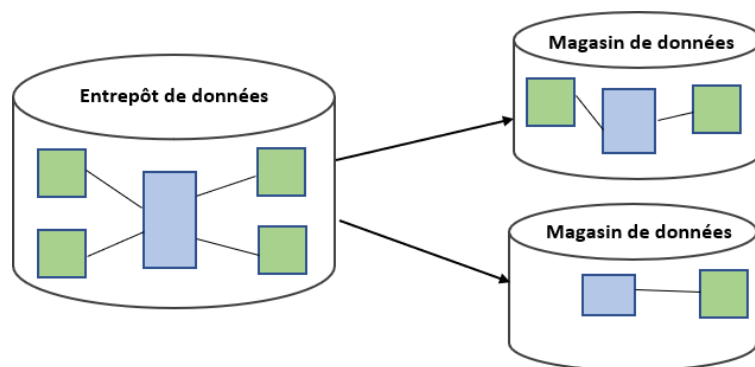


Figure 9 : Entrepôt de données et magasins de données

#### 4. Modélisation des entrepôts de données :

##### 4.1 Définition de la modélisation multidimensionnelle :

La modélisation multidimensionnelle souvent appelée modélisation OLAP se présente comme une alternative au modèle relationnel. Elle correspond mieux aux besoins du décideur tout en intégrant la modélisation par sujet. En effet, la modélisation multidimensionnelle est une méthode de conception logique qui vise à présenter les données sous une forme standardisée intuitive et qui permet des accès hautement performants.

Elle aboutit à présenter les données non pas sous forme de table mais de cubes centrés sur une activité. Ce type de modélisation considère les données comme des points dans un espace à plusieurs dimensions. Ces points représentent les centres d'intérêt décisionnel (sujets) analysés en fonction des différents axes d'analyse. Conceptuellement, une modélisation multidimensionnelle a donné naissance aux concepts de fait et de dimension [Kimball 1996].

##### 4.2 Concepts de base de la modélisation multidimensionnelle :

**Un fait** : C'est un sujet d'analyse, c'est-à-dire, un centre d'intérêt décisionnel. Il regroupe un ensemble d'attributs numériques représentant les mesures d'activité ainsi que les identifiants associés aux dimensions [Kimball 1996].

**Une mesure** : est un indicateur d'analyse de type numérique et cumulable. Une mesure est accompagnée d'un ensemble de fonctions d'agrégation qui permettent de l'agréger en fonction des axes d'analyse. Les mesures peuvent être additives (additionnable suivant toutes les dimensions), semi-additives (additionnable suivant certaines dimensions) et non additives (non additionnable quel que soit la dimension) [Kimball 1996].

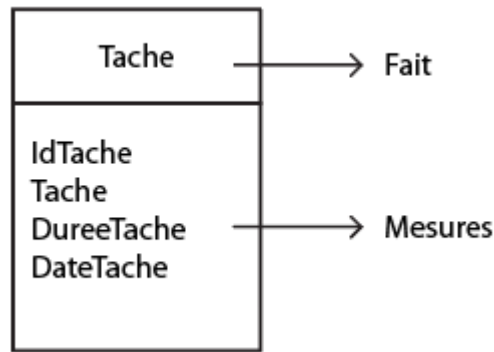


Figure 10 : Exemple d'un fait

**Une dimension** : c'est un axe d'analyse selon lequel sont visualisées les mesures d'activité d'un sujet d'analyse [Kimball 1996].

**Un paramètre** : c'est un attribut appartenant à une dimension. Il représente un niveau de détail selon lequel sont visualisées les mesures d'activité d'un sujet d'analyse. Les paramètres d'une dimension peuvent être accompagnés de descripteurs appelés attributs faibles qui ne sont pas utilisés dans les calculs de regroupement [Kimball 1996].

**Une hiérarchie** : c'est une perspective d'analyse définie dans une dimension. Elle regroupe un ensemble de paramètres organisés de la granularité la plus fine vers la granularité la plus générale [Kimball 1996].

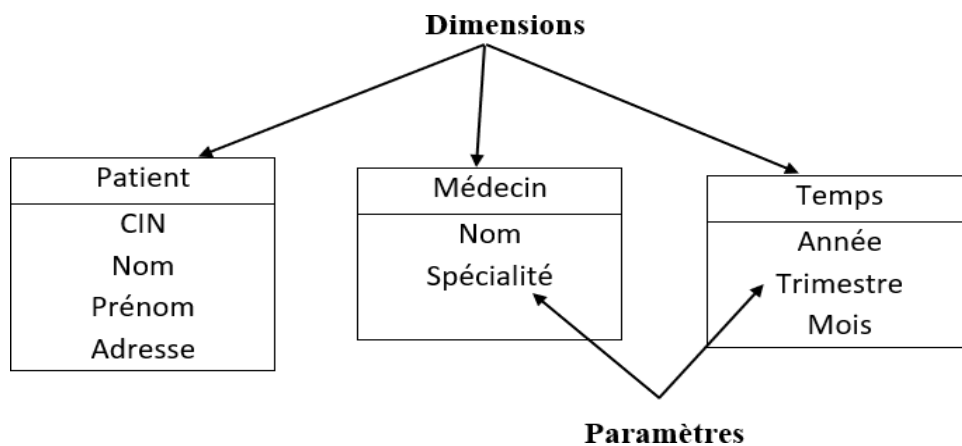


Figure 11 : Exemples de dimensions

### 4.3 Les schémas multidimensionnels :

Dans l'entrepôt de données, les bases de données relationnelles sont organisées spécifiquement pour pouvoir être exploitées de manières multidimensionnelles suivant soit le modèle étoile, soit le modèle en flocons de neige ou le modèle de constellation [Kimball 1996].

**Schéma en étoile :** Il comporte un seul sujet d'analyse (fait), englobant un ou plusieurs indicateurs (mesures). Ce schéma opère selon plusieurs axes d'analyse (dimensions) comportant la description des paramètres selon plusieurs perspectives d'analyse (hiérarchie) organisant les paramètres de dimension.

Le modèle en étoile est une représentation fortement dénormalisée qui assure un haut niveau de performance des requêtes même sur de gros volumes de données.

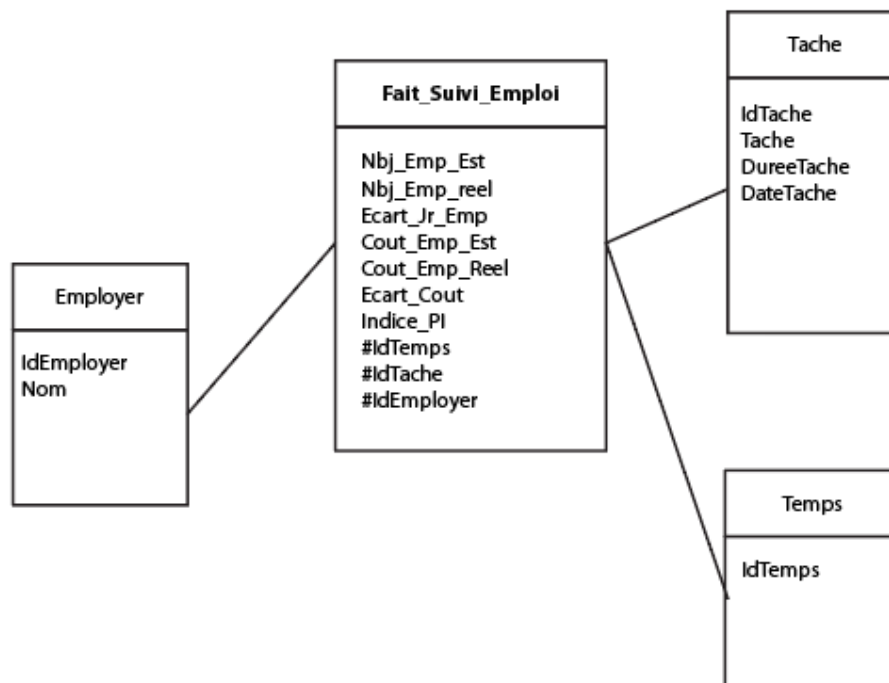


Figure 12 : Exemple d'un schéma en étoile

**Schéma en flocon de neige :** C'est un raffinement du schéma en étoile. En effet c'est le modèle en étoile avec une normalisation des dimensions lorsque celles-ci s'avèrent trop volumineuses. La différence réside dans le fait que les dimensions sont décrites par une succession de tables représentant la granularité de l'information. Ce schéma évite les redondances d'information mais nécessite des jointures lors de l'agrégation de ces dimensions.

Le modèle en flocon est aussi un modèle dénormalisé, mais un peu moins que le modèle en étoile : il conserve un certain niveau de décomposition pour chaque dimension prise isolément.

Donc c'est un dérivé du schéma en étoiles où chacune de ses dimensions est normalisée.

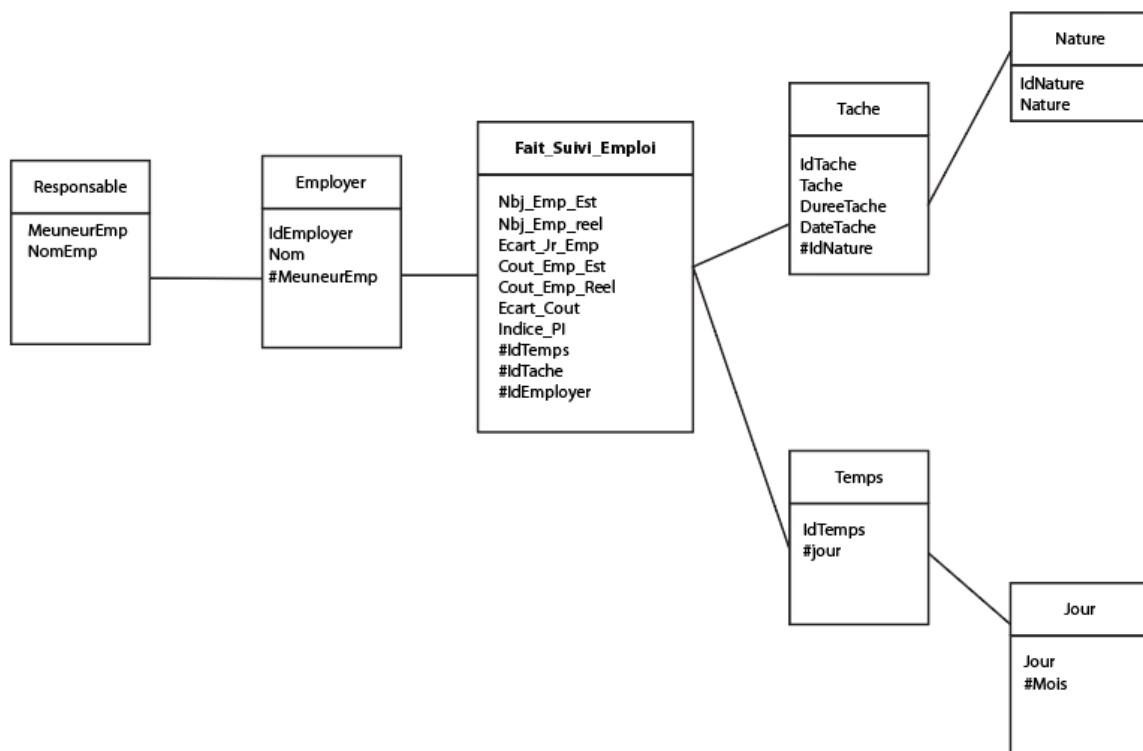


Figure 13 : Exemple d'un schéma en flocon de neige

**Schéma en constellation** : est une généralisation du schéma en étoile. Il comporte plusieurs sujets d'analyse (fait) étudiés selon plusieurs axes d'analyse (dimensions) pouvant être partagés entre les différents faits. La figure 14 montre un exemple d'un schéma en constellation selon la notion de Golfarelli.

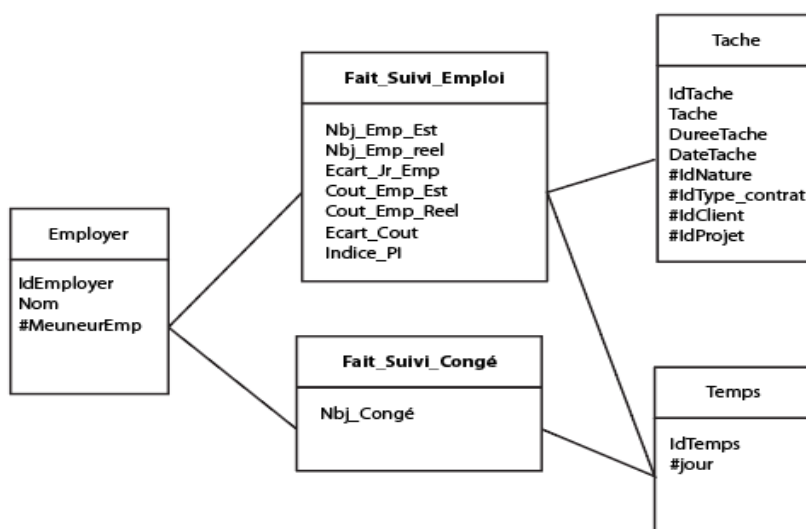


Figure 14 : Exemple d'un schéma en constellation [Golfarelli 2009]

## 5. L'alimentation de l'entrepôt de données :

L'ETL (Extract-Transform-Load) permet de concevoir et d'organiser les processus de migration du système transactionnel vers le système décisionnel. C'est le processus de copie des données depuis les tables des systèmes transactionnels vers les tables de l'entrepôt.

**Extraction** : il s'agit de la première étape de récupération des informations dans l'environnement de l'entrepôt de données. L'extraction comprend la lecture et la compréhension de la source de données, ainsi que la copie des parties nécessaires à une exploitation ultérieure dans la zone de préparation.

**Transformation** : une fois que les données extraites dans la zone de préparation des données, on rencontre plusieurs étapes de transformation :

- Nettoyage des données, à savoir correction des fautes d'orthographe, résolution des conflits de domaine (noms de villes incompatibles avec le nom de la région), résolution des cas d'informations manquantes et conversion en format standard.
- Combinaison des sources de données par mise en correspondance exacte avec des valeurs clés ou par mise en correspondance approximative d'attributs hors clé, y compris la recherche d'équivalent textuelle des codes des systèmes source.
- Création de clés de substitution pour chaque enregistrement dimensionnel afin d'éviter de dépendre des clés définies dans le système source. Dans ce cas, c'est le processus de génération des clés qui assure l'intégrité référentielle entre les tables dimensionnelles et les tables des faits.
- Construction d'agrégats pour optimiser les performances des requêtes les plus courantes.

**Alimentation et indexation** : le processus de chargement des données dans l'entrepôt s'effectue par réplication des tables des faits et des tables dimensionnels, qui seront ensuite présentées aux services de chargement en masse de chaque magasin de données destinataires. Les magasins de données doivent indexer les données afin d'optimiser les performances de requête.

## 6. Modélisation logique :

On distingue plusieurs variations de la modélisation logique d'un cube OLAP selon le type de données à analyser .

**Approche relationnelle (R-OLAP) :** Il renvoie à une base relationnelle classique stockée dans un SGBD relationnel (SGBDR) sous la forme de schéma en étoile ou en flocons pour réagir à la manière d'une base OLAP, donc un moteur OLAP permet de simuler le comportement d'un SGBD multidimensionnel. Ce schéma présente plusieurs avantages tels que la facilité de mise à jour des données le stockage de gros volumes de données avec un coût moins cher. En revanche, il est moins efficace et inefficace pour les calculs complexes. En effet, ces calculs sont limités par les fonctionnalités du SQL où l'on doit traduire les requêtes OLAP en SQL.

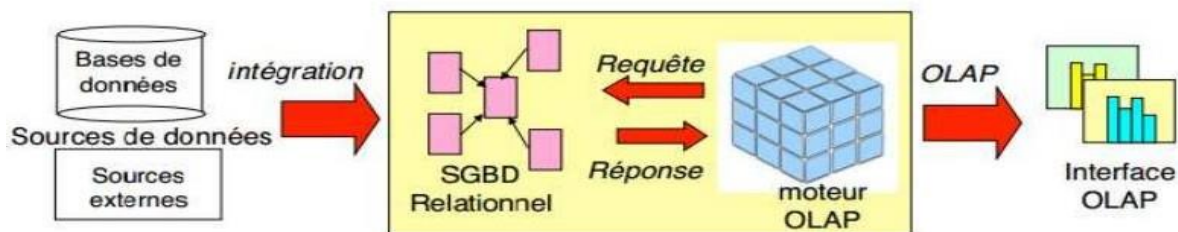


Figure 15 : Approche relationnelle (R-OLAP)

**Approche multidimensionnelle (M-OLAP) :** Le multidimensionnel OLAP consiste à utiliser un SGBD multidimensionnel pur, qui gère des structures multidimensionnelles natives. Ces structures utilisent des tableaux à n dimensions. Les données sont accessibles directement à partir du cube. Cela permet un accès rapide à l'information mais augmente le temps de mise à jour. Le multidimensionnel OLAP est efficace en matière de temps d'exécution mais il est un peu délicat de mettre en place ainsi qu'il ne soutenu pas un grand volume de données.

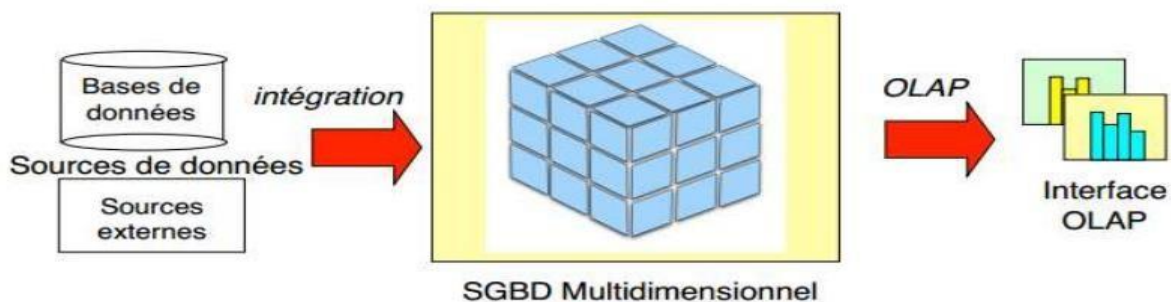


Figure 16 : Approche multidimensionnelle (M-OLAP)

**Approche hybride OLAP (H-OLAP) :** Les systèmes HOLAP essaient de combiner les bons côtés des systèmes ROLAP et MOLAP. En stockant les données détaillées de l'entrepôt de données dans une base relationnelle standard tandis que le reste des données (les agrégées) de l'entrepôt de données, souvent des magasins de données, dans une base multidimensionnelle. Ils permettent ainsi d'avoir des entrepôts de données de taille importante tout en ayant des temps

de réponse satisfaisants.

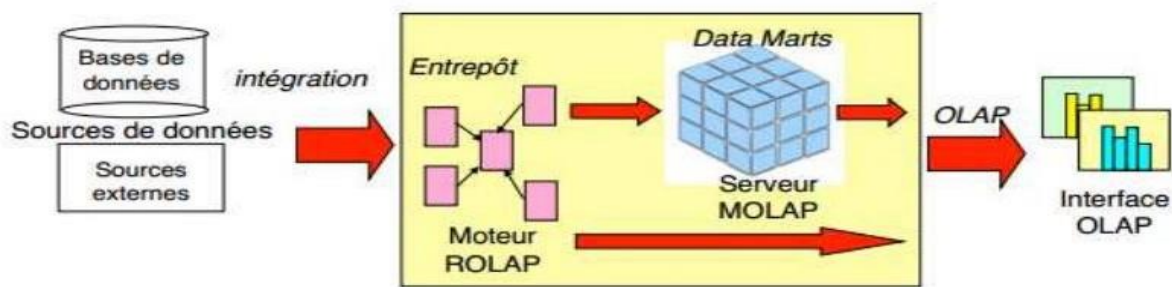


Figure 17 : Approche hybride (H-OLAP)

Dans notre perspective de travail, nous nous appuyons sur l'utilisation de ROLAP pour répondre aux besoins spécifiques de notre société, ainsi qu'à la méthode d'intégration associée.

### 7. Gestion de projet décisionnel :

Le projet décisionnel est un système qui permet aux dirigeants d'analyser et d'interpréter des données complexes d'une entreprise. Le projet décisionnel sert à améliorer la performance opérationnelle de l'entreprise. Il sert également à programmer l'activité à partir du système décisionnel. L'approche est illustrée par la figure 18.

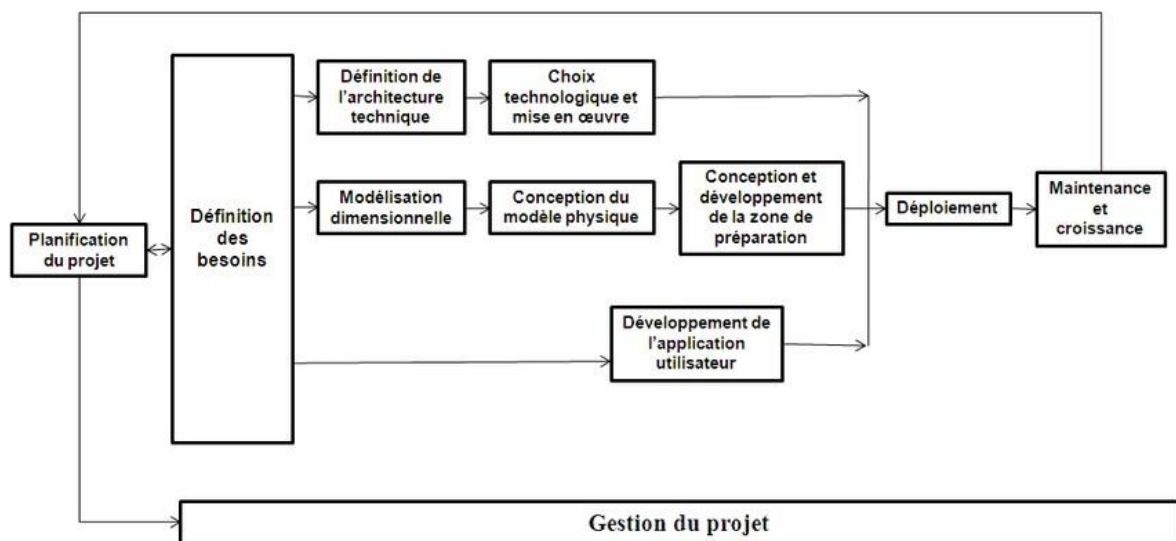


Figure 18 : Schéma de cycle de vie dimensionnel [Kimball 2008]

### 8. Langage DAX :

Le langage DAX, acronyme anglais de "Data Analysis eXpression", est un langage d'expression de formules utilisé dans les vues données et rapport de Power BI reposant sur des fonctions qui permettent d'effectuer des requêtes et des calculs complexes sur des données en colonnes et des tables mises en relation dans un modèle de données.

DAX se retrouve également dans d'autres produits Microsoft, tels que Power Pivot et le modèle tabulaire SSAS. DAX est conçu pour fonctionner avec des tables.

Il utilise la surcharge d'opérateur, ce qui signifie que des types de données peuvent être

combinés dans les calculs, et que les résultats changent en fonction du type de données utilisé dans les entrées.

Il s'agit d'une collection de fonctions, d'opérateurs et de constantes qui peuvent être utilisées dans une formule ou une expression pour calculer et retourner une ou plusieurs valeurs.

Il s'agit d'un langage inspiré en partie des fonctions Excel et du langage SQL mais adapté à la réalité d'analyse de bases de données. Il permet notamment la création d'indicateurs sur mesure répondant aux problématiques des utilisateurs métiers.

## 9. Conclusion :

Tout au long de ce chapitre nous avons essayé d'expliquer les notions de base des systèmes décisionnels, telles que l'architecture décisionnelle, la modélisation multidimensionnelle, et la construction de l'entrepôt à travers les processus ETL.

La maîtrise de telles notions favorise le bon déroulement du projet.



# Chapitre 3 :

---

Modélisation conceptuelle :

### 1. Introduction :

Dans ce chapitre, nous présentons les étapes d'une partie importante du volet décisionnel de notre projet. Il s'agit de la conception multidimensionnelle de l'entrepôt de données CLINISYS que nous visons construire.

### 2. Processus de modélisation d'un ED :

La modélisation d'un entrepôt de données est constituée de quatre phases. La figure 19 montre le processus de la modélisation.



Figure 19 : Processus de modélisation d'un entrepôt de données

### 3. Méthodes de conception d'un ED :

Nous distinguons trois approches de conception d'un schéma multidimensionnel.

**Méthode descendante :** La méthode descendante commence par l'identification des besoins des utilisateurs. En premier lieu, elle permet de collecter les différents sujets d'intérêt dans l'entreprise. Par la suite, de les combiner avec les différents axes d'analyse (les dimensions). Cette méthode utilise la matrice des besoins qui englobe plusieurs sujets d'analyse et l'ensemble de leurs dimensions.

**Méthode ascendante :** La méthode ascendante fournit un ensemble d'étapes pour la définition des faits, des dimensions et des hiérarchies à partir du schéma de la source. Le fait représente les événements fréquents dans le modèle de l'entreprise. Les dimensions sont formées à partir d'une portion du schéma source qui dépend de la relation représentant le fait. Les attributs de cette portion sont extraits puis réorganisés et affinés pour former les hiérarchies.

**Méthode mixte :** La méthode mixte est la fusion des deux approches : ascendante et descendante. Elle intègre à la fois les données des sources opérationnelles et les besoins des décideurs dans la définition du modèle dimensionnel. En effet, elle résout les problèmes rencontrés dans les méthodes précédentes pour satisfaire les différents besoins utilisateurs tout en respectant les données sources.

#### 4. Modélisation conceptuelle :

##### 4.1 Méthode descendante :

La figure 20 montre les différentes étapes de la méthode descendante. Dans ce qui suit nous détaillons chaque étape.

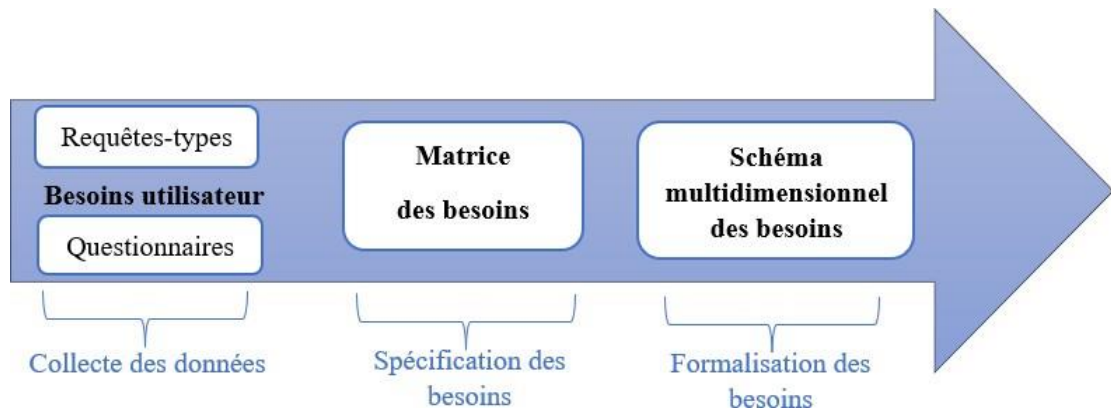


Figure 20 : Les étapes de la modélisation descendante

##### 4.1.1 Collecte des données :

Pour mieux caractériser et identifier les besoins des décideurs, nous avons effectué des questionnaires selon les clauses suivantes :

- La clause « Analyser » répond à la question Quoi ?  
→ Elle définit les données que les décideurs souhaitent analyser.
- La clause « En fonction de » répond aux questions Qui ? Où ? Et quand ?  
→ Elle indique les paramètres d'analyse des données décrites par la clause analyser.

Ainsi, nous avons déterminé les requêtes-types suivantes :

- R1 : Analyser le nombre jours emploi estimé  
En fonction de Client, Employer, Projet, Tache, Type\_contrat, Nature, Responsable, Région.
- R2 : Analyser le nombre jours d'emploi réel  
En fonction de Client, Employer, Projet, Tache, Type\_contrat, Nature, Responsable, Région.
- R3 : Analyser le nombre jours de congé  
En fonction de Client, Employer, Projet, Tache, Type\_contrat, Nature, Responsable, Région.

- R4 : Analyser l'écart du jours d'emploi  
En fonction de Client, Employer, Projet, Tache, Type\_contrat, Nature, Responsable, Région.
- R5 : Analyser l'indice de performance individuel  
En fonction de Client, Employer, Projet, Tache, Type\_Contrat, Nature, Responsable, Région.
- R6 : Analyser le cout estimé  
En fonction de Client, Employer, Projet, Tache, Type\_Contrat, Nature, Responsable, Région.
- R7 : Analyser le cout réel  
En fonction de Client, Employer, Projet, Tache, Type\_Contrat, Nature, Responsable, Région.

#### 4.1.2 Spécification des besoins :

##### ❖ Construction de la matrice des besoins :

Après l'étape de collecte des données, nous abordons l'étape de définition des données, qui commence par la construction de la matrice des besoins présentée dans le tableau 1. La matrice des besoins relie les mesures aux paramètres qui les analysent. Après avoir défini la liste des propriétés extraites des requêtes, nous avons construit une matrice carrée dont les entêtes des lignes et des colonnes sont composés de la liste des propriétés. Chaque case cochée (X) dans la matrice indique que la propriété en ligne est analysée en fonction de celle en colonne.

	Nbj_Emp_Est	Nbj_Emp_reel	Nbj_Congé	Ecart_jr_Emp	Cout_Emp_jr	Cout_Emp_réel	Indice_PI	Employer	Client	Tache	Projet	Nature	Type_contrat	Region	Jour	Mois	Année
Nbj_Emp_Est								X	X	X	X	X	X	X	X	X	X
Nbj_Emp_reel								X	X	X	X	X	X	X	X	X	X
Nbj_Congé								X	X	X	X	X	X	X	X	X	X
Ecart_jr_Emp								X	X	X	X	X	X	X	X	X	X
Cout_Emp_Est								X	X	X	X	X	X	X	X	X	X
Cout_Emp_réel								X	X	X	X	X	X	X	X	X	X
Indice_PI								X	X	X	X	X	X	X	X	X	X
Employer																	
Client																	
Tache																	
Projet																	
Nature																	
Type_contrat																	
Region																	
Jour																	
Mois																	
Année																	

Tableau 1: La matrice des besoins

❖ Simplification de la matrice des besoins :

Ensuite, nous simplifions la matrice des besoins en éliminant les lignes et les colonnes vides (voir tableau 2).

	Employer	Client	Tache	Projet	Nature	Type_contrat	Region	Jour	Mois	Année
Nbj_Emp_Est	X	X	X	X	X	X	X	X	X	X
Nbj_Emp_reel	X	X	X	X	X	X	X	X	X	X
Nbj_Congé	X	X	X	X	X	X	X	X	X	X
Ecart_jr_Emp	X	X	X	X	X	X	X	X	X	X
Cout_Emp_jr	X	X	X	X	X	X	X	X	X	X
Cout_Emp_réel	X	X	X	X	X	X	X	X	X	X
Indice_PI	X	X	X	X	X	X	X	X	X	X

Tableau 2 : Simplification de la matrice des besoins

#### 4.1.3 Formalisation des besoins :

##### ❖ Définitions des dimensions :

Cette étape réalise le regroupement des paramètres dans des dimensions (axes d'analyse) puis l'enrichissement de ces dimensions à travers de nouvelles propriétés.

##### ● **Regroupement des attributs :**

Après l'analyse des dépendances fonctionnelles entre les paramètres nous avons déterminé les dimensions suivantes :

- Client (NomClient)
- Employeur (Nom)
- Responsable (MeneurEmp)
- Projet (Nom)
- Tache (Tache)
- Nature (Nature)
- Type\_Contrat (type\_Contrat)
- Région (designation)
- Temps (jour, mois, année)

##### ● **Enrichissement des dimensions :**

Après enrichissement des dimensions avec de nouveaux attributs obtenus pour donner suite aux questionnaires réalisés, nous définissons les dimensions comme suit :

- Client (IdClient, NomClient, Region)
- Employeur (IdEmployer, Nom,)
- Responsable (MeneurEmp, NomEmp)
- Projet (IdProjet, Nom)
- Tache (IdTache, Tache, Source, DateTache, Duree)
- Nature (IdNature, Nature)
- Type\_Contrat (IdType\_contrat, Type\_contrat, source)
- Région (Code, designation)
- Temps (clé\_date, mois, trimestre, année)

- **Définition des Faits :**

Dans cette étape, il s'agit de regrouper les mesures (indicateurs) analysées par les mêmes paramètres identiques dans les mêmes faits. De plus, nous définissons les fonctions d'agrégation compatibles avec chaque mesure.

➤  $\text{Fait\_Suivi\_Emploi} = \{(\text{Nbj\_Emp\_Est}, \{ \text{SUM}, \text{AVG} \}), (\text{Nbj\_Emp\_Reel}, \{ \text{SUM}, \text{AVG} \}), (\text{Nbj\_Congé}, \{ \text{SUM}, \text{AVG} \}), (\text{Ecart\_Jr\_Emp}, \{ \text{SUM}, \text{AVG} \}), (\text{Cout\_Emp\_Est}, \{ \text{SUM}, \text{AVG} \}), (\text{Cout\_Emp\_Reel}, \{ \text{SUM}, \text{AVG} \}), (\text{Indice\_PI}, \{ \text{SUM}, \text{AVG} \})\}$ .

- **Définition des hiérarchies :**

Dans cette partie, il s'agit d'organiser les paramètres de chaque dimension dans des hiérarchies comme suit : (" → " signifie " agrégation "). Le tableau 3 montre ces hiérarchies.

H Client : IdClient → All

H Emp : IdEmployer → All

H Res: MeneurEmp → All

H Projet : IdProjet → All

H Tache : IdTache → All

H Nature: IdNature → All

H Type\_Contrat : IdType\_contrat → All

H Region: Code → All

H Temps : jour → Mois → Année → All

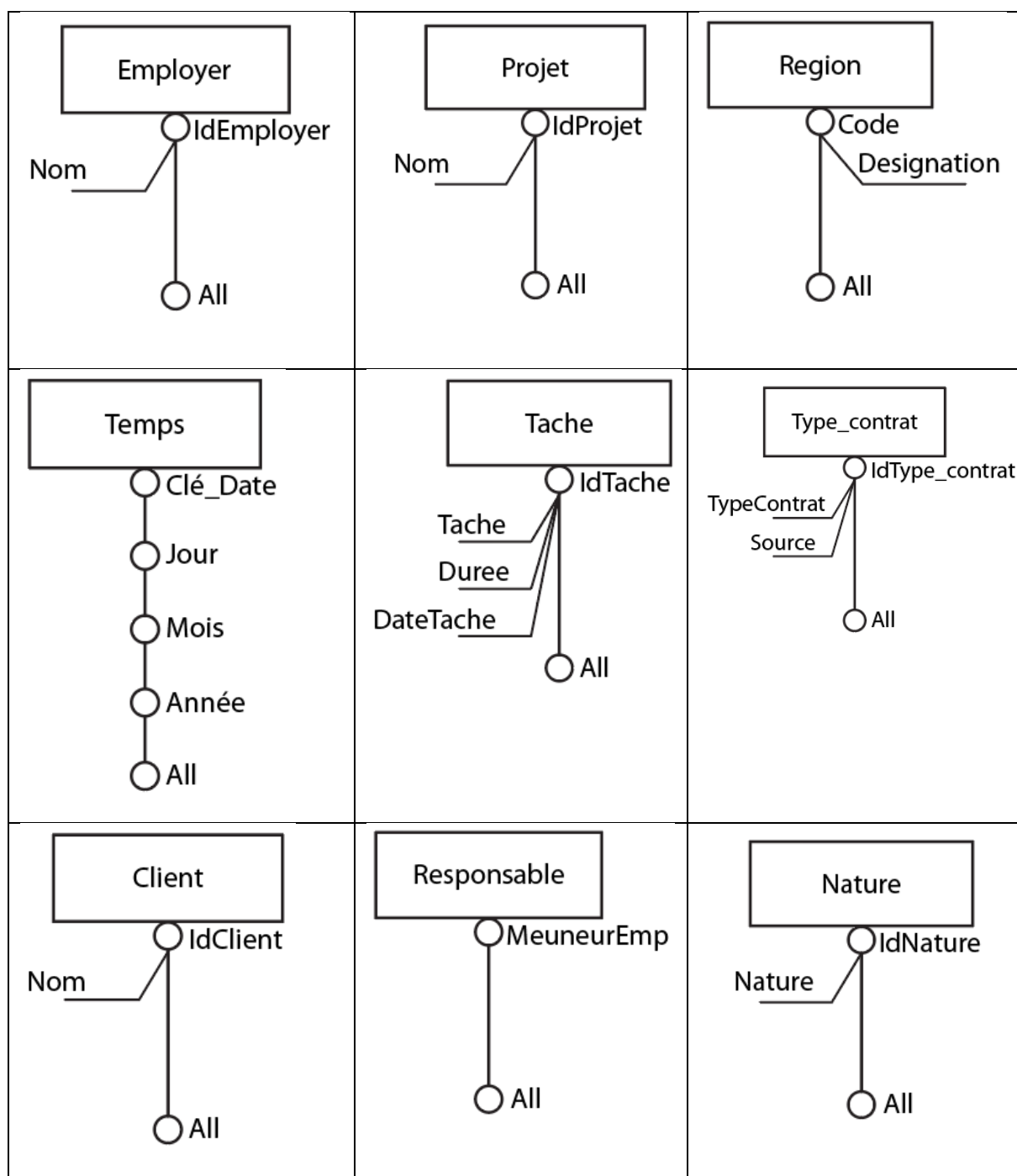


Tableau 3: Hiérarchies des dimensions (démarche descendante)



- **Schéma multidimensionnel en étoile :**

En se basant sur la matrice des besoins, nous réalisons l'affectation des dimensions au fait. Chaque dimension comportant des paramètres qui ont des intersections avec une mesure d'un fait est affectée à ce fait.

La figure 21 montre le schéma multidimensionnel obtenu en appliquant la démarche descendante.

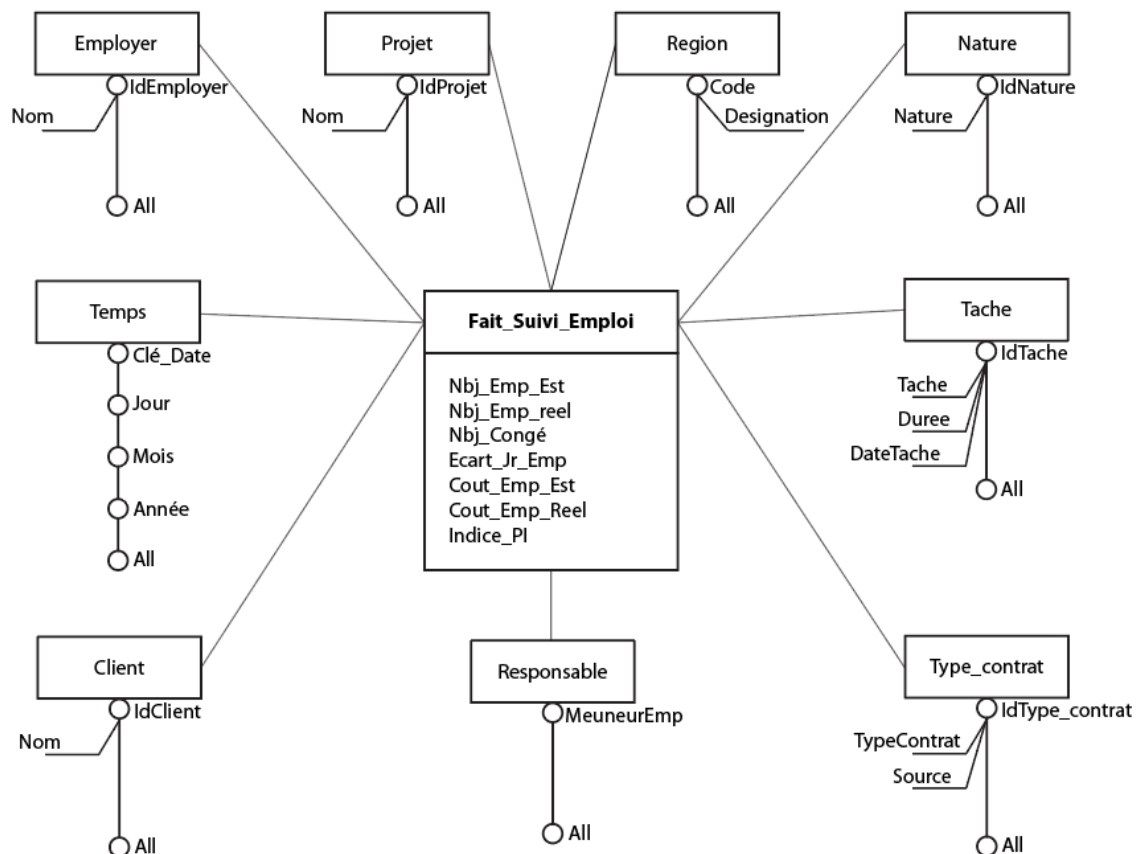


Figure 21 : Schéma multidimensionnel en étoile (démarche descendante)

## 4.2 Méthode ascendante :

La figure 22 montre les différentes étapes de la méthode ascendante. Dans ce qui suit, nous détaillons ces étapes à partir de l'étape de détermination des faits jusqu'à obtenir le schéma multidimensionnel.

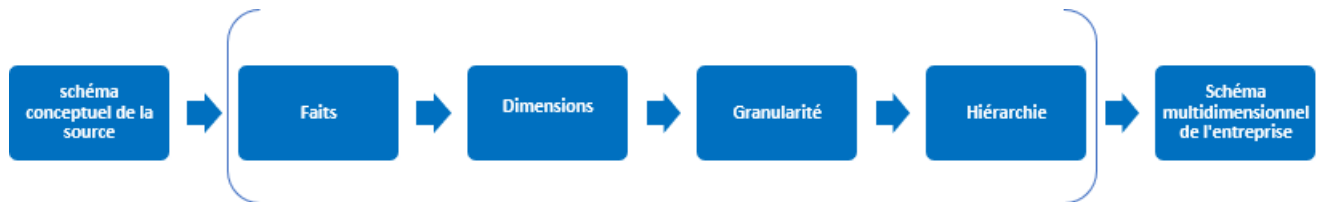


Figure 22 : Les étapes de la modélisation ascendante

### 4.2.1 Détermination des faits :

Avant de commencer la conception de notre schéma multidimensionnel, nous introduisons le diagramme de classe de nos sources de données. En effet, la source dans notre contexte est constituée de plusieurs tables différentes qui ne coexistent jamais pour une série de données à un instant donné, ce qui présente de nombreux problèmes d'hétérogénéité.

L'analyse de ces tables a permis d'établir le diagramme de classe suivant afin de pouvoir appliquer notre approche de conception ascendante. La Figure 23 montre le schéma de la base de données source. Il faut noter que le schéma n'est pas dans sa taille réelle car il contient des données confidentielles de la société.

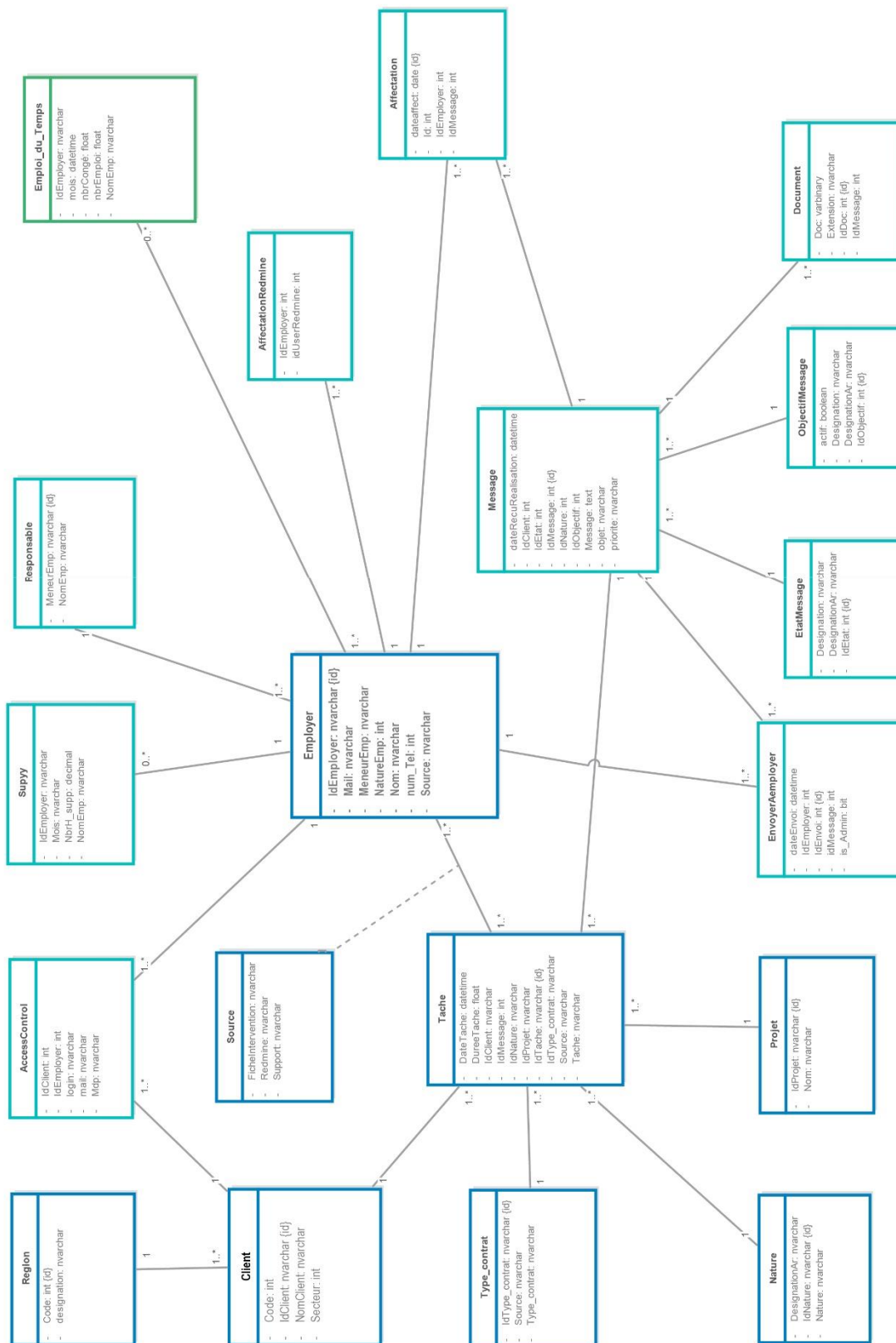


Figure 23 : Diagramme de classe de la source de données

#### ❖ Analyse du schéma de la source globale

Pour déterminer le schéma physique de la base de données source illustrée à la figure 23, nous avons effectué une analyse de la base de données source en l'interrogeant via des requêtes SQL.

#### ❖ Détection des classes représentatives de l'analyse

Une classe représentative (CR) décrit un événement qui se produit à un instant donné, elle contient les mesures d'analyse. Dans notre cas, nous avons détecté les classes représentatives suivantes : CR= {Tache, Emploi\_du\_temps} Nous définissons deux faits : Fait\_Suivi\_Emploi et Fait\_Suivi\_Congé comme suit. Notons que "S" signifie notre source de données.

#### ❖ Choix des mesures à analyser

Fait\_Suivi\_Emploi = (  
{Nbj\_Emp\_Est= (S.Emploi\_du\_Temps.nbrEmploi) (SUM, AVG)},  
{Nbj\_Emp\_Reel = (S.Tache.DureeTache) (SUM, AVG)},  
{Ecart\_Jr\_Emp=( (S.Emploi\_du\_Temps.nbrEmploi/S.Tache.DureeTache)\*100)(SUM,AVG),  
{Cout\_Emp\_Est = ((S.Emploi\_du\_Temps.NbrEmploi)\*206) (SUM,AVG)},  
{Cout\_Emp\_Reel = ((S.Tache.DureeTache)\*206) (SUM,AVG)},  
{Indice\_PI= ((S.Tache.DureeTache /S.Emploi\_du\_Temps.nbrEmploi )\*100 ) (SUM,AVG)})

Fait\_Suivi\_Congé = (  
{NbjCongé= (S.Emploi\_du\_temps.NbjCongé) (SUM,AVG)})

### 4.2.2 Détermination des dimensions :

#### ❖ Détection des classes déterminantes des classes représentatives

Une classe Ci est déterminante d'une classe représentative CR si :

$C_i = CR$

CR est reliée à Ci par un lien de dépendance fonctionnelle directement ou indirectement

Le graphe suivant représente les liens de dépendances fonctionnelles entre les classes représentatives et les classes déterminantes dans la source de données.

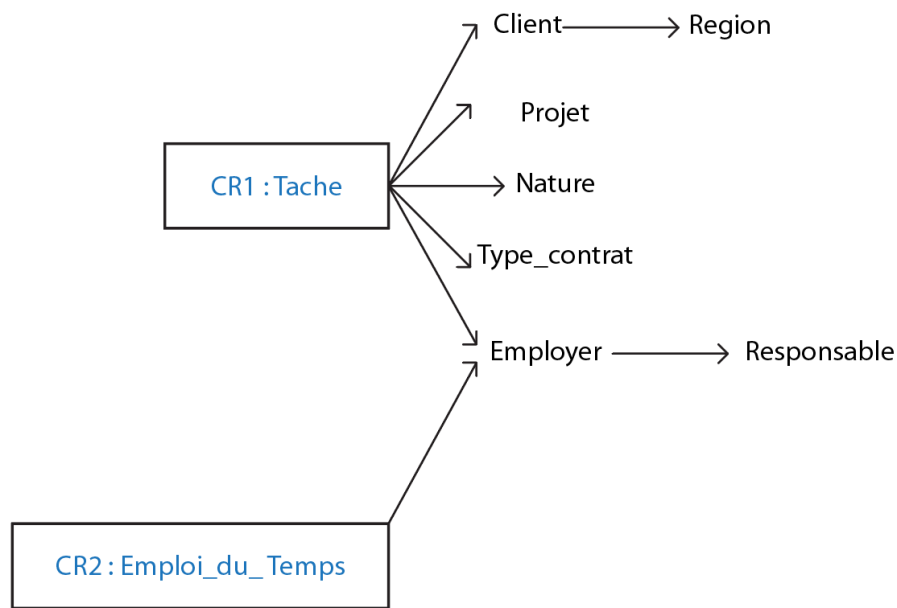


Figure 24 : Graphes de dépendances entre classes

Après une analyse sémantique, nous avons dégagées les dimensions présentées ci-dessous :

```

Employer (
  {IdEmployer = S.Employer. IdEmployer },
  {Nom = S.Employer.Nom},
  {MeneurEmp=S.Employer.Responsable.MeneurEmp},
  {NomEmp= S.Employer.Responsable.NomEmp })

Tache (
  {IdTache= S.Tache. IdTache},
  {Tache= S.Tache. Tache},
  {Source= S.Tache. Source},
  {DateTache= S.Tache. DateTache},
  {DureeTache= S.Tache. DureeTache},
  {IdNature= S.Tache.Nature. IdNature},
  {Nature= S.Tache.Nature. Nature},
  {IdProjet= S.Tache.Projet. IdProjet},
  {Nom= S.Tache.Projet. Nom},
  {IdType_contrat= S.Tache. Type_contrat. IdType_contrat},
  {Type_contrat= S.Tache. Type_contrat. Type_contrat},

```

{Source= S.Tache.Projet. Type\_contrat. Source},  
 {IdClient= S.Tache. Client. IdClient},  
 {NomClient= S.Tache.Client. NomClient},  
 {Code= S.Tache. Client.Region. Code},  
 {Designation= S.Tache. Client.Region. Designation})

#### ❖ Définition de la dimension temporelle

Dans un système décisionnel, la dimension Temps revêt une importance particulière et doit faire l'objet de la plus grande attention. La dimension Temps est unique et sa présence est obligatoire dans toute table de faits, quelle qu'elle soit.

Temps (  
 {Jour= S.Temps.Jour},  
 {Mois = S.Temps.Mois},  
 {Annee = S.Temps.Annee})

#### 4.2.3 Hiérarchisation des dimensions :

Dans cette étape, nous définissons les hiérarchies complètes des dimensions en fonction du schéma de la source.

$H^1_{\text{Employer}} : \text{IdEmployer} \rightarrow \text{MeneurEmp} \rightarrow \text{All}$

$H^2_{\text{Tache}} : \text{IdTache} \rightarrow \text{IdNature} \rightarrow \text{All}$

$H^3_{\text{Tache}} : \text{IdTache} \rightarrow \text{IdType\_contrat} \rightarrow \text{All}$

$H^4_{\text{Tache}} : \text{IdTache} \rightarrow \text{IdProjet} \rightarrow \text{All}$

$H^5_{\text{Tache}} : \text{IdTache} \rightarrow \text{IdClient} \rightarrow \text{Code} \rightarrow \text{All}$

$H^6_{\text{Temps}} : \text{Clé date} \rightarrow \text{Jour} \rightarrow \text{Mois} \rightarrow \text{Année} \rightarrow \text{All}$

Le tableau 4 montre la schématisation des différentes hiérarchies de dimensions.

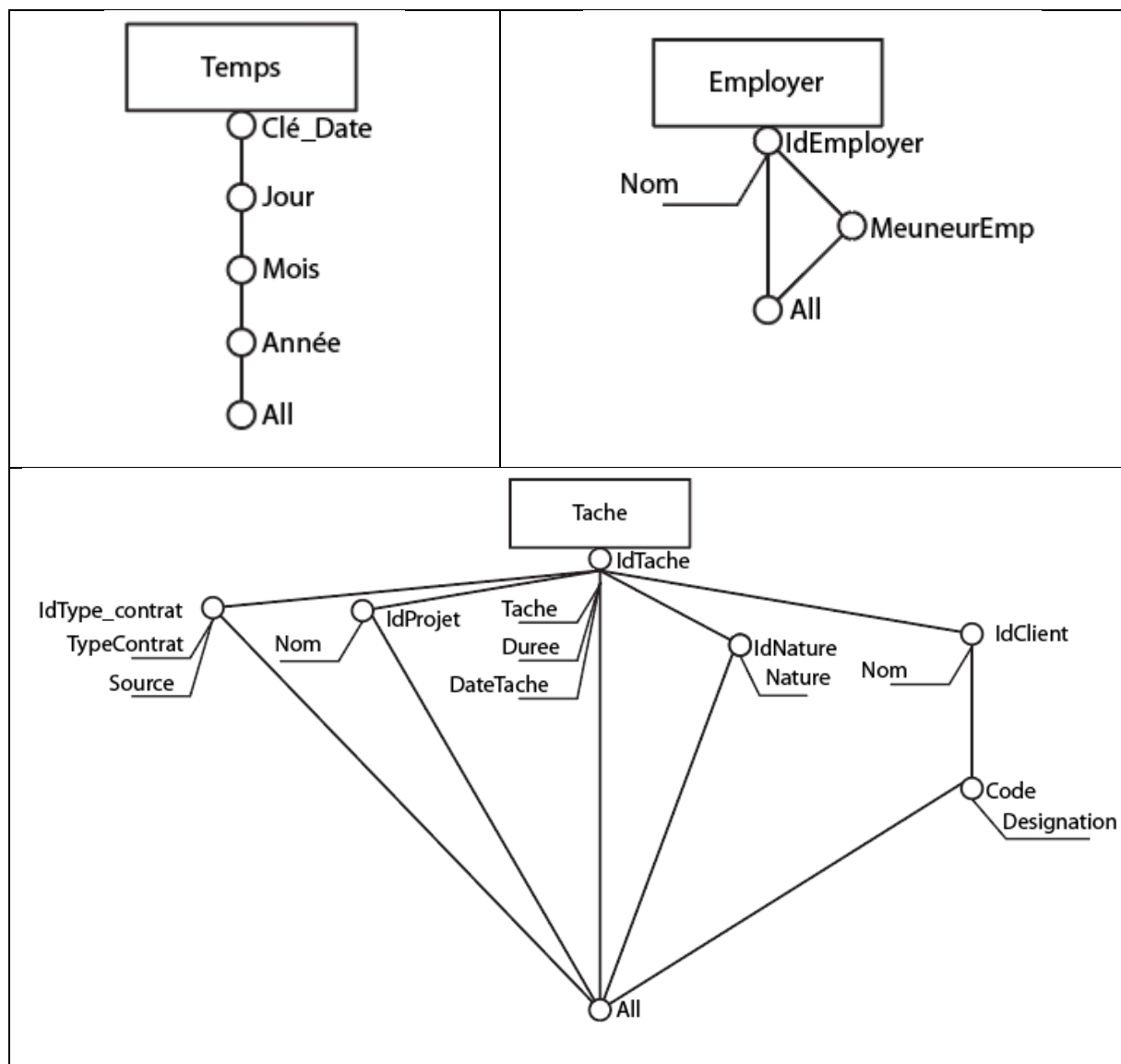


Tableau 4: Hiérarchies des dimensions (démarche ascendante)

#### 4.2.4 Schéma multidimensionnel en constellation :

La figure 25 montre le schéma multidimensionnel que nous avons obtenu en appliquant la méthode ascendante.

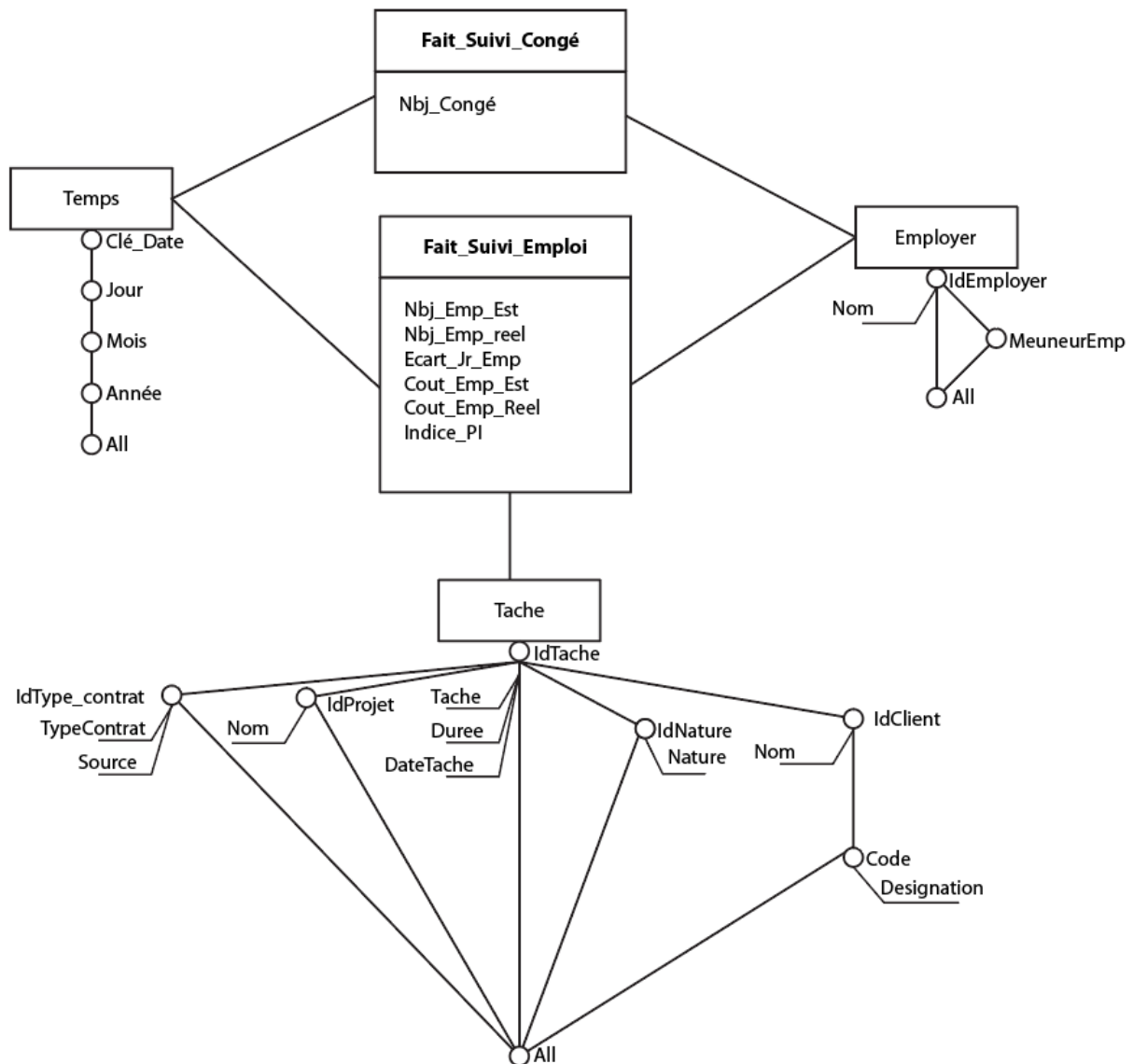


Figure 25 : Schéma multidimensionnel en constellation (démarche ascendante)

#### 4.2.5 Méthode mixte :

Dans l'approche descendante On a commencé à partir des objectifs et des besoins stratégiques de l'entreprise pour définir les indicateurs clés de performance (KPI) pertinente. Puis, on décompose ces KPI en dimensions et en mesures spécifiques. Par la suite, dans l'approche ascendante on effectue des analyses exploratoires sur les données source dans cette étape on a isolé la mesure "Nbj\_Congé" des autres mesures et à organiser les données dans deux tables de faits distinctes, une fois que les mesures sont réparties dans les tables de faits appropriées, on a créé des hiérarchies entre les dimensions. Les hiérarchies permettent de structurer les dimensions de manière logique, en établissant des relations entre les niveaux de granularité. En établissant ces hiérarchies, on facilite l'analyse des données en permettant des regroupements selon différents niveaux de granularité.



Pour ce cela on a choisi l'approche ascendante comme une méthode mixte ou on a dégagé deux Faits le premier Se concentre sur le suivi du travail des employés, et le deuxième met l'accent sur le suivi de congé des employés. Par la suite, nous avons attribué certaines modifications telles que l'ajout d'une nouvelle mesure.

La figure 26 montre le schéma multidimensionnel que nous avons obtenu en appliquant la méthode mixte.

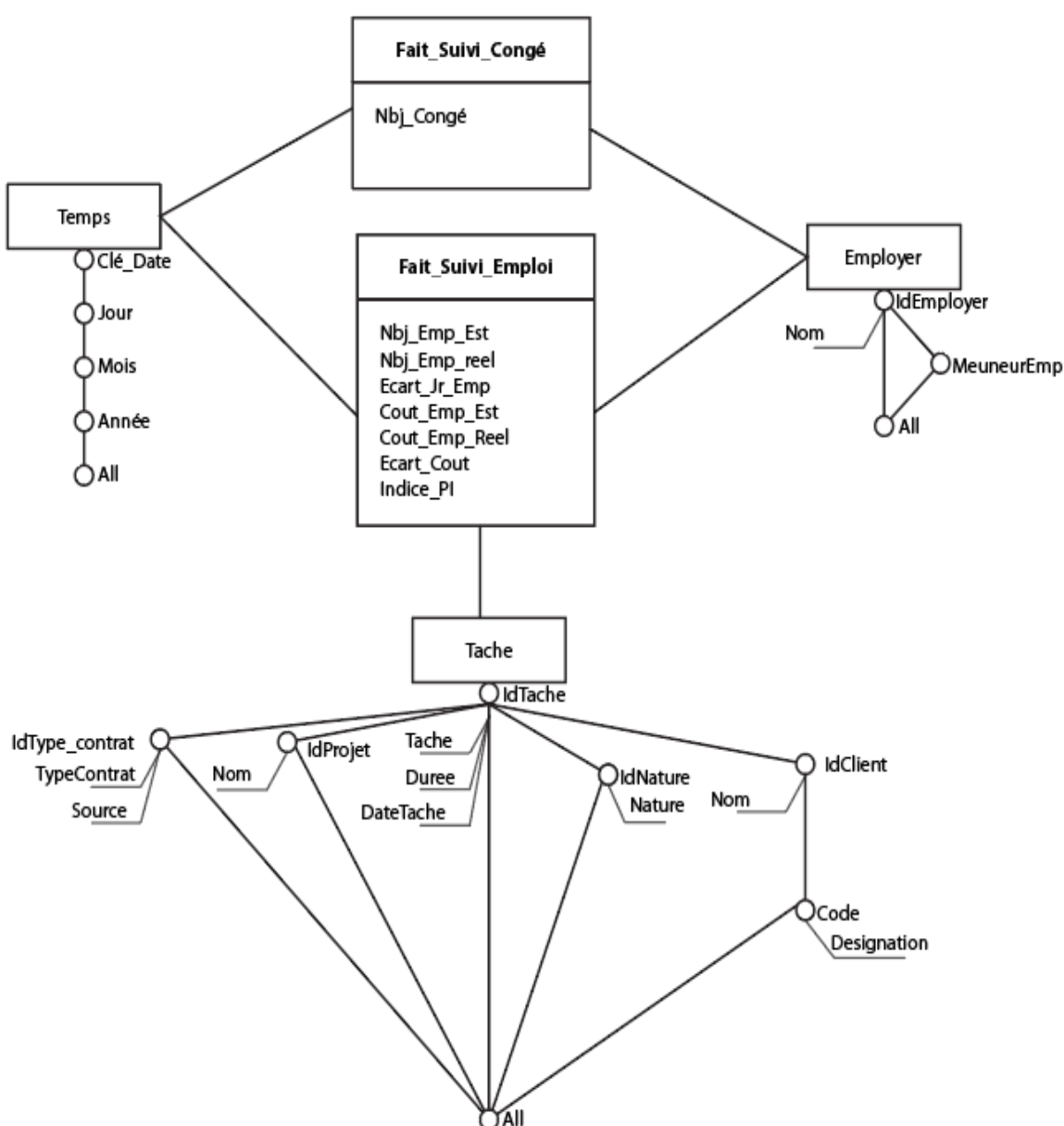


Figure 26 : Schéma multidimensionnel en constellation (démarche Mixte)

## 5. Conclusion :

Dans ce chapitre nous avons présenté une modélisation conceptuelle de l'entrepôt de données selon deux approches ascendante et descendante. Ensuite nous avons obtenu le schéma conceptuel final en utilisant l'approche mixte. Il ne reste que le développement des processus ETL, c'est ce que nous allons traiter dans le chapitre suivant.

# Chapitre 4 :

---

Construction de l'entrepôt de données :

## 1. Introduction :

Dans ce chapitre, nous définissons, en premier lieu, les outils utilisés pour la réalisation de notre entrepôt de données ainsi que l'étape de reporting. En deuxième lieu, nous définissons les différentes phases de la construction de l'entrepôt de données en commençant par la modélisation logique et physique pour achever à la phase d'intégration des données où nous présentons la conception et le développement de l'ETL.

## 2. Outils utilisés :

Les outils de développement utilisés lors de la phase d'intégration des données et la phase de présentation de l'application reporting sont :

### 2.1 Microsoft SQL Server Intégration Services :



Dans un projet décisionnel, SQL Server Intégration Services (SSIS) est le premier service à être activé, car il gère l'extraction et l'enregistrement des données en dehors de l'outil de production. Il joue donc le rôle d'un ETL (Extract, Transform, Load) en assurant la performance des processus d'extraction et de transformation des données. Ces données sont alors stockées dans un entrepôt de données sous SQL Server Management Studio (SSMS).

Figure 27 : Logo Microsoft SQL Server Intégration Services

### 2.2 Adobe Illustrator :



est un logiciel de création graphique et de dessin vectoriel il permet de créer des illustrations, des logos, des graphiques avec une grande précision. Illustrator offre une multitude d'outils de dessin, de peinture et de manipulation des couleurs, ainsi que des fonctionnalités avancées telles que la création de motifs répétitifs et la gestion des calques. C'est un outil essentiel pour exprimer la créativité de manière professionnelle et percutante

Figure 28 : Logo Adobe Illustrator

### 2.3 Microsoft SQL Server Management Studio 19 :



L'implémentation de la base de données a été assurée à travers le serveur de base de données relationnelle Microsoft SQL Server Management Studio (SSMS). Ce choix de SSMS est justifié par son flexibilité et sa particularité par rapport à ses concurrents qui sont Oracle, MySQL ou PostgreSQL. SSMS est un environnement intégré pour configurer, gérer et administrer la gestion de tous les composants de Microsoft SQL Server. SSMS se distingue par le fait que c'est un SGBDR multi-bases et multi-schéma.

Figure 29 : Logo Microsoft SQL Server Management Studio 19

### 2.4 Visual Studio 2019:



Microsoft Visual Studio 2019 est un IDE complet idéal pour les développeurs. C'est le principal environnement employé pour développer des solutions de gestion incluant des projets Analysis Services et Intégration Services. Il s'agit d'un logiciel qui permet de gérer les SQL Server Integration Services (SSIS) représentés par des flux de données et des flux de contrôle.

Figure 30 :Logo Visual Studio 2019

### 2.5 Microsoft Excel:



Depuis plusieurs années, Excel ne se contente plus d'être un simple tableur. Il est devenu un outil d'analyse de données à part entière : il intègre de véritables fonctions d'analyse de données comme les tableaux croisés dynamiques.

Figure 31 :Logo Microsoft Excel

### 2.6 Microsoft Power BI :



Microsoft Power BI est une solution d'analyse de données de Microsoft. Il permet de créer des visualisations de données personnalisées et interactives avec une interface suffisamment simple pour que les utilisateurs finaux créent leurs propres rapports et tableaux de bord. Pour cela, nous utilisons Power BI pour charger les données qui sont stockées dans notre entrepôt de données puis pour produire les mesures et générer les différents visuels.

Figure 32 :Logo Microsoft Power BI

## 2.7 Entreprise Architect :



Enterprise Architect (EA) est un logiciel de modélisation et de conception UML, il supporte plusieurs techniques de modélisation standard. Couvrant, par ses fonctionnalités, l'ensemble des étapes du cycle de conception d'application, il est l'un des logiciels de conception et de modélisation les plus reconnus. Il permet d'analyser, de dessiner, de coder, de tester et de déployer. Il offre en outre un environnement graphique très intuitif et très facile à aborder.

Figure 33 :Logo Entreprise Architect

## 2.8 Redmine :



Est un système de gestion de projets open source et web. Il offre une plateforme centralisée pour gérer les projets, les tâches, les problèmes, les ressources et les documents associés à un projet il est largement utilisé par les équipes de développement logiciel, les équipes de gestion de projets et les organisations pour gérer efficacement leurs projets et collaborer sur des tâches et des problèmes.

Figure 34 :Logo Redmine

## 3. Conception détaillé de l'ETL :

Les processus ETL représentent le cœur du projet décisionnel. La conception de ce processus permet d'avoir une vision claire de cette étape et d'éviter les erreurs d'implémentation. Pour cela, nous avons choisi pour la description de ce processus d'utiliser le diagramme d'activités puisqu'il modélise un flux et décrit la logique d'une opération. Le diagramme est illustré par la figure 35.

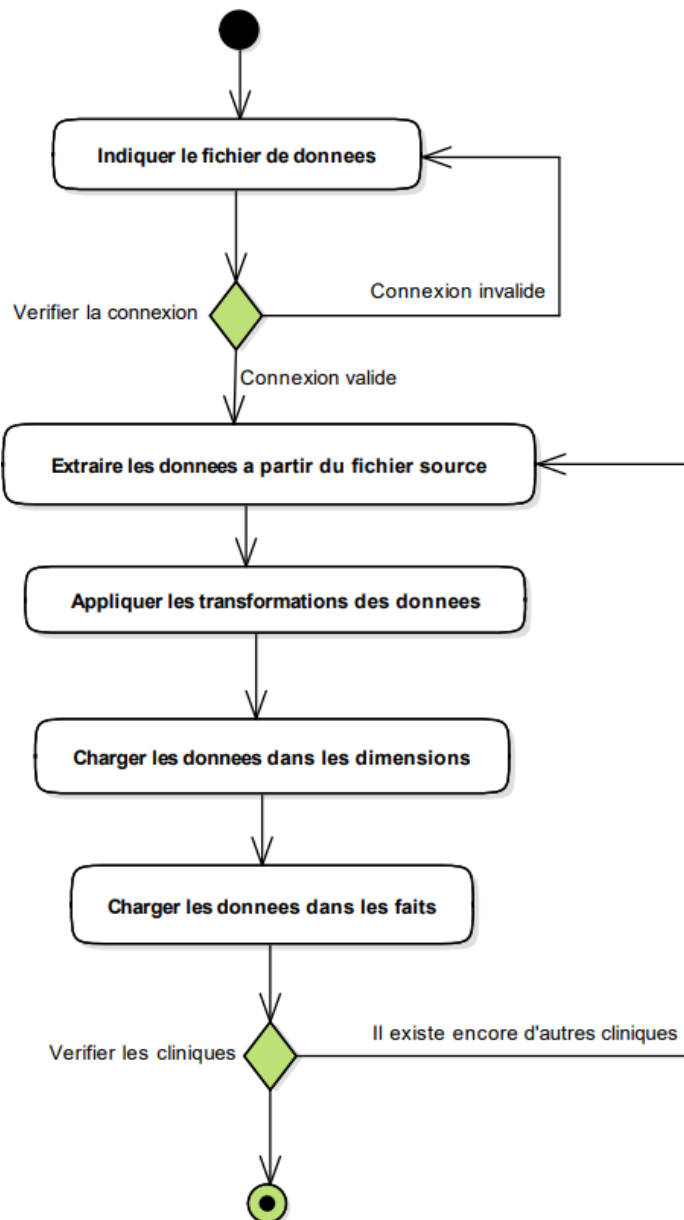


Figure 35 : Diagramme d'activités d'alimentation des tables de dimensions et de faits

Ce diagramme présente la phase d'intégration générale qui consiste à récupérer les données des différentes cliniques à partir de la source et à les charger dans la base de destination. Cette fonctionnalité sera assurée par les activités citées ci-dessous :

- Vérification de l'existence des données dans la source.
- Vérification de la validité des données.
- Transformation des données si nécessaire.
- Alimentation de l'entrepôt de données.

#### 4. Développement d'ETL :

La première étape de l'implémentation de l'ETL repose sur la construction de notre entrepôt de données cible dans SSMS. Nous avons donc créé une base de données et ensuite nous avons créé les squelettes de toutes les tables comme indiqué sur la figure 36.

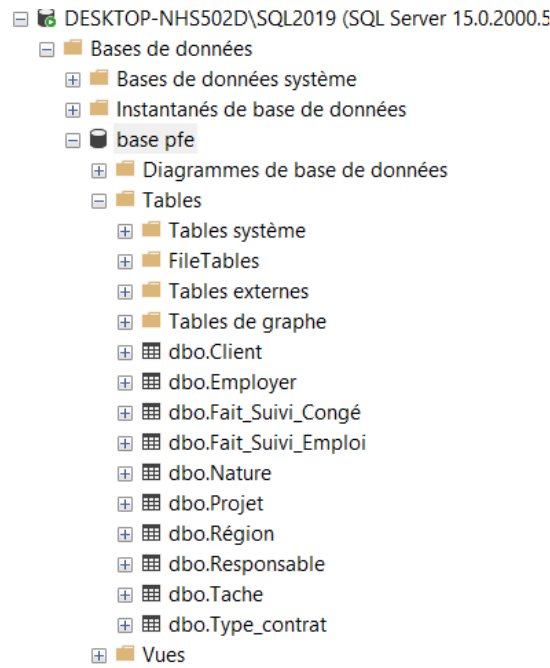


Figure 36 : Tables de l'entrepôt de données.

Après avoir terminé la création de la base de données nous allons passer à la deuxième étape qui consiste à créer un nouveau projet dans le visuel studio nommé "ETL" dans lequel nous avons créé deux nouveaux paquets SSIS appelé « dimension.dtsx » et « faits.dtsx » comme le montre la figure 37.

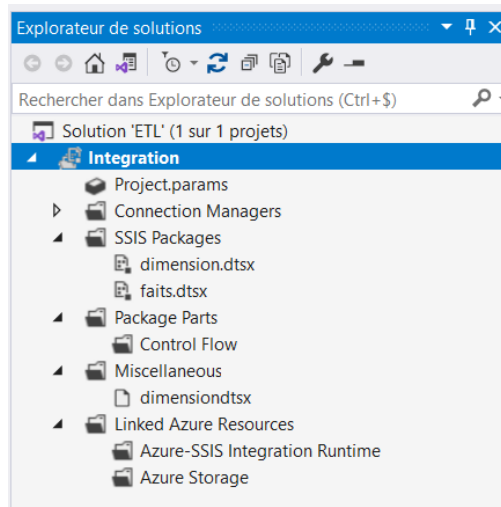


Figure 37 : Le projet SSIS.



Dans notre cas, nous travaillons sur 3 sources pour cela voici un exemple des principales étapes de la phase de connexion de bases de données Redmine dans Visual Studio.

Dans la première étape sélectionner un type de connexion « OLEDB » et l'ajouter en cliquant sur « Add ».

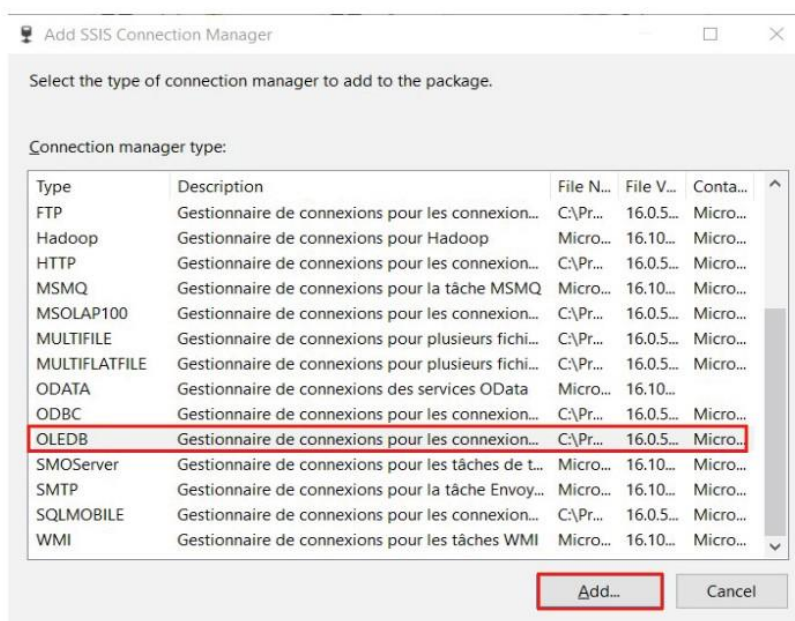


Figure 38 :1er Étape de configuration de la connexion PostgreSQL.

Dans la deuxième étape appuyer sur « Nouveau » pour choisir la source de données

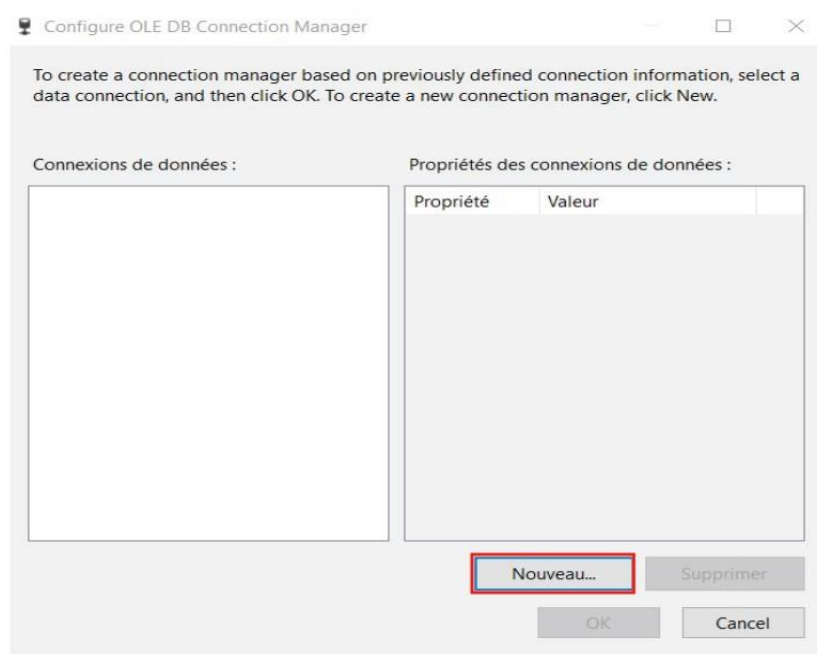


Figure 39 :2eme Étape de configuration de la connexion PostgreSQL.

Par la suite sélectionner le fournisseur OLEDB approprié pour la base de données à laquelle nous avons connecté « Intellisoft OLE DB provider for PostgreSQL »

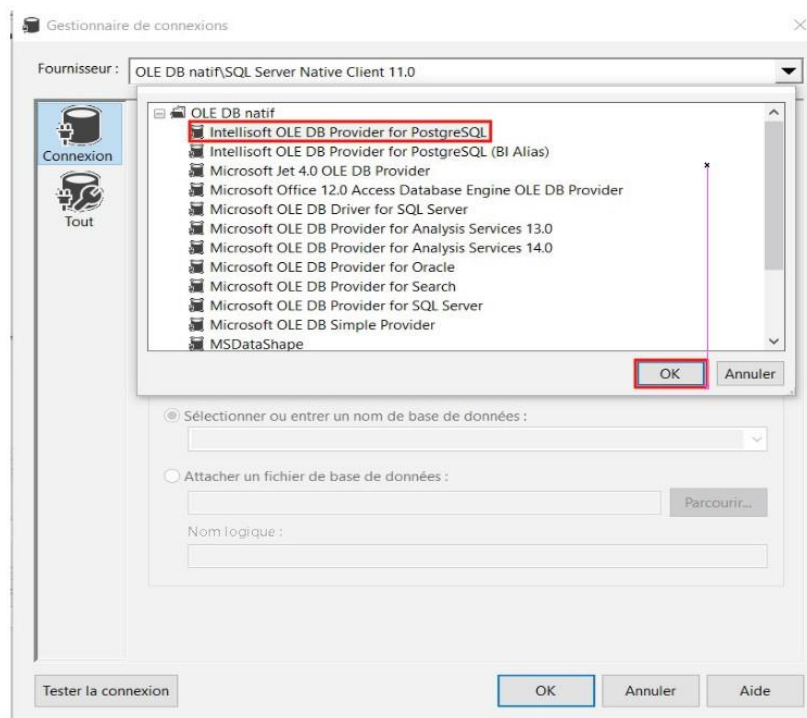


Figure 40 : 2eme Étape de configuration de la connexion PostgreSQL.

Finalement spécifier les détails de connexion nécessaires. Cela comprend généralement des informations telles que le nom du serveur, le nom d'utilisateur, le nom de la base de données et mot de passe. Enfin tester la connexion pour assurer qu'elle fonctionne correctement et cliquer sur Ok comme le montre la figure 41.

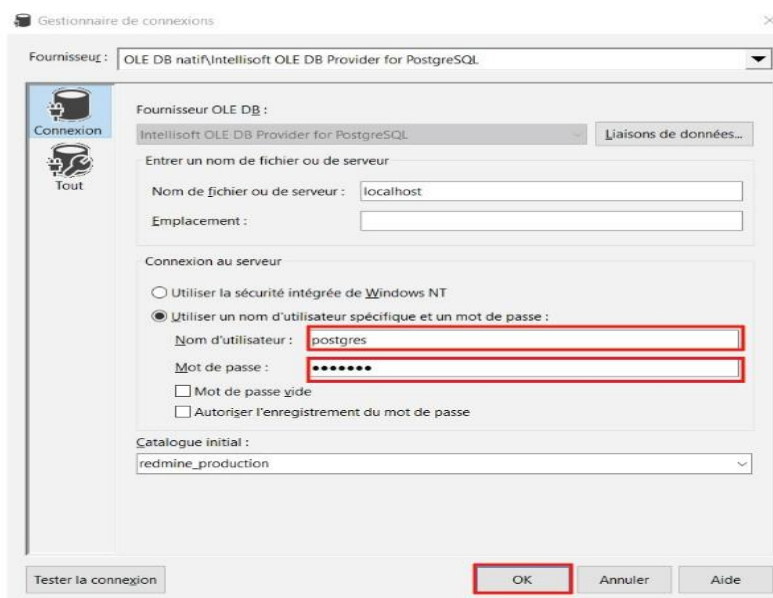


Figure 41 : Étapes de configuration de la connexion PostgreSQL.

Nous obtenons les connexions ci-dessous :

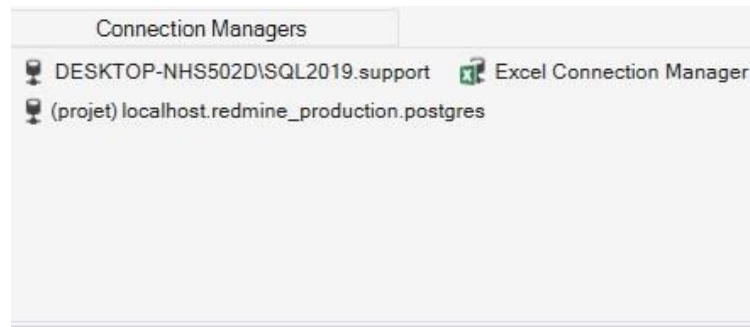


Figure 42 : Gestionnaires de connexions.

Notre ETL consiste en une première étape d'extraction, de transformation et de chargement des tableaux de dimensions. Ensuite, traiter les tables des faits. Cela est dû au fait que les tables de faits contiennent des clés étrangères vers les dimensions.

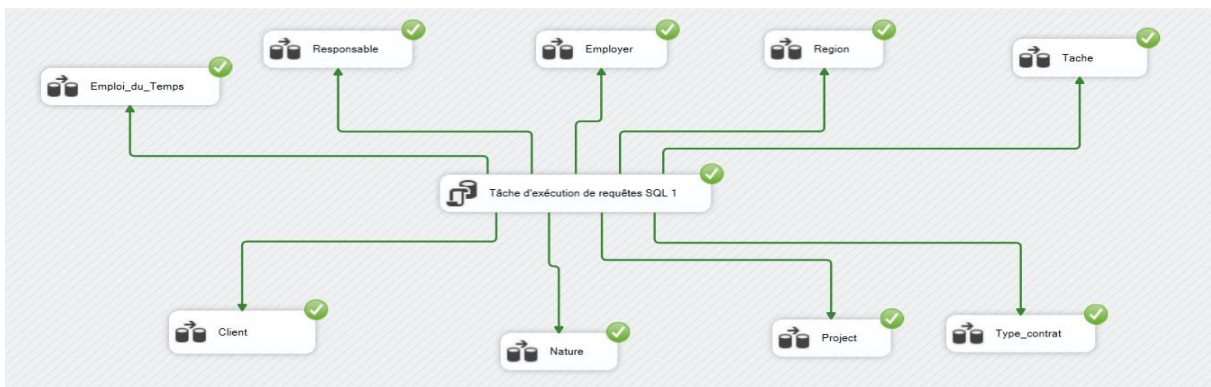


Figure 43 : Flux de contrôle du paquet « dimensions »

Ensuite, traiter les tables des faits. Cela est dû au fait que les tables de faits contiennent des clés étrangères vers les dimensions.

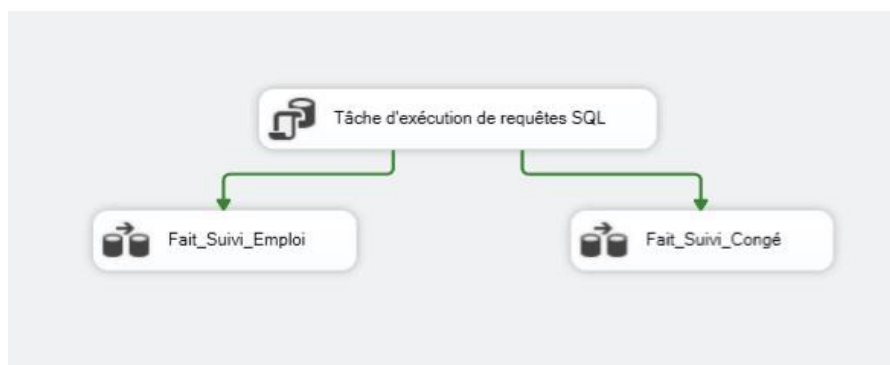


Figure 44 : Flux de contrôle du paquet « dimensions »

➤ Composition d'une dimension :

Ce composant contient un ensemble des tâches de flux de données relatives à l'extraction des nouvelles données, leur transformation et leur chargement dans les dimensions. Nous détaillons dans la figure 45 suivante le processus d'extraction, de transformation et de chargement des nouvelles données d'une tâche de flux de données relatives à la dimension Tache.

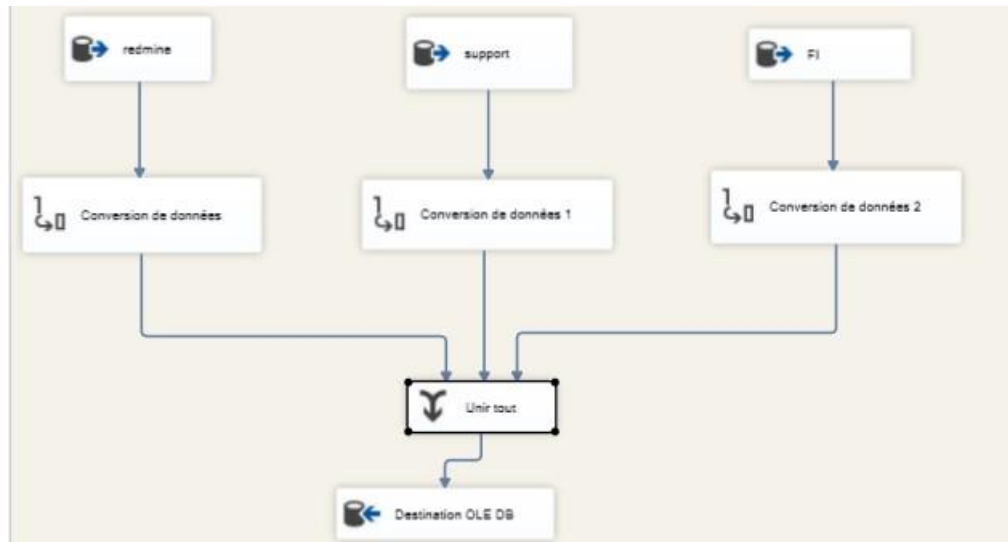


Figure 45 : Flux de données relative à la dimension « Tache »

Dans un premier temps, nous définissons les trois sources de données de la table Tache dans les bases opérationnelles pour récupérer les données qui vont être acheminées vers la dimension Service à travers le composant "Source OLE DB". Nous détaillons dans la figure 43 un exemple de chargement de données de la source Support à travers SQL Commande.

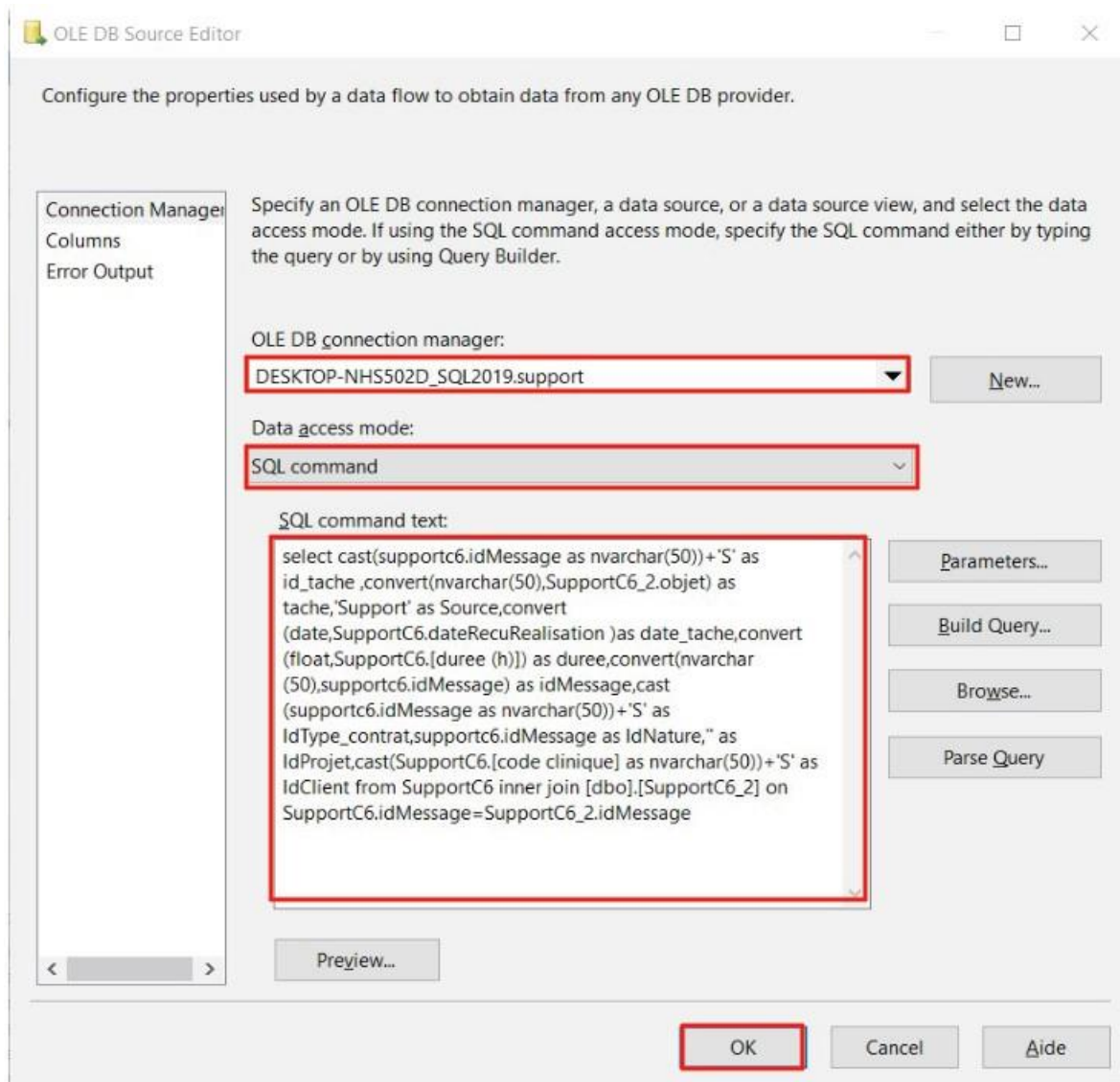


Figure 46 : Chargement de données de la source Support.

Ensuite, nous avons converti les données requises pour les mettre dans le même type, puis, à travers le composant "unir tout" nous a rassemblé et unifié les colonnes. Enfin, on a utilisé l'élément « Destination OLE DB » pour charger les données dans notre entrepôt. Comme l'illustre la figure 47.

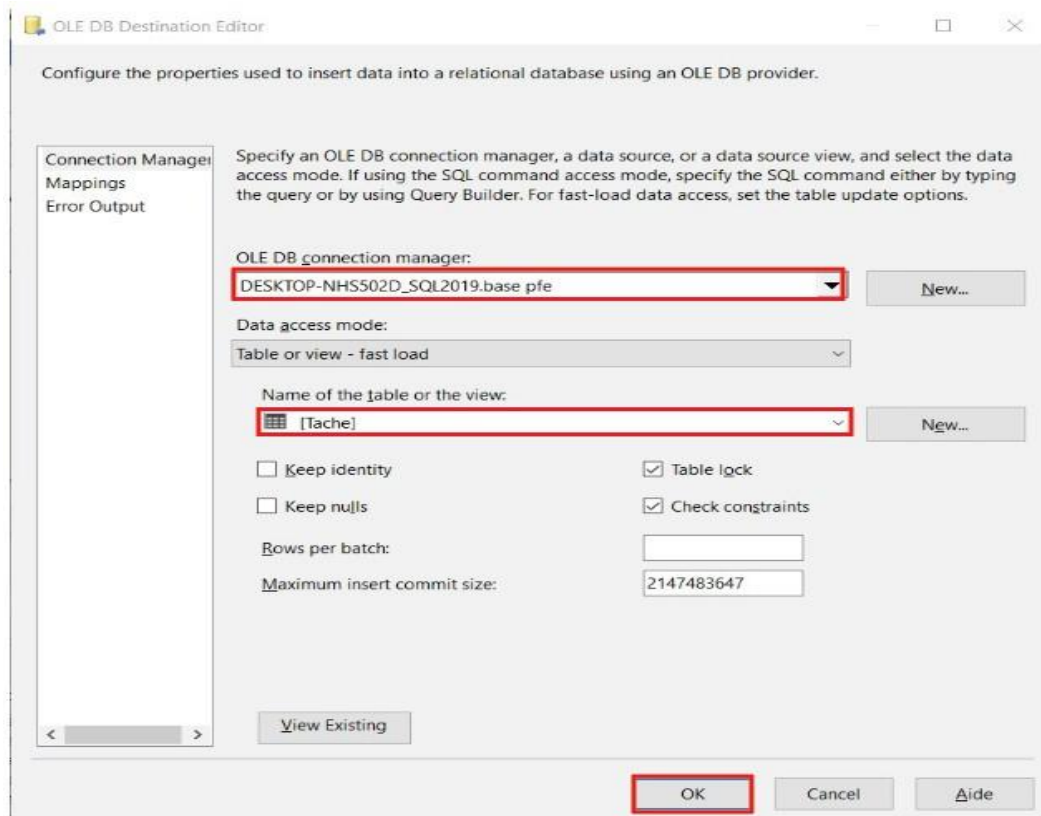


Figure 47 : Destination de la dimension Tache.

#### ➤ Composition d'un Fait :

Ce composant contient un ensemble de tâches de flux de données relatives chacune au traitement d'un fait. Nous détaillons dans la figure 48 le parcours de récupération des clés étrangères des dimensions et le processus de chargement des nouvelles données d'une tâche de flux de données relatives à la fait Suivi\_Emploi.

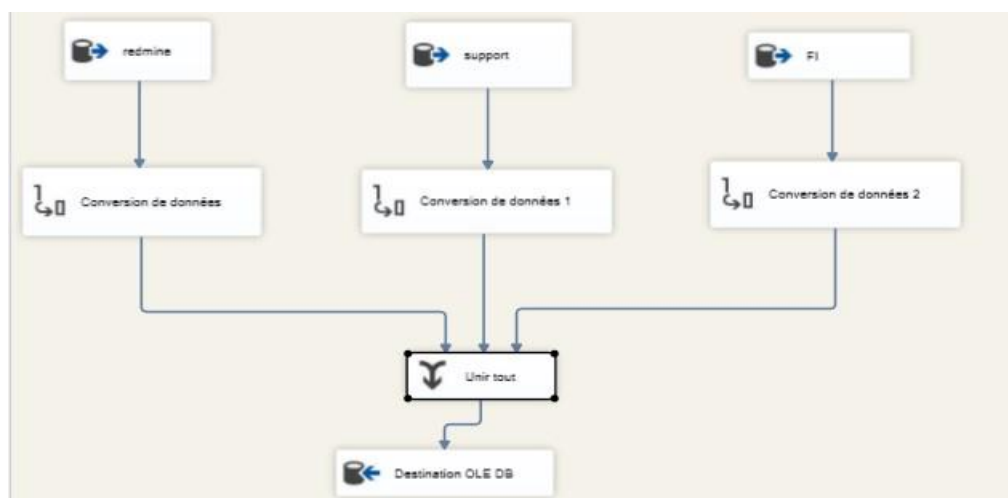


Figure 48 : Flux de données « Fait\_Suivi\_Emploi »

Le chargement de la table de fait se faire en trois grandes étapes :

Nous devons tout d'abord importer les données à partir de trois sources de données « Support », « Redmine » et « FicheInterv » qui contiennent les colonnes demandées. Puis, nous passons par une jointure de fusion qui nous permet de charger toutes les colonnes dans la table de fait à travers des requêtes SQL comme montre un exemple dans la figure suivante.

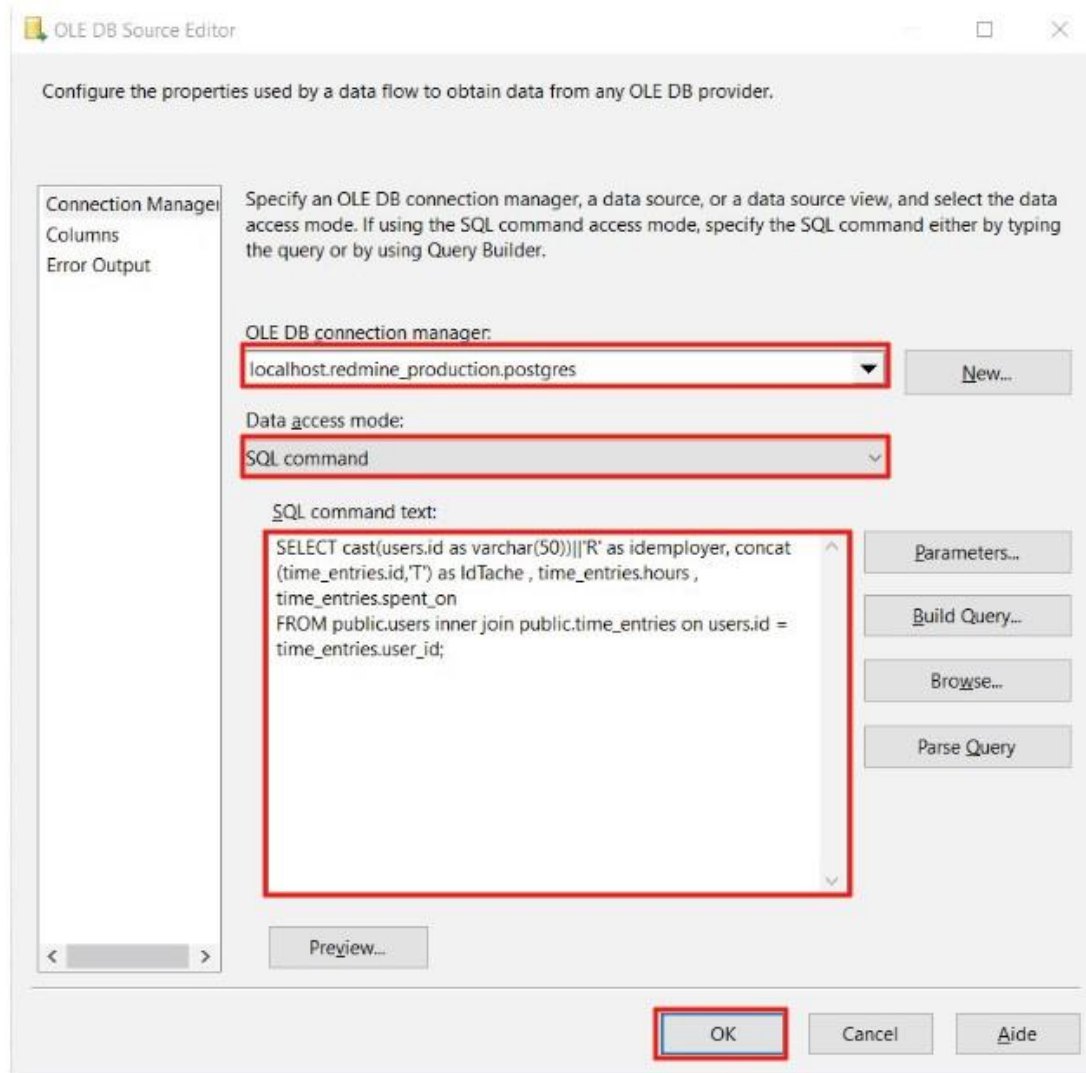


Figure 49 : Chargement de données de la source Redmine.

## Conclusion :

Dans ce chapitre, nous avons présenté la partie ETL du projet. En effet, grâce à ce dernier, les données ont été extraites des systèmes sources, nettoyées, normalisées et transformées en structures cohérentes et exploitables, à l'aide des outils techniques.

# Chapitre 5 :

---

Restitution des données de l'ED :



## 1. Introduction :

Dans ce chapitre, on va présenter la modélisation de l'application business intelligence (BI) en définissant ses différents types, usages et utilisateurs. Dans une deuxième étape, on va traiter des interfaces montrant la mise au point de cette application.

## 2. Modélisation de l'application BI :

Au moment de l'identification des besoins où nous avons maintenu les membres de la société cliente de ce système décisionnel, nous avons défini le type d'application d'interrogation de l'entrepôt de données attendu, pour quel usage, et pour quel type d'utilisateur. Nous détaillons dans la suite le type d'application BI attendue, l'utilisation de cette dernière et ses utilisateurs.

### 2.1 Type d'application BI :

Il existe différents types d'application BI, parmi lesquels nous distinguons :

- ❖ **Tableaux de bords** : c'est un outil de pilotage. Il s'agit d'un ensemble d'indicateurs de pilotage, établis périodiquement, à l'intention du responsable, afin de guider ses décisions et ses actions en vue d'atteindre les objectifs de performance.
- ❖ **Reporting** : c'est un outil de contrôle. Il s'agit d'un ensemble d'indicateurs de résultats, construit de façon périodique, afin d'informer la hiérarchie des performances de l'unité. Il permet de vérifier que les centres de responsabilité respectent leurs engagements contractuels.
- ❖ **Applications analytiques** : C'est la représentation des données sous la forme d'un cube multidimensionnel où chaque côté représente une dimension d'analyse et chaque case un indicateur permettant d'appliquer les opérations sur les données, la navigation dans les hiérarchies dimensionnelles et les visualisations des données sous la forme de tableaux croisés.

Nous avons choisi dans notre cas d'utiliser des tableaux de bord comme type d'application pour offrir une interface visuelle concise et informative, permettant aux utilisateurs de visualiser et de gérer facilement les données clés liées à leur domaine spécifique.

### 2.2 Type d'analyse BI :

Il existe différents types d'analyses BI, parmi lesquels nous distinguons :

- ❖ **BI stratégique** : il s'agit de l'analyse des tendances d'affaires selon une métrique. Il concerne les objectifs à long terme de l'entreprise. En effet, il permet à une entreprise de prendre des décisions éclairées concernant les conditions futures de sa place dans un

marché ou dans un secteur particulier.

- ❖ **BI tactique** : fournit aux décideurs des métriques de performance et les informations nécessaires pour surveiller les changements en temps réel de leur environnement et les aide à découvrir de nouvelles opportunités. Il concerne les objectifs à court terme de l'entreprise. De plus, il s'applique principalement aux gestionnaires et aux cadres supérieurs.
- ❖ **BI opérationnel** : fournit un support informatique aux points d'affaires de l'entreprise. Il concerne l'état opérationnel de l'entreprise. En effet, il utilise les données issues en temps réel des applications pour les croiser avec les données historiques de l'entreprise afin de fournir un support informationnel aux points d'affaire de l'entreprise.

Nous avons choisi d'utiliser le type d'analyse stratégique dans notre cas, ce qui signifie que nous avons concentré sur l'évaluation globale de l'entreprise et sur la prise de décision clés pour atteindre nos objectifs à long terme.

### 3. Développement de l'application :

#### 3.1 Projet Power BI :

La figure 47 présente le projet Power BI.

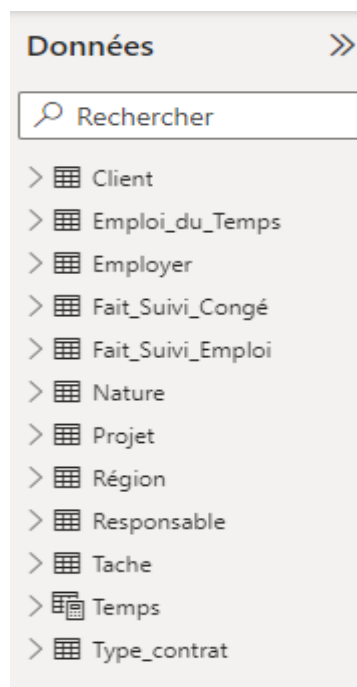


Figure 50 : Le projet Power BI.

Dans notre projet Power BI et grâce au langage DAX, nous avons mis en place les étapes suivantes :

❖ Nous avons créé la dimension date comme le montre la figure 51.

```
1 Temps = ADDCOLUMNS( CALENDARAUTO(), "Jour", DAY([Date]), "Mois", MONTH([Date]), "Annee", YEAR([Date]))
2
```

Date	Jour	Mois	Annee
31/12/2023 00:00:00	31	12	2023
30/12/2023 00:00:00	30	12	2023
29/12/2023 00:00:00	29	12	2023
28/12/2023 00:00:00	28	12	2023
27/12/2023 00:00:00	27	12	2023
26/12/2023 00:00:00	26	12	2023
25/12/2023 00:00:00	25	12	2023
24/12/2023 00:00:00	24	12	2023
23/12/2023 00:00:00	23	12	2023
22/12/2023 00:00:00	22	12	2023
21/12/2023 00:00:00	21	12	2023
20/12/2023 00:00:00	20	12	2023
19/12/2023 00:00:00	19	12	2023
18/12/2023 00:00:00	18	12	2023
17/12/2023 00:00:00	17	12	2023
16/12/2023 00:00:00	16	12	2023
15/12/2023 00:00:00	15	12	2023

Figure 51 : La requête chargement de la dimension date.

❖ Nous avons défini et crée les mesures nécessaires dans chaque fait comme illustre la figure 52.

```
1 Nbj_Emp_saisie =
2 var duree_fiche=SUMX(
3     Fait_Suivi_Emploi,
4     IF(
5         RIGHT(Fait_Suivi_Emploi[IdTache], 1) = "F",
6         IF(
7             Fait_Suivi_Emploi[Duree] = 0,
8             0,
9             DIVIDE(Fait_Suivi_Emploi[Duree], 60)
10        )
11    )
12 ))
13
14 return CALCULATE(sum(Fait_Suivi_Emploi[Duree])+duree_fiche,filter(Fait_Suivi_Emploi,RIGHT(Fait_Suivi_Emploi[IdTache], 1) <> "F"))
```

Figure 52 : Requête de la mesure Nbj\_Emp\_saisie avec DAX.

### 3.2 Présentation de l'application :

Le tableau de bord est un outil puissant qui permet de regrouper plusieurs graphiques et tableaux afin d'offrir une vision globale de l'activité de l'entreprise. Power BI contient de nombreux graphiques qui nous permettent de visualiser nos données, nous devons donc choisir des visuels qui correspondent exactement à nos besoins et qu'ils soient compatibles avec les données présentes dans notre entrepôt de données, c'est pour cette raison que nous avons effectué une explication sur le choix des graphiques utilisés après chaque tableau de bord.

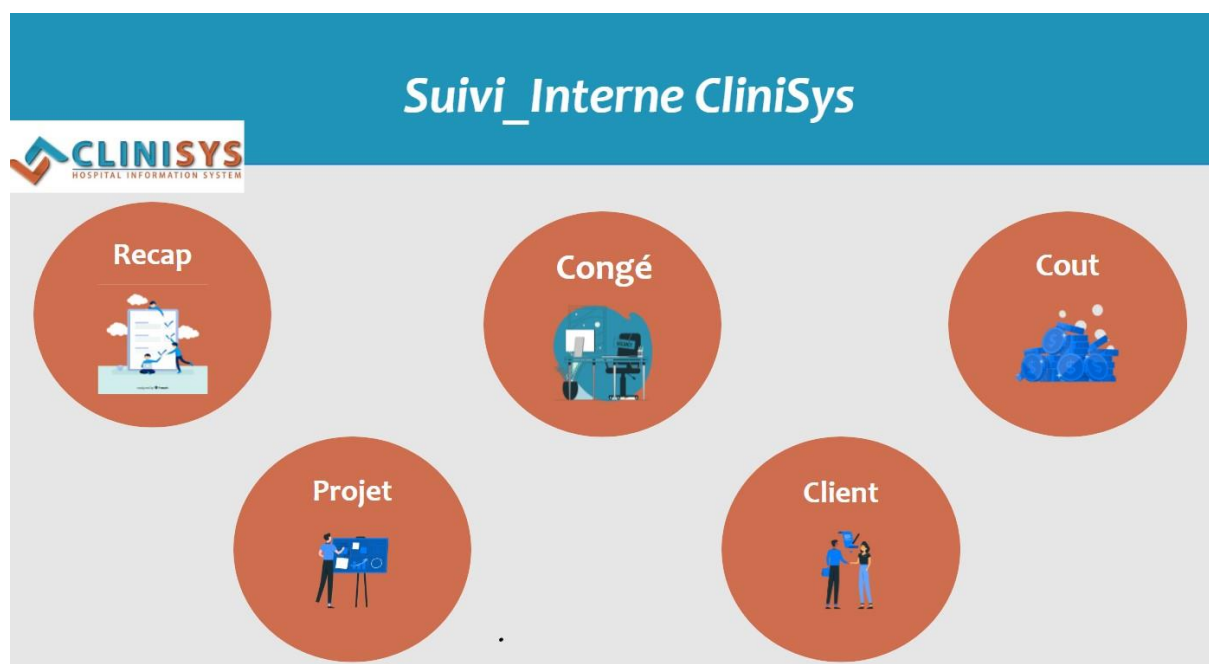


Figure 53 : Tableau de bord « Suivi Interne ».

En premier lieu, nous avons présenté le tableau de bord de menu avec des boutons liés entre les pages qui offrent une interface intuitive pour naviguer facilement dans un système ou une application. Les boutons sont disposés de manière efficace, permettant aux utilisateurs d'accéder rapidement aux différentes sections. Ce tableau de bord facilite la navigation en réduisant le nombre de clics nécessaires pour atteindre une page spécifique.



Figure 54 : Tableau de bord « Vue d'ensemble des ressources et des performances ».

En second lieu, nous avons exposé le tableau de bord récapitulatif des informations clés telles que le mois, l'année, la source de données, le nombre de clients, l'effectif des employés, le nombre des projets, les coûts réels et estimés des employés, ainsi que les jours de travail estimés et réels des employés. Cela fournit une vue concise et informative de l'état global des ressources et des performances du projet en trois lignes.

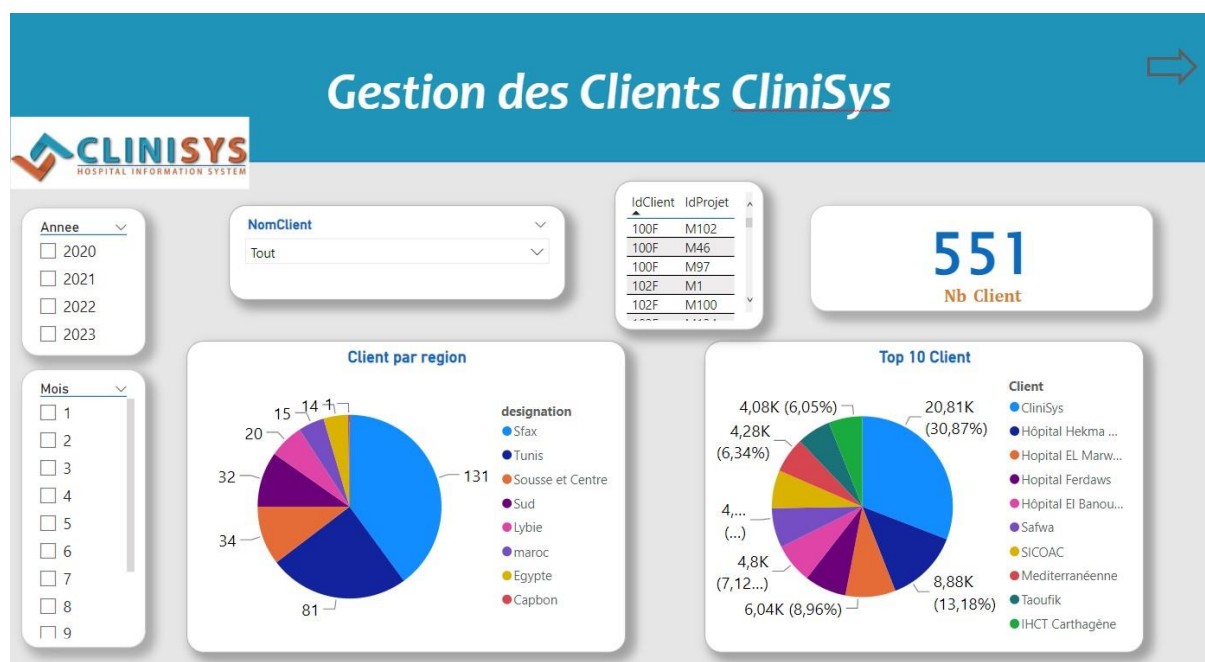


Figure 55 : Tableau de bord « Gestion des clients Clinisys».

Par la suite, le tableau de bord de gestion des clients comprend plusieurs visuels clés qui permettent de suivre et de gérer efficacement les informations relatives aux clients :

- ❖ Graphique en secteur - Clients par région : Ce graphique met en évidence la répartition des clients par région. Il permet de visualiser rapidement la proportion de clients dans chaque région, ce qui peut aider à identifier les marchés clés ou les régions où l'activité est plus concentrée.
- ❖ Graphique en secteur - Top 10 clients : affiche les 10 clients les plus importants ou les plus rentables. Il illustre visuellement la part de chaque client dans l'ensemble des activités commerciales, offrant ainsi une vue rapide des clients les plus influents.
- ❖ Tableau des informations clients : contient des informations clés sur les clients, telles que leur nom, leur identifiant client et le nombre de projets auxquels ils sont associés. Il permet de visualiser et de filtrer les données clients, facilitant ainsi la gestion et le suivi des activités spécifiques à chaque client.

Ce dernier offre une vue globale des informations clés relatives aux clients.

Les visuels en secteur permettent d'identifier les tendances régionales et les clients les plus importants. Cela facilite la gestion des clients.

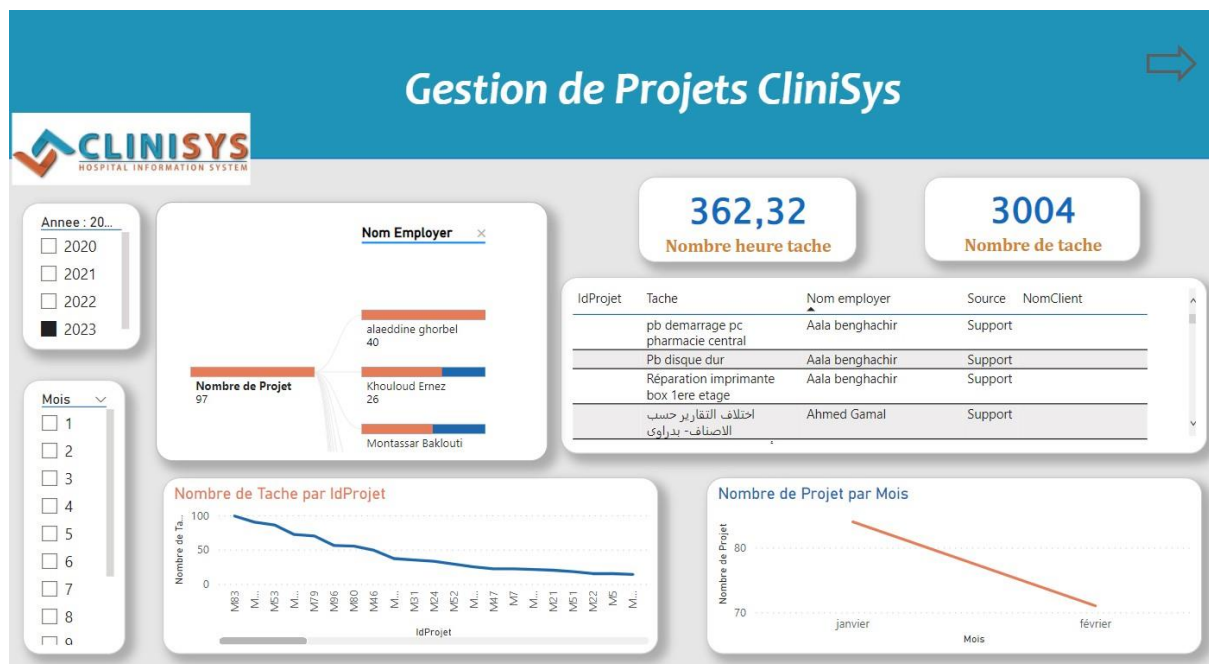


Figure 56 : Tableau de bord « Gestion de projet Clinisys».

Ce tableau de bord dispose comme visuels :

- ❖ Graphique en courbe - Nombre de projets par mois : Ce graphique en courbe donner le privilège de visualiser l'évolution du nombre de projets au fil des mois. Il met en évidence les périodes d'activité intense ou de faible charge de travail, ce qui permet de planifier les ressources en conséquence.
- ❖ Tableau des informations clés de la tâche : Ce tableau comprend des informations essentielles sur les tâches, telles que le nom de tâche, l'IdProjet, le nom du client et la source de données. Il offre une vue d'ensemble des attributions, des responsabilités et des parties prenantes impliquées dans le projet.
- ❖ Arbre de composition - Nombre de projets par nom d'employé : Cet arbre de composition présente une visualisation hiérarchique du nombre de projets attribués à chaque employé. Il permet d'évaluer la répartition de la charge de travail entre les membres de l'équipe et d'identifier les employés les plus sollicités.

Ce tableau de bord de gestion de projet offre une vue globale des projets, des tâches, des employés. Il permet de surveiller l'avancement du projet, de gérer les ressources efficacement et de prendre des décisions éclairées pour assurer le succès du projet.

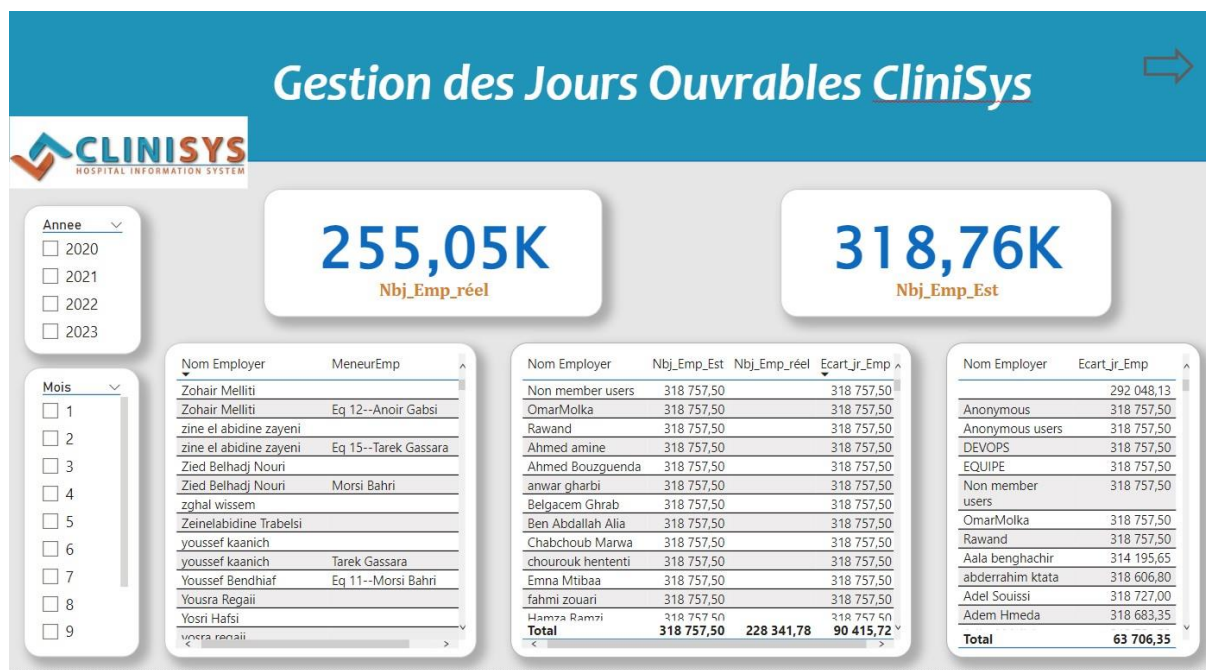


Figure 57 : Tableau de bord « Gestion des jours ouvrables Clinisys ».

Puis, le tableau de bord de gestion des jours ouvrables est conçu pour aider les décideurs à suivre et gérer les jours d'emploi. Il contient les visuels suivants :

- ❖ Tableau des employés et de leurs responsables : Ce tableau affiche les noms des employés ainsi que les noms de leurs responsables hiérarchiques
- ❖ Tableau des employés avec les jours d'emploi estimés et réels et l'écart entre eux Il permet de comparer les estimations avec les résultats réels et d'analyser les écarts.
- ❖ Pour le tableau de l'écart des jours d'emploi Ce dernier présente les noms des employés avec leurs écarts respectifs. Il met en évidence les employés qui ont les écarts les plus significatifs, écarts positifs signifie manque de saisie, écart négatifs c'est à dire qu'il y a des jours supplémentaires.

En utilisant ce tableau de bord, les responsables et les gestionnaires peuvent avoir une vue d'ensemble de l'utilisation des jours de travail, identifier les employés qui ont des écarts importants entre les estimations et les réalisations, et prendre des mesures pour améliorer l'efficacité et la planification des ressources.



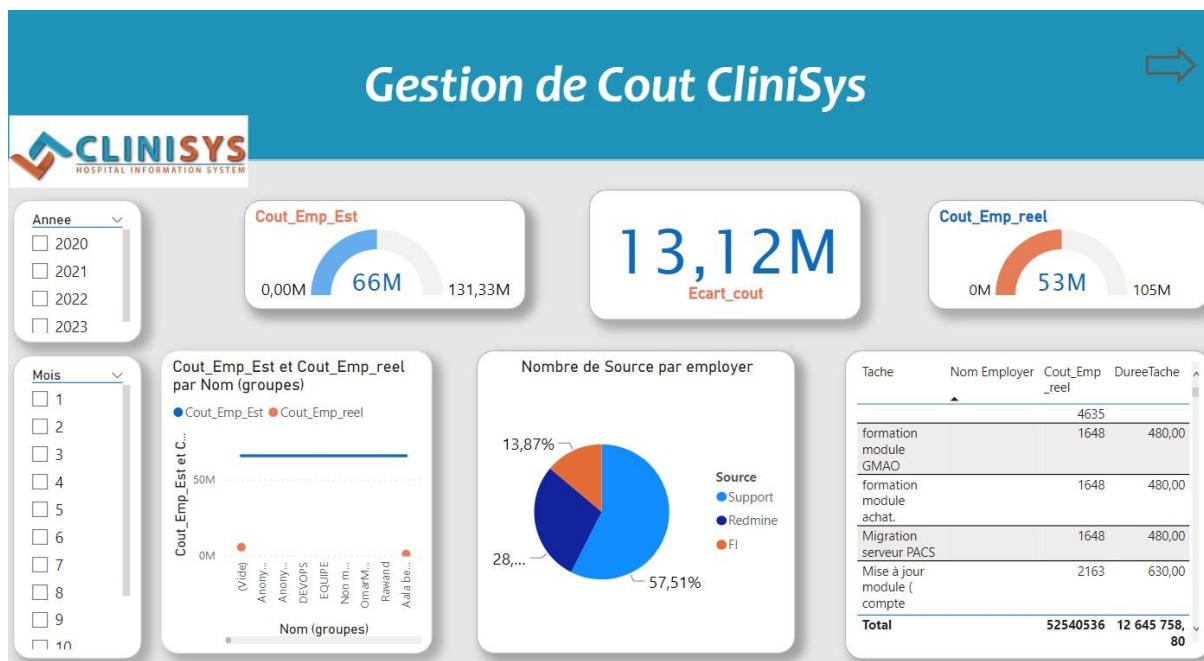


Figure 58 : Tableau de bord « Gestion de cout Clinisys ».

En passe au tableau de bord suivante comprend des visuels clés permettant de suivre et d'analyser les données relatives aux coûts des employés et à l'utilisation des sources de données:

- ❖ **Coût employé estimé** : Ce visuel affiche le coût estimé des employés pour une période donnée. Il permet de visualiser les dépenses prévues liées aux employés, fournissant ainsi une estimation des coûts associés à la main-d'œuvre.
- ❖ **Coût employé réel** : Ce visuel présente les coûts réels des employés pour la même période. Il permet de comparer les dépenses réelles par rapport aux estimations, offrant ainsi une évaluation précise des coûts effectifs liés aux employés.
- ❖ **Écart de coût** : Ce visuel met en évidence l'écart entre les coûts estimés et les coûts réels des employés. Il permet de visualiser les différences entre les prévisions et les dépenses réelles, aidant ainsi à identifier les écarts et à prendre des mesures correctives si nécessaire.
- ❖ **Nombre de sources par employé** : Ce visuel montre le nombre de sources de données utilisées par chaque employé. Il permet de quantifier l'utilisation des différentes sources de données par les membres de l'équipe, fournissant ainsi des informations sur la diversité et l'étendue des données consultées.
- ❖ **Graphique en courbe** : Ce graphique en courbe présente l'évolution des données clés au fil du temps. Il peut être utilisé pour représenter différents paramètres tels que les coûts, les performances ou les utilisations des ressources, offrant

ainsi une vue temporelle des tendances et des variations.

Ce tableau de bord fournit une vision complète des données liées aux coûts des employés et à l'utilisation des sources de données. Les visuels permettent d'évaluer les coûts estimés et réels, de suivre les écarts.

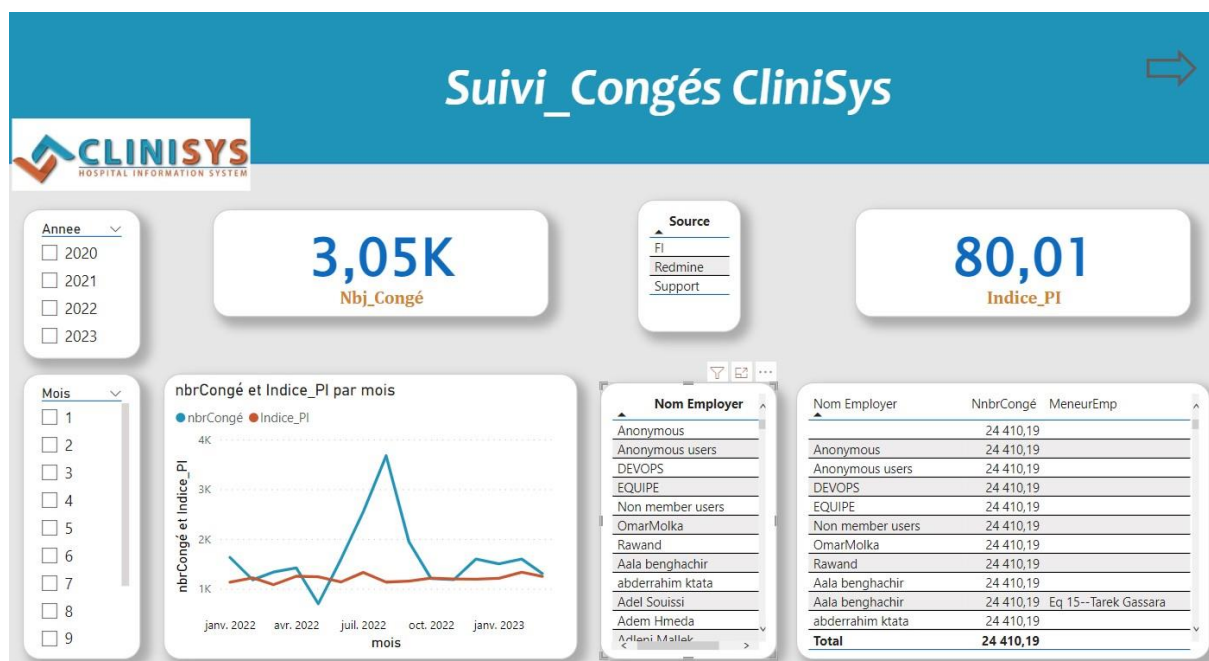


Figure 59 : Tableau de bord « suivi\_congés Clinisys ».

Finalement, le tableau de bord de suivi des congés qui comprend des visuels clés pour suivre les informations relatives aux congés des employés :

- ❖ Graphique en courbe - Nombre de congés par mois : Ce graphique en courbe affiche le nombre de congés pris par mois. Il permet de visualiser les variations dans les congés au fil du temps, ce qui peut aider à identifier les périodes de forte ou de faible demande de congés.
- ❖ Graphique en courbe - Indice de performance individuel par mois : Ce graphique en courbe présente l'indice de performance individuel des employés pour chaque mois. Il permet de suivre la performance individuelle au fil du temps, en mettant en évidence les périodes de performances exceptionnelles ou en deçà des attentes.
- ❖ Tableau des informations des employés : Ce tableau contient les informations clés sur les employés, tels que leur nom, leur responsable hiérarchique et le nombre de jours de congé pris. Il offre une vue détaillée des employés, de leurs responsables et de leur historique de congés, facilitant ainsi le suivi et la gestion des congés.

Le tableau de bord de suivi des congés permet de suivre de manière efficace les congés des employés, d'évaluer la performance individuelle et de faciliter la gestion des ressources.

# Conclusion générale :

De nos jours, le business intelligence évolue sans cesse, ce qui nécessite une attention toute particulière pour se tenir informé des dernières tendances. De plus, la BI est un domaine où les enjeux sont considérables et pour lequel il est très important d'anticiper sur l'avenir. En effet, les technologies BI deviennent une nécessité pour les entreprises modernes afin qu'elles gardent leurs évolution et existence sur le marché.

Dans ce contexte, convaincue de cette nécessité, la société CLINISYS nous a accueilli au sein de sa direction informatique dans le cadre de notre stage de fin d'étude en vue de l'obtention du diplôme national en licence en Science Informatique. CLINISYS nous a confié la construction d'un entrepôt de données pour l'analyse de ses activités interne de la société.

La phase de conception pourrait bien être considérée comme la phase de décision la plus importante, puisqu'on s'y est strictement engagé à planifier la réponse à tous les besoins aussi bien que de définir le statut de l'application apte à satisfaire les estimations de l'entreprise.

Tout au long de notre travail de conception et de réalisation, nous avons utilisé la méthode ascendante comme une méthode mixte. Cette démarche a permis de répondre aux attentes et besoins des utilisateurs tout en exploitant au mieux les données générées par le système opérationnel de manière à anticiper sur des besoins non exprimés. Dans une deuxième étape, la modélisation de la zone de stockage des données a été réalisée à l'aide des principes de la modélisation multidimensionnelle en adoptant un schéma en flocon de neige.

La partie ETL, a été la partie la plus ennuyeuse et la plus longue du projet. Cette étape nous a permis de concevoir et de réaliser les routines d'extraction, transformation et chargement des données sous l'outil SSIS de SQL Server Management Studio.

En ce qui concerne les rapports, cette étape nous a permis de concevoir et de réaliser les différents visuels et tableaux de bord de notre projet à l'aide de l'outil Power BI.

Fondé sur un entrepôt de données, notre système d'aide à la décision est ouvert aux futures évolutions :

- ❖ Extension de l'entrepôt par la prise en compte des autres sources de données.
- ❖ Règlementation sur la confidentialité des données

# Bibliographie :

**[Inmon 1996]** Inmon W. H, “Building the Data Warehouse”. John Wiley & Sons, deuxième édition, ISBN 04771-14161-5, 1996.

**[Kimball, 1996]** Kimball R., “The data warehouse toolkit: practical techniques for building dimensional data warehouses”. John Wiley & Sons, 1996.

**[Kimball 2001]** Ralph Kimball, Laura Reeves. Concevoir et déployer un datawarehouse. Eyrolles, 2001.

**[Kimball et al., 2002]** Ralph Kimball, Margy Ross. The datawarehouse toolkit. Wiley Compute, 2002.

**[Kimball et al., 2008]** Ralph Kimball, Laura Reeves, Margy Ross, Warren Thornthwaite. Le datawarehouse guide de conduit de projet.2008.

**[Golfarelli 2009]** Matteo Golfarelli, “Data Warehouse Design: Modern Principles and Methodologies”. New York McGraw-Hill 2009.

**[Ghozzi 2004]** Ghozzi F., “Conception et manipulation de bases de données dimensionnelles à contraintes”. Thèse de doctorat, Université Paul Sabatier, Toulouse, France, 2004.

# Netographie :

- [1] <https://www.lebigdata.fr/business-intelligence-definition>
- [2] <https://docs.microsoft.com/en-us/dax/>
- [3] <https://waytolearnx.com/2018/09/difference-entre-rolap-et-molap>
- [4] <https://tssperformance.com/difference-entre-tableau-de-bord-et-reporting-ou-rapport/>
- [5] <https://asana.com/fr/resources/strategy-vs-tactics>
- [6] <https://www.microsoft.com /BI>
- [7] [www.tableau-de-bord.org](http://www.tableau-de-bord.org)
- [8] <https://www.decivision.com/blog/power-bi>