Sujet

Mise en œuvre d'un système décisionnel pour les processus RH

Liste des abréviation

Abréviation	Désignation
BI	Business Intelligence
PowerBI	Power bi
DRH	Direction des Ressources humaines
ETL	Extraction Transformation Loading
HTML	HyperText Markup Language
OLAP	On Line Analytical Processing
PAQ	Plan Assurance Qualité
RH	Ressources Humaines
ROLAP	Relational On Line Analytical Processing
SCD	Slowly Changing Dimension
SGBD	Système de Gestion des Bases de Données
SIRH	Système d'Information Ressources Humaines
SQL	Structred Query Language

Table des figures

Figure 1: Schéma du cycle de vie dimensionnel	15
Figure 2: Diagramme GANT du projet	17
Figure 3: Flux d'alimentation de la dimension Employée	42
Figure 4: Mapping dimension Employée	43
Figure 5: TF_absentéisme	44
Figure 6: Lookup	45
Figure 7: Table de fait mapping	46
Figure 8: Connexion à la base de données	47
Figure 9: Vue de source de donnée	48
Figure 10: Creation d'un Cube	48
Figure 11: Déploiement et insertion des mesures souhaité	49
Figure 12: Connexion base de données (PowerBI)	50
Figure 13: Cube de données (PowerBI)	50
Figure 14: Création rapport Absentéisme par Recruteurs, Département	t et Année
	51
Figure 15: Création de connexion et commande pour le kmeans	51

Liste des tableaux

Tableau 1: Livrables du projet	18
Tableau 2: Indicateurs "effectifs"	23
Tableau 3: Indicateurs "flux"	24
Tableau 4: Indicateurs "absentéisme"	25
Tableau 5: Indicateurs "Finance"	26
Tableau 6: Indicateurs "plaintes et erreur de praoduction"	27
Tableau 7: Indicateurs "heures supplémentaires"	27
Tableau 8: Axes d'analyse	28
Tableau 9: Matrice d'expression des besoins	35
Tableau 10: Tableaux de bord "effectifs"	38
Tableau 11: Tableaux de bord "rémunération"	38
Tableau 12: Tableaux de bord "conditions de travail"	39

Table des matières

Introduction générale	12
Chapitre I : Contexte général du projet	12
Motivation et objectifs du projet	14
Conduite du projet	14
1.1. Démarche	14
1.2. Planning	16
1.3. Livrables	
1.4. Plan Assurance qualité	18
Conclusion	19
Chapitre II : Analyse et spécifications	20
2. Étude de l'existant	21
2.1. Description de l'existant	21
2.2. Critique de l'existant	21
3. Recueil des besoins	21
3.1. Besoins principaux	22
4. Cadrage fonctionnel du projet	22
4.1. Description des indicateurs	22
4.1.1. Indicateurs "emploi"	22
4.1.2. Indicateurs "Finance"	25
4.1.3. Indicateurs "emploi"	26
4.2. Description des axes d'analyse	27
5. Cadrage technique	28
5.1. Macro Architecture	28
5.2. Architecture applicative	29
Conclusion	30
Chapitre III : Conception	31
1. Généralités sur le Datawarehouse	32
2. Modélisation du datawarehouse	33
2.1. Matrice d'expression des besoins	33
2.2. Schéma dimensionnel	36
2.3. Modélisation des datamarts	36
3. Conception des tableaux de bord	
3.1. Tableaux de bord "Employées"	37
3.2. Tableaux de bord "Finance"	
3.3. Tableaux de bord "Employées (Production)"	39

Conclusion	39
Chapiter IV : Mise en Oeuvre	40
1. Réalisation de l'ETL	41
1.1. Alimentation des tables de dimension	41
1.2. Alimentation des tables de faits	43
2. Cubes OLAP	46
2.1. Création des cubes OLAP	46
3. Rapports	49
3.1. Création des rapports	
Conclusion	52
Conclusion générale	53
Bibliographies/Webographie	54
Index	

Introduction générale

Le Système d'Information Ressources Humaines (SIRH) est devenu un outil incontournable des professionnels de la gestion des ressources humaines. Avec un grand nombre de processus automatisés, le SIRH permet d'effectuer de très importants gains de productivité. Il est également une opportunité pour améliorer certains services que la DRH propose à ses clients internes et ainsi crédibiliser son apport de valeur ajoutée au sein de l'entreprise.

Aujourd'hui, les entreprises sont dans la recherche de l'efficience de leurs RH. Elles ont besoin de paramètres qui aident à prévenir les climats sociaux et à se prémunir contre d'éventuels dérapages. D'où la quête incessante des entreprises d'enrichir leurs SIRH de solutions de Business Intelligence (BI). En effet, mettre en place une plateforme décisionnelle RH c'est mettre à la disposition de ses gestionnaires les informations pertinentes et précises afin de leur permettre d'analyser toutes les composantes RH de leur organisation, de comprendre le présent et d'organiser efficacement l'avenir de leur capital humain. C'est dans ce cadre nous avons lancé notre projet de fin de module pour s'offrir un outil de pilotage et d'aide à la décision. Notre projet consiste donc en la mise en œuvre d'un système décisionnel pour la fonction RH. Notre champ d'implication couvre deux thèmes : emploi et finance. Ce mémoire présente les différentes étapes de réalisation de notre projet. Il se compose de quatre chapitres dont la description est comme suit. Le premier chapitre aborde le contexte général du projet. puis expose la motivation et les objectifs du projet. Ensuite, on passe à la démarche suivie, le planning et les livrables attendus avant de terminer avec le PlanAssuranceQualité.

Le deuxième chapitre est dédié à l'analyse et spécifications des besoins. Ce qui se traduit par une étude de l'existant et ses limites et le recensement des différents besoins de la direction des ressources humaines , pour arriver à un cadrage fonctionnel et technique de la solution proposée. Quant au troisième chapitre, il traite la conception du modèle multidimensionnel de notre Datawarehouse et des tableaux de bord sur lesquels se base le reporting. Le quatrième chapitre détaille la phase de la mise en œuvre du projet dans ses différentes étapes à savoir l'alimentation du Datawarehouse et la création des cubes OLAP et des rapports. Pour clôturé ce mémoire, nous résumons les résultats et acquis réalisés lors de ce projet, les difficultés rencontrées et des perspectives à ce travail.

Chapitre I: Contexte général du projet

Ce chapitre a comme objectif de décrire le contexte général du projet, qui consiste en la mise en œuvre d'un système décisionnel pour la fonction RH. Ainsi que nous présenterons ensuite notre projet et ses objectifs. Enfin, nous décrirons la démarche suivie, le planning et le plan assurance qualité adoptés, et les livrables attendus.

Motivation et objectifs du projet

Dans le but d'étendre et d'améliorer notre SIRH, on vise réaliser de nouvelles briques qui sont le module de l'Hygiène Sécurité Qualité et un module décisionnel pour les briques actuelles et celles en vue. Et la mise en œuvre du lot décisionnel pour les briques actuelles. Ce travail pourra servir aussi pour les briques décisionnelles futures. En effet vu la criticité et l'importance des RH, le besoin de ce lot décisionnel est devenu une nécessité à laquelle il faut donner de la priorité.

Le SIRH actuel ne fournit aucun outil pour la réalisation de tableaux de bord et la génération de rapports synthétisés pour les briques RH déjà couvertes. Notre projet vient pour répondre à ces besoins. Son but est de définir les indicateurs et de générer les tableaux de bord et les rapports qui offrent une meilleure visibilité sur la fonction RH et qui mettent à disposition l'information essentielle pour la prise de décision relativement aux thèmes emploi et Finance.

Les tableaux de bord à générer visent :

le suivi des du personnel afin de mesurer l'évolution des effectifs et de déduire le degré de satisfaction professionnelles des salariés;

le suivi des absences afin de les diminuer :

le suivi des salaire pour savoir combien consomme chaque service ou chaque opération

suivi d'exercice-compta depuis 2011;

le suivi des plainte et des erreur de production afin de les diminuer suivi des heures supplémentaire

Conduite du projet

1.1. Démarche

Notre projet étant à caractère décisionnel, pour sa conduite, nous allons suivre le cycle de vie dimensionnel. Ce cycle commence par la planification du projet puis passe à la définition des besoins qui constituent le point de départ de trois trajectoires parallèles qui sont la technologie, les données et l'interface utilisateur. Ces trajectoires se rejoignent en phase de déploiement qui est suivie par la phase de maintenance et croissance.

Le schéma ci-dessous représente la succession des tâches de haut niveau nécessaires à la conception, au développement et au déploiement d'entrepôts de données efficaces. [Kimball 2001]

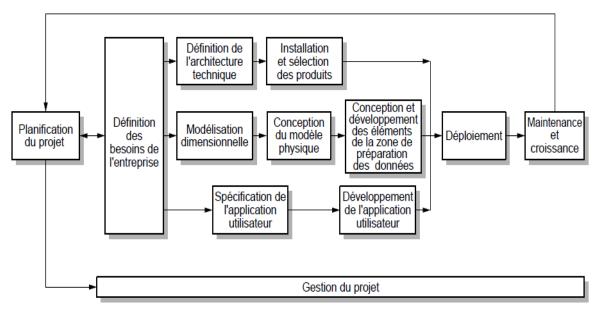


Figure 1: Schéma du cycle de vie dimensionnel

Les étapes de la figure ci-dessus sont :

Planification du projet : aborde la définition et l'étendue du projet. Elle permet l'affectation des tâches à leurs durées et à leur séquencement.

Définition des besoins : permet d'identifier les besoins de l'entreprise et détermine les données requises pour répondre aux besoins d'analyse des utilisateurs.

Modélisation dimensionnelle : commence par la construction d'une matrice qui représente les processus positions clé et leurs dimensionnalités. A partir de cette matrice, on élabore une analyse plus détaillée des données des systèmes sources, puis on développe le modèle dimensionnel.

Conception du modèle physique: définit les structures nécessaires pour l'implémentation du modèle dimensionnel. Cette étape nécessite la détermination des règles de nommage des objets et la mise en place de l'environnement de la base de données.

Conception et le développement de la zone de préparation des données : se déroule en trois phases majeures : Extraction, Transformation et Chargement.

Définition de l'architecture technique : permet d'avoir une vision globale de l'architecture technique à mettre en œuvre. Dans cette phase, trois facteurs doivent être pris en compte : les besoins, l'environnement existant et les orientations techniques stratégiques planifiées.

Sélection des produits et l'installation : permet la sélection des composants spécifiques tels que la plateforme matérielle, le SGBD (Système de Gestion de Bases de Données), et les outils de préparation et d'accès aux données. Une fois évalués et sélectionnés, ceux-ci devront être installés et testés minutieusement.

Spécifications de l'application utilisateur : décrit les maquettes d'états, les critères de sélection laissés à l'utilisateur et les calculs nécessaires.

Développement de l'application utilisateur : commence par une compréhension commune par l'équipe de développement et les utilisateurs finaux des spécifications établies.

Déploiement : est le point de convergence de la technologie, des données et des applications utilisateur. Une formation des utilisateurs est nécessaire. Elle intègre tous les aspects de cette convergence avant de permettre à ceux-ci d'accéder à l'application. La prise en compte des demandes d'évolution et de correction est également indispensable.

Maintenance et croissance: permet de s'occuper des utilisateurs en leur procurant un service de support et une formation continue.

Gestion du projet : garantit que les activités du cycle de vie dimensionnel restent sur la bonne voie et sont bien synchronisées.

1.2. Planning

La planification du projet découle du cycle de vie dimensionnel, décrit auparavant, afin de placer les éléments du projet dans le bon ordre au bon moment.

Le projet est découpé en phases chronologiques et organisationnelles, qui elles aussi sont découpées en tâches pour faciliter l'estimation du temps de réalisation du projet

La répartition des tâches à accomplir, leurs durées et leur succession sont décrites dans le diagramme de GANTT présenté ci-dessous.

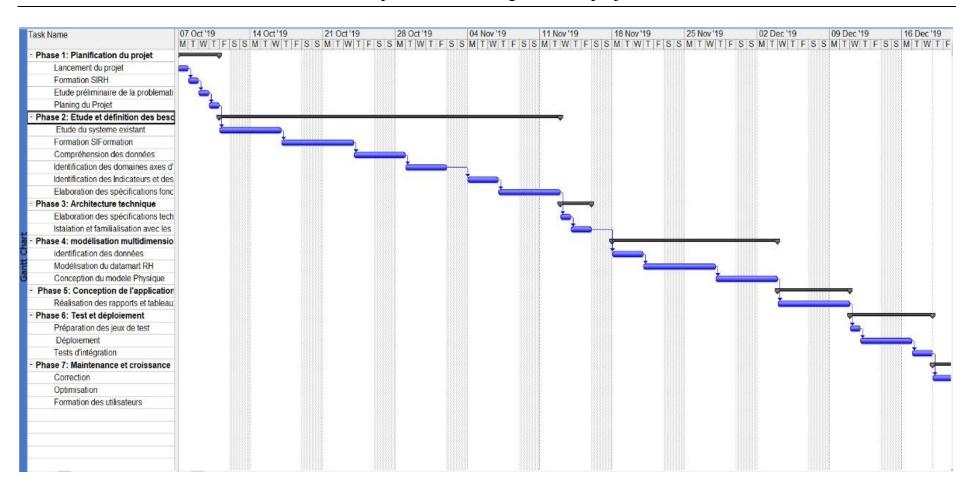


Figure 2: Diagramme GANT du projet

1.3. Livrables

Pour chaque étape du diagramme de GANTT, une remise de livrables est nécessaire, afin d'assurer la traçabilité du projet et de bien mener la communication avec les utilisateurs pour pouvoir valider chaque étape. Le tableau ci-dessous montre les livrables de chaque phase.

Phase	Livrables	
Planification du projet	Planning et méthodologie de conduite du	
	projet	
Etude et définitions des besoins	Cahier des charges	
	Dossiers des spécifications fonctionnelles	
	détaillées	
	PAQ (Plan Assurance Qualité)	
Architecture technique	Dossier des spécifications techniques	
	détaillées	
	PAQ mis à jour	
Modélisation multidimensionnelle	Dossier conceptuel	
	Matrice mesure/dimension	
	Modèle logique du Datawarehouse	
	Modèle physique du Datawarehouse	
	PAQ mis à jour	
Conception de l'application	Dossier de description des cubes OLAP	
utilisateur	Dossier de description des rapports	
Test et déploiement	Rapports et scénarii des tests	
	Manuel d'utilisation	
Maintenance et croissance	Manuel d'utilisation mis à jour	

Tableau 1: Livrables du projet

1.4. Plan Assurance qualité

Le PAQ définit les méthodes, l'organisation et les activités d'assurance et de contrôle de la qualité spécifique au projet.

L'utilisation du PAQ doit permettre d'atteindre les objectifs suivants :

- ✓ Constituer une référence commune à tous les intervenants du projet. Il permettra d'assurer une bonne cohérence et une homogénéité dans les méthodes de travail.
- ✓ Garantir la qualité du produit. Cette qualité s'exprime par des critères à respecter dans le cadre de ce projet.

✓ Définir les procédures à suivre, les outils à utiliser, les normes à respecter, la méthodologie de développement du produit et les contrôles prévues pour chaque activité.

Le PAQ est utilisé tout au long du projet. Il est donc destiné à être constamment ajusté en fonction des circonstances, selon une procédure d'actions correctrices définies dans ce document.

Ce document est élaboré dans la phase d'étude et recueil des besoins et il est mis à jour par la suite à la fin de chaque phase du projet.

Conclusion

Dans ce chapitre nous avons donné une présentation générale de notre projet de fin d'étude qui consiste en la mise en œuvre d'un système décisionnel concernant les fonctions RH. Il s'agit précisément de définir les indicateurs et de générer les tableaux de bord et les rapports qui offrent une meilleure visibilité sur la fonction RH et qui mettent à disposition de la DRH. l'information essentielle pour la prise de décision relativement aux thèmes emploi, et Finance. Vu le caractère décisionnel du projet, nous avons aussi expliqué dans ce chapitre la démarche à suivre qui respecte les différentes étapes du cycle dimensionnel, depuis l'analyse et spécifications des besoins jusqu'au déploiement en passant par l'alimentation du Datawarehouse. Enfin nous avons proposé un planning pour le déroulement de notre projet, ainsi qu'un plan assurance qualité.

Il convient maintenant d'attaquer le premier volet cité dans notre démarche à savoir l'analyse et spécifications des besoins ce qui est l'objet du chapitre suivant.

Chapitre II: Analyse et spécifications

Le but de ce chapitre est d'aboutir à un cadrage fonctionnel et technique de notre projet. Ainsi, nous allons d'abord commencer par une description de l'existant en termes de sources de données et de reporting utilisé. Une critique de cet existant permettra de situer les problématiques et limites relativement à ces deux derniers volets. Nous passerons ensuite au recueil des besoins afin de remédier à ces défaillances et de garantir la couverture de l'ensemble des processus positions inclus dans le périmètre de notre projet notamment en matière d'indicateurs et d'axes d'analyse de la solution décisionnelle envisagée. Nous terminerons ce chapitre en précisant une architecture applicative adéquate à notre solution.

2. Étude de l'existant

2.1. Description de l'existant

Dans cette section, nous décrivons l'existant en termes de sources de données ainsi que l'outil de reporting utilisé auparavant par la DRH.

✓ Sources de données :

L'étude des différentes sources de données est une phase très importante afin de choisir les tables qui contiennent les données nécessaires à la conception de notre système décisionnel.

Les sources de données sont :

- La base de données EFM_RH du SIRH qui contient 18 tables. Cette base est gérée par le SGBD SQL.
- Le grand livre de la paie qui est un fichier Excel généré, à la fin de chaque mois, à partir d'un logiciel permettant la gestion de la paie des salariés.

✓ Aucun reporting utilisé :

2.2. Critique de l'existant

La phase de l'étude de l'existent a permis d'examiner de près les données manipulées par la DRH, Ces données sont multiples, éparpillées et non consolidées. Elles présentent aussi une certaine incohérence due à la redondance de certains enregistrements. Par conséquent, on peut dire que la direction des ressources humaines ne dispose pas d'une base de données fiable, contenant toute l'information nécessaire à la production de rapports d'aide à la décision pour la DRH.

Notre travail vient pour remédier à ces problèmes en proposant une solution décisionnelle adéquate que nous détaillerons dans ce qui suit.

3. Recueil des besoins

Le but de la phase de recueil des besoins est de garantir la couverture fonctionnelle de l'ensemble des processus positions inclus dans le périmètre du projet. Il importe donc de faire un balayage global des besoins définis dans le cahier des charges. En effet, ces besoins ont été complétés et modifiés tout au long des réunions effectuées entre notre équipe

3.1. Besoins principaux

✓ Centralisation des sources de données dans un Data Warehouse :

Puisque les données sont stockées dans différents emplacements, leur consolidation dans un Data Warehouse s'avère primordiale. Ce qui permettra de répondre aux besoins de facilité et de rapidité d'accès aux informations. Ce Data Warehouse constituera une archive riche des données de la direction des ressources humaines.

✓ Génération des tableaux de bord selon les critères des utilisateurs :

Une fois les données stockées dans le Data Warehouse et les indicateurs sont chargés, les données seront ensuite publiées sous forme de tableaux de bord. Afin d'avoir des informations précises, les utilisateurs demandent la génération de rapports en précisant la distribution des indicateurs et des axes d'analyse sur les tableaux de bord.

√ Facilité d'accès aux données :

Les tableaux de bord seront générés dans le même serveur utilisé par le SIRH. Ainsi, les utilisateurs auront un environnement auquel ils sont habitués.

4. Cadrage fonctionnel du projet

En décisionnel, Le but du cadrage fonctionnel est de combler les besoins fonctionnels de notre projet en termes d'indicateurs et d'axes d'analyse. Il s'agit de décrire les indicateurs et axes d'analyse fixés par la DRH relativement aux thèmes emploi et Finance.

4.1. Description des indicateurs

Suite aux réunions entre notre equipe, nous avons pu fixer les principaux indicateurs ayant un enjeu relatif à la prise de décision.

4.1.1.Indicateurs "emploi"

Les indicateurs relatifs à l'emploi peuvent être divisés en cinq classes :

✓ Indicateurs "effectifs"

Ces indicateurs permettent de mesurer l'évolution des effectifs et d'avoir une vue globale sur les ressources. Le tableau 3 présente les différentes propriétés des trois indicateurs "effectifs".

	Effectif	% Effectif	Effectif moyen	Taux de
Indicateurs				racisme
Propriétés				
Définition	Le nombre de salariés de l'entreprise	Le pourcentage des salariés de l'entreprise selon l'axe d'analyse	La moyenne des effectifs de deux périodes (mois, année) successives	Le taux de racisme par rapport au position
Enjeu	Mesurer l'évolution de l'effectif de l'entreprise	Mesurer la répartition des salariés de l'entreprise selon chaque axe d'analyse	Mesurer l'évolution des effectifs moyens de l'entreprise	Mesurer si l'entreprise est raciste au niveau de métier est échelles
Périodicité	Annuel	<u> </u>		
Règle de gestion	Calculer le nombre des salariés	Diviser un effectif correspondant à un ou plusieurs axes d'analyse par l'effectif total	Diviser la somme des effectifs de deux périodes (mois, année) successives par deux	Calculer le nombre des gens ayant un couleur de peau par leur grade
Source	Base de données : EFM_RH			
	Tables : employée, Position, nationalité			
Axes	Temps, sexe, statut marital, tranche d'âge, position, RaceDesc			
d'analyse nationalité,Departement				

Tableau 2: Indicateurs "effectifs"

Le tableau 4 présente en détail les indicateurs des flux qui permettent de suivre l'évolution des embauches et des départs

[✓] Indicateurs "flux"

Indicateurs Propriétés	Nombre embauches	des	Nomb	re des départs	
Définition	Le nombre personnes embauchées		Le nombre des salariés qui ont quitté l'entreprise		
Enjeu	Suivre l'évolution nombres d'embauche		Mesurer le nombre des démissions, abandons de poste, licenciements et ruptures des périodes d'essai en vue de le diminuer		
Source	Base de données : EFM_RH Table : employée, nationalité, position, employé status				
Axes d'analyse	Temps, position, sexe, nationalité , opération	positi sexe,		Temps, position, sexe	

Tableau 3: Indicateurs "flux"

✓ Indicateurs "absentéisme"

Cette série d'indicateurs, concernant l'absentéisme, permet de le diminuer et d'en analyser les causes.

	Nombre des jours d'absence		Taux d'assiduité
Indicateurs Propriétés			
Définition		Le rapport du nombre de jours d'absence et du nombre de jours travaillés	d'assiduité du

Enjeu	Mesurer le	Evaluer	Mesurer
	nombre de jours	l'absentéisme du	l'assiduité du
	d'absence et en	personnel par	personnel
	analyser les	rapport au nombre	
	motifs	de jours travaillés	
		afin de le	
		diminuer	
Périodicité	Mensuel, annuel		*
Règle de gestion	Calculer la	Diviser le nombre	Diviser la
	somme des	de jours d'absence	différence du
	nombres de jours	par le nombre de	nombre de jours
	d'absence des	jours travaillés de	travaillés et le
	salariés	la période choisie	nombre de jours
			d'absence par le
			nombre de jours
			travaillés de la
			période choisie
Source	Base de données: E	FM_RH	
	Tables : absence, I	Département employ	ée, position
Axes d'analyse	Temps, âge sexe statut, position, salarié	Temps, position, sa	larié

Tableau 4: Indicateurs "absentéisme"

4.1.2.Indicateurs "Finance"

Ces indicateurs permettent d'évaluer et d'améliorer l'efficacité de la gestion de la rémunération.

Indicateurs	Mass salarial	Salaire	Exercice
Propriétés		moyenne	Comptable
Définition	L'ensemble de toutes les charges se rattachant à la	des rémunérations	le rendement par an

	rémunération		
	des salariés		
Enjeu	Mesurer l'évolution de la masse salariale pour concourir à sa	Evaluer la rémunération moyenne	
	maîtrise		
Périodicité	Annuel	d	ba
Règle de gestion	coûts totaux entreprise de	Diviser la somme des nets à payer par l'effectif de la période choisie	
Source	Fichier Excel: Le grand livre de la paie Base de données : EFM_RH		
	Tables : Fond, Département, Employée ,Position		
Axes d'analyse	Temps, position, sexe, Temps,		

Tableau 5: Indicateurs "Finance"

4.1.3.Indicateurs "emploi"

✓ Indicateurs "plainte et erreur de production"

Cette série d'indicateurs a comme objectif d'évaluer le nombre de plainte et le taux d'erreur quotidien. Ils sont décrits dans le tableau ci-dessous.

Indicateurs Propriétés	Nombre de plainte	Taux d'erreur quotidien			
Définition	Total des plaintes	Taux d'augmentation des erreurs quotidien			
Enjeu	Mesurer le nombre des	Mesurer le Taux des erreurs			

	plainte	fait par jour
Périodicité	Quotidien, 90-jours	
Source Base de données : EFM_RH		
	Tables: employée, fonction,	, Position
Axes d'analyse	Temps, position, Manager	

Tableau 6: Indicateurs "plaintes et erreur de praoduction"

Ces indicateurs permettent de connaître le coût des heures supplémentaires.

	Coût des heures supplémentaires			
Indicateurs				
Propriétés				
Définition	Montant d'indemnisation des heures supplémentaires			
Enjeu	Mesurer les coûts des heures supplémentaires			
Périodicité	Annuel			
Règle de gestion	Sommer les montants d'indemnisation des heures supplémentaires à 25%			
Source	Base de données : EFM_RH Tables : employée, fonction, position			
Axes d'analyse	Temps, position,			

Tableau 7: Indicateurs "heures supplémentaires"

4.2. Description des axes d'analyse

Chacun des indicateurs relevés est analysé selon un certain nombre d'axes. Le tableau ci-dessous définit l'ensemble de ces axes d'analyse en précisant pour chacun les différents niveaux d'hiérarchie.

Axe	Description	Niveaux
Temps	La dimension temps	Année
		Mois
		Jour

[✓] Indicateurs "heures supplémentaires"

Sexe	Le sexe d'un salarié Homme/Femme	Sexe
Statut marital	Célibataire, marié, divorcé	Statut marital
Tranche d'âge	Moins de 20 ans, DE 20 ans à 25ans, DE 25 ans à 30ans, DE 30 ans à 40ans, DE 40 ans et plus	Tranche d'âge
Position	Le position du salarié	Poste
Nationalité	La nationalité du salarié	Nationalité
fonds	Capital servant au financement.	fonds
Département	Division territoriale	Service Catégorie
Prime	Le type de la prime	Prime
Salarié	Le nom du salarié	Manager Salarié

Tableau 8: Axes d'analyse

5. Cadrage technique

5.1. Macro Architecture

Cette architecture respecte la décomposition habituelle de l'architecture constitutive d'un système décisionnel, qui comporte :

✓ la phase d'extraction et de transformation des données :

Périodiquement et sur la base des flux de données reçus, des extractions seront réalisées par l'ETL (Extract Transform Load). Ces données extraites subiront par la suite un ensemble de transformations (nettoyage, unification, mise en forme, etc.).

✓ la phase de chargement et stockage des données :

Après avoir effectué les traitements nécessaires, les données manipulées seront ensuite chargées grâce à l'ETL pour être stockées dans notre Data Warehouse.

✓ La phase de restitution et d'analyse :

C'est à ce point que sont calculés les indicateurs, fournissant ainsi des informations de valeur. Ces informations sont ensuite restituées sous forme de tableaux de bord.

5.2. Architecture applicative

✓ Alimentation

Pour la préparation de données, nous avons choisi le composant SSIS de l'outil MSBI.

En fait, il permet de créer graphiquement des processus de manipulation et de transformation de données puis de générer l'exécutable correspondant Périodiquement, des scripts ETL sont exécutés afin d'alimenter notre Data Warehouse à partir des sources de données.

✓ Stockage

Les données seront stockées dans le SGBD relationnel SQL selon le mode de stockage ROLAP. Le principe de ROLAP est de conserver toutes les données brutes dans une base de données relationnelles et de calculer les agrégats par des requêtes SQL. ROLAP permet le stockage d'un grand volume de données.

✓ Analyse

Une fois notre Data Warehouse alimenté par l'outil ETL, on peut créer des cubes OLAP qui permettent de faire des analyses multidimensionnelles et dynamiques. Pour la création des cubes, nous allons utiliser le composant SSAS de la suite MSBI aussi.

✓ Restitution

L'outil de restitution utilisé dans notre projet est le logiciel POWERBI Nous avons choisi cet outil car La création de rapports et tableaux de bord interactifs et très visuels sous powerbi. En effet, POWERBI permet La création de rapports à partir de différentes sources de données

Conclusion

Dans ce chapitre, nous avons pu voir que la direction des ressources humaines ne dispose pas d'une base de données fiable, contenant toute l'information nécessaire à la production de rapports d'aide à la décision. Les rapports actuels utilisés au sein de la DRH sous le tableur Excel, exigent une longue démarche très coûteuse en termes de temps et ne répondent pas aux besoins des utilisateurs en termes d'analyse et de pertinence des données.

Suite à plusieurs réunions, nous avons pu recueillir les besoins en termes de données et de reporting. Nous avons ensuite arrêté en particulier les indicateurs ayant un certain enjeu décisionnel relativement aux thèmes emploi, et finance, tout en précisant leurs axes d'analyse. Enfin, nous avons défini la macro architecture de notre solution empreinte du cycle dimensionnel et précisé les outils décisionnels adoptés.

Les tâches précisées précédemment constituent l'étape d'analyse et de collecte des besoins, étape cruciale pour le bon déroulement du projet. La définition précise, la collecte fidèle des données lors de ces tâches nous a facilités grandement la phase de conception présentée dans le chapitre suivant.

Chapitre III: Conception

Une fois les besoins définis et analysés, la phase de conception est l'étape suivante dans le cycle dimensionnel. Le but de cette phase est de modéliser les différentes couches du système décisionnel cible. Ainsi, dans ce chapitre, nous présenterons des généralités sur le Data Warehouse et la démarche suivie pour le concevoir. Nous détaillerons ensuite les différents Datamarts conçus. Nous y décrirons enfin la conception de nos tableaux de bord en spécifiant pour chacun les indicateurs et les axes d'analyse mis en jeu.

1. Généralités sur le Datawarehouse

Le Data Warehouse est une collection de données intégrées, orientées sujet, non volatiles, historisées, résumées et disponibles pour l'interrogation et l'analyse.

Il permet de stocker des données nécessaires à la prise de décision. Les données d'un Entrepôt de données respectent donc les caractéristiques suivantes:

- -Intégrées: Les données de l'entrepôt proviennent de différentes sources de données et peuvent exister sous différentes formes. Ainsi, il faut les intégrer afin de les homogénéiser.
- -Orientées sujet: Les données de l'entrepôt sont structurées par thèmes (les sujets majeurs de l'entreprise).
- -Historisées: Le Data Warehouse stocke l'historique des valeurs que la donnée aura prises au cours du temps.
- -Non volatiles: Les données de l'entrepôt sont très rarement modifiées, la non volatilité des données est en quelque sorte une conséquence de l'historisation.
- -**Résumées**: Les informations qui proviennent des sources de données doivent être agrégées (ou résumées) afin de faciliter le processus de prise de décision.
 - ✓ Approche de conception du Datawarehouse

Les deux approches les plus connues en conception de Data Warehouse sont :

- L'approche basée sur les besoins d'analyse
- L'approche basée sur les sources des données

L'approche la plus convenable à notre projet est celle basée sur les besoins d'analyse. Elle est aussi appelée « approche descendante » (*Top-Down Approach*) et est proposée par Ralph KIMBALL.

Pour Ralph KIMBALL, le data warehouse est constitué de l'agrégation des différents Datamarts. Les Datamarts sont les subdivisions logiques du Data Warehouse. KIMBALL met en avant le développement itératif et incrémental du Data Warehouse, chaque itération devant correspondre au traitement d'un sujet particulier et donc à la création d'un nouveau Datamart.

Les datamarts sont placés au centre de l'architecture. Le reste sera composé d'un *staging area* temporaire. Ce dernier est la zone où la plupart des opérations de nettoyage et de préparation des données ont lieu avant leur chargement dans le Datawarehouse. Le staging area est dit temporaire car les données sont détruites une fois le chargement des datamarts terminé.

2. Modélisation du datawarehouse

2.1. Matrice d'expression des besoins

A partir des besoins répertoriés dans le deuxième chapitre, nous avons lié chaque indicateur à un certain nombre d'axes d'analyse recensés. Le *tableau 10* présente la matrice des liaisons relevées. Cette matrice permet de faciliter la conception du Datawarehouse.

Chapitre III : Conception

Axes d'analyse	Sexe	Statut marital	Tranche d'âge	Position	Nationalité	Département	Employement source	Salarié	Fond	Temps
Indicateurs										
Effectif	X	X	X	X	X	X	X	X		X
% Effectif	X	X	X	X	X	X	X	X		X
Effectif moyen	X	X	X	X	X					X
Nombre d'embauches	X			X	X		X			X
Nombres de départs	X			X						X
Nombre de jours d'absence				X		X		X		X
Taux d'absentéisme				X				X		X
Taux d'assiduité				X				X		X
Masse salariale	X			X			X			X
Salaire moyenne	X			X			X			X

Projet de Fin d'Etudes ISMAGI 2019-2020

Chapitre III : Conception

Exercice comptable			X X
Coût heures supplémentaires	X	X	X
Coût heures supplémentaires à 100%	X	X	X
Plainte	X	X	X
Taux erreur quotidien	X	X	X

Tableau 9: Matrice d'expression des besoins

35

2.2. Schéma dimensionnel

Un datamart, ou magasin de données, permet de stocker toute l'information utile, relative à un thème précis. Avant d'être intégrée dans le magasin de données, l'information est tout d'abord extraite des bases sources et nettoyée. Puis, elle est mise en forme de manière à l'adapter aux besoins de l'utilisateur final. Cette mise en forme passe par la modélisation des données en schémas dimensionnels plus proches des utilisateurs finaux et adaptés à leurs besoins.

Pour cela, nous passons par une modélisation des données en un schéma dimensionnel, car elle est la plus appropriée aux requêtes et analyses effectuées sur les entrepôts de données. Elle est simple à créer, stable et intuitivement compréhensible par les utilisateurs finaux. Un schéma dimensionnel consiste en une ou plusieurs « tables de faits » qui contiennent les faits ou les mesures à calculer. Ces mesures serviront plus tard à calculer les indicateurs dont nous avons besoin. Ces tables de faits sont entourées par d'autres tables, appelées « dimensions », qui contiennent les éléments descriptifs des faits. En effet, le modèle ressemble à une étoile.

2.3. Modélisation des datamarts

Toutes les mesures de nos tables de faits sont calculées par rapport aux axes d'analyse recensés dans la phase d'étude des besoins et qui sont modélisés sous forme de tables de dimension. Pour chaque table de dimension, nous avons créé une clé primaire, appelée clé technique, différente de celle de la table source, appelée clé fonctionnelle. La clé technique est une séquence qui s'auto-incrémente et sert à indexer les enregistrements de la dimension et à améliorer la performance de la recherche. Nous avons toutefois gardé la clé fonctionnelle pour permettre les futures jointures.

On va présenter dans ce qui suit les datamarts conçus dans notre projet

3. Conception des tableaux de bord

Le modèle du datamart établi ne suffit pas pour clôturer la phase de conception et passer à la réalisation. En effet, il nous faut décider de la meilleure manière de présenter les indicateurs relevés à l'utilisateur final. Cette présentation doit se faire de la façon la plus simple et la plus expressive possible Ainsi, nous avons

conçu plusieurs tableaux de bord dont chacun a un objectif précis. Les données, relatives à un indicateur, sont affichées dans la plupart du temps sous forme de tableaux croisés, mais nous avons également introduit des graphiques.

Les tableaux ci-dessous résument la conception des tableaux de bord à réaliser. Pour chaque tableau de bord, nous avons indiqué l'indicateur et les axes d'analyse concernés. Nous avons également explicité l'objectif et les résultats attendus de chacun.

3.1. Tableaux de bord "Employées"

Le tableau ci-dessous présente le tableau de bord des effectifs.

N° du tableau de bord	Intitulé du tableau de bord	Indicateurs	Axes d'analyse	Objectifs et résultats attendus
1	Ventilation des effectifs	Effectif % Effectif Effectif moyen	Temps Sexe Statut marital Niveau d'études Tranche d'ancienneté Tranche d'âge Position Nationalité	 ✓ Connaitre le nombre et la répartition des salariés pour avoir une vue globale sur les ressources que la société possèdent ✓ Suivre le comportement des effectifs
2	Embauches	Nombre d'embauches	Temps Position Sexe Nationalité	✓ Suivre les embauches par position, sexe et nationalité
3	Nombre de départs	Nombre des départs	Temps Position Sexe Motif de sortie Tranche d'ancienneté	 ✓ Connaitre le nombre de départs subis et des départs voulus. ✓ Evaluer les taux de départs subis et voulus afin de les diminuer

4	Nombres des absences	Nombre des absences	Temps Décision absence Motif d'absence Position	✓ Connaitre le nombre de jours d'absence et en analyser les motifs
5	Taux d'absentéisme	Taux d'absentéisme	Temps Position	✓ Evaluer l'absentéisme du personnel par rapport au nombre de jours travaillés

Tableau 10: Tableaux de bord "effectifs"

3.2. Tableaux de bord "Finance"

Le tableau 5 liste les deux tableaux de bord de salaire et Exercice Comptable, en présentant les indicateurs traités par chacun et les axes d'analyses utilisés.

N° du tableau de bord	Intitulé du tableau de bord	Indicateurs	Axes d'analyse	Objectifs et résultats attendus
7	Evaluation de la rémunération moyenne et la masse salariale	Masse salariale Salaire moyenne	Temps Position Sexe Client	✓ Mesurer l'évolution de la masse salariale et la rémunération moyenne
8	Exercice Compta	Exercice Comptable	Temps Fond Departement	✓ Suivre l'exercice Comptable

Tableau 11: Tableaux de bord "rémunération"

3.3. Tableaux de bord "Employées (Production)"

Le tableau 6 présente les tableaux de bord traitant la partie Employée (Production)

N° du tableau de bord	Intitulé du tableau de bord	Indicateurs	Axes d'analyse	Objectifs et résultats attendus
10	Suivi des Plainte et Taux erreur	Plainte Taux erreur	Temps Position	✓ Suivre le total des plainte dans le département ✓ Connaitre le taux erreur dans le département production
11	Suivi des heures supplémentaires	Coût des heures supplémentaires	Client	✓ Analyser les coûts des heures supplémentaires

Tableau 12: Tableaux de bord "conditions de travail"

Conclusion

Grâce aux indicateurs et axes d'analyse identifiés dans le chapitre précédent, nous avons pu ressortir la matrice d'expression des besoins et modéliser par la suite notre datawarehouse selon la démarche descendante de Kimball. Nous avons détaillé les différents datamarts conçus dans notre projet : effectifs, embauches, départs, rémunération, absentéisme, heures supplémentaires et congés. Nous avons également abordé la conception de différents tableaux de bord (statiques et dynamiques) relativement aux thèmes emploi et Finance en précisant l'objectif, les indicateurs et les axes d'analyse pour chacun d'entre eux. Après cette phase de conception, nous passons à la phase de mise en œuvre, exposée en détail dans le chapitre suivant.

Chapitre IV : Mise en oeuvre

Dans ce chapitre, nous présentons la phase de la réalisation qui concrétise la conception exposée lors du chapitre précédant. Nous décrivons la mise en œuvre du datawarehouse, en explicitant les étapes suivies pour notre processus ETL. Nous illustrons également la réalisation de nos cubes OLAP et nos rapports.

1. Réalisation de l'ETL

Comme déjà expliqué, l'ETL est le processus qui nous permet d'extraire les données depuis les sources, les transformer et les charger dans le datawarehouse. Avant de s'y attaquer, nous avons créé la base de données sur le serveur de test dédié. L'ETL est un processus qui sera lancé plusieurs fois, de façon périodique, pour que les nouvelles données issues des sources soient incluses dans notre datawarehouse.

Pour notre réalisation, nous avons procédé par étapes. En effet, nous avons alimenté d'abord les tables de dimension, puis les tables intermédiaires et les tables de faits. Ces étapes sont détaillées dans ce qui suit.

1.1. Alimentation des tables de dimension

Il s'agit de remplir nos tables de dimensions à partir des sources de données. Dans cette étape, nous effectuons en même temps et les transformations nécessaires à nos dimensions et leur chargement. Pour le rafraichissement des données de ces tables, tout en insérant que les données non existantes dans les tables de dimensions.

Prenons l'exemple d'alimentation de la table de dimension Employée illustré dans la figure 15.

Chapitre IV : Réalisation et mise en œuvre

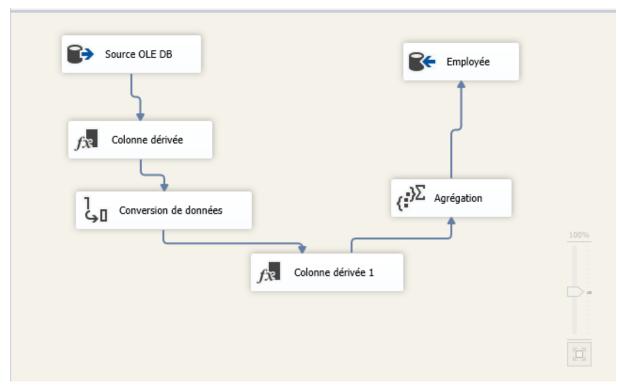


Figure 3: Flux d'alimentation de la dimension Employée

Pour alimenter la dimension Employée, on commence par récupérer lesdonnées employé à partir de l'ancienne base de données du SIRH. Pour ce faire, on utilise le composant Source OLE DB. Puis, on utilise le composant colonne derivée pour fusionner entre les colonnes nom et prenom après en a converti les donnée et en a fait une aggregation pour definir les mesures et pour lier les données en entrées avec les données que l'on veut avoir en sortie en utilisé destination OLE DB voila comme illustré dans la figure 16.

Chapitre IV : Réalisation et mise en œuvre

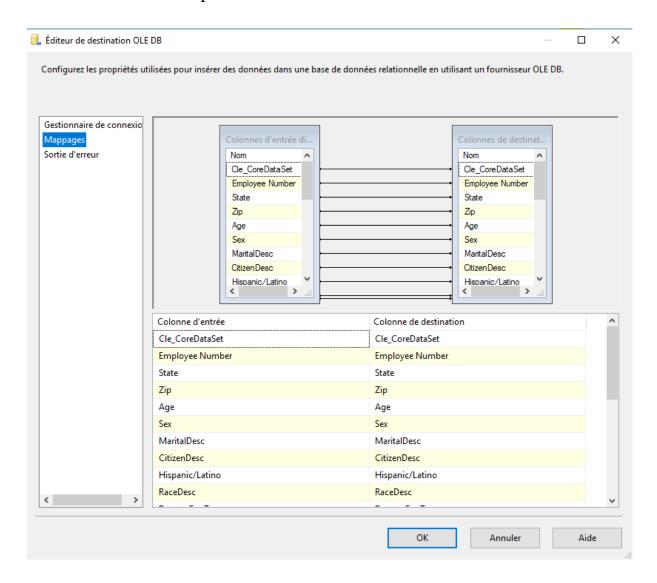


Figure 4: Mapping dimension Employée

1.2. Alimentation des tables de faits

L'alimentation des tables de faits commencent tous par « TF ». En outre, on a suivi la même démarche pour alimenter toutes les tables de faits.

Pour ce faire, on commence par alimenter les tables intermédiaires. Ces dernières servent à stocker toutes les données nécessaires à l'alimentation des tables de faits.

Prenons comme exemple la table de faits tf_absenteisme. La figure ci-dessous l'illustre.

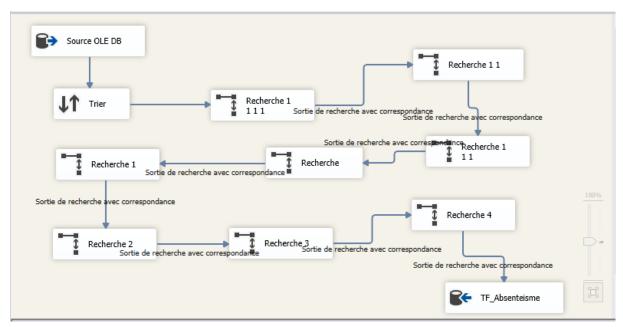


Figure 5: TF_absentéisme

Pour alimenter la table de fait absentéisme, on commence par récupérer les données employé à partir des dimensions liées. Pour ce faire, on utilise le composant Source OLE DB. Puis, on utilise le composant Recherche appelé aussi (lookup) pour récupère les mesure et les identificateurs comme illustré dans la figure 16.

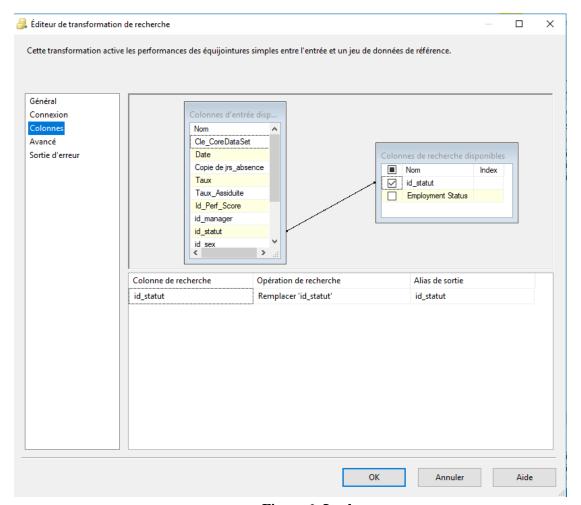


Figure 6: Lookup

On utilise ensuite Destination OLE DB pour lier les données en entrées avec les données que l'on veut avoir en sortie.

Chapitre IV : Réalisation et mise en œuvre

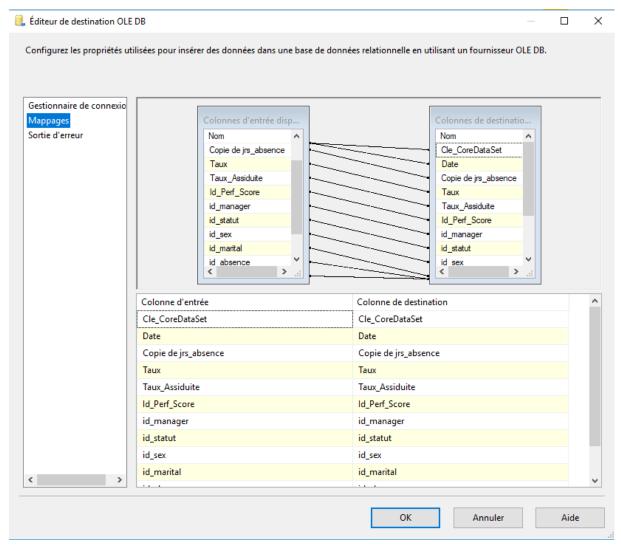


Figure 7: Table de fait mapping

2. Cubes OLAP

Dans SSAS, un cube OLAP est un rapport dynamique basé sur les croisements des axes. Il permet de naviguer au sein de grands volumes de données des plus agrégées au plus détaillées.

2.1. Création des cubes OLAP

Tout d'abord, nous commençons par configurer la connexion à la base de données.

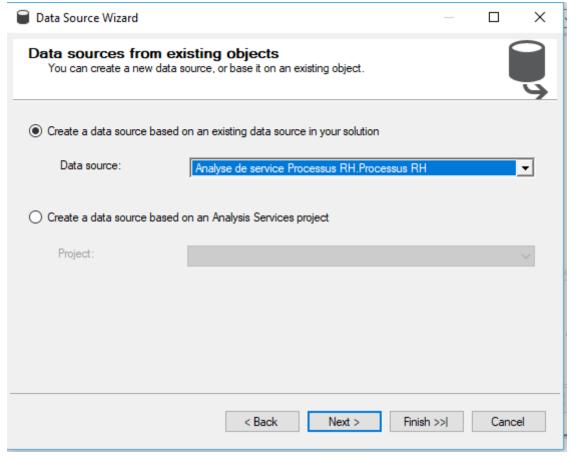
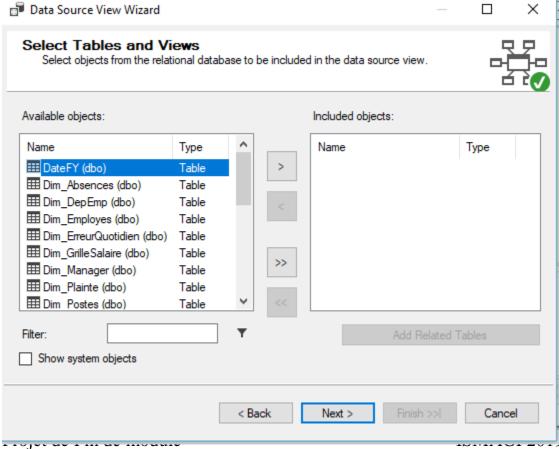


Figure 8: Connexion à la base de données

Ensuite, on crée un vue ou bien schema de source de donnée.



9-2020

Figure 9: Vue de source de donnée

Pour créer un cube, on doit d'abord indiquer sa table de faits et les dimensions partagées qu'il contient. On doit également créer ses propres dimensions, c.-à-d. les dimensions non existantes dans les autres cubes. Déployer et Enfin, il ne nous reste plus qu'à insérer la mesure que nous souhaitons calculer.

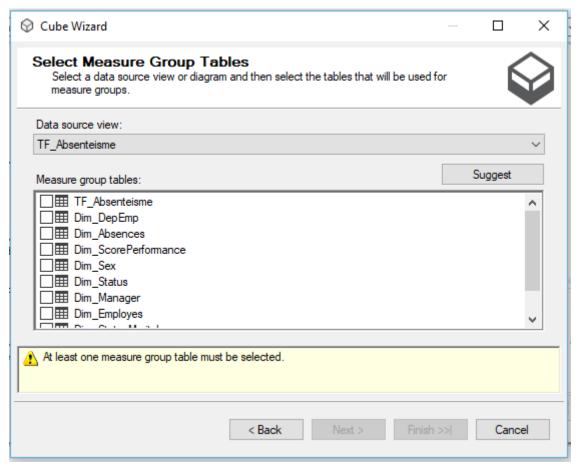


Figure 10: Creation d'un Cube

Chapitre IV : Réalisation et mise en œuvre

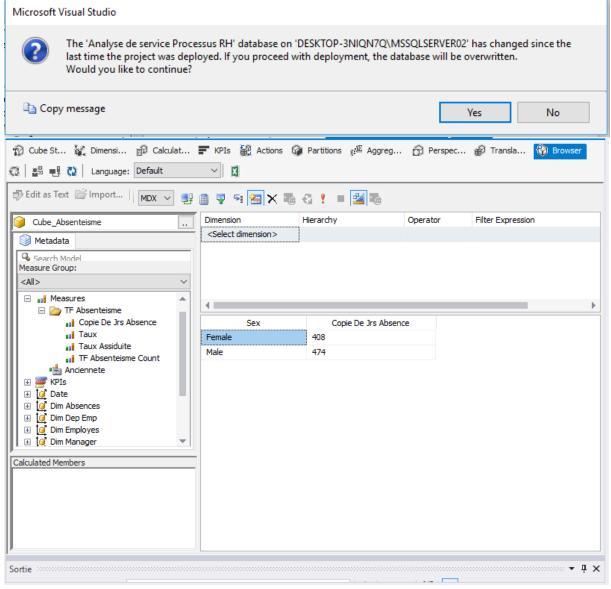


Figure 11: Déploiement et insertion des mesures souhaité

3. Rapports

3.1. Création des rapports

Les rapports statiques sont des rapports clairs et précis. Ils sont destinés à des utilisateurs qui n'ont pas besoin de faire des analyses en navigant dans le cube. La création de ces rapports est faite via l'outil POWERBI.

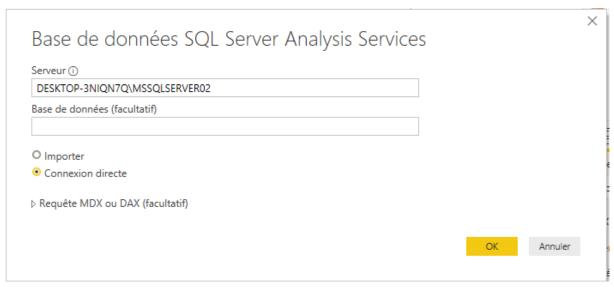


Figure 12: Connexion base de données (PowerBI)

Comme dans SSAS, on doit d'abord créer la connexion à la base de données. Après la connexion à la base données on doit choisir le projet SSAS qui contient nos cubes

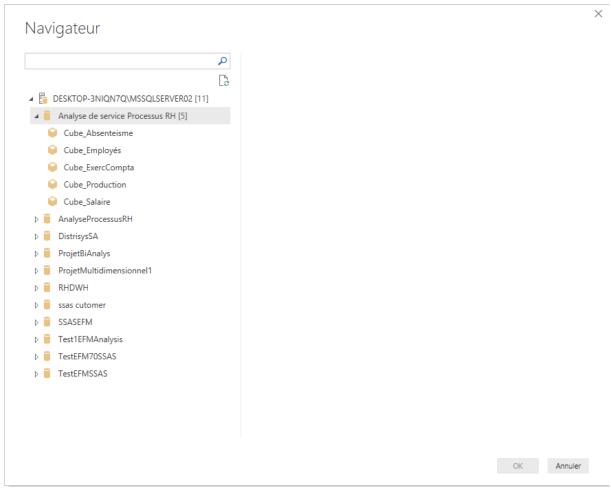


Figure 13: Cube de données (PowerBI)

Dès le choix de cube de données, on peut commencer notre rapport

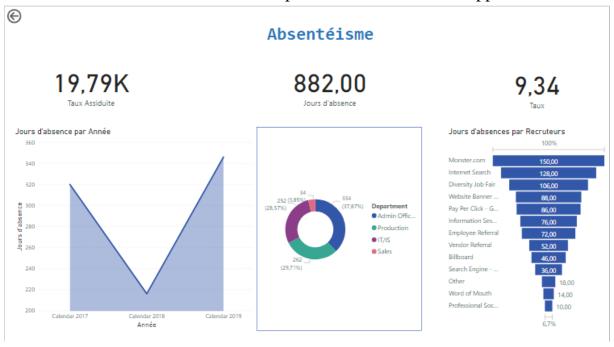


Figure 14: Création rapport Absentéisme par Recruteurs, Département et Année

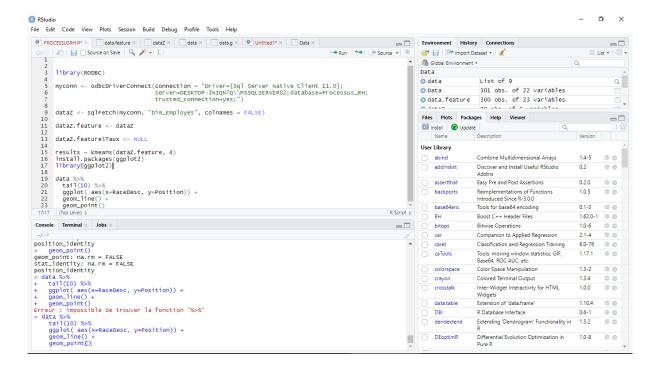


Figure 15: Création de connexion et commande pour le kmeans

Chapitre IV : Réalisation et mise en œuvre

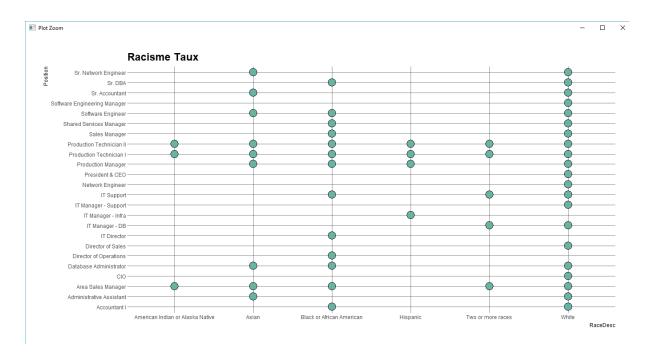


Figure36: affichage du suivi de taux de racisme

Conclusion

Ce chapitre a concerné l'aboutissement au résultat attendu de notre projet à savoir la création d'un système décisionnel pour la fonction RH. Il a détaillé d'abord le processus ETL que nous avons élaboré avec SSIS. Ce processus est une composante clé de notre plateforme vu qu'il assure l'alimentation du datawarehouse conçu et qu'il prévoie son évolutivité. Ce datawarehouse a été ensuite exploité pour la création et le déploiement de nos cubes OLAP et nos rapports statiques. Les cubes OLAP présentent un outil de pilotage et d'aide à la décision dans la mesure où ils permettent selon des techniques appropriées de navigation un filtre sur les données à générer, donc une information précise, et offrent des tableaux croisés dynamiques pour d'éventuels besoins avancés d'analyse. Les rapports statiques comportent des graphes et des tableaux croisés qui permettent à l'utilisateur une vue globale mais aussi synthétique des données. Dans ce cadre, Nous avons donné, dans ce chapitre, quelques exemples cubes OLAP des et des rapports générés.

Conclusion générale

Notre projet consistait à mettre en œuvre un système décisionnel pour la fonction RH.

Il s'agit précisément de générer des tableaux de bord qui offrent une meilleure visibilité sur la fonction RH et qui mettent à disposition de la DRH l'information essentielle pour la prise de décision relativement aux thèmes emploi et Finance. Vu le caractère décisionnel du projet, nous avons suivi les différentes étapes du cycle de vie dimensionnel. Après avoir étudié l'existant et recueilli les besoins des utilisateurs finaux, nous avons défini les indicateurs et les axes d'analyse et abouti à l'architecture technique de la solution. Grâce aux indicateurs et axes d'analyse identifiés, nous avons pu dresser une conception du modèle dimensionnel du datawarehouse et de la présentation des tableaux de bord. Nous nous sommes ensuite consacrés à la réalisation de la solution en alimentant le datawarehouse grâce à l'outil SSIS et en générant les états de restitution ciblés sous forme de rapports dynamiques sur la base de cubes OLAP avec SSAS ou de rapports statiques à l'aide de BIRT.

Au terme de ces étapes, nous pouvons affirmer que notre plateforme a satisfait les clauses du cahier de charges de notre projet et permet des possibilités intéressantes d'analyse et d'accès à l'information offertes aux utilisateurs finaux. Par ailleurs, ce projet m'a été bénéfique aussi bien au niveau informatique qu'au niveau professionnel. Il m'a permis de raffiner mes capacités d'abstraction et de conception en modélisation décisionnelle et offert l'occasion de découvrir et manipuler plusieurs outils Open Source de la BI et enfin de m'ouvrir plus sur le position RH.

En perspective, on peut envisager d'intégrer d'autres thèmes de la fonction RH dans notre système décisionnel.

Bibliographies/Webographie

Bibliographies

[INMO] W.H Inmon, Building the Data Warehouse, second edition, Wiley Computer Publishing USA, 1996

[KIMBALL 2001] R. Kimball, L. Reeves, M. Ross, W. Thornthwaito, Concevoir et déployer un Data Warehouse, édition Eyrolles, 2001

[KIMBALL 2002] R. Kimball, M. Ross - The Data Warehouse Toolkit, 2nd edition, 2002

[LBSMILE] Décisionnel Solutions open source, SMILE, 2010 [TOS] Guide référence de Talend open studio

Webographie

[CNAM] Le Datawarehouse et les systèmes multidimensionnels, http://ww1.cnam.fr/maths/IMG/pdf/wattiau 09-10 3.pdf, Mai 2011

[JNET] Journal du Net, SIRH: l'informatique au service des Ressources Humaines, http://www.journaldunet.com/solutions/intranet-extranet/sirh/, Avril 2011

[Smile] les composants décisionnels / Talend Open Studio, http://decisionnels-pen-source.smile.fr/Les-composants-decisionnels/Talen-Open-Studio, Juin 20

Index

Α

Alimentation

5, 14, 23, 32, 47, 48, 50,

52, 66, 71 73

technique

Architecture 20, 71

Axe

25, 26, 27, 28, 29, 30, 31, 34, 35, 36, 37, 38, 43, 44, 45, 54,

59, 61, 74, 75

В

Besoins

5, 18, 19, 20, 22, 23, 24, 26, 27, 34, 35, 36, 37, 38, 45, 57,

63, 70, 71, 74

Business Intelligence

6, 7,

13, 33, 74, 76

C

Cadrage

13, 24, 27, 32

Cahier

charges

22, 26, 74

Conception

14, 16, 17, 19, 20, 22, 25, 33, 34, 35,

36, 43, 45, 71

Conditions de travail

13, 19, 23, 27, 30, 34, 41,

45, 67, 81, 83

Cycle 6, 14

de

de

vie

D

Datamart

5, 6, 36, 38, 39, 40,

41, 45, 74, 75

Annexe B : Manuel de l'utilisateur

Datawarehouse

5, 6, 22, 23, 36, 45, 46, 47, 54,

70, 71, 72, 74

7, 22, 31, 33, 38, 47, 48, 49, 52, 56,

Dimension 57, 71, 74, 75

Е

Effectif

27, 28, 30, 37, 38, 43, 44, 45, 57, 60, 62,

63, 64, 82, 92

ETL

7, 32, 33, 46,

47, 70, 71, 75

Emploi

45, 55, 56, 71

5, 13, 19, 23, 27, 34, 38,

F

Fait

38, 41, 50, 52, 53,

56, 57, 74, 75

G

Gestion 13, 17

des

ressources

humaines

I

Indicateur

19, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 43,

44, 45, 74, 75

M

Maintenance

19, 20, 22

Modélisation

36, 38, 71, 75

20, 22,

O

OLAP

5, 6, 7, 14, 22, 33, 46, 54,

58, 70, 71, 75

Outil

5, 13, 18, 20, 23, 25, 32, 33, 34, 47, 54, 63,

70, 71, 73, 75

P

Plan Assurance Qualité

7,

13, 15, 22, 23

Planning

13, 15,

18, 20, 22, 23

Plateforme décisionnelle

5, 13,

20, 33, 65, 70

R

Règles

de

gestion

28, 29, 30, 31

Rémunération

5, 13, 19, 23, 27, 29, 34, 37, 41, 44,

45, 63, 64, 71

Reporting

33, 34, 71, 72

5, 6, 7, 14, 24, 25, 26,

S

SGBD

7,

20, 25, 33, 83

Sources de données

5, 24, 25,

26, 33, 36, 47 SQL

~ ~_

7, 33, 53, 72

T

Tableaux de bord

5, 43, 44, 45