

République Algérienne Démocratique et Populaire
Université Abou Bakr Belkaid– Tlemcen
Faculté des Sciences
Département d'Informatique

Mémoire de fin d'études

Pour l'obtention du diplôme de Master en Informatique

Option: génie logiciel (GL)

Thème

Développement d'une solution Business Intelligence pour la scolarité de l'Université de Tlemcen

Réalisé par :

- Abdallah BENSLIMANE
- Abdel Hafid BOUKLIHACENE

Présenté le 22 Juin 2017 devant le jury composé de MM.

- Mr Abdelkrim BENAMAR *(Président)*
- Mr Azzedine CHIKH *(Examinateur)*
- Mr Houcine MATALLAH *(Encadrant)*
- Mr Mohammed TADLAOUI *(Encadrant)*

Remerciements

Nos remerciements vont tout spécialement à nos familles, qui ont sus nous supporter et encourager tout au long de notre vie, ainsi que pour leur aide inestimable, leur patience et leur soutien indéfectible.

Nous tenons aussi, à remercier tous les enseignants qui ont contribué de près ou de loin à notre formation.

Nous remercions Mr Mohammed TADLAOUI, Mr Houcine MATALLAH et Mm Amel HALFAOUI pour avoir assuré l'encadrement de ce projet.

On remercie vivement Messieurs les membres du jury d'avoir accepté d'évaluer ce travail.

Pour finir, et afin de n'oublier personne (amis, membre de la famille et tous ceux qui nous sont chers) nous utiliserons la formule : « Merci à... ».

Dédicaces

Je dédie ce modeste travail à :

Mes parents, qui n'ont jamais cessé de m'encourager et me soutenir,

Mon frère : Ahmed, et ma sœur : Rania

Mon binôme et ami Abdallah,

Mesamis :Norreddine, Yassine, Djawed, Choukri, houssam, Selim, Fayçal,

Ilyes, Abdelkadder, Walid, Imed...

Tous les membres de ma famille,

A tous ceux qui me sont chers

Hafid

Dédicaces

Je dédie ce modeste travail

À mes parents

À ma très chère famille

À toute mes amis

À tous ceux qui m'ont encouragé

Abdallah

Table des matières

Introduction générale.....	3
Motivation.....	4
Problématique	4
Objectifs du projet	4
Organisation du mémoire	5
Chapitre 1 : Informatique Décisionnelle	6
I Introduction	6
II Les systèmes décisionnels (Business Intelligence)	6
II.1 Les avantages d'une solution BI	7
III Décisionnel vs transactionnel (OLAP vs OLTP).....	7
IV Data Warehouse	8
IV.1 Qu'est Ce qu'un Data Warehouse.....	8
IV.2 Historique des Data Warehouse.....	9
IV.3 Modélisation de données d'un DW.....	10
V Concept OLAP	12
VI Architecture et composants d'un système BI	13
VII Solutions BI présentes sur le marché	15
VIII Conclusion	18
Chapitre 2 Analyse et conception	19
I. Introduction	19
II. Planification et conduite du projet	19
III. Analyse de l'existant	21
III.1. Identification de l'environnement.....	21
III.2. Retro-Ingénierie	21
III.3. Problèmes et obstacles rencontrés	27
IV. Etude des besoins	27
IV.1. Rapports de gestion	27
IV.2. Utilisateurs du système.....	29
V. Conception de la solution	30
V.1. Diagramme des cas d'utilisation global	30

V.2.	Conception multidimensionnelle	30
V.3.	Mapping des données	34
V.4.	Architecture Globale de l'application	36
VI.	Conclusion	37
Chapitre 3 Réalisation.....	38	
I.	Introduction	38
II.	Environnement logiciel	38
II.1.	SGBD Microsoft SQL Server.....	38
II.2.	Microsoft Visual Studio	38
II.3.	SQL Server Integration Services (SSIS)	38
II.4.	SQL Server Analysis Services (SSAS)	38
II.5.	SQL Server Reporting Services (SSRS)	38
III.	Processus d'extraction, de transformation et de chargement	39
IV.	Génération du cube OLAP	44
V.	Génération des rapports	44
VI.	Nettoyage des données	47
VII.	Gestion des droits d'accès	48
VIII.	Organisation des fichiers de la plateforme.....	49
IX.	Architecture technique	49
Conclusion générale	51	
Références bibliographiques	53	
Liste des figures	55	
Liste des tableaux	56	
Liste des abréviations	57	
Annexe1 : table de mapping.....	54	

Introduction générale

Introduction générale

Les établissements d'enseignement supérieur du monde entier exercent aujourd'hui dans un environnement dynamique et complexe. Le processus de mondialisation et le développement rapide des systèmes d'information et de la communication, ont conduit à une forte compétition, non seulement entre les compagnies et les entreprises, mais aussi entre les universités.

Les grandes universités du monde aux Etats-Unis, en Europe ou en Australie ont été contrariées par un tel changement qui les a poussées à chercher à être proactifs, en introduisant de nouvelles approches de gestion et en utilisant des méthodes progressives et plus performantes, afin de rester compétitif. Les universités modernes sont déjà conscientes du besoin urgent d'analyser parfaitement les données disponibles dans leurs différentes bases de données des systèmes IT, afin de tirer les connaissances approfondies des étudiants, pour mieux comprendre leurs caractéristiques d'apprentissage individuel et leurs besoins éducatifs.

La disponibilité des technologies de l'information(IT) à tous les niveaux ces dernières années, a amené les universités à collecter de gros volumes de données se référant à leur étudiants, enseignants, personnel administrative, ...etc. Cependant, il n'existe toujours pas de système automatique ou semi-automatique universitaire développé qui permet de ressortir des données propres, utiles et adaptées à l'analyse et la prise de décision. Généralement, les éducateurs interagissent directement avec les étudiants, surveillent le processus d'apprentissage, reçoivent une réaction en fonction des observations personnelles et à partir des documents papier, puisqu'ils sont dépourvus de moyens efficaces pour extraire les données pertinentes qui pourrait être utile pour améliorer la livraison du matériel pédagogique qui tient en compte des capacités d'apprentissage individuelles et des besoins des élèves.

En réponse à ce besoin, une multitude de nouvelles technologies de l'information, méthodes et outils ont été introduit et largement utilisés au cours des 25 dernières années, connus sous différents termes : système d'aide à la décision, les systèmes d'information exécutifs, ...etc., connu maintenant sous le nom de l'informatique décisionnelle ou Business Intelligence (BI). La BI fusionne un ensemble d'architectures, des entrepôts de données et des outils analytiques, qui vise de faciliter la réduction de l'écart entre la performance actuelle et celle désirée d'une organisation.

Les universités ont compris que pour rester compétitif et pouvoir saisir de nouvelles opportunités, elles doivent être capables de prendre des décisions rapides et compétentes. Les systèmes de BI largement utilisés par les entreprises à caractère commercial ou industriel ont pu attirer l'attention des gestionnaires des établissements d'enseignement. De

nombreuses universités ont déjà anticipé en introduisant de tels systèmes; D'autres sont actuellement en cours de conception ou de réalisation de ces projets.

Motivation

En étant des étudiants en master informatique option « génie logiciel », à l'université Abou Bakr Belkaid Tlemcen, on a voulu contribuer à l'évolution de notre université dans le domaine d'automatisation des différentes tâches, même pour celles qui étaient réservées uniquement à des utilisateurs humains. L'université algérienne et en particulier celle de Tlemcen souffre du manque d'outils décisionnels qui fournissent les différents éléments d'information nécessaires pour d'éventuelles prises de décision.

D'autre part, vu que ce type de sujets fait partie d'une nouvelle tendance dont on a jamais exploré concrètement, et vu les notions de base introduites au sujet du business intelligence durant notre première année master dans le module bases de données avancées, on a été très motivé pour travailler sur un vrai projet BI dans une entreprise ou une organisation.

Ainsi la motivation est double, d'un côté, vouloir explorer un domaine informatique d'actualité et très dynamique, et d'un autre côté, participer au développement et l'évolution de l'université algérienne et en particulier l'université de Tlemcen pour qu'elle soit compétitive et ayant un rang mondial.

Problématique

La décision nécessite une matière première : la donnée ou l'information. Si l'information est incomplète, floue, imprécise..., probablement, les décisions seront approximatives, inadaptées, voire fatales; par contre si l'information est exacte, complète et pertinente, les décisions ont davantage de chances d'avoir une efficacité garantie. Justement, le manque de cette matière première sous forme d'information utile ou bien la difficulté de la faire ressortir, constitue aujourd'hui, une vraie problématique pour les responsables au sein de l'université de Tlemcen pour l'élaboration des rapports de gestion, l'édition des graphes et statistiques et les problèmes de prise de décisions en général.

Objectifs du projet

Ce projet a pour but d'introduire, les processus d'informatique décisionnelle au sein de l'université Abou Bakr Belkaid, en mettant à la disposition des décideurs, un support fiable pour une meilleure prise de décision. Ainsi, les principaux objectifs de ce PFE sont :

- La réalisation d'une plateforme dynamique qui permette la visualisation instantanée des rapports via une interface unique avec la possibilité de générer des rapports automatiques personnalisés selon le besoin de l'utilisateur.

- Offrir des informations fiables et cohérentes contenant les règles de gestion souhaitée.
- L'amélioration des performances et la réduction de la durée globale de l'élaboration des rapports.

Organisation du mémoire

Le présent rapport de projet de fin d'étude s'articule autour de 3 chapitres :

- **Chapitre1** : est destiné à la présentation des concepts de base de l'informatique décisionnelle ainsi que les entrepôts de données.
- **Chapitre2** : est consacré à l'étude de l'existant, la définition des besoins et enfin la conception de la solution proposée.
- **Chapitre3** : est le dernier chapitre qui décrit toutes les phases du processus de réalisation de l'application.

Nous terminerons ce document par une conclusion générale qui résume le travail effectué avec des perspectives pour des futurs travaux en relation de ce projet.

Chapitre1 : Informatique Décisionnelle

Chapitre 1 : Informatique Décisionnelle

I Introduction

Toutes les entreprises du monde disposent d'une masse de données plus ou moins considérable. Ces informations proviennent soit de sources internes (générées par leurs systèmes opérationnels au fil des activités journalières) ou bien de sources externes (web, fichiers plats, ... etc.).

Cette surabondance des données d'une part, et les limites des systèmes opérationnels pour exploiter ces volumes de données à des fins d'analyse d'une autre part, ont conduit et pousser les entreprises à tourner vers une nouvelle ère informatique dite décisionnelle (Business Intelligence). La BI se concentre sur la compréhension de l'environnement de l'entreprise et offre aux décideurs une meilleure vision de leur environnement et son évolution pour une bonne prise de décision.

II Les systèmes décisionnels (Business Intelligence)

La BI est un terme générique qui fait référence à un ensemble se composant d'applications, d'infrastructures, d'outils et de bonnes pratiques qui permettent l'analyse de données dans le but d'améliorer les performances et aider à la prise de meilleures décisions au sein d'une organisation, sur la base des données qui ont été analysé et fourni par le système BI.

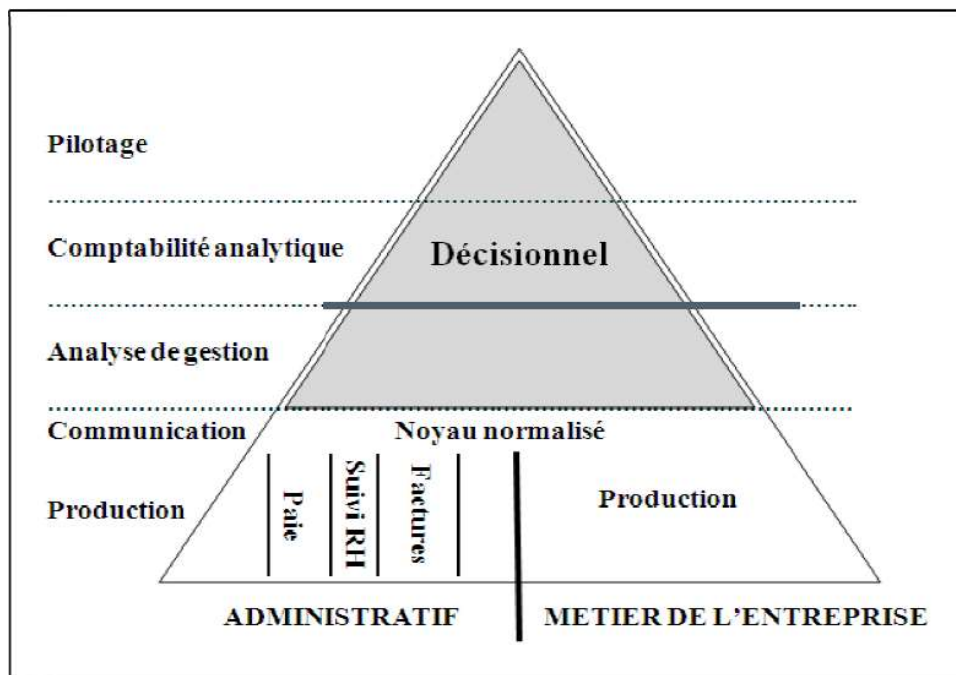


Figure 1-1 Le décisionnel au sein du système d'information [1]

La figure ci-dessus illustre parfaitement la place d'un système BI au sein de l'organisation d'une entreprise. Ce dernier assume plusieurs fonctions clés de l'entreprise.

II.1 Les avantages d'une solution BI

Le terme « BI » est devenu sans aucun doute, l'un des mots du buzz les plus utilisés dans le monde du business. Les nouvelles solutions proposées sont facile à utiliser, dotées d'une simple interface ergonomique. En revanche, la question qui doit être posée est :« est-ce que j'ai vraiment besoin d'une solution BI dans mon entreprise ? ». Si la réponse est oui : « Comment ça peut influencer mes décisions au sein de mon entreprise ? »

Voilà les bénéfices apportés couramment, en implémentant une solution BI au sein d'une organisation :

- Rapidité de la prise de décision fondée sur des faits ;
- Présentation uniforme d'informations fiables pour tous les membres de l'organisation (« Une seule version de la vérité ») ;
- Capacité d'intégration de données, provenant de diverses sources hétérogènes (ERP, CRM, systèmes comptables, feuilles de calcul, paie...) dans le même système;
- Amélioration de l'efficacité et identification de nouvelles opportunités ;
- Meilleure compréhension du business « passé, présent, futur » ;
- Elaboration des rapports nécessaires au bon moment, à tous les niveaux organisationnels, et dans les meilleurs formats possibles. [2]

III Décisionnel vs transactionnel (OLAP vs OLTP)

On peut diviser les systèmes informatiques en transactionnel (OLTP) et analytique (OLAP) : en général, on peut considérer que les systèmes OLTP sont les sources de données pour un système BI, alors que les systèmes OLAP aident à les analyser.

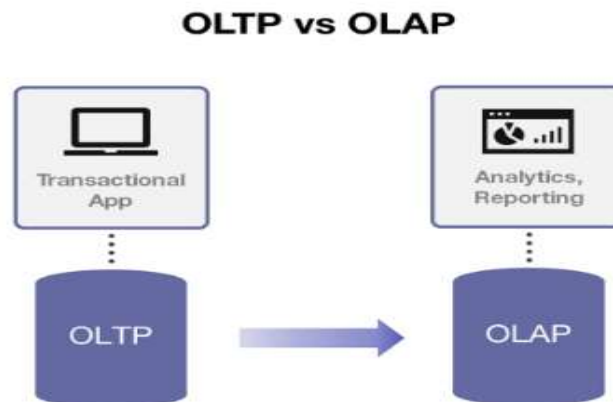


Figure 1-2 différence entre système OALP et OLTP [3]

Le tableau suivant, rappelle les différences principales existantes entre les deux types de systèmes :

Différence	Systèmes transactionnels	Systèmes décisionnels
Données	Orienté applications	Orienté thèmes et sujets
	Situation instantanée	Situation historique
Requêtes	Relativement standardisées	Souvent des requêtes complexes
	Requêtes simples	
Design de la base de données	Temps d'exécution très rapide	Dépend de la quantité des données
	Très normalisé avec beaucoup de tables	Souvent dé-normalisé avec moins de tables
Usage	Ex : « Qui as acheté x ? »	Ex : « Combien de personnes ont acheté x ? »

Table 1-1 Tableau comparatif entre les systèmes transactionnels et les systèmes décisionnels [3]

Ces nouvelles exigences des entreprises ont motivé la nécessité de mettre en place un système répondant aux besoins décisionnels. Ce système n'est rien d'autre que le « **Data Warehouse** ».

IV Data Warehouse

IV.1 Qu'est Ce qu'un Data Warehouse

Un Data Warehouse ou un entrepôt de données est une vision centralisée et universelle de toutes les informations de l'entreprise. C'est une structure (comme une base de données) qui a pour but, contrairement aux bases de données, de regrouper les données de l'entreprise pour des fins analytiques et pour aider à la décision stratégique.

Tous les types de Data Warehouse partagent les cinq caractéristiques suivantes :

- Données organisées ;
- Données cohérentes ;
- Données évolutives dans le temps (les données sont conservées sur une période plus longue, cela permet la comparaison et le suivi de l'évolution des valeurs dans le temps) ;
- Non volatiles (les données ne peuvent pas être mises à jours, les données sont historisées) ;
- Structure relationnelle. [4]

IV.2 Historique des Data Warehouse

L'origine du concept « *Data Warehouse* » *DW* (entrepôt de données en français) remonte aux années 80, durant lesquelles un intérêt croissant au système décisionnel a vu le jour, dû essentiellement à l'émergence des SGBD relationnels avec leur simplicité et la puissance offerte par le langage SQL.

Au début, un Data Warehouse n'était rien d'autre qu'une copie des données d'un système opérationnel prise de façon périodique, dédiée à un environnement de support à la prise de décision. Ainsi, les données étaient extraites du système opérationnel, stockées dans une nouvelle base de données « concept d'infocentre ». Le motif principal étant de répondre aux requêtes des décideurs sans affecter les performances des systèmes opérationnels.

Le Data Warehouse, tel qu'on le connaît actuellement, n'est plus vu comme une copie ou un cumul de copies prises de façon périodique des données du système opérationnel mais une nouvelle source d'information, alimentée avec des données recueillies et consolidées des différentes sources internes et externes. [5]

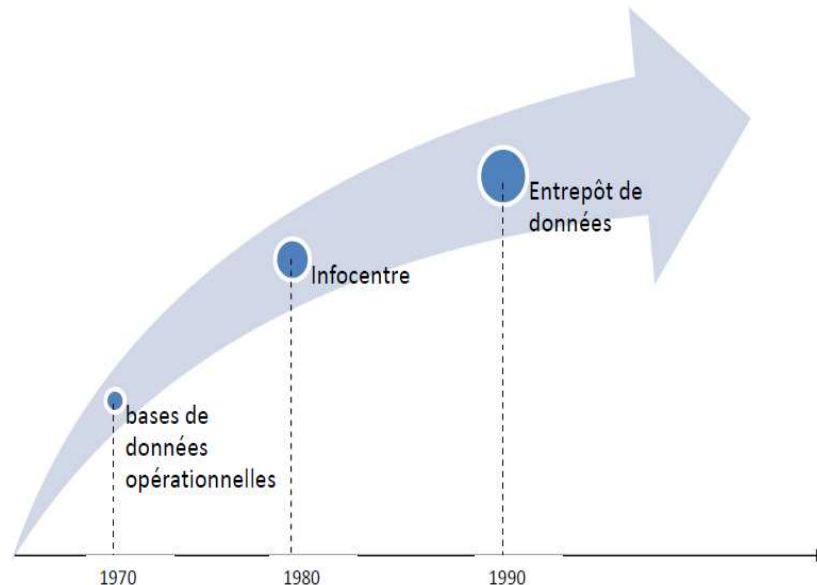


Figure 1-3 Evolution des bases de données décisionnelles [6]

IV.3 Modélisation de données d'un DW

La modélisation dimensionnelle d'un DW s'appuie sur les deux concepts suivants :

IV.3.1 Concepts

Concept de fait

Une table de fait est la table centrale d'un modèle dimensionnel, contenant des informations opérationnelles vitales pour l'entreprise, comme par exemple la table de faits des ventes (chiffre d'affaire net, quantités et montants commandés, facturés ou retournés, volumes des ventes, etc.) ou peut être celle des ressources humaines (performances des employés, nombre de demandes de congés, nombre de démissions, taux d'absence des employés, etc.). Ces tables contiennent généralement des clés étrangères qui sont des clés primaires des tables de dimension, et des mesures qui donnent les valeurs numériques du fait. [7]

Concept de dimension

Les tables de dimension sont les tables qui accompagnent 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 peut être compréhensible et utilisable. [7]

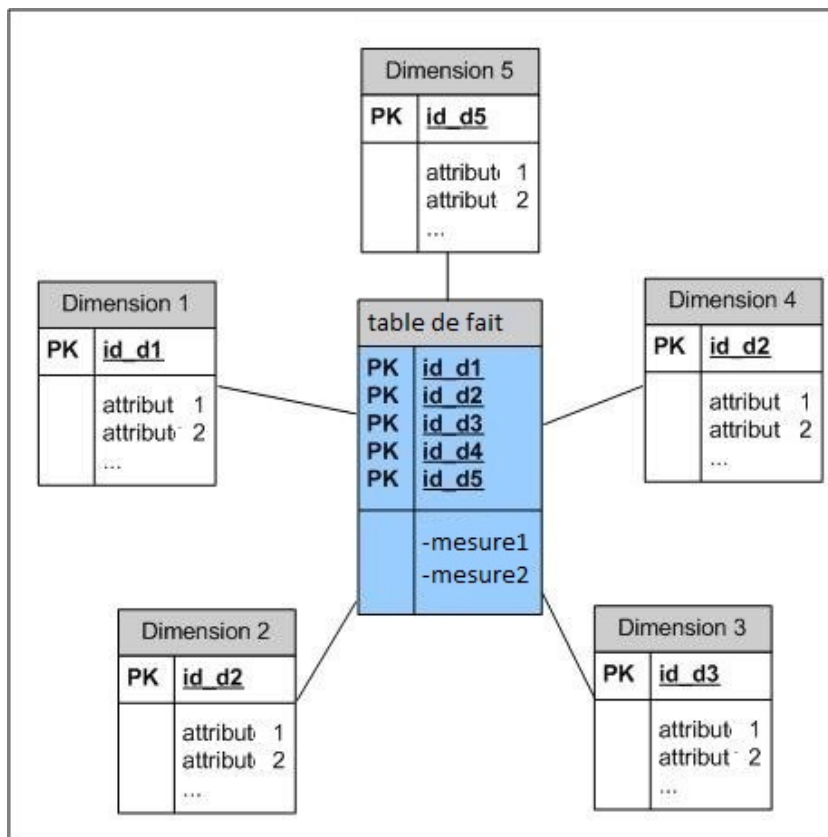


Figure 1-4 Architecture en étoile qui illustre le concept de fait et le concept de dimension [8]

Le tableau suivant résume les différences entre les tables de faits et les tables de dimensions au niveau des données :

	Tables de faits	Tables de dimensions
Structure	Peu de colonnes beaucoup de lignes	Peut 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 a des calculs	Participe à des contraintes
Signification	Valeur de mesures	Descriptive
Rôle	Assure les relations entre les dimensions	Assure l'interface homme/entrepôt de données

Table 1-2Tableau comparatif entre les tables de faits et les tables de dimensions [7]

IV.3.2 Différents modèles de la modélisation dimensionnelle

Principalement, les trois modèles, couramment employés pour la modélisation dimensionnelle sont :

Modèle en étoile

Ce modèle se présente comme une étoile dont le centre n'est autre que la table des faits et les branches sont les tables de dimension. La force de ce type de modélisation est sa lisibilité et sa performance. [9]

Modèle en flocon

Identique au modèle en étoile, sauf que ses branches sont éclatées en hiérarchies. Cette modélisation est généralement justifiée par l'économie d'espace de stockage, cependant elle peut s'avérer moins compréhensible pour l'utilisateur final et très coûteuse en termes de performances. [9]

Modèle en constellation

Ce n'est rien d'autre que plusieurs modèles en étoiles liés entre eux par des dimensions en commun. [9]

V Concept OLAP

C'est une catégorie de logiciels axés sur l'exploration et l'analyse rapide des données selon une approche multidimensionnelle à plusieurs niveaux d'agrégation.

Le noyau d'un système OLAP est son serveur. Ces serveurs sont classés en trois architectures systèmes qui peuvent être distinguées comme suit :

- Les systèmes à architecture **MOLAP** « **M**ultidimensional **O**n-line **A**nalytical **P**rocessing »
- Les systèmes à architecture **ROLAP** « **R**elationnel **O**n-line **A**nalytical **P**rocessing »
- Les systèmes à architecture **HOLAP** « **H**ybride **O**n-line **A**nalytical **P**rocessing »

	ROLAP	MOLAP	HOLAP
Stockage des données de base	BD relationnelle	BD multidimensionnelle (hyper cube)	BD relationnelle
Structure de la BD	Modèle particulier (étoile, flocon, etc)	Similaire à ROLAP la différence est au niveau d'entrepasage des données	Croisement des architectures ROLAP et MOLAP
Fonctionnement	Le serveur extrait les données par des requête SQL et interprète les données selon une vue multidimensionnelle avant de les présenter au module client.	Le serveur MOLAP extrait les données de l'hyper cube et les présente directement au module client.	Accède aux deux BD et les présente au module client selon leur méthode respective
Performance des requêtes	Le moins performant	Le plus performant	Performance moyenne

Table 1-3Tableau comparatif des differentes architectures OLAP [10]

Le langage MDX (Multidimensional Expressions)

Le MDX est un langage de requête sur les bases de données multidimensionnelles, il permet d'explorer les cubes OLAP, interroger et manipuler les données qui y sont contenues. [11]

VI Architecture et composants d'un système BI

Après avoir défini la BI, on va découvrir ses composants et expliquer comment ça fonctionne les uns avec les autres. Le schéma suivant illustre l'architecture et les composants principaux d'un système BI :

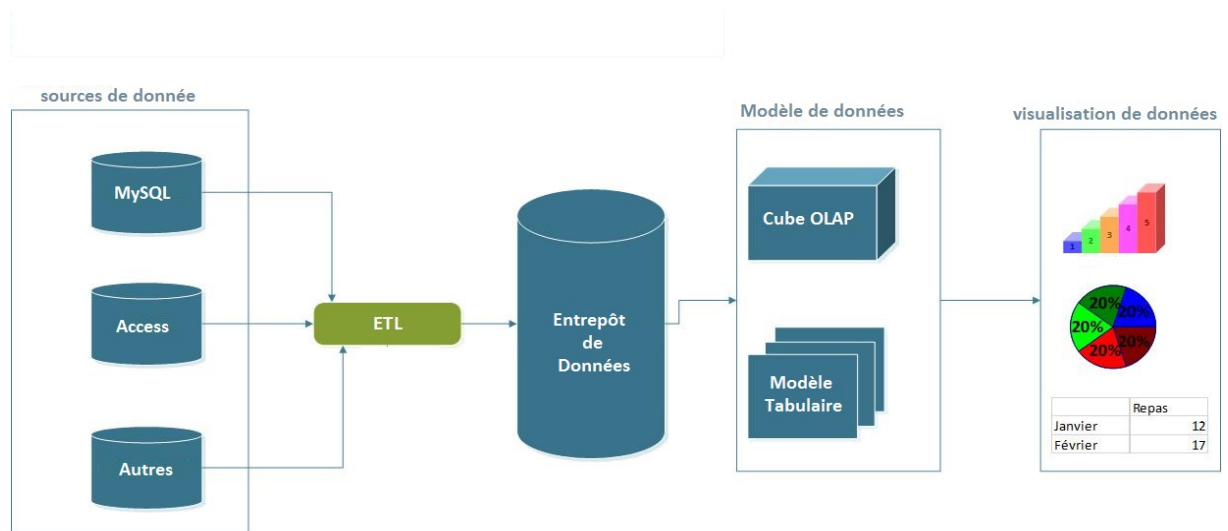


Figure 1-5 vue d'ensemble de l'architecture d'un système BI [12]

L'architecture présentée dans le schéma ci-dessus contient les composants les plus courants dans la plupart des systèmes BI, puisque l'architecture avec ses composants peuvent changés en fonction des outils, de l'environnement...etc. Chacun des composants présents sur le schéma sera abordé avec plus de détails dans les sections qui suivent.

Sources de données

Les sources de données sont souvent diverses et variées ex : les bases de données relationnelles (MySQL, Oracle, SQL Server), les fichiers plats Excel, XML, les services web... etc. On ne peut pas construire un système BI, sans source de données.

ETL

Dans la majorité des cas, les sources de données pour un système BI sont nombreuses. Ainsi, on va procéder à une opération de consolidation qui consiste à extraire les données de leurs sources, de les transformer et ensuite de les déverser dans l'entrepôt de données tout en respectant les spécifications du Data Warehouse. Cette opération est appelée **Extract Transform Load (ETL)**.

Entrepôt de données

Considéré comme étant le cœur du système vu son importance, il sert à conserver toutes les données de l'entreprise provenant de ses différents services, dans le but de faciliter l'analyse des données et le Reporting.

Modèle de données

Un Entrepôt de données est conçu pour être la source d'analyse et de rapports, de sorte qu'il soit plus rapide que les systèmes opérationnels pour la production de rapports. Cependant, il n'est pas aussi rapide qu'il prétend, pour pouvoir couvrir tous les besoins, puisque lui-même est une base de données relationnelle qui est définie avec un ensemble de contraintes qui réduisent le temps de réponse d'une requête. L'exigence d'un traitement plus rapide en ayant un temps de réponse plus faible d'une part, et avoir des informations agrégées d'autre part, entraîne donc la création d'une autre couche dans les systèmes de BI. Cette couche, appelé **Modèle de données**, contient un modèle de données basé sur des fichiers ou basé sur la mémoire vive, afin de produire rapidement les réponses aux requêtes pour l'édition des rapports.

Le **Modèle de données** s'appuie sur deux technologies: **Cube OLAP** et le modèle en mémoire de tableau (**Modèle tabulaire**).

L'**Cube OLAP** est un stockage de données basé sur un fichier qui charge les données à partir d'un entrepôt de données dans un modèle de cube. Le cube contient des informations descriptives telles que les dimensions (par exemple, clients et produits) et des cellules (par exemple, des faits et des mesures, telles que les ventes et les remises).

Le **Modèle tabulaire** est basé sur un nouveau moteur en mémoire pour les tables. Le moteur en mémoire charge toutes les données rangées de tables dans la mémoire et répond à des questions directement à partir de la mémoire. C'est très rapide en termes de temps de réponse.

Visualisation de données

L'interface d'un système BI est la visualisation de données. En d'autres termes, la visualisation de données est la partie du système de BI que les utilisateurs peuvent voir. Il existe différentes méthodes pour la visualisation d'informations, tels que des tableaux de bord stratégiques et tactiques, indicateurs clés de performance ou **Key performance Indicators (KPI)**, des rapports consolidés.

C'est donc la partie qui intéresse le plus les décideurs d'une entreprise. De ce fait, il existe un grand nombre d'outils de visualisation de rapport sur le marché, telle que **SharePoint, PowerBI, IDashBoards**, etc. [12]

VII Solutions BI présentes sur le marché

Beaucoup de fournisseurs de solutions BI tentent de convaincre les clients des performances de leurs applications. Néanmoins, ces solutions BI peuvent ne pas convenir aux différents environnements des entreprises. Dans la liste ci-dessous vous pouvez voir les grands fournisseurs BI sur le marché. Ces produits sont principalement destinés aux entreprises et compagnies comprenant un paquet complet d'applications, des bases de données, jusqu'aux outils d'intégration et de visualisation.

- SAP Business Objects
- IBM Cognost
- Pantaho BI
- SpagoBI
- JasperSoft
- SAS BI
- Oracle BI
- Microsoft BI

A cet égard, nous avons choisi dans le cadre de notre travail, de mener une étude comparative entre deux suites propriétaires à savoir Microsoft et Oracle et une autre suite Open Source qui est Pentaho.

Microsoft BI Microsoft

Cette suite développée par Microsoft, intègre l'ensemble des briques nécessaires à la construction d'un système d'information décisionnel. En effet, cette plateforme s'accompagne à de nombreux outils pour intégrer, analyser et afficher les données, tout en garantissant une haute performance. La plateforme repose sur les outils suivants : SQL Server Integration services (SSIS), SQL Server Analysis Services (SSAS) et SQL Server Reporting Services (SSRS).

Oracle BI

Concurrente de Microsoft, la solution Business Intelligence d'Oracle est une plateforme complète de système décisionnel qui fournit toute la gamme de fonctionnalités d'analyse et de production de rapports. La plateforme contient une base de données, un outil d'intégration de données (ETL), des raquetteurs, et un portail permettant de publier des tableaux de bord.

Pentaho

Cette suite est considérée comme le leader des suites de BI dans le monde de l'Open Source. C'est une plateforme décisionnelle extrêmement complète et auto-suffisante, qui

permet non seulement d'utiliser les différents outils décisionnels open source depuis une interface unique simple à utiliser, mais aussi d'étendre et de combiner leurs fonctionnalités grâce à l'utilisation d'un moteur de Workflow. Elle repose sur JFreeReport pour les rapports, Weka pour le Datamining, PDI pour l'ETL, et Mondrian pour le serveur OLAP.

Le tableau suivant oppose les fonctionnalités des trois suites décisionnelles :

Fonctionnalités	Microsoft	Oracle	Pentaho
Alimentation	-Chargement des données externes par fichier plat -Chargement de toute source de données	-Chargement des données externes par fichier plat -Chargement de toute source de données	-Connexion à n'importe quelle base de données à travers d'un driver JDBC.
Stockage	-Cube OLAP -Entrepôt de données -Datamarts métiers	-Cube OLAP -Entrepôt de données -Datamarts métiers	-SGBD avec l'option OLAP
Analyse	-Statistiques -Analyse multidimensionnelle -Extraction des données à un format de type tableur	-Statistiques -Analyse multidimensionnelle -Extraction des données à un format de type tableur	-permet la manipulation des données selon plusieurs axes d'analyse
Reporting	-Tableaux de bord personnalisables et interactifs -Mise à disposition d'une interface graphique de construction de tableaux de bord personnalisés et de restitutions dans un portail décisionnel	-Tableaux de bord standards et interactifs -Mise à disposition d'une interface graphique de construction de tableaux de bord personnalisés	-Support de sources de données multiples dans les mêmes rapports - Support de sous rapports

Table 1-4 Tableau comparatif des suites BI [13]

Le carré magique de GARTNER ci-dessous offre une vue claire du positionnement des divers solutions BI présent sur le marché, pour l'année 2017. Ce quadrant repose sur deux axes d'analyse, qui sont : la capacité d'exécution et la complétude de la vision. On distingue bien que La suite Microsoft prend l'ascendant pour l'année 2017 sur les différentes autres suites telles qu'Oracle et Pentaho, vu qu'elle est moins chère par rapport a Oracle, facile à intégrer avec SharePoint et Microsoft Office et elle offre un niveau de sécurité plus élevé.



Figure 1-6 BI quadrant magique de Gartner 2017

Suite à cette comparaison nous pouvons conclure que certes, les outils Open Source offrent un ensemble de fonctionnalités assez performantes mais restent dépassés par les suites propriétaires telles que Microsoft, Cette dernières proposent plusieurs composants qui répondent aux besoins des entreprises et couvrent l'ensemble de la chaîne décisionnelle avec un niveau de performance assez élevé.

VIII Conclusion

Dans ce chapitre, nous avons commencé par une présentation des concepts de base des systèmes décisionnels(BI) et leurs divergences avec les systèmes transactionnels. Par la suite, on a détaillé l'architecture globale d'un cycle BI et en particulier ses différents composants. Pour en conclure, on s'est intéressé aux différentes solutions BI présentes dans le marché en présentant une étude comparative des trois suites les plus répandues actuellement, à savoir Microsoft, Oracle ainsi que Pentaho.

Chapitre2 : Analyse et Conception

Chapitre 2 Analyse et conception

I. Introduction

Ce chapitre est consacré en premier lieu à une analyse de la situation actuelle en étudiant les systèmes qui sont actuellement en exploitation. Les différentes difficultés rencontrées ainsi les besoins exprimés par les différents utilisateurs ont été recensés.

II. Planification et conduite du projet

Le développement de tout projet nécessite une phase de planification qui permet de définir les tâches à réaliser, maîtriser les risques et rendre compte de l'état d'avancement du projet. Dans le but d'une meilleure organisation du projet et gestion efficace des délais disponibles, nous avons présenté sous forme d'un tableau récapitulatif l'ensemble des tâches à réaliser avec leurs durées de réalisation, puis nous avons établi un diagramme de GANTT qui illustre cet ordonnancement.

Tâches	Date début	Date fin	Durée
Identification	01/02/2017	11/02/2017	11j
Etude de l'environnement	01/02/2017	11/02/2017	11j
Formation	12/02/2017	15/04/2017	25j
Formation SQLServer	12/02/2017	16/02/2017	5j
Formation SSIS	17/02/2017	21/02/2017	5j
Formation SSAS	22/02/2017	26/02/2017	5j
Formation SSRS	06/04/2017	10/04/2017	5j
Formation Sharepoint	11/04/2017	15/04/2017	5j
Conception	27/02/2017	15/03/2017	17j
Analyse de l'existant et des besoins	27/02/2017	05/03/2017	7j
Modélisation multidimensionnelle	06/03/2017	15/03/2017	10j
Mise en œuvre	16/03/2017	22/05/2017	58j
Processus ETL	16/03/2017	31/03/2017	16j
Cube OLAP	1/04/2017	05/04/2017	5j
Reporting	16/04/2017	25/04/2017	10j
Déploiement des rapports sur la plateforme Web	26/04/2017	30/04/2017	5j
Développements des modules Data Cleansing (DW)	1/05/2017	07/05/2017	7j
Gestion des droits d'accès	08/05/2017	22/05/2017	15j
Amélioration	01/04/2017	14/06/2017	73j
Rédaction du rapport	01/4/2017	29/05/2017	59j
Test complet du système	23/05/2017	01/06/2017	10j

Présentation de la solution aux utilisateurs finals	10/6/2017	10/06/2017	1j
Formation de l'administrateur de l'application	12/06/2017	14/06/2017	3j

Table 2-1 Tableau récapitulatif de l'ensemble des tâches du projet

Le diagramme ci-dessous décrit la planification ainsi que le l'ordonnancement prévu des différentes phases du projet :

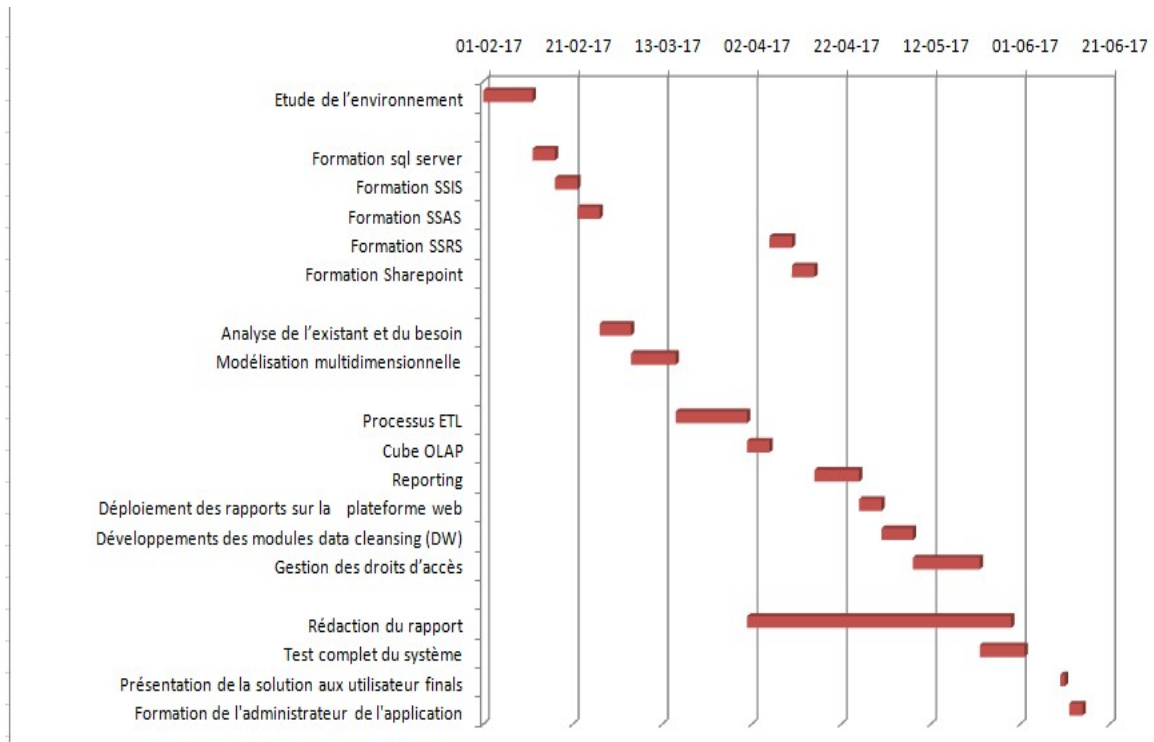


Figure 2-1 Diagramme de Gantt de l'avancement du projet

III. Analyse de l'existant

Cette phase consiste à identifier l'environnement logiciel dans lequel notre étude va être menée et analyser les structures des données des logiciels de gestion de scolarité de deux facultés à savoir la faculté des sciences et la faculté de technologie au sein de l'université de Tlemcen.

III.1. Identification de l'environnement

En analysant la situation actuelle, on a pu distinguer trois systèmes sources qui sont :

Logiciel S.E.E.S : Logiciel fourni par le ministère de l'enseignement supérieur pour la gestion des inscriptions, suivi des notes, etc., des étudiants durant leurs cursus académique. Ce logiciel fonctionne sur une base de données Microsoft Access.

Logiciel des emplois du temps : Logiciel pour la planification et la distribution des heures et des salles pour les étudiants et les enseignants au sein de la faculté. Ce logiciel est basé sur une base de données MySQL.

Logiciel d'affectation des choix : Logiciel déployé sur une plateforme Web qui permet aux étudiants de faire leurs choix de spécialité. Ce logiciel s'occupe par la suite du traitement des choix selon des contraintes prédéfinies et d'affecter les étudiants aux différentes spécialités. Ce logiciel s'appuie sur une base de données MySQL.

III.2. Retro-Ingénierie

La rétro-ingénierie ou reverse engineering, également appelée rétro-conception, est l'activité qui consiste à étudier un objet pour en déterminer son fonctionnement interne ou sa méthode de fabrication [15].

Pour réaliser notre rétro-ingénierie, il a fallu extraire les relations existantes entre les tables des bases de données, pour cela nous avons utilisé les outils : MySQL Workbench pour les bases de données MySQL et Microsoft Management Studio pour la base de données Access.

Le processus de rétro-ingénierie a généré les diagrammes suivants :

Le modèle ci-dessous a été inféré de la rétro-ingénierie de la base de données Microsoft Access du logiciel S.E.E.S. Le modèle faisant objet de notre étude, se constitue de 38 différentes tables :

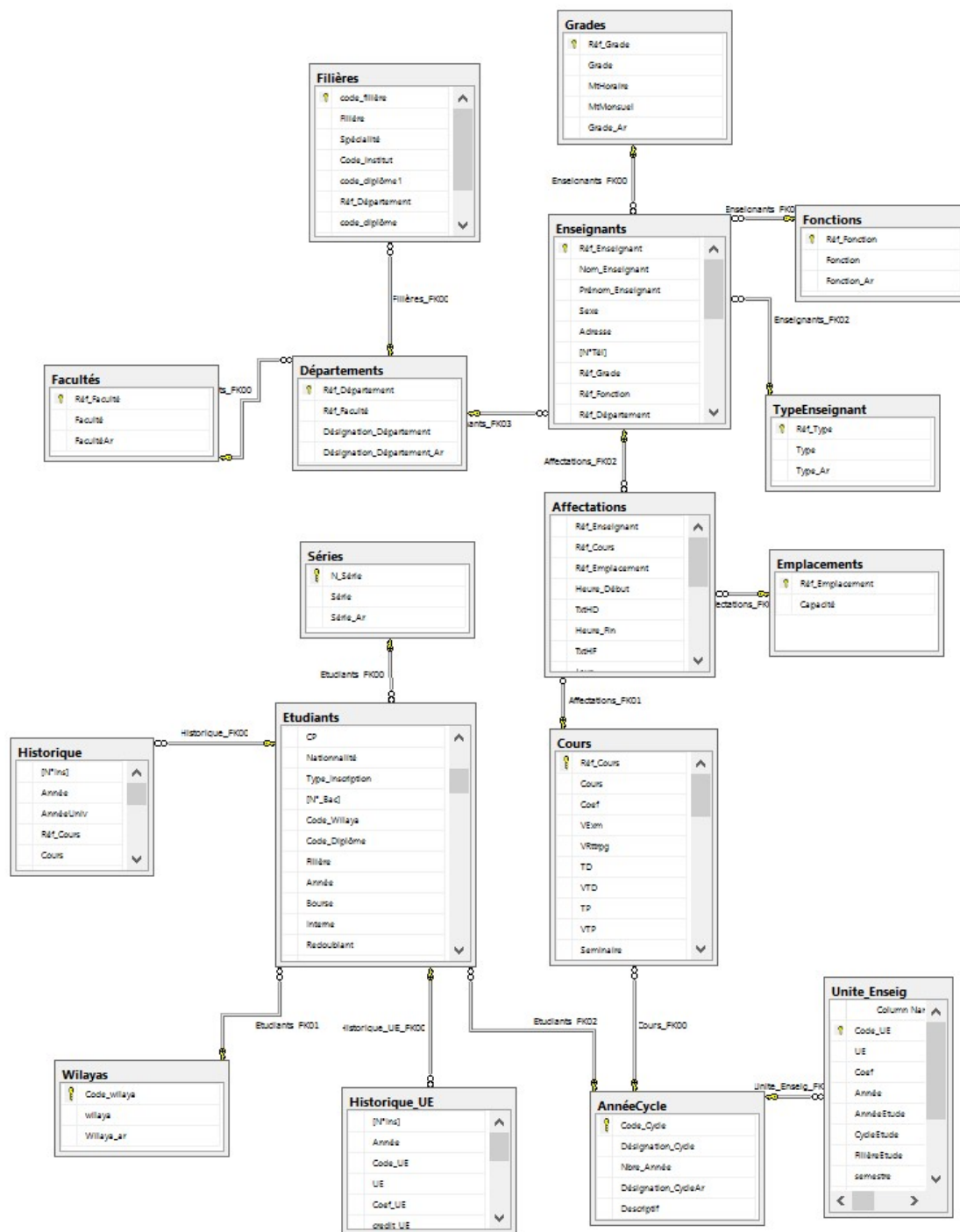


Figure 2-2Modèle relationnel de données (Base de données Access) partie 1

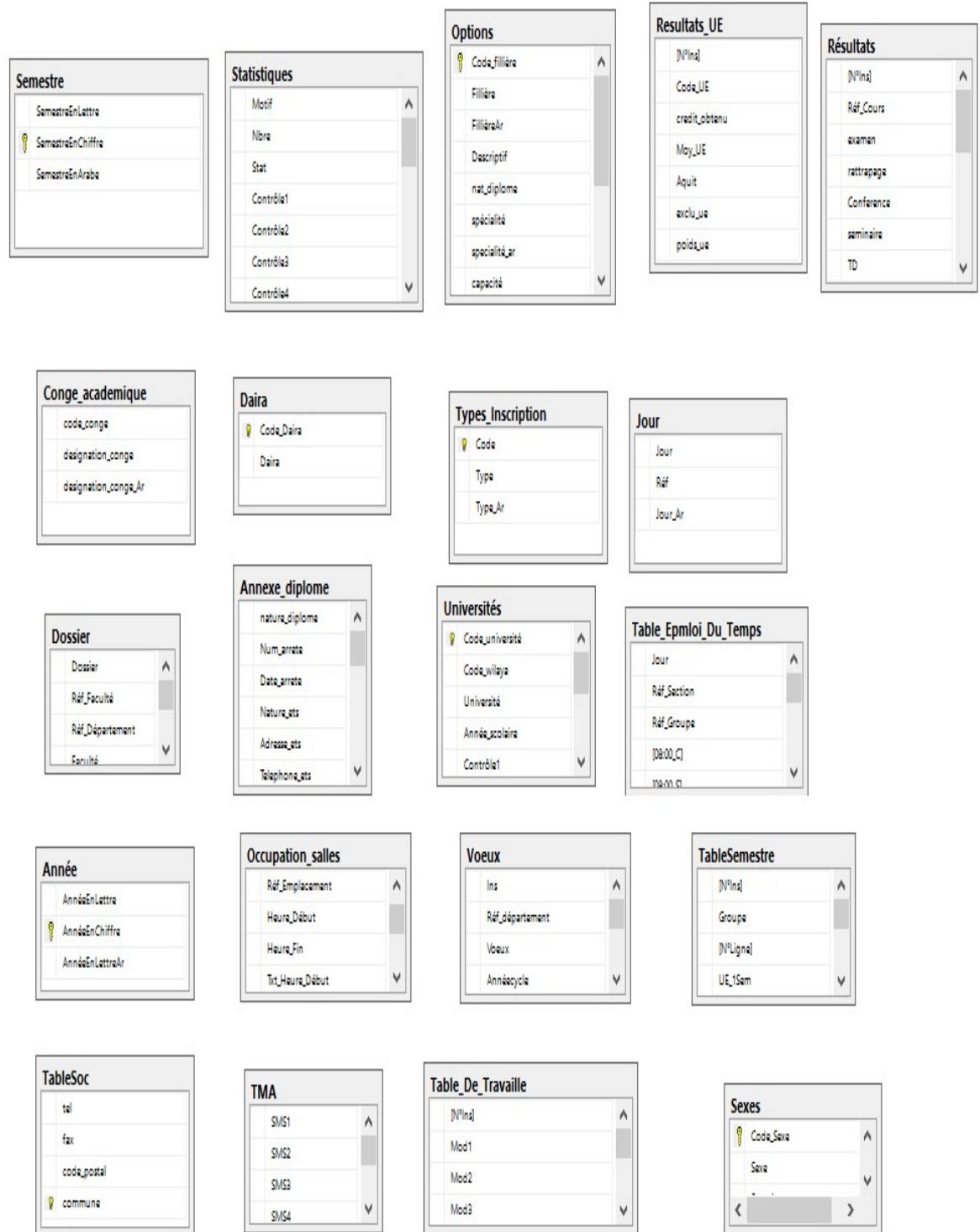


Figure 2-3Modèle relationnel de données (Base de données Access) partie 2

Le modèle ci-dessous est le résultat du processus de rétro-ingénierie de la base de données MySQL utilisée dans le logiciel des emplois du temps. Ce modèle se compose de 7 tables différentes, qui sont :

- Table EDT : Contient les différentes informations sur un emploi du temps semestriel (semestre, jour, salle, séance, option, groupe,etc)
- Table EDT_examen : Contient les différentes informations sur un emploi du temps d'examen (semestre, jour, salle, séance, option, groupe,etc)
- Table formations : Comprend les différentes informations sur les spécialités disponibles dans une faculté.
- Table filière : Contient les différentes informations des filières des départements.
- Table départements : Stocke les différentes informations d'un département de la faculté.
- Table modules : Détient les différentes informations concernant les modules des différentes spécialités.
- Table enseignants : Contient les différentes informations sur les enseignants de la faculté.

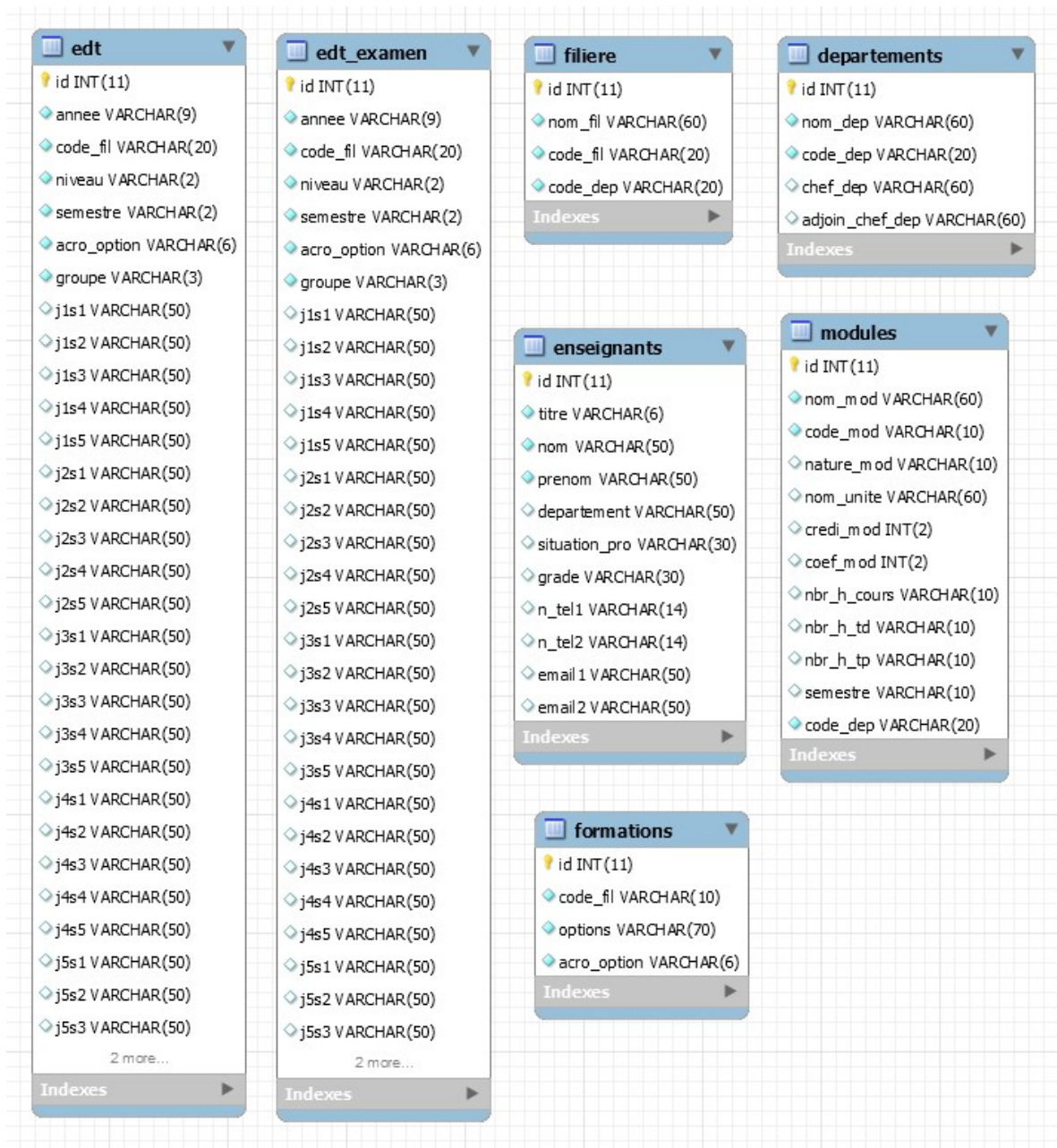


Figure 2-4Modèle relationnel de données (Base de données MySQL) : Logiciel des emplois du temps

Chapitre2: Analyse et conception

La rétro-ingénierie de la base MySQL employée pour le logiciel d'affectation des choix aux étudiants, a généré le modèle ci-dessous constitué des 5 tables différentes suivantes :

- Table étudiant : Contient les différentes informations d'un étudiant (nom, prénom, etc..), ses différents choix ainsi que le choix qui lui été affecté.
- Table spécialité : Détient les différentes informations d'une spécialité (intitulé, nombre des étudiants par spécialité, etc..)Ainsi que son niveau.
- Table niveau : Contient les différents niveaux possibles d'une filière.
- Table filière : Comprend la liste des filières d'une faculté avec leurs informations respectives.
- Table faculté : Comprend la liste des facultés avec leurs informations respectives.

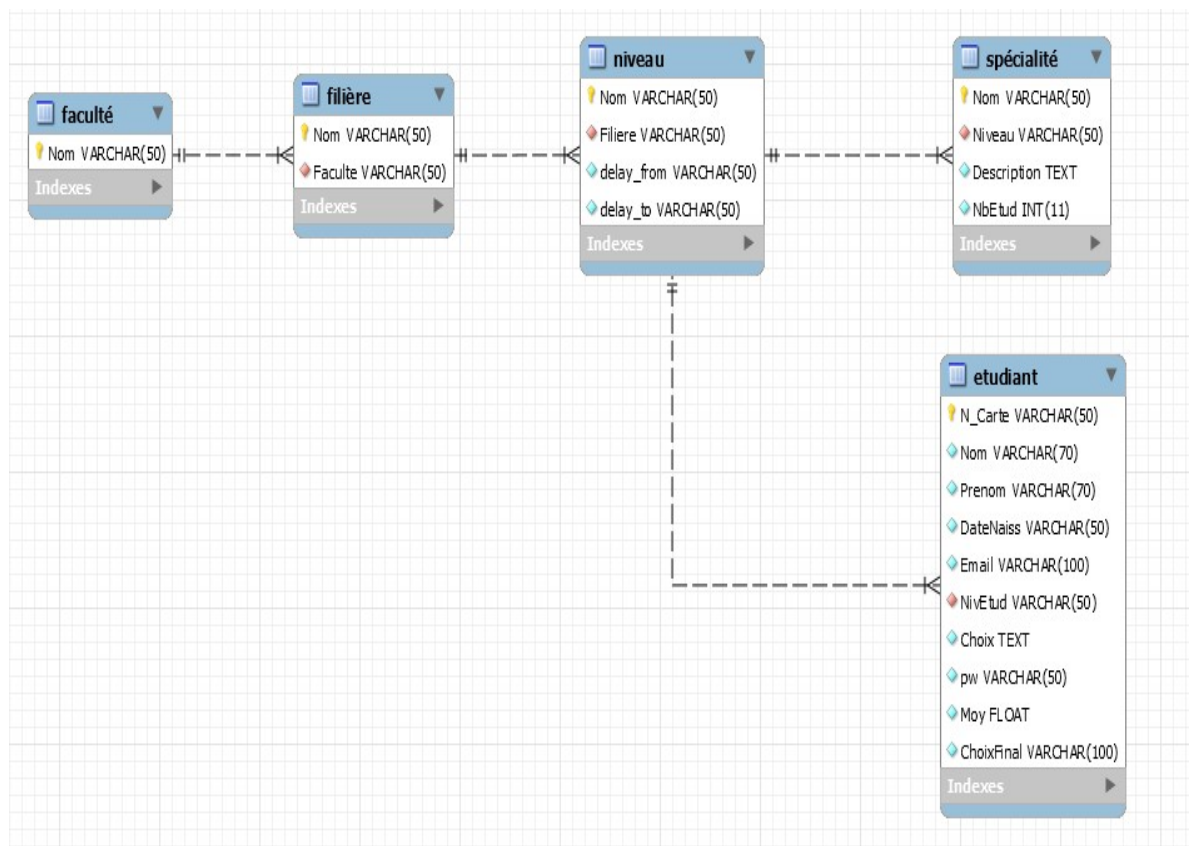


Figure 2-5Modèle relationnel de données (Base de données MySQL) : Logiciel d'affectation des choix

III.3. Problèmes et obstacles rencontrés

Au cours de l'étude menée sur les 3 modèles employés actuellement, on s'est opposé à une multitude de problèmes et difficultés majeures qui nous ont retardés dans notre démarche et particulièrement dans le modèle Microsoft Access.

La liste suivante contient une liste non exhaustive des obstacles et des problèmes les plus importants :

- Bases de données non structurées ;
- Tables non utilisées ;
- Tables non référencées (pas de relations entre les tables) ;
- Champs de tables non renseignés (vide) ou partiellement renseignés ;
- Champs renseignés dans quelques bases de données et non renseignés dans d'autres ;
- Champs remplis de manière erronée ;
- Champs référencés dans les tables fils ne respectant pas les valeurs des tables pères. (les entrées sont saisies on hardcode).
- Manque de données pour faire exécuter les requêtes.

- Champs redondants (Même champs retrouvé plusieurs fois dans la base de données)

- Plusieurs informations intégrés dans le même champ séparées par des #

IV. Etude des besoins

La conception et la réalisation d'un Data Warehouse doit être en mesure de répondre aux attentes des utilisateurs, qui ne peut, évidemment, se faire sans une étude approfondie de leurs besoins exprimés.

IV.1. Rapports de gestion

L'étude faite au sein des services de scolarité des deux facultés des sciences et de technologie, a permis de recenser les différents rapports dont les décideurs ont besoin, les plus importants d'entre eux, sont classés dans le tableau suivant :

Nom du rapport	Catégorie	Description
Nombre des étudiants qui ont abandonnés	Histogramme	Cet histogramme représente le nombre des étudiants qui ont abandonné leur filière d'étude sur l'axe y et les différents filières sur l'axe x
Nombre des étudiants diplômés par département	Histogramme	Cet histogramme montre le nombre des étudiants qui ont obtenu leurs diplôme, soit dans la série licence ou master : l'axe y représente le nombre des étudiants, l'axe x représente les différents départements et la série dont l'étudiant est diplômé (licence, master).
Nombre des enseignants par département	Histogramme	Cet histogramme expose le nombre des enseignants sur l'axe y et les différents départements sur l'axe x.
Nombre des enseignants par grade	Histogramme	Dans cet histogramme, on peut distinguer le nombre des enseignants selon leurs grades dans un département, l'axe y représente le nombre des enseignants et l'axe x représente les différents grades.
Moyenne générale des modules	Histogramme	Cet histogramme affiche la moyenne générale d'une promo dans un module, l'axe y affiche la moyenne générale, l'axe x affiche les différents modules d'une promotion.
Moyenne générale par option	Histogramme	Cet histogramme montre la moyenne générale de l'ensemble des étudiants d'une même promotion, sur l'axe y, on a la moyenne générale de la promotion, sur l'axe x les différentes spécialités.
Nombre des étudiants par département	Histogramme	Cet histogramme permet aux décideurs de visualiser les effectifs des étudiants par département, sur l'axe y on trouve le nombre des étudiants, sur l'axe x les différents départements.
Nombre des étudiants par observation	Histogramme	Cet histogramme montre le nombre d'étudiants selon leur observation d'admission ou non admission, durant l'année universitaire (admis, admis session 2, admis dettes, ajourné).
Nombre des étudiants par option	Histogramme	Cet histogramme permet aux décideurs de visualiser les effectifs des étudiants par spécialité, sur l'axe y on trouve le nombre des étudiants et sur l'axe x les différentes spécialités.
Effectif des étudiants par cycle	Camembert	Ce camembert affiche l'effectif des étudiants dans les départements ou bien dans la faculté selon leurs cycles d'étude (licence, master).
Evolution du nombre d'étudiants par faculté	Ligne	Ce diagramme en ligne montre l'évolution des effectifs des étudiants d'une faculté répartis sur les différents années, sur l'axe y on trouve le nombre des étudiants, sur l'axe x on y trouve les différentes années universitaires.
Nombre des étudiants par willaya	Carte	Ce rapport sous forme d'une carte géographique expose la population des étudiants de l'université selon leurs wilayas d'origine.
Nombre des étudiants par pays	Carte	Ce rapport sous forme d'une carte géographique expose la population des étudiants de l'université selon leurs pays d'origine.

Table 2-2 Tableau récapitulatif des différents rapports requis

IV.2. Utilisateurs du système

L'analyse des besoins exprimés nous a permis d'identifier huit (8) utilisateurs potentiels classés en cinq (5) catégories :

- **Administrateur (1 utilisateur) :** C'est l'ingénieur principal de la faculté, il a l'accès à tout le système, le droit de modifier tous les utilisateurs (login, mot de passe, ex...) ainsi que d'affecter les rapports aux utilisateurs, ...etc.
- **Recteur (1+1 utilisateurs) :** C'est monsieur «Mustapha Djaâfour» agissant en qualité de recteur de l'université responsable de toutes les activités. Il a accès à tous les sources de données pour avoir une vision globale sur le déroulement des activités scolaires.
 - Le vice-recteur responsable de la pédagogie peut accéder aussi avec les mêmes privilèges du recteur.
- **Doyen (1+1 utilisateurs) :** C'est le seul utilisateur dans la faculté qui a la possibilité d'accéder à toutes les informations concernant cette faculté, pour prendre des décisions à propos de ces départements.
 - Le vice-Doyen peut accéder aussi avec les mêmes privilèges du doyen.
- **Chef du département (1+1 utilisateurs) :** Ce dernier a la possibilité de prendre des décisions qu'au niveau de son département, donc il peut voir tout ce qui concerne son département avec ses différentes spécialités.
 - Son adjoint peut aussi accéder avec les mêmes privilèges.
- **Responsable de la spécialité (1 utilisateur) :** Chaque spécialité au niveau d'un département a son responsable. Ce dernier n'a la possibilité de voir que les informations concernant les étudiants de la spécialité dont il est responsable.

V. Conception de la solution

V.1. Diagramme des cas d'utilisation global

Dans ce diagramme nous présentons tous les cas d'utilisation global de notre système ainsi que les acteurs du système :

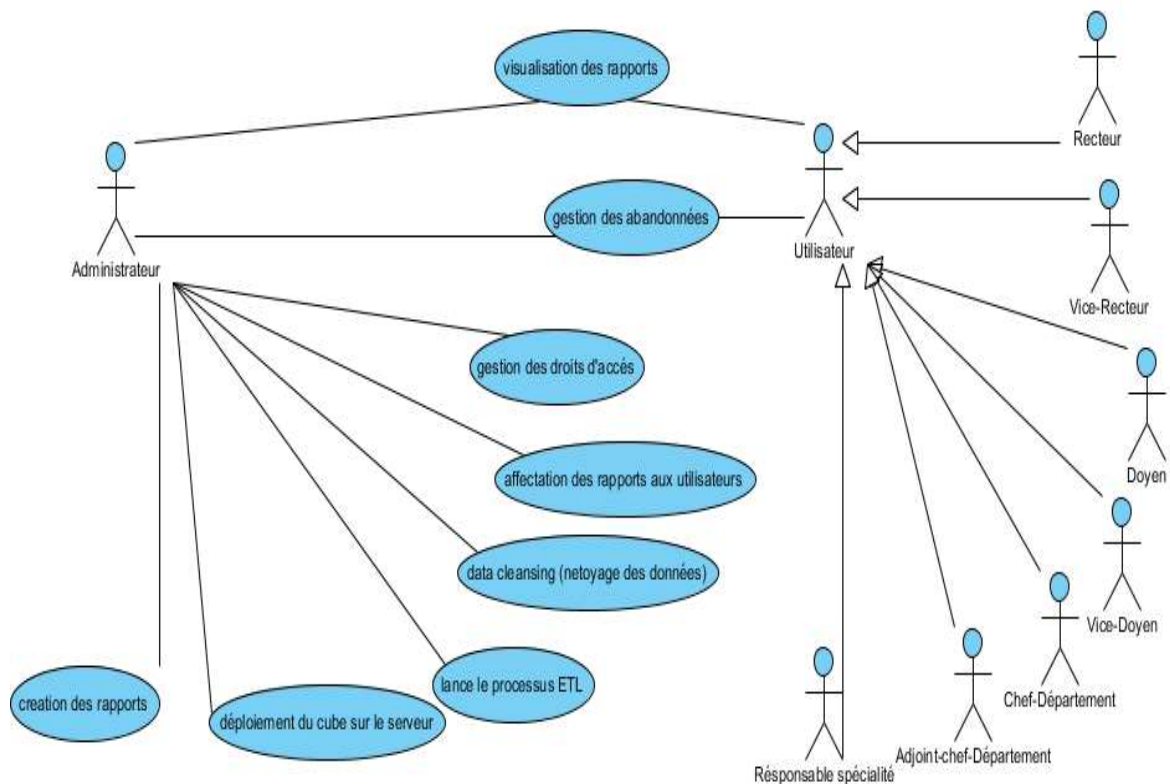


Figure 2-6 Diagramme des cas d'utilisations global

V.2. Conception multidimensionnelle

Dans cette phase, nous avons adopté une approche ascendante pour la construction de notre entrepôt de données (DW). Cette approche nommée « Bottom-Up », consiste à commencer par concevoir les datamarts et les regrouper par la suite, jusqu'à l'obtention du DataWarehouse (DW) [16].

Dans le cas étudié, on a trois modèles de datamarts :

- Le datamart qui concerne les étudiants et leurs inscriptions.
- Le datamart qui concerne les résultats des étudiants.
- Le datamart qui concerne la répartition des enseignants sur les séances.

Nous avons présenté dans le premier chapitre les différents schémas d'un DW (en étoile, en flocon, en constellation). Pour réaliser notre solution, nous avons opté pour le modèle flocon de neige jugé comme le modèle le plus adapté à notre application puisqu'il est basé sur le principe de création des hiérarchies de dimensions.

Datamart « inscription » : Ce datamart est constitué d'une table de fait et neuf tables de dimension, la table de fait (fact_enrollment) ne contient que des clés étrangères des différents tables de dimension ainsi que des mesures.

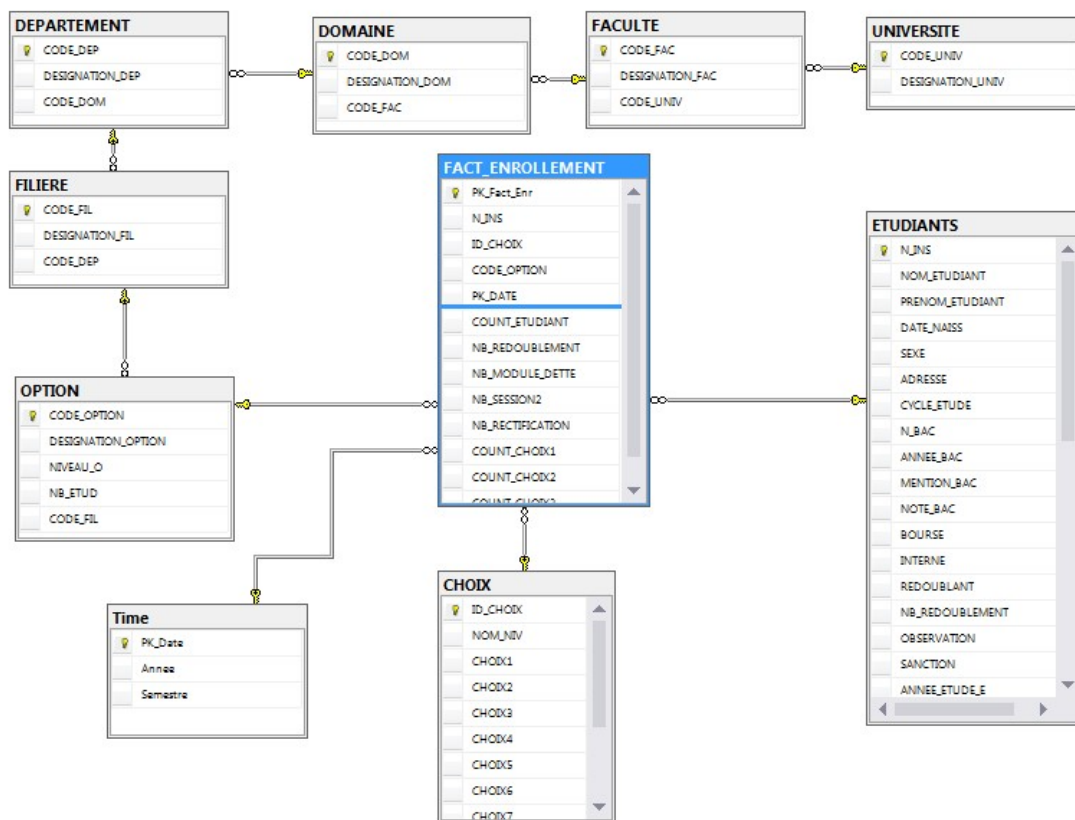


Figure 2-7 Datamart des inscriptions

Datamart « résultats »: ce dernier est constitué de la table de fait « fact_results » et 13 autres tables de dimension. La table de fait contient les clés primaires de toutes les tables de dimensions avec lesquelles elle est liée, aussi les différents mesures.

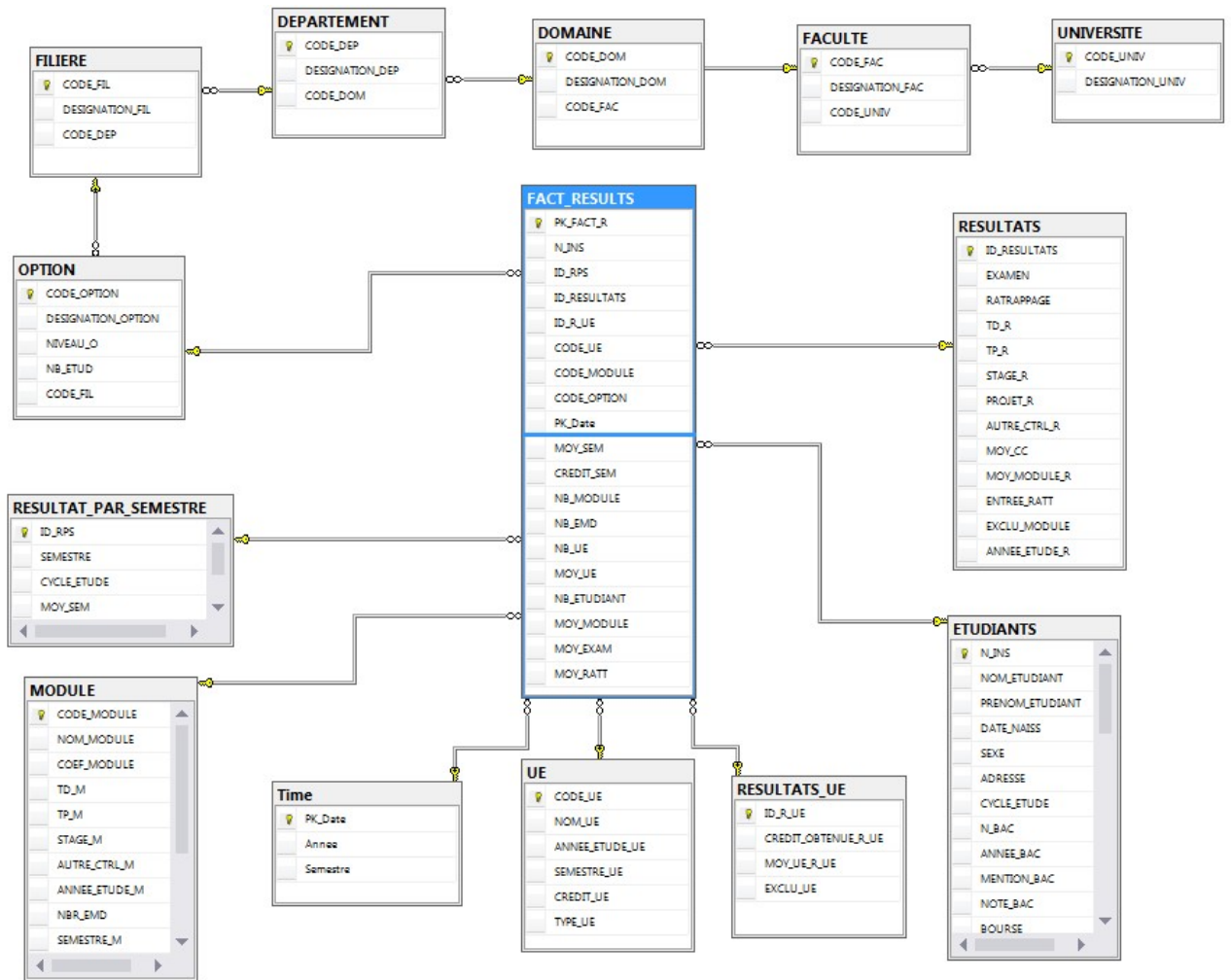


Figure 2-8 Datamart des résultats

Datamart « répartition_enseignants_séance » : Ce dernier est constitué de la table de fait « fact_repartition_ens_seance » qui représente l'axe d'analyse, et 11 autres tables qui jouent le rôle des dimensions.

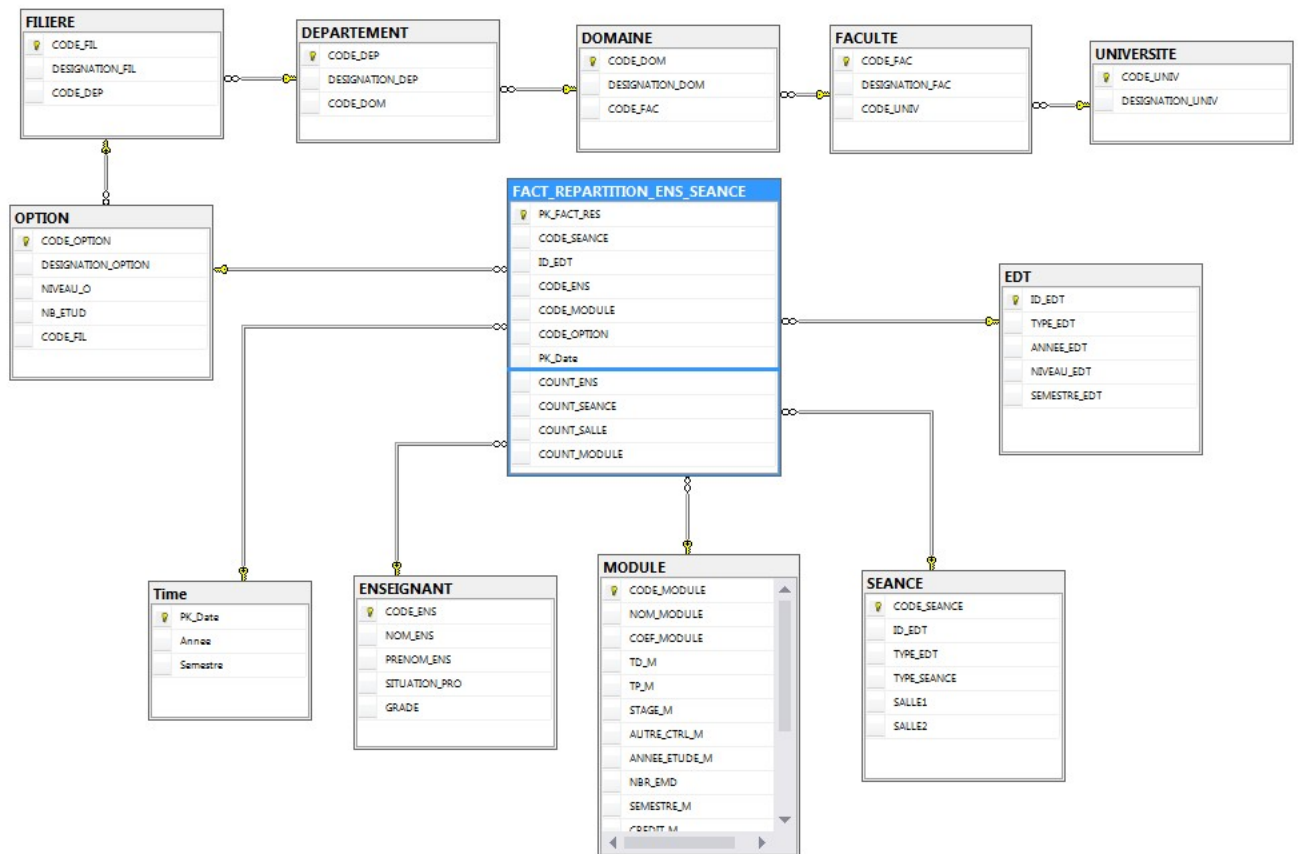


Figure 2-9 Datamart répartition_enseignants_séance

V.3. Mapping des données

L'opération de mapping consiste à faire la correspondance entre les données sources et les champs du Data Warehouse afin de charger correctement ce dernier.

Vu le volume et la dimension de la table de mapping, on ne va présenter qu'un seul exemple, le reste des tables sont reportés dans l'annexe1.

Le mapping de la table « Etudiants » est le suivant :

Table origine	Champs origine	Table destination	Champs destination	Type champ	Table référencé	Champs référencé	Type de la modification
Etudiant		Etudiant					
	N°Ins		n_ins	Int(11)			One to One
	Nom_Etudiant		nom_etudiant	Varchar(50)			One to one
	Prénom_Etudiant		Prenom_etudiant	Varchar(50)			One to one
	Nom_Arab		nom_arab	Varchar(50)			One to one
	Prenom_Arab		Prenom_arab	Varchar(50)			One to one
	Date_naissance		Date_naiss	Varchar(50)			One to one
	Lieu_naissance						Suppression
	Sexe		Sexe	Varchar(1)			One to one
	Adresse		Adresse	Varchar(100)			One to one
	CP		Cp	Varchar			One to one
	Nationalité		Nationalité	Varchar(20)			One to one
	Type_Inscription						Suppression
	N°Bac		N_bac	Varchar(20)			One to one
	Code_Wilaya		Code_wilaya	Varchar(20)			One to one
	Code_Diplôme						Suppression
	Filière						Suppression
	Année						Suppression
	Bourse		Bourse	Bit			One to one
	Interne		Interne	Bit			Ont to one
	Redoublant		Redoublant	Bit			One to one
	Nbre		Nb_redoublement	Int(11)			One to one
	Année_Bac		Annee_bac	Int(11)			One to one
	N_Série_Bac		N_serie_bac	Varchar(20)	Séries	N_Série	One to one
	Mention_Bac		Montion_bac	Varchar(20)			One to one
	Note_Bac		Note_bac	Double			One to one
	Wilaya_Bac		Wilaya_bac	Int(11)	Wilayas	Code_wilaya	One to one
	Réf_Group						Suppression
	Réf_Section						Suppression
	Observation		Observation	Varchar(20)			One to one
	Année_Cycle						Suppression

Chapitre2: Analyse et conception

Sanction	Sanction	Varchar(20)		One to one
AnnéeEtude	Annee_etude_	Int(11)		One to one
CycleEtude	Cycle_etude	Varchar(50)	annéeCycle Code_cyle	Traitement du contenu
Orienté				Suppression
Gradé	Inscrit_master	Bit		One to one
AnnéeOrientation				Suppression
CycleOrientation	Cycle_orientation	Varchar(50)		One to one
FilièreOrientation				Suppression
pre_nom_pere				Suppression
nom_prenom_mere				suppression
Nature_Abondance	Nature_abondance	Int(11)	Conde_academique Code_conge	One to one
Transféré	Transfere	Bit		One to one
Code_Daira				Suppression
Université				Suppression
Code_Institut				Suppression
Cycle				Suppression
Controle1				Suppression
Controle2				Suppression
ChampSupText1				Suppression
ChampSupText2				Suppression
ChampSupNum1				Suppression
ChampSupNum2				Suppression
Moy1Sem	Resultat_sem	Moy_sem	Double	One to one
Moy2Sem	Resultat_sem	Moy_sem	Double	One to one
Moy3Sem	Resultat_sem	Moy_sem	Double	One to one
Moy4Sem	Resultat_sem	Moy_sem	Double	One to one
Moy5Sem	Resultat_sem	Moy_sem	Double	One to one
Moy6Sem	Resultat_sem	Moy_sem	Double	One to one
Crédit1Sem	Resultat_sem	Crédit	Double	One to one
Crédit2Sem	Resultat_sem	Crédit	Double	One to one
Crédit3Sem	Resultat_sem	Crédit	Double	One to one
Crédit4Sem	Resultat_sem	Crédit	Double	One to one
Crédit5Sem	Resultat_sem	Crédit	Double	One to one
Crédit6Sem	Resultat_sem	Crédit	Double	One to one
Rachetable	Rachetable	Bit		One to one
Orient				Suppression
Photo				Suppression
presume				Suppression

Nbre_Total	Nb_redoublement_total	Int(11)	One to one
N°Ins_Anc	N_ind_anc	Int(11)	One to one
Lieu_Naissance_FR			Suppression
Nationalité_FR	Nationalité	Varchar(50)	Suppression
Nb_dettes	Nb_modules_en_dettes	Int(11)	One to one
Nb_session2	Nb_session2	Int(11)	One to one
reintegre			Suppression
rec	Nb_rectification	Int(11)	One to one

Table 2-3 Table de mapping de la table source « Etudiants »

V.4. Architecture Globale de l'application

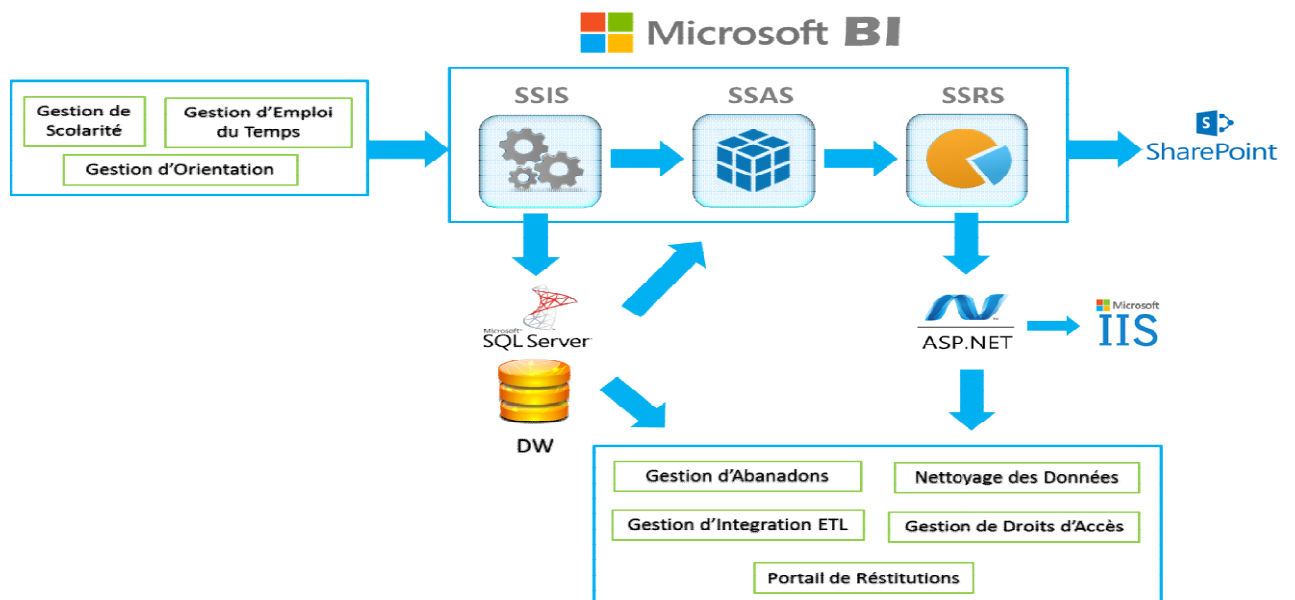


Figure 2-10 architecture Globale de l'application

Le processus de fonctionnement de l'application commence par le chargement des données à partir de plusieurs bases de données source de type Access et MySQL. Ces bases de données entre dans le cycle MSBI (Microsoft Business Intelligence) qui commence par l'intégration de données en utilisant l'outil SSIS intégré dans Visual Studio. Le résultat de cette intégration est enregistré dans une base de données SQL Server (Data Warehouse), par la suite on génère le Cube OLAP avec l'outil SSAS et finalement le cycle BI qui s'achève par la création des rapports avec SSRS. Une fois qu'on a abouti, nous avons créé une application ASP qui contrôle l'accès aux pages et aux rapports, les pages ASP sont déployés sur un serveur IIS. La dernière phase consiste à fusionner cette page ASP avec

une autre application plate-forme Web développée avec une multitude de technologies telles que PHP, Material Design, Bootstrap, HTML5, CSS3, JS5, Materialize, jQuery, etc... Cette plate-forme possède toutes les fonctionnalités évoquées dans ce chapitre.

VI. Conclusion

Dans cette section, on a pu diagnostiquer l'état actuel des deux services de scolarité des facultés des sciences et des technologies dans l'optique de proposer et concevoir une solution BI sous forme d'outil d'aide à la décision, qui fournit un ensemble de rapports et tableaux de bords nécessaires pour le contrôle et le suivi des différentes activités pédagogiques des facultés.

Chapitre3 : Réalisation

Chapitre 3 Réalisation

I. Introduction

Dans cette section, nous nous intéressons à la description détaillée des différentes phases de développement de notre système, en commençant par les différents outils logiciels employés pour réaliser la solution proposée. Nous allons présenter par la suite, la succession d'étapes de développement qui s'entame par la mise en œuvre du processus d'alimentation (ETL) des datamarts et qui s'achève par l'élaboration de l'architecture globale du système.

II. Environnement logiciel

Pour aboutir notre projet, nous avons dû employer un certain nombre d'outils et logiciels :

II.1. SGBD Microsoft SQL Server

Vu que notre solution est purement Microsoft, on a utilisé le SGBD Microsoft SQL Server pour le stockage du DataWarehouse.

II.2. Microsoft Visual Studio

C'est l'outil principal utilisé dans le développement, puisque notre solution BI est basée sur les trois modules (SSIS, SSAS, SSRS) intégrés dans le logiciel.

II.3. SQL Server Integration Services (SSIS)

Ce module est le responsable de la phase ETL, il permet d'extraire les données à partir d'une ou plusieurs bases de production sources, les transformer à l'aide d'une série d'opérations et de transformations, et finalement les charger dans le DataWarehouse.

II.4. SQL Server Analysis Services (SSAS)

C'est la solution proposée par la suite Microsoft qui permet de générer des cubes OLAP, puis analyser les données grâce à des fonctions d'analyse multidimensionnelle.

II.5. SQL Server Reporting Services (SSRS)

C'est le générateur de rapports entièrement intégré dans Visual Studio qui met à la disposition toutes les fonctionnalités de création de rapports.

III. Processus d'extraction, de transformation et de chargement

La phase ETL est la première phase de la réalisation d'un projet décisionnel, sa finalité est de construire l'entrepôt de données (DW).

Pour le faire, nous avons appliqué, aux données extraites des différents bases de production source, les transformations et les opérations nécessaires telles que les conversions de types jointures, ainsi que d'autres opérations. On a aussi respecté les contraintes d'intégrité lors du chargement des clés étrangères.

L'intégration a été divisé sur 4 packages :

La figure ci-dessous représente l'intégration des deux dimensions facultés et départements. En premier lieu, on a fait un test de type de bases de données par le module (check DB type) puisqu'on a deux sources de données (MySQL et Access) :

- Si la base de données est de type MySQL(EDT), on fait l'intégration de la dimension département ;
- Si c'est la base Access, on fait d'abord l'intégration des facultés puis celle des départements pour qu'on respecte les contraintes d'intégrité relatives aux clés étrangères.

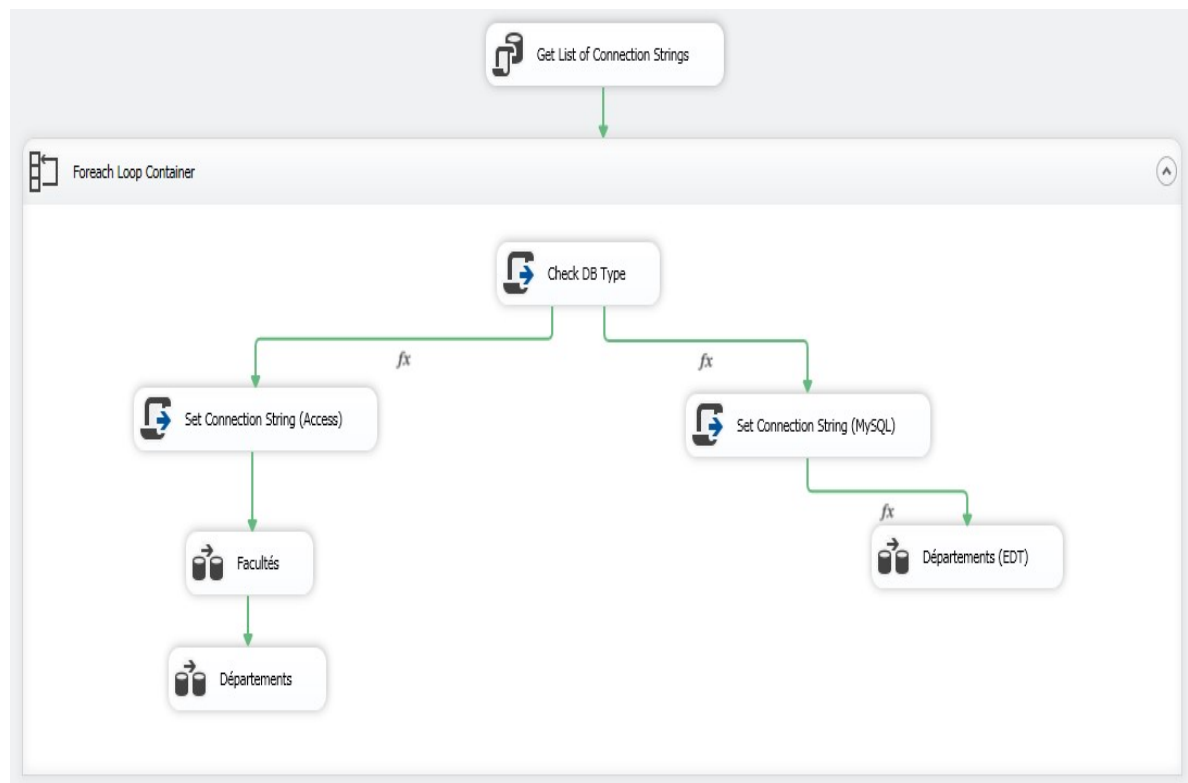


Figure 3-1 Processus ELT du premier package

La figure ci-dessous représente le processus d'intégration des différentes filières provenant des deux bases de données MySQL (EDT) et Microsoft Access.

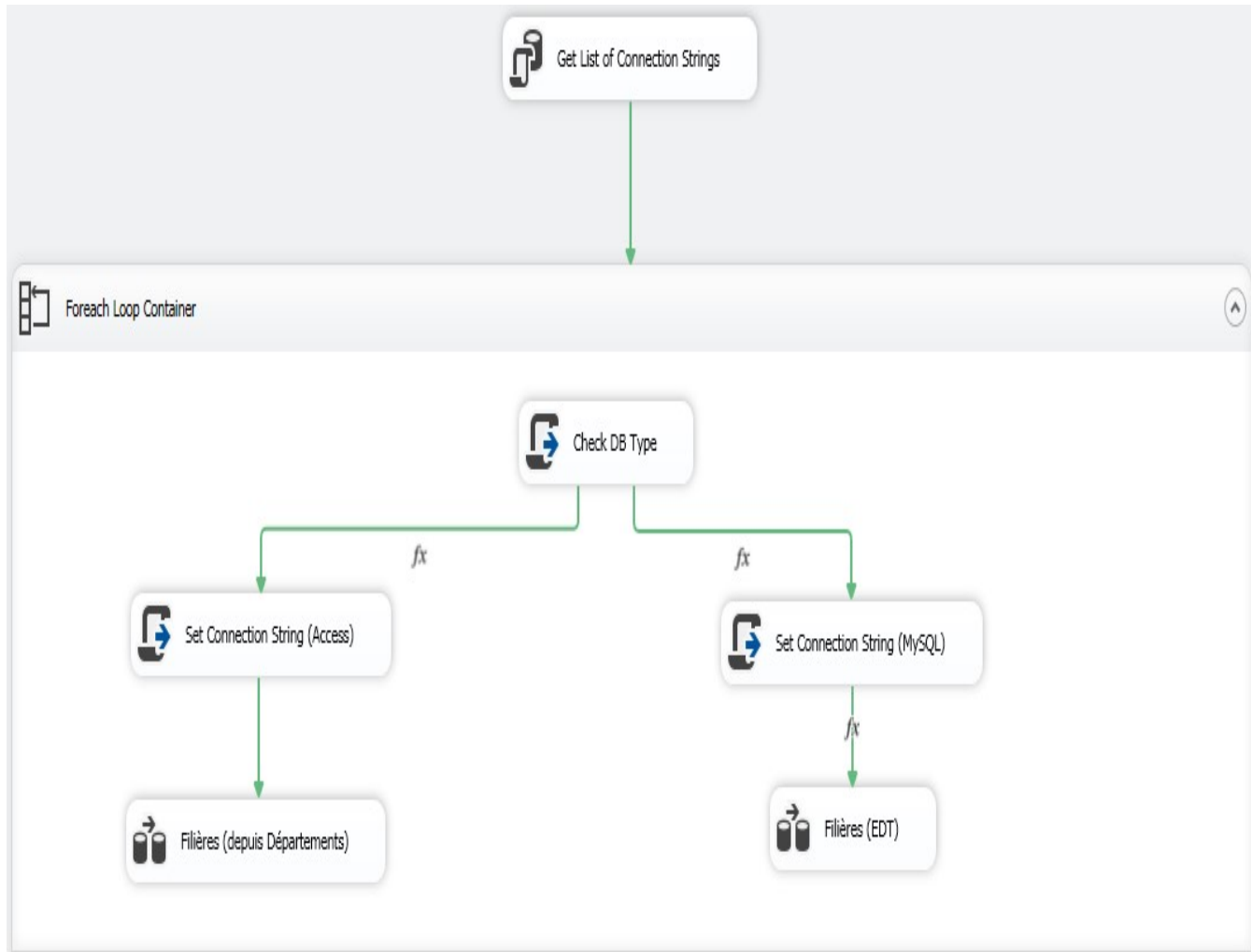


Figure 3-2 Processus ELT du 2ème package

Vu l'ambiguïté et la redondance de la table option issue de trois sources de données : MySQL(EDT), MySQL (fichevoeux) et Access, on a pris tout un package pour faire l'intégration de cette table. Le processus d'intégration est illustré dans la figure ci-dessous :

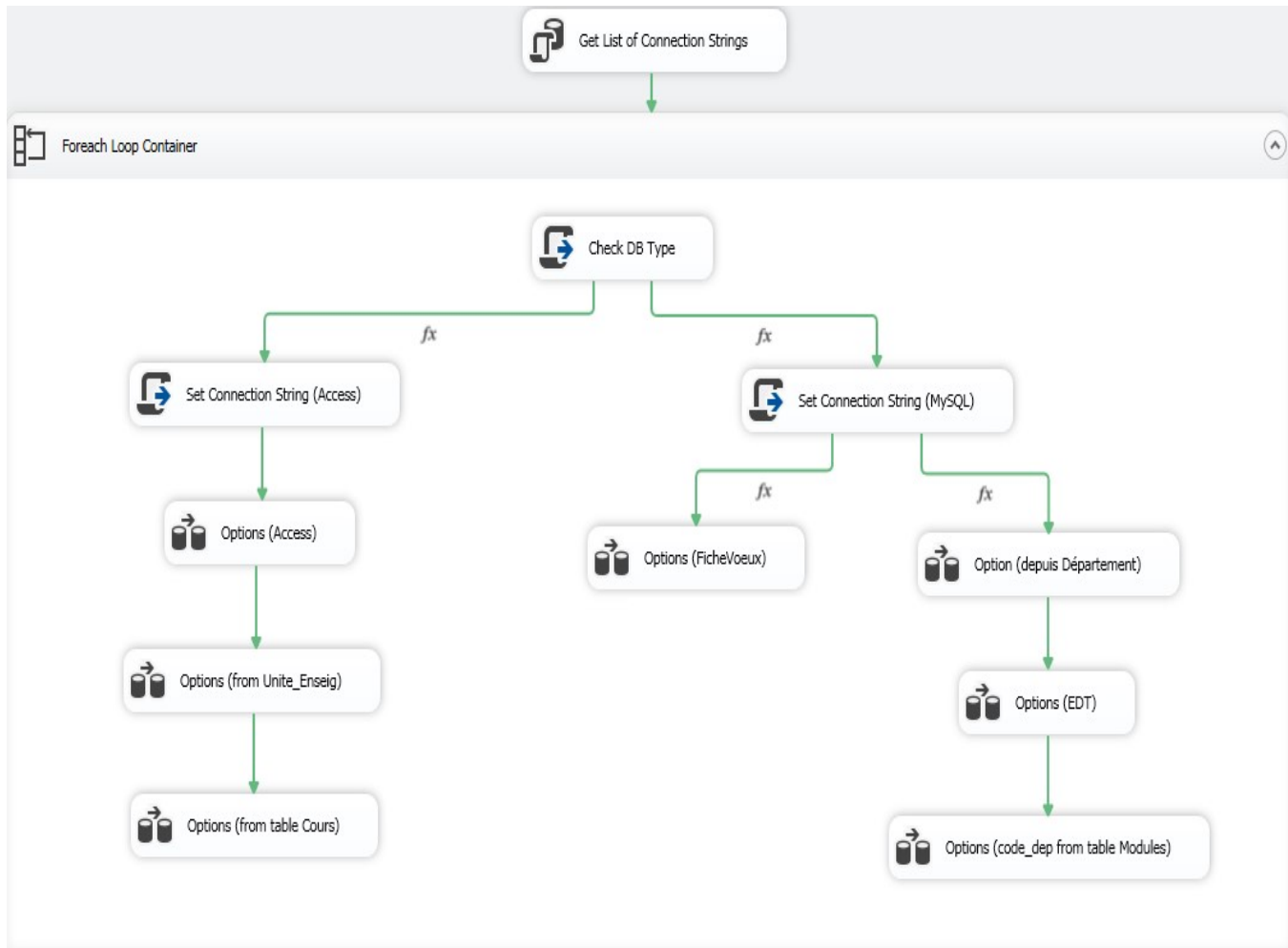


Figure 3-3 Processus ELT du 3ème package

La figure ci-dessous représente un package récapitulatif du processus d'intégration du reste des tables en respectant un schéma et un ordre d'intégration relativement compliqué comme ci-illustré dans la figure :

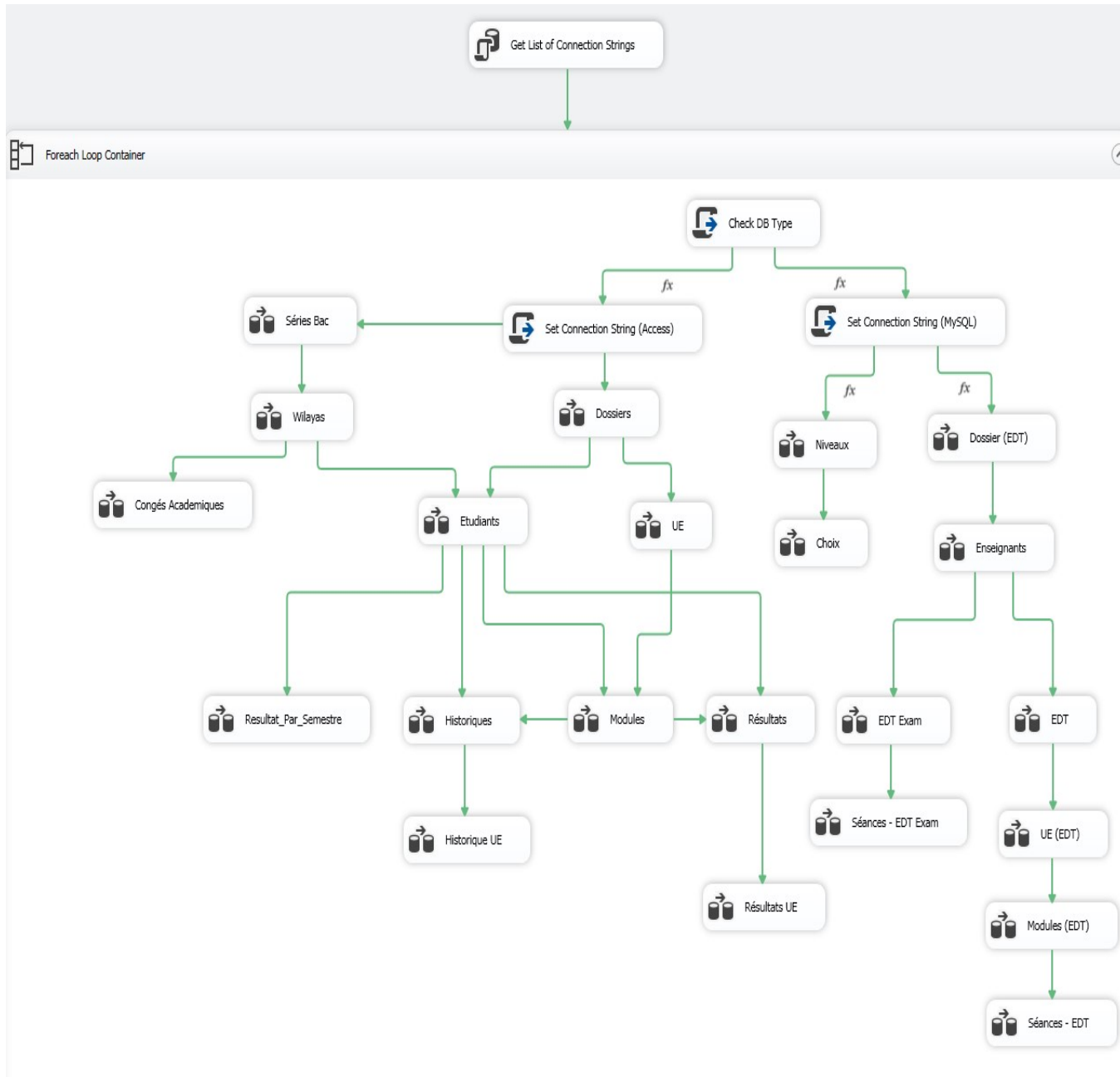


Figure 3-4 Processus ELT du 4ème package

Chaque module est constitué d'un certain nombre de tâches d'intégration. Pour donner des explications plus détaillées sur le processus d'intégration, on a choisi la tâche « Etudiants » :

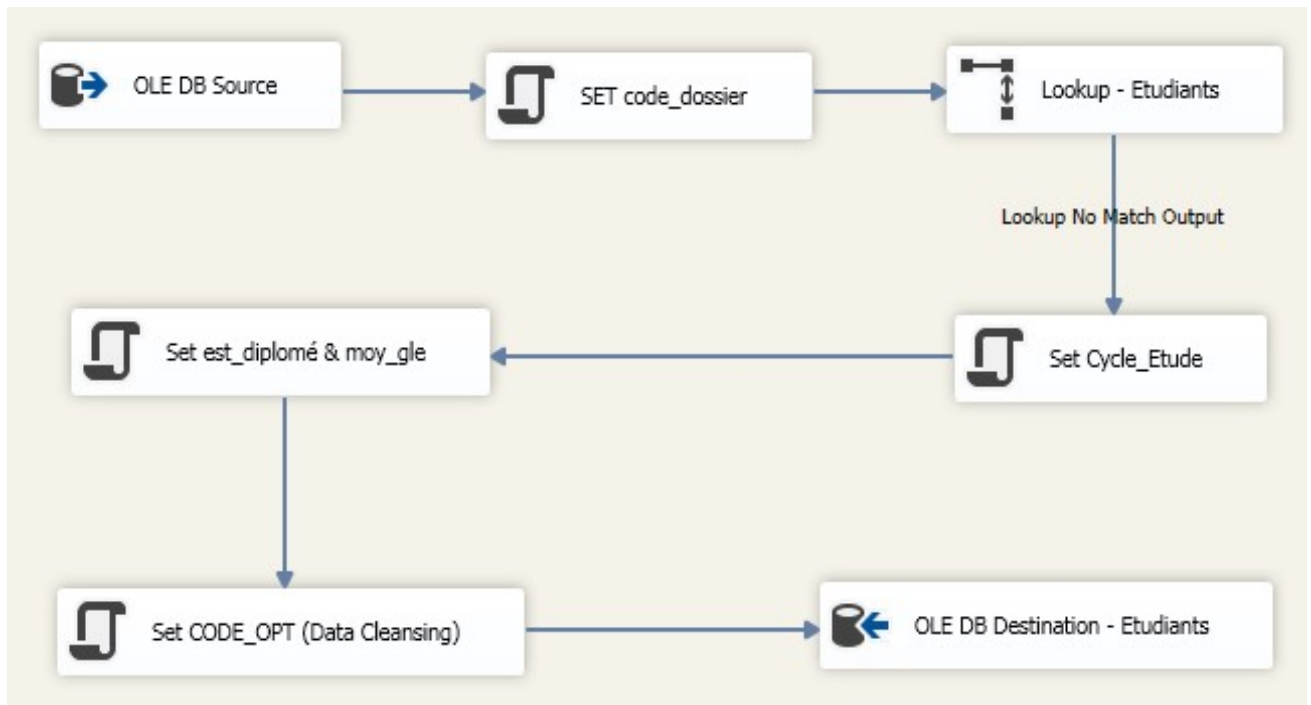


Figure 3-5Exemple détaillé d'une tâche ELT (Table Etudiants)

- **OLE DB Source** : Ce module sert à extraire les données depuis le Data Source (Access), table « Etudiant ».
- **Script (SET code_dossier)** : Un script C# qui établit une connexion avec le Data Source courant (Access) et récupère le champ « code_dossier » depuis la table Dossier.
- **Lookup** : Un composant qui cherche l'existence de l'enregistrement courant (étudiant courant) dans la table « Etudiant » du DW qui se fait par le mapping des clés primaires(utile pour les nouveaux enregistrements qui vont être insérés).
- **Script (Set Cycle_Etude)** : Un script C# qui récupère le champ « CycleEtude » depuis la table « Etudiant » et sélectionne le nombre d'années correspondant dans la table « AnnéeCycle », et selon ce nombre, on va décider si le cycle est « Licence » ou « Master » (3 années = Licence, 2 années = Master).
- **Script (Set est_diplomé&moy_gle)** : Un autre script qui calcule la moyenne générale de l'étudiant pour chaque année universitaire à partir de la somme des moyennes des semestres. Ensuite, on récupère l'année d'études, le cycle d'études et tester si l'étudiant a une moyenne supérieure à 10 pour chaque année pour en décider s'il va être diplômé ou non.

- **Script (Set CODE_OPT)** : Ce script est spécifique pour le nettoyage des données, indispensable dans toutes les tâches d'intégration. Il sert à traiter les données des tables qui ont une relation (clé étrangère) avec une des tables suivantes : Université, Faculté, Domaine, Département, Filière et Option. Dans le cas de l'exemple précédent ou la table « Etudiant » quia une clé étrangère (CODE_OPT) provenant de la table « Option », le script va sélectionner l'option d'étudiant depuis le data source (Access), et la remplacer avec sa « valeur réelle » depuis la table Option (on va revenir au nettoyage des données dans la section suivante).
- **OLE DB Destination** :Ce module sert à charger les données dans le DW avec le mapping des champs.

Notons qu'il existe d'autres composants plus complexe comme la création des nouvelles colonnes, la conversion des types des données,...etc. qui sont utilisés dans d'autres cas différents.

IV. Génération du cube OLAP

Après avoir construit et alimenté notre DW, on procède à l'élaboration des cubes OLAP. Cette étape est primordiale dans tout projet décisionnel.

Le moteur SSAS offert par Microsoft nous a permis la génération du cube OLAP à partir du data source (DataWarehouse). Ensuite, on a fait quelques modifications sur ce dernier, comme la création des hiérarchies et l'ajout des membres calculés, ...etc. Une fois la création et les modifications sont terminées, on a déployé le cube dans le serveur SQL Server Analysis services 2014, pour qu'on puisse analyser nos indicateurs selon les axes d'analyses.

V. Génération des rapports

La génération des rapports est considérée comme la phase finale dans un cycle BI, dans laquelle, on s'intéresse à la création des différents rapports recensés lors de l'étude des besoins. Pour cela, on va utiliser le générateur de rapport SSRS fourni par Microsoft. La création d'un rapport nécessite une bonne connaissance du langage MDX. Ce dernier est un langage de requête sur les bases de données multidimensionnelles qui permet d'explorer les cubes OLAP, interroger les données qui y sont contenues et les manipuler.

Parmi les rapports générés, dans cette étape, on cite à titre d'exemple les rapports suivants :

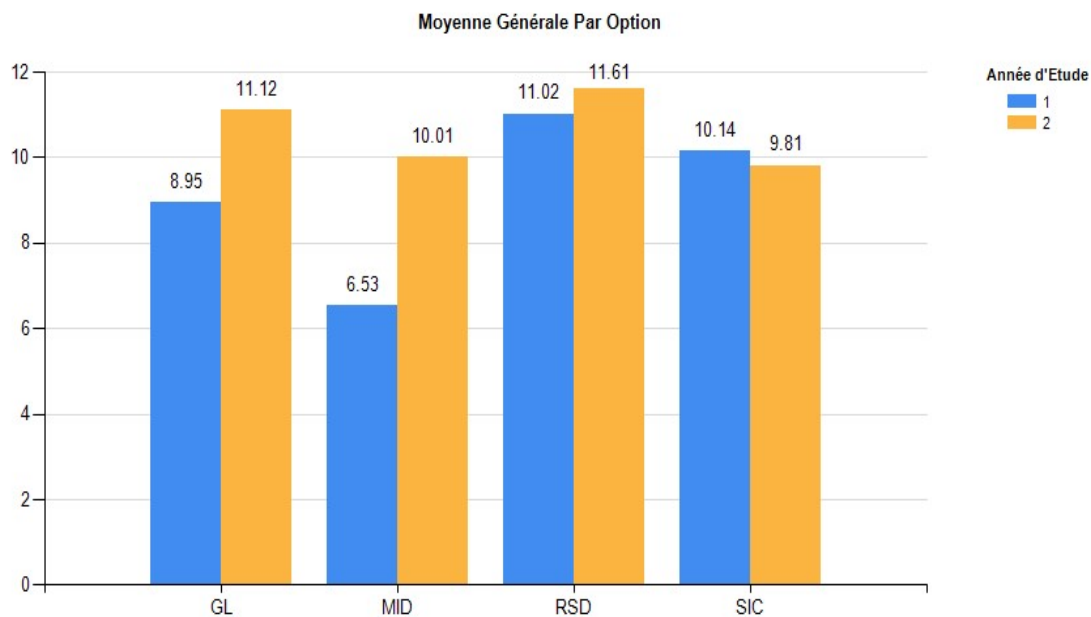


Figure 3-6 Exemple1 : Rapport moyenne générale des étudiants par spécialité

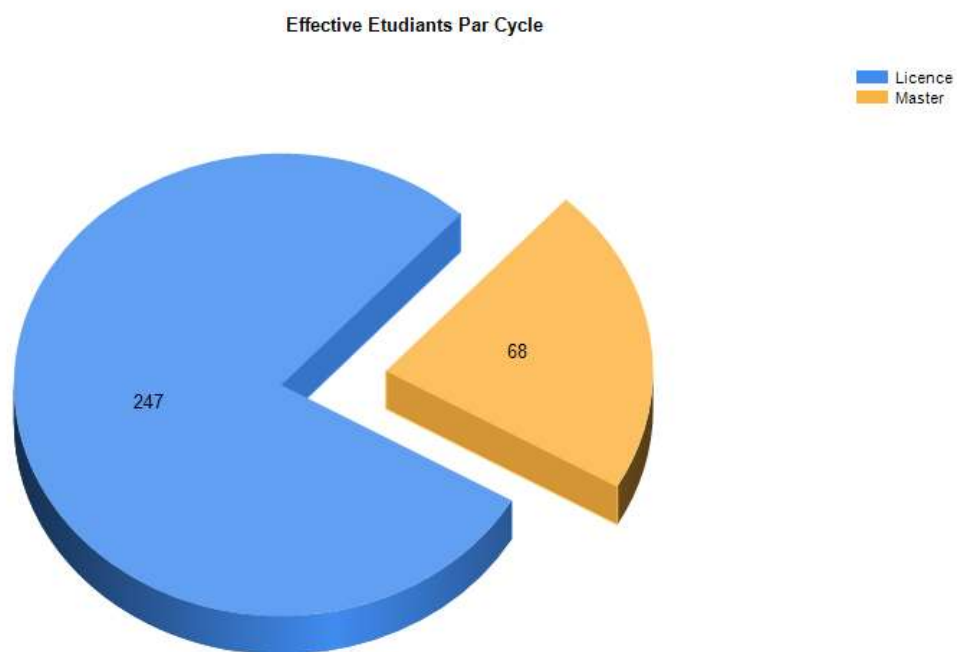


Figure 3-7 Exemple2 : Rapport effectif des étudiants par cycle d'étude (licence, master)

Evolution de la faculté

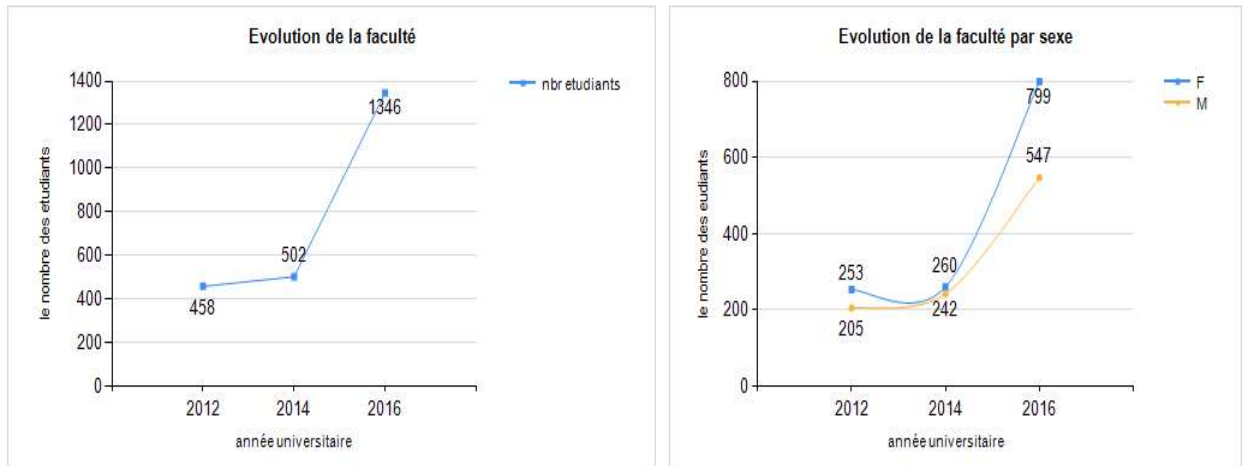


Figure 3-8 Exemple3 : Rapport évolution des effectifs des étudiants de la faculté par année

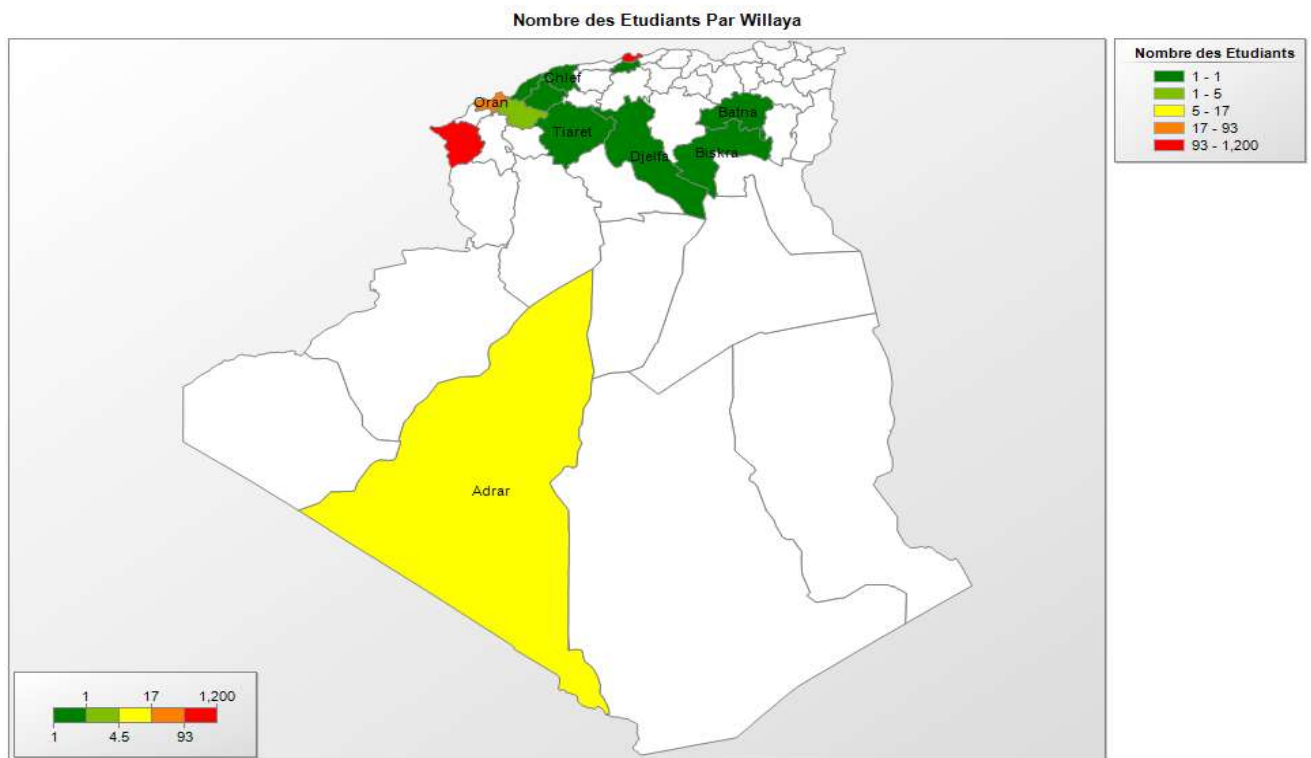


Figure 3-9 Exemple4 : Rapport de distribution des étudiants de l'université par wilaya

VI. Nettoyage des données

Vu que les data sources sont de types différents d'une part, d'autre part, la saisie des données qui diffère d'un département à l'autre ; Comme par exemple, la saisie des données dans la scolarité est complètement différente que celle des départements alors pour la saisie du code de l'option « Génie Logiciel », certains saisissent « GL », d'autres « M_GL » et d'autres « GENLO »...), donc dans le but d'avoir des rapports fiables, nous avons réalisé une application web, faisant partie de notre plateforme, qui fait la gestion du nettoyage des données, selon le scénario suivant :

1. L'administrateur ajoute les universités, les facultés ainsi que les domaines puisqu'on n'a pas ces informations dans les data sources. Par la suite, il associe les uns avec les autres : les domaines aux facultés et les facultés à l'université pour construire une hiérarchie.
2. Ensuite, l'administrateur lance le processus du chargement des départements (avec SSIS). Après que l'intégration des départements s'est terminée, il associe un domaine pour chaque département. Pour résoudre le problème de redondance des noms du même département, l'administrateur va unifier ces différents noms sous un seul nom choisi, qui va être utilisé dans nos rapports.
3. Le même principe va être appliqué pour les filières et options.

Le chargement des données sera dans des tables temporaires, en revanche, l'assemblage final va être dans les tables « originaux » du DW. A titre d'exemple : on commence par charger les filières depuis les data sources puis on les stocke dans la table « TMP_FILIERE ». Les filières qui vont être utilisées dans les rapports, vont être créées par la suite ; à ce niveau, deux cas de figures peuvent exister : soit, on va commencer la création de zéro, soit on va convertir les filières chargées directement comme des filières fiables et prêtes à être utilisées. Enfin, on associe chaque filière de la table « TMP_FILIERE » avec une filière de la table « FILIERE ».

La plupart des bases de données Access du logiciel « SEES », installées au niveau des départements, ont été remplies d'une manière standard, sauf celle du département d'informatique. Alors pour y remédier, on était dans l'obligation de développer une autre application Web (DataCleansing) pour corriger tous ces défauts et procéder à l'uniformisation de ces bases, qui consiste au changement des colonnes, le transfert des données d'une table vers une autre, la mise à jour des références, ...etc.

Un autre problème de pertinence de données est en relation avec les étudiants qui ont abandonné leurs études puisqu'ils figurent encore dans les bases de données source. Pour y remédier, une autre application pour le nettoyage des données a été réalisée à savoir la gestion des abandons, où l'utilisateur va sélectionner les étudiants qui ont abandonnés leurs études et procéder à leur nettoyage du DW, pour avoir des rapports plus fiables.

VII. Gestion des droits d'accès

Vu que notre application est multi-utilisateurs avec différents niveaux d'accès aux données, nous avons créé un système de gestion d'accès aux données et aux rapports, qui consiste à :

- Création des rôles : Pour contrôler et filtrer les données par visibilité (Université, Faculté, Domaine, Département, Filière, Option), donc les données disponibles pour le doyen ne sont pas les mêmes pour le chef du département.
- Création des profils : Pour contrôler l'accès aux rapports, donc le chef de département n'a pas l'accès aux rapports de la faculté par exemple.
- Création des utilisateurs : chaque utilisateur a un rôle et un profil.

Pour garantir le contrôle des accès aux données (filtrage), il faut disposer de 6 paramètres de filtrage dans chaque rapport qui seront reliés avec les dimensions : Université, Faculté, Domaine, Département, Filière et Option. Ces paramètres sont visibles et modifiables selon le rôle d'utilisateur ; par exemple, le doyen peut sélectionner n'importe quel département qui fait partie de sa faculté, par contre le chef du département n'a l'accès qu'à son département et aux filières faisant partie de ce dernier.

Le contrôle d'accès aux données ou aux rapports est effectué par le biais d'une page ASP hébergée dans un serveur IIS, utilisant le langage de programmation C#. Cette page est affichée dans une autre page PHP en tant que<iframe>, cette dernière reçoit 2 paramètres (GET) qui sont l'URL et le titre du rapport. Un composant d'affichage des rapports est créé dans cette page avec les paramètres reçus, ou le filtrage des données a été appliqué.

VIII. Organisation des fichiers de la plateforme

La plateforme est une application web qui utilise 2 types de serveurs : Apache et IIS, elle contient un ensemble de pages PHP et ASPX, qui s'articule sur un ensemble de fichiers organisé comme suit :

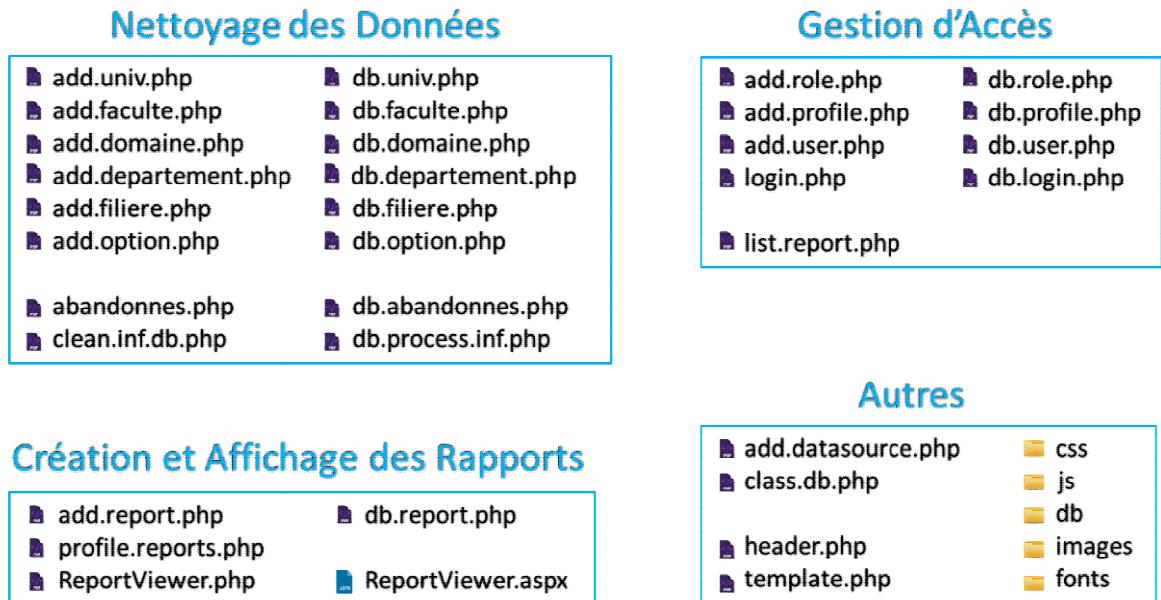


Figure 3-10 architecture des fichiers de notre application

IX. Architecture technique

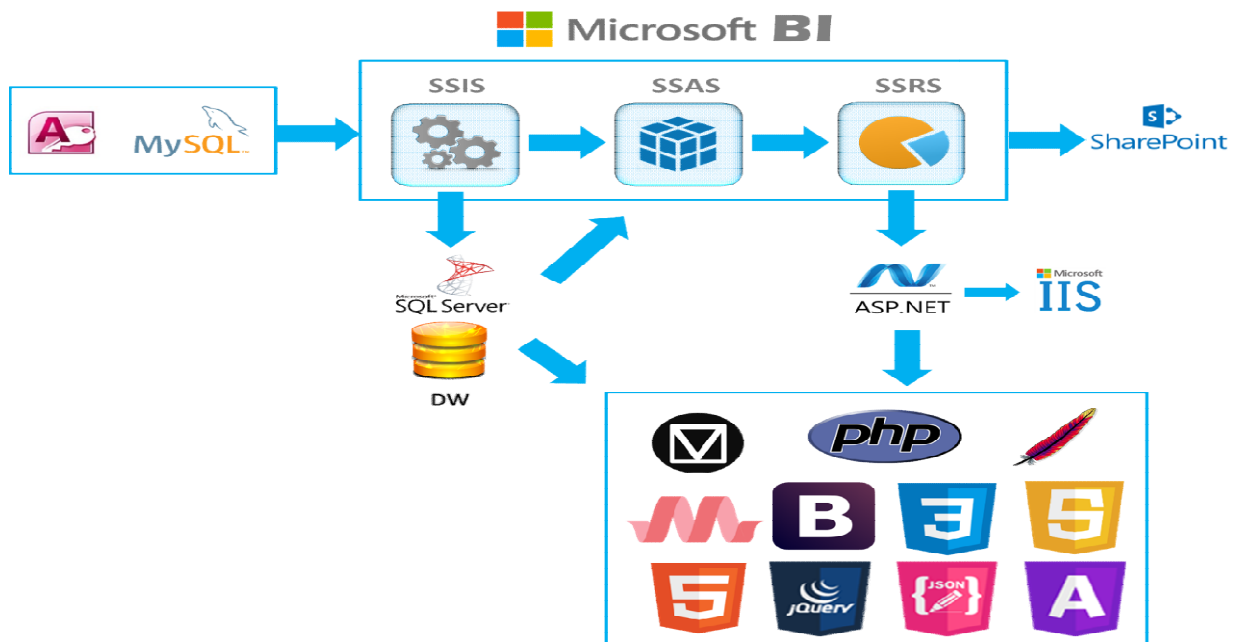


Figure 3-11 architecture technique de l'application

La figure ci-dessus montre l'architecture technique de notre application qui se décompose comme suit :

Base de données sources :

- base de données MySQL (logiciel d'orientation)
- base de données MySQL (logiciel d'emplois du temps)
- base de données Access (logiciel S.E.E.S gestion de la scolarité)

La suite Microsoft BI :

- SSIS le module Microsoft responsable du processus ETL
- SSAS le module Microsoft responsable de la création des cubes OLAP et de les analyser
- SSRS le module responsable de la Création et la restitution des différents rapports

Portail web :

- Sharepoint : c'est le portail web développé par Microsoft qui permet la gestion de contenu, la gestion électronique des documents, la possibilité de créer des formulaires et des statistiques décisionnelle, ...etc.

SGBD :

- Microsoft SQL server qui est le système de gestion de base de données développé et commercialisé par la société Microsoft

Serveur Web :

- IIS : le serveur web de Microsoft qui permet l'hébergement d'un site web en ASP
- Apache : le serveur web qui permet l'hébergement d'un site web en PHP

Langage et framework :

- PHP, Material Design, Bootstrap, HTML5, CSS3, JS5, Materialize, jQuery.

Conclusion générale

Conclusion générale

En conclusion, nous rappelons que notre projet de fin d'études s'est déroulé au sein des services de scolarité des deux facultés, à savoir la faculté des sciences et la faculté des technologies de l'université Abou Bekr Belkaid Tlemcen. L'objectif principal ciblé était la mise en place d'une solution décisionnelle qui permet aux décideurs de suivre et contrôler les activités pédagogiques comme les inscriptions des étudiants et leurs résultats ainsi que la répartition des tâches des enseignants et des salles.

Tout au long du processus de conception et de réalisation, nous avons suivi une démarche mixte, alliant de ce fait entre deux approches connues dans le domaine du « Data Warehousing » à savoir l'approche « Besoins d'analyse » et l'approche « Sources de données », afin de mieux formaliser et réunir les conditions de réussite de notre projet décisionnel. Pour cela on a défini un planning qui nous a permis de fixer les tâches nécessaires à la réalisation de notre projet.

En respectant les étapes de cette méthodologie, nous avons commencé, par définir les besoins et les objectifs fixés pour aboutir notre projet, après une série de réunions de travail avec les responsables des fonctions concernées. Ensuite une analyse de l'existant a été effectuée, cette étape est considérée l'une des étapes les plus dures dans notre projet car elle exige la compréhension de la structure des sources de données. En effet, nous avons dû surmonter cette structure complexe composée de plusieurs tables, chacune pouvant provenir d'un système différent et comportant plusieurs dizaines de champs. L'absence et l'indisponibilité de la documentation sur la structure des sources de données nous a obligé en conséquence à passer beaucoup de temps pour comprendre les relations entre les tables ainsi que la signification de chaque attribut. Pour en finir par la conception des modèles multidimensionnels.

La réalisation du projet et la concrétisation de la solution proposée a consisté à l'alimentation du DataWarehouse, la création du cube OLAP ainsi que la création des différents rapports et tout cela en utilisant les modules BI offerts par la suite Microsoft. En dernier lieu, nous avons développé des modules de nettoyage de données pour avoir des résultats de rapports plus fiables, ainsi qu'une gestion des droits d'accès aux données et aux rapports a été mise en place. Toutes ces fonctionnalités ont été déployées sur une plateforme web.

Pour en conclure, on peut affirmer que ce projet original et prometteur nous a permis d'approfondir nos connaissances acquises en matière de l'informatique décisionnelle ainsi que d'acquérir une première expérience professionnelle dans ce domaine. On peut considérer qu'on a largement abouti aux objectifs fixés et souhaités par les utilisateurs. Le nouveau système développé, a été validé et même très apprécié par les deux structures concernées et promis d'être adopté immédiatement.

Néanmoins quelques perspectives et améliorations peuvent être envisagées pour perfectionner le travail effectué, citons principalement :

- Assister le déploiement du nouveau système et recueillir les correctifs et les remarques éventuelles des utilisateurs.
- Etendre le déploiement de manière à couvrir tous les services de scolarité de l'université
- Continuer le développement du portail de restitution des autres rapports non programmés.
- Intégrer des données d'autres applications de l'université

Références bibliographiques

- [1] J.F. Goglin; « La Construction du Datawarehouse : du Datamart au Dataweb »; Hermes 1998
- [2] Benjamin Goewey “the 10 most important benefits of business intelligence” 31/10/2015 <http://www.datamensional.com/the-10-most-important-benefits-of-business-intelligence/> visité le: 10/06/2017
- [3] Data warehouse4You “OLTP vs OLAP” 2008/2009 <http://datawarehouse4u.info/OLTP-vs-OLAP.html> visité le: 11/06/2017
- [4]Wikipedia “DataWarehouse” 16/05/2017 https://en.wikipedia.org/wiki/Data_warehouse visité le 11/06/2017
- [5] Shannon kemp “a short history of DataWarehousing” 23/08/2015 <http://www.dataversity.net/a-short-history-of-data-warehousing/> visité le : 11/06/2017
- [6] http://pagesperso.lina.univ-nantes.fr/~prie-y/archives/VEILLE-2009-2012/2010/bi-nouvtech/le_MDM.html
- [7] techdifferences “The difference between fact and dimension tables” 26/11/2016 <http://techdifferences.com/difference-between-fact-and-dimension-table.html> visité le : 11/06/2017
- [8] Igm « la méthodologie OLAP » avril 2014 http://www-igm.univ-mlv.fr/~dr/XPOSE2009/informatique_decisionnelle_olap/md.html visticé le 11/06/2017
- [9]blog.developper « starschema vs snowflake » 12/03/2015 https://blog.developpez.com/jmalkovich/p8718/modelisation/modele_en_etoile_ou_en_flocons visité le 11/06/2017
- [10] Rugved Darwhekar « MOLAP vs ROLAP vs HOLAP » 12/04/2017 <http://onlydifferencefaqs.blogspot.com/2012/09/difference-between-molap-rolap-and.html> visité le 11/06/2017
- [11]msdn « multidimensional expressions » 2017 [https://msdn.microsoft.com/fr-fr/library/ms145506\(v=sql.90\).aspx](https://msdn.microsoft.com/fr-fr/library/ms145506(v=sql.90).aspx) visité : juin 2017.
- [12] cheick sanogo « vue d’ensemble de l’architecture d’un système BI » 30/10/2015 <http://www.supinfo.com/articles/single/1182-vue-ensemble-architecture-system-bi> visité le 11/06/2017
- [13] FILALI ABDERRAHMANE,KEDJNANE SOFIANE : « Conception et réalisation d’un Data Warehouse pour la mise en place d’un système décisionnel » mémoire de fin d’étude 2015

[14] Thomas becquet « Gartner Magic Quadrant BI 2017 » 1/05/2017
<https://blog.censio.fr/bi/magic-quadrant-bi-2017/> visité le : 17/06/2017

[15] wikipedia « rétro-ingénierie » avril 2017 <http://fr.wikipedia.org/wiki/Rétroingénierie>
visité: juin 2017

[16] R. Kimball et J. Caserta; « The Data warehouse ETL Toolkit» publié 2004

Liste des figures

FIGURE 1-1 LE DECISIONNEL AU SEIN DU SYSTEME D'INFORMATION [1]	6
FIGURE 1-2 DIFFERENCE ENTRE SYSTEME OALP ET OLTP [3]	7
FIGURE 1-3 EVOLUTION DES BASES DE DONNEES DECISIONNELLES [6]	9
FIGURE 1-4 ARCHITECTURE EN ETOILE QUI ILLUSTRE LE CONCEPT DE FAIT ET LE CONCEPT DE DIMENSION [8]	10
FIGURE 1-5 VUE D'ENSEMBLE DE L'ARCHITECTURE D'UN SYSTEME BI [12]	13
FIGURE 1-6 BI QUADRANT MAGIQUE DE GARTNER 2017	17
FIGURE 2-1 DIAGRAMME DE GANTT DE L'AVANCEMENT DU PROJET	20
FIGURE 2-2 MODELE RELATIONNEL DE DONNEES (BASE DE DONNEES ACCESS) PARTIE 1	22
FIGURE 2-3 MODELE RELATIONNEL DE DONNEES (BASE DE DONNEES ACCESS) PARTIE 2	23
FIGURE 2-4 MODELE RELATIONNEL DE DONNEES (BASE DE DONNEES MYSQL) : LOGICIEL DES EMPLOIS DU TEMPS	25
FIGURE 2-5 MODELE RELATIONNEL DE DONNEES (BASE DE DONNEES MYSQL) : LOGICIEL D'AFFECTATION DES CHOIX	26
FIGURE 2-6 DIAGRAMME DES CAS D'UTILISATIONS GLOBAL	30
FIGURE 2-7 DATAMART DES INSCRIPTIONS	31
FIGURE 2-8 DATAMART DES RESULTATS	32
FIGURE 2-9 DATAMART REPARTITION_ENSEIGNANTS_SEANCE	33
FIGURE 2-10 ARCHITECTURE GLOBALE DE L'APPLICATION	36
FIGURE 3-1 PROCESSUS ELT DU PREMIER PACKAGE	39
FIGURE 3-2 PROCESSUS ELT DU 2EME PACKAGE	40
FIGURE 3-3 PROCESSUS ELT DU 3EME PACKAGE	41
FIGURE 3-4 PROCESSUS ELT DU 4EME PACKAGE	42
FIGURE 3-5 EXEMPLE DETAILLE D'UNE TACHE ELT (TABLE ETUDIANTS)	43
FIGURE 3-6 EXEMPLE1 : RAPPORT MOYENNE GENERALE DES ETUDIANTS PAR SPECIALITE	45
FIGURE 3-7 EXEMPLE2 : RAPPORT EFFECTIF DES ETUDIANTS PAR CYCLE D'ETUDE (LICENCE, MASTER)	45
FIGURE 3-8 EXEMPLE3 : RAPPORT EVOLUTION DES EFFECTIFS DES ETUDIANTS DE LA FACULTE PAR ANNEE	46
FIGURE 3-9 EXEMPLE4 : RAPPORT DE DISTRIBUTION DES ETUDIANTS DE L'UNIVERSITE PAR WILAYA	46
FIGURE 3-10 ARCHITECTURE DES FICHIERS DE NOTRE APPLICATION	49
FIGURE 3-11 ARCHITECTURE TECHNIQUE DE L'APPLICATION	49

Liste des tableaux

TABLE 1-1 TABLEAU COMPARATIF ENTRE LES SYSTEMES TRANSACTIONNELS ET LES SYSTEMES DECISIONNELS [3]	8
TABLE 1-2 TABLEAU COMPARATIF ENTRE LES TABLES DE FAITS ET LES TABLES DE DIMENSIONS [7]	11
TABLE 1-3 TABLEAU COMPARATIF DES DIFFERENTES ARCHITECTURES OLAP [10].....	12
TABLE 1-4 TABLEAU COMPARATIF DES SUITES BI [13]	16
TABLE 2-1 TABLEAU RECAPITULATIF DE L'ENSEMBLE DES TACHES DU PROJET	20
TABLE 2-2 TABLEAU RECAPITULATIF DES DIFFERENTS RAPPORTS REQUIS	28
TABLE 2-3 TABLE DE MAPPING DE LA TABLE SOURCE « ETUDIANTS »	36

Liste des abréviations

BI : Business Intelligence

CRM: Customer Relationship Management

DW: Data Warehouse

EDT: Emplois Du Temps

ERP: Enterprise Resource Planning

ETL: Extract Transform Load

HOLAP: Hybrid On Line Analytical Process.

IT: Information Technology

KPI: Key Performance Indicator

MDX: Multidimensional Expression

MOLAP: Multidimensional On Line Analytical Process.

OLAP: On Line Analytical Process.

OLTP: On Line Transactional Process

PFE : Projet Fin Etude

ROLAP: Relational On Line AnalyticalProcess.

SGBD: Système de Gestion de Base de Données.

SQL: Structured Query Language

SSAS: Sql Server Analysis Services.

SSIS: Sql Server Integration Services.

SSRS: Sql Server Reporting Services.

Annexes

Annexe1 : table de mapping

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Historique		Historique	id	int(11)			Ajoute
	N°ins		n_ins	int(11)	Etudiant	n_ins	
	Année		annee_etude	int(11)			Traitement
	AnnéeUniv		annee_univ	varchar(20)			Traitement
	Réf_Cours		code_module	varchar(50)	Cours	code_module	
	Cours						Suppression
	Coef						Suppression
	Moy_Module		moy_module	double			Traitement
	Moy_Gle						Suppression
	Observation						Suppression
	Université_Origine		universite_origine	varchar(50)			
	Faculté_Origine		faculte_origine	varchar(50)			
	Département_Origine-		specialite_origine	varchar(50)			
	Cycle_Origine-		domaine_origine	varchar(50)			
	Filière_Origine-		specialite_origine				
	credit_matière						Suppression
	Code_UE						Suppression
	credit_UE						Suppression
	Aquit		aquit_module	bit			
	An_d_obtent						Suppression
	Faculté_OrigineAR						Suppression
	Département_OrigineAR						Suppression
	Filière_OrigineAR						Suppression
	Université_OrigineAR						Suppression
	Cycle_OrigineAR						Suppression
	specialité_origine						Suppression
	specialité_origine_ar						Suppression
	poids						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Historique_UE		Historique_UE	id	int(11)			Ajoute
	N°ins		n_ins	int(11)	Etudiant	n_ins	
	Année						Suppression
	Code_UE						Suppression
	UE						Suppression
	Coef						Suppression
	credit_UE						Suppression
	AnnéeEtude						Suppression
	type_UE						Suppression
	Moy_UE		moy_UE	double			Traitement
	credit_obtenue		credit_obtenue	double			
	Aquit		aquit_UE	bit			
	An_d_obtent_UE						Suppression
	poids_ue						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Résultats		RESULTATS	id	int(11)			Ajoute
	N°ins		n_ins	int(11)	Etudiants	n_ins	
	Réf_Cours		code_module	varchar(50)	Cours	code_module	
	examen		examen	double			Traitement
	rattrapage		rattrapage	double			Traitement
	Conference						Suppression
	seminaire						Suppression
	TD		td_r	double			Traitement
	TP		tp_r	double			Traitement
	stage		stage_r	double			Traitement
	Projet		projet	double			Traitement
	autre_Ctl		autre_ctrl	double			Traitement
	Moy_CC		moy_cc	double			Traitement
	Moy_Module		moy_module_r	double			Traitement
	Aquit						Suppression
	Option						Suppression
	ratt		entree_ratt	bit			
	exclu_matière		exclu_module	bit			
	poids						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Résultats_UE		Resultats_UE	id	int(11)			Ajoute
	N°ins		n_ins	int(11)	Etudiant	n_ins	
	Code_UE		code_ue	varchar(20)	Unite_Enseig	Code_UE	
	credit_obtenu		credit_obtenue_r	double			Traitement
	Moy_UE		moy_ue_r	double			Traitement
	Aquit						Suppression
	exclu_ue		exclu_ue	bit			
	poids_ue						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Unite_Enseig		UE					
	Code_UE		code_UE	varchar(20)			
	UE		nom_ue	varchar(50)			Traitement
	Coef		coef	double			Traitement
	Année						Suppression
	AnnéeEtude		annee_etude	int(11)			Traitement
	CycleEtude		specialite	varchar(50)			Traitement
	FilièreEtude		specialite	varchar(50)			Traitement
	Semestre		semestre	int(11)			
	credit		credit	int(11)			
	type_UE		type_ue	varchar(50)			
	PwUE						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Cours		Module					
	Réf_Cours		code_module	varchar(20)			Traitement
	Cours		nom_module	varchar(50)			Traitement
	Coef		coef	double			Traitement
	VExam						Suppression
	VRttrpg						Suppression

	TD		td	double			Traitement
	VTD						Suppression
	TP		tp	double			Traitement
	VTP						Suppression
	Seminaire						Suppression
	Vsem						Suppression
	Conférence						Suppression
	Vconf						Suppression
	Stage		stage	double			Traitement
	VStg						Suppression
	Projet						Suppression
	VPrj						Suppression
	Autre_CTL						
	VCTL						Suppression
	Année						Suppression
	AnnéeEtude		annee_etude	int(11)			Traitement
	CycleEtud		specialite	varchar(50)			Traitement
	FilièreEtud		specialite	varchar(50)			Traitement
	Nbreemd		nbr_emd	int(11)			
	Semestre		semestre	int(11)			
	credit		credit	int(11)			
	coefExam		pourcentage_exam	double			Traitement
	coefCC		pourcentage_cc	double			Traitement
	CodeUE		code_UE	varchar(20)	UE	code_UE	Traitement
	Obligatoire		obligatoire	bit			
	PWCours						Suppression
	MatSaisie		mat_saisie	bit			
	mode_calcul_rat		mode_calcul_rat	double			Traitement

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Année_Cycle		Cycle_Etude					
	Code_Cycle		code_cycle	varchar(20)			Traitement
	Désignation_Cycle		designation_cycle	varchar(50)			Traitement
	Nbre_Année		nbr_annee	int(11)			Traitement
	Désignation_CycleA r						Suppression
	Descriptif						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Conge academique		Conge_Academique					
	code_conge		code_conge	int(11)			
	designation_conge		designation_conge	varchar(50)			
	designation_conge_ar						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Dossier		Dossier					
	Dossier		code_dossier	varchar(20)			
	Réf_Département		code_dep	varchar(20)	Departement	code_dep	Traitement
	Faculté						Suppression
	Département						Suppression
	Année						Suppression
	Semestre						Suppression
	Cycle						Suppression
	AnnéeUniv		annee_univ	varchar(20)			Traitement
	Clôture						Suppression
	Note						Suppression
	Délibération						Suppression
	Nb						Suppression
	Université						Suppression
	PasswordApp						Suppression
	PasswordBDD						Suppression
	AnnéeEnLettre						Suppression
	SemestreEnLettre						Suppression
	DésignationCycle						Suppression
	Filière						Suppression
	C_TD						Suppression
	C_TP						Suppression
	ChargeMoyCours						Suppression
	ChargeMoyTD						Suppression
	ChargeMoyTP						Suppression
	CôutHeure						Suppression
	NbreSemAnne						Suppression
	FacultéAR						Suppression

	DépartementAR						Suppression
	FilièreAR						Suppression
	AnnéeEnLettreAr						Suppression
	UniversiteAr						Suppression
	CycleAr						Suppression
	note_sup						Suppression
	Abattement						Suppression
	Descriptif						Suppression
	label_fac_ar						Suppression
	label_fac_fr						Suppression
	LouM						Suppression
	Nature_diplome						Suppression
	specialite						Suppression
	specialite_ar						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Faculté		Faculte					
	Réf_Faculté		code_fac	varchar(20)			
	Faculté		designation_fac	varchar(50)			
	FacultéAr						Suppression
			code_univ	Varchar(20)	Universite	Code_univ	Ajoute

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Département		Departement					
	Réf_Département		code_dep	varchar(20)			Traitement
			code_dom	varchar(20)	Domaine	code_dom	Ajoute
	Désignation_Département		designation_dep	varchar(50)			Traitement
	Désignation_Département_Ar						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Séries		Bac_Series					
	N_Série		n_serie	varchar(20)			
	Série		designation_serie	varchar(50)			
	Série_Ar						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Option	Code_filière	Option	code_option	varchar(20)			Traitement
	Filière		designation_option	varchar(50)			Traitement
			code_cycle	varchar(20)	Cycle_Etude	code_cycle	Ajoute
			niveau	varchar(50)			Ajoute
			nb_etud	int(11)			
	Suppression du reste...						
			code_fil	varchar(20)	Filiere	code_fil	Ajoute

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
wilayas	Code_wilaya	Wilaya	code_wilayas	varchar(20)			Traitement
	wilaya		nom_wilayas	varchar(50)			
	Wilaya_Ar						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
edt		EDT					
	id		id	int(11)(ai)			
	annee		annee	varchar(20)			
	code_fil		code_fil	varchar(20)	filiere	code_fil	
	niveau		niveau	varchar(20)			
	Semestre		semestre	varchar(20)			
	acro_option		code_option	varchar(20)	option	code_option	
	groupe						Suppression

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
edt		Seance					Création
	j1s1...j5s5		code_seance	varchar(20)			
	j1s1...j5s5		code_module	varchar(20)	module	code_module	
	j1s1...j5s5		type_seance	varchar(20)			
	j1s1...j5s5		code_ens1	varchar(20)			

	j1s1...j5s5		code_ens2	varchar(20)			
	j1s1...j5s5		code_ens3	varchar(20)			
	j1s1...j5s5		code_ens4	varchar(20)			
	j1s1...j5s5		salle1	int(11)			
	j1s1...j5s5		salle2	int(11)			

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
enseignant		Enseignant					
	id		code_ens	int(11)			
	nom		nom	varchar(20)			
	prenom		prenom	varchar(20)			
	departement		Code_dep	Varchar(20)	Departement	Code_dep	
	situation_pro		situation_pro	varchar(20)			
	grade		grade	varchar(20)			

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
		Filiere					Création
			code_fil	varchar(20)			
			designation_fil	varchar(50)			
			code_dep	varchar(20)	Departement	code_dep	

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
		Domaine					Création
			code_dom	varchar(20)			
			designation_dom	varchar(50)			
			code_fac	varchar(20)	Faculte	code_fac	

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
		Universite					Création
			code_univ	varchar(20)			
			designation_univ	varchar(50)			

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Etudiant		Choix					Création
			id_choix	int(11)			Ajoute
	N_Carte		n_ins	int(11)	Etudiant	n_ins	
	NivEtud		niveau	varchar(50)	Niveau	nom_niv	
	Choix		choix1	varchar(50)			
	Choix		choix2	varchar(50)			
	Choix		choix3	varchar(50)			
	Choix		choix4	varchar(50)			
	Choix		choix5	varchar(50)			
	Choix		choix6	varchar(50)			
	Choix		choix7	varchar(50)			
	Choix		choix8	varchar(50)			
	Choix		choix9	varchar(50)			
	Choix		choix10	varchar(50)			
	ChoixFinal		choix_final	varchar(50)			

Table Original	Champ Original	Table Destination	Champ Destination	Type Champ	Références (Table)	Références (Champ)	Type Modification
Niveau	nom	Niveau	nom_niv	varchar(50)			
	filiere		code_dom	varchar(20)	Domaine	code_dom	

Résumé

Les établissements algériens, en général, et les universités en particulier, souffrent du manque de moyens et d'outils d'aide à la décision. Pour y apporter l'assistance nécessaire pour une gestion efficace des activités de scolarité et le recueil des informations fiables pour l'élaboration des différents rapports et statistiques au sein de l'université de Tlemcen, notre thèse a été proposée. L'objectif principal de ce projet a été la mise en place d'une solution d'informatique décisionnelle au sein de notre faculté ainsi que celle des technologies, en s'appuyant sur les différentes sources de données hétérogènes disponibles. Grâce à la solution « Business Intelligence » développée, les différents décideurs au sein de l'université, seront en mesure de prendre des décisions meilleures et intelligentes pour améliorer les activités pédagogiques.

Mot clés: Informatique décisionnelle, Entrepôt de données, Magasin de données, OLAP, Retro-Ingénierie.

Abstract

Generally, Algerian institutions and especially universities suffer of the lack of tools for decision-making. In order to provide the necessary assistance for an efficient management of the university activities, and collecting trusty information for making various reports and statistics within the University of Tlemcen, our thesis was proposed. the main aim of our project was implementing a business intelligence solution within both Sciences and Technologies faculties based on different available heterogeneous data sources. With the "Business Intelligence" solution developed, the decisions-makers will be able to make better and smarter decisions to improve teaching activities.

Keywords: Business Intelligence, Data Warehouse, Datamart, OLAP, Reverse engineering.

ملخص

المؤسسات الجزائرية عموما، والجامعات على وجه الخصوص تعاني من نقص كبير في أدوات اتخاذ القرارات. من أجل توفير دعم فعال لتسيير نشاطات الجامعة والحصول على معلومات موثوقة لاستخدامها في مختلف الإحصائيات والتقارير على مستوى جامعة تلمسان، قد تم اقتراح هذه الأطروحة. الهدف الرئيسي من هذا المشروع هو إنشاء نظام دعم واتخاذ القرارات لكل من كليتي العلوم والتكنولوجيا بالاعتماد على مختلف قواعد البيانات الغير متجانسة. يساعد و يُمكن نظامنا المنجز المسؤولين من اتخاذ قرارات أفضل وأذكى، وبالتالي تسيير أفضل للمنظومة الجامعية.

الكلمات المفتاحية: معلوماتية اتخاذ قرار، مستودع البيانات، مركز البيانات، أولاب، الهندسة العكسية.