

# Mémoire de fin d'études

---

**Master 2 Conduite de Projet Informatique**  
**2013/2014**



**adservio**

**Université Mouloud Mammeri Tizi-Ouzou**

**Tuteur enseignant : Mr Mohamed OUAMRANE**

**Enseignant-chercheur**

**Adservio**

**Maître de stage : Mr Ahcène OUGUENOUNE**

**Responsable de l'équipe Business Intelligence**

**Tinhinan SEDKAOUI**  
**Consultante Business Intelligence stagiaire**  
**Stage en entreprise**  
**Du 15/04/2014 au 15/09/2014**

# *Remerciement*

Je tiens à remercier dans un premier temps, toute l'équipe pédagogique de l'université Mouloud Mammeri de Tizi-Ouzou et les responsables de la formation Master 2 CPI, pour avoir assuré la partie théorique de celle-ci.

Je remercie particulièrement et je tiens à témoigner toute ma reconnaissance aux personnes suivantes :

Mon tuteur pédagogique Mr Mohamed OUAMRANE « Enseignant à l'université Mouloud Mammeri de Tizi-Ouzou » pour sa disponibilité, son écoute, son soutien et ses encouragements.

Mon tuteur de stage Mr Ahcène OUGUENOUE « Responsable de l'équipe Business Intelligence » pour sa confiance, la formation qu'il m'a prodiguée avec pédagogie, son aide et ses précieux conseils.

Toute l'équipe Adservio, pour l'expérience enrichissante et pleine d'intérêt qu'elle m'a fait vivre durant ce stage.

Je remercie tout particulièrement mes parents, ma sœur, mes frères et mes amis(es) pour leurs encouragements au quotidien.

Je vous remercie tous chaleureusement.

# Sommaire

---

RésuméP4

---

AbstractP7

---

Introduction P9

---

## Chapitre 1 : Présentation de l'entreprise Adservio

---

- I.1. IntroductionP11
- I.2. Présentation généraleP11
- I.3. Projets et clients d'AdservioP12
- I.4. Affiliations et partenariatsP13
- I.5. ServicesP14
- I.6. Organisation d'AdservioP16
- I.7. ConclusionP17

## Chapitre 2: Présentation du Business Intelligence

---

- II.1. IntroductionP19
  - II.2. Définition du BIP19
  - II.3. Data-WarehouseP23
  - II.4. DatamartP25
  - II.5. ETL/OutilsP25
  - II.6. Reporting/OutilsP26
  - II.7. Tableau de bordP27
  - II.8. ConclusionP28
-

## Chapitre 3: Travaux réalisés

---

III.1. IntroductionP30

III.2. Mon arrivé à « Adservio »P30

III.3. Formations effectuéesP31

III.4. Environnement de travail/Outils utilisésP32

III.5. Projet APMP Disneyland ParisP41

III.5.1. a/ Intégration des donnéesP54

III.5. 1. b/ Extraction des donnéesP64

III.5.2. Qualité de donnéesP67

III.5.3. QlikviewP72

III.6. ConclusionP76

## Chapitre 4. Les apports du stage

---

IV. Les apports pour l'entrepriseP78

IV. Les apports personnelsP78

ConclusionP80

---

BibliographieP82

---

# Résumé

## Résumé

---

Du 01 Mai au 15 Septembre 2014, j'ai effectué mon stage pour la validation de la formation de « Master 2 Conduite de Projet Informatique » au sein de l'entreprise « Adservio », basée à 12 Rue Portalis - 75008 Paris.

J'ai intégré le pôle « Expertise Décisionnelle » en tant que consultante Business Intelligence stagiaire, pour une formation dans le décisionnel ainsi que sur les outils qu'on utilise afin de mener à bien une démarche décisionnelle pour intégrer l'équipe réalisant les projets de cas réels.

Plus largement, ce stage était pour moi l'opportunité d'appréhender des solutions métier dans plusieurs secteurs, avoir un double enrichissement à la fois fonctionnel et technique en participant à la réalisation d'un projet de cas réel, ainsi que l'utilisation des nouvelles technologies décisionnelles tel que « Talend DI & DQ » pour l'Extraction la transformation, le chargement et l'analyse de qualité de données ainsi que « QlikView » pour le reporting. D'autre part ce stage m'a permis de mettre en œuvre mes connaissances acquises durant mon parcours universitaire.

J'ai eu également la chance de me voir confier un certain nombre de responsabilités et d'être associée à part entière dans des processus décisionnels, où j'ai eu l'occasion de contribuer à des projets réels en prenant en compte les exigences et les contraintes clients et où j'ai pu apprendre beaucoup sur les rapports humains et fonctionnels d'une entreprise.

Ce mémoire est organisé en quatre parties :

- Présentation d'Adservio, ses activités et son organisation
- Présentation du Business Intelligence
- Présentation des travaux réalisés : formations effectuées et contribution à la réalisation du projet APMP Disneyland Paris

- Les apports du stage

# Abstract

## Abstract

---

From 01 May to 15 September 2014, I did my internship for the validation of the formation of "Master 2 Driving IT Project Manager" in the business "Adservio" based at 12 Rue Portalis - 75008 Paris.

I joined the "BI Expertise" division as Business Intelligence consultant trainee for training in the decision as well as the tools used to carry out a decision-making to join the team carrying out projects actual cases.

More broadly, this course was for me the opportunity to understand business solutions in many sectors have a dual enrichment both functional and technical participation in the realization of a real project cases and the use of new decision-making technologies such as "Talend DI & DQ" for the extraction, processing, loading and analysis of quality data and "QlikView" for reporting. On the other hand this internship allowed me to implement my knowledge acquired during my academic career.

I have also been fortunate to have been given a number of responsibilities and be fully associated in the decision-making process, where I had the opportunity to contribute to real projects taking into account the requirements and customers constraints and where I could learn a lot about human and functional business reports.

This report is organized into four parts:

- Introducing Adservio, activities and organization
- Presentation of Business Intelligence
- Presentation of the work performed: training completed and contribution to the achievement of the APMP Disneyland Paris project



- The contributions of the course

# Introduction

## Introduction

---

La Business Intelligence (BI) est un sujet en pleine évolution, s'adressant à la direction générale tout comme aux métiers. Outil d'aide à la décision, la BI permet d'avoir une vue d'ensemble des différentes activités de l'entreprise, et de son environnement. Cette vue transversale nécessite de connaître les différents Métiers de l'entreprise et implique certaines spécificités organisationnelles et managériales.

L'organisation du BI dans l'entreprise est fortement dépendante de l'organisation de l'entreprise elle-même. Cependant, la BI peut avoir un impact structurant pour l'entreprise, notamment par la formalisation de référentiels de données et par la mise en place de centre de compétences. La BI a aussi un rôle fédérateur entre DSI, Métiers, DG, et renforce le rôle du DSI par rapport à la performance globale de l'entreprise.

La mise en place de projets BI ne peut se faire sans avoir défini préalablement une stratégie décisionnelle globale. Plus que pour des projets SI classiques, les projets BI nécessitent une grande maturité dans les relations DSI-Métiers, et s'inscrivent dans une démarche d'amélioration continue. De ce fait, les méthodologies utilisées se veulent agiles et itératives, pour coller au plus près de la demande client.

Aujourd'hui, le marché BI propose des solutions assez complètes concernant les aspects de reporting et de consolidation de données, tant du domaine propriétaire que de l'open source. Les évolutions possibles, à court-moyen terme, porteraient sur les outils d'analyse proactive et de simulation, ainsi que sur l'interactivité et la convivialité des accès sources de données internes et

externes. Mais le manque de visibilité des clients sur le marché BI (fusions-acquisitions, absence de roadmap, méconnaissance des stratégies des éditeurs, ...) crée une grande incertitude quant à la pérennité des offres actuelles et nuit grandement à la définition d'une stratégie décisionnelle à moyen-long terme.

# **Présentation de l'entreprise**

## I. Présentation de l'entreprise

---

### I.1. Introduction

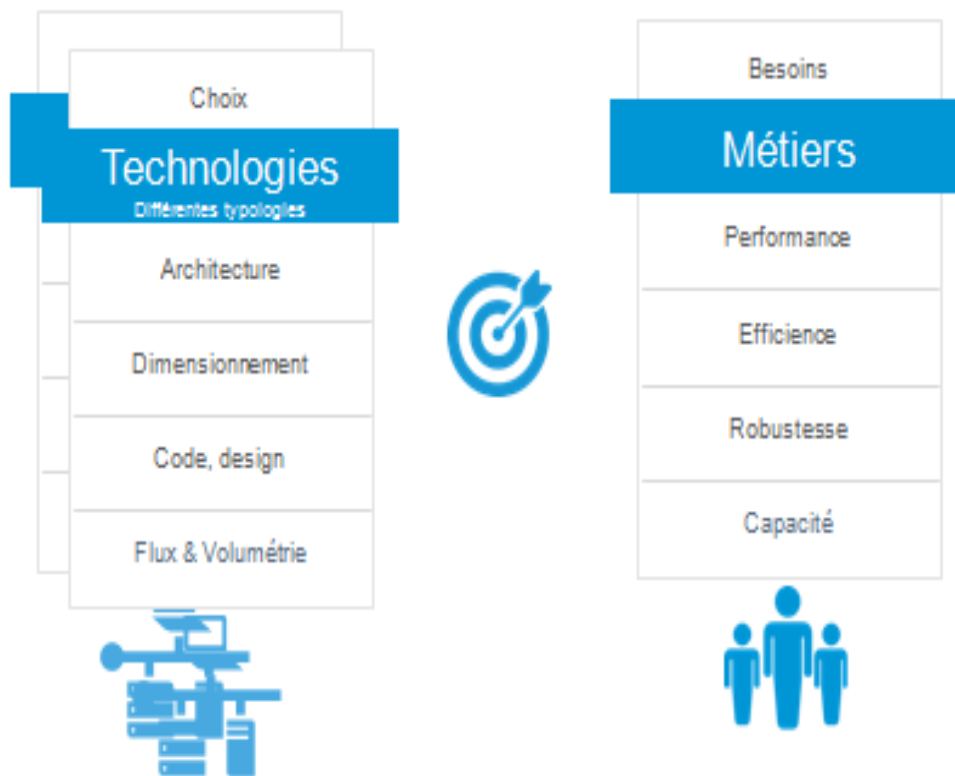
Le travail élaboré dans ce projet de fin d'études est réalisé pour « Adservio ». Il consiste à instaurer le programme de paiement mensuel des passeports annuels de « Disneyland Paris », un des clients de l'entreprise ainsi que la détection des périodes, régions et type de passeport les meilleurs afin de faire des promotions ... et d'apporter des solutions à la problématique afin de garder et d'attirer le plus de clientèle.

Dans ce 1<sup>er</sup> chapitre nous commencerons par la présentation de l'entreprise d'accueil ainsi que son organisation.

### I.2. Présentation générale [1]

« Adservio » est une société de conseil en systèmes et logiciels informatique dans l'optimisation de la performance des directions financières et opérationnelles.

La société « Adservio » a été fondée en 2010. Elle est constituée d'une trentaine de salariés. Le rôle des équipes d'Adservio est d'aligner les ressources du SI, les technologies IT par rapport aux besoins métiers non-fonctionnels. Elles accompagnent leurs clients vers l'excellence opérationnelle aux différentes étapes de leurs projets.



*Figure. Présentation générale d'Adservio*

Jusqu'à aujourd'hui, « Adservio » est implantée en deux pays, en France à 90% et en Allemagne à 10%.

### I.3. Clients et projets d'Adservio

Avec plus de 200 projets dans son actif, « Adservio » travaille pour ses différents clients de plusieurs domaines.

Les plus connus ; SFR, RATP, BNP PARIBAS, DISNEYLAND PARIS ...



*Figure. Clients d'Adservio*

#### I.4. Affiliations et partenariats

Les affiliations et partenariats d'Adservio sont comme suit :

### Affiliations



Figure. Affiliations d'Adservio

### Partenariats



Figure. Partenariats d'Adservio

## I.5. Services

« Adservio » est composée de trois pôles :

- I. *Pôle Expertise Technique et Nouvelles Technologies*
- II. *Pôle Expertise Décisionnelle*
- III. *Pôle Assurance de Qualité Technique*

Le pôle expertises techniques et nouvelles technologies concerne:

- L'expertise intégration et production
- L'expertise outils de profiling
- L'expertise Java/J2EE, PHP ...

Le pôle expertises assurance de qualité technique concerne :

- La qualité logicielle
- La qualité des données
- La qualité des processus



Le pôle expertises décisionnels concerne :

- L'expertise BI
- L'analyse, traitement, restauration des BDDs
- La migration des BDDs

Le schéma suivant énumère les activités de chaque pôle :

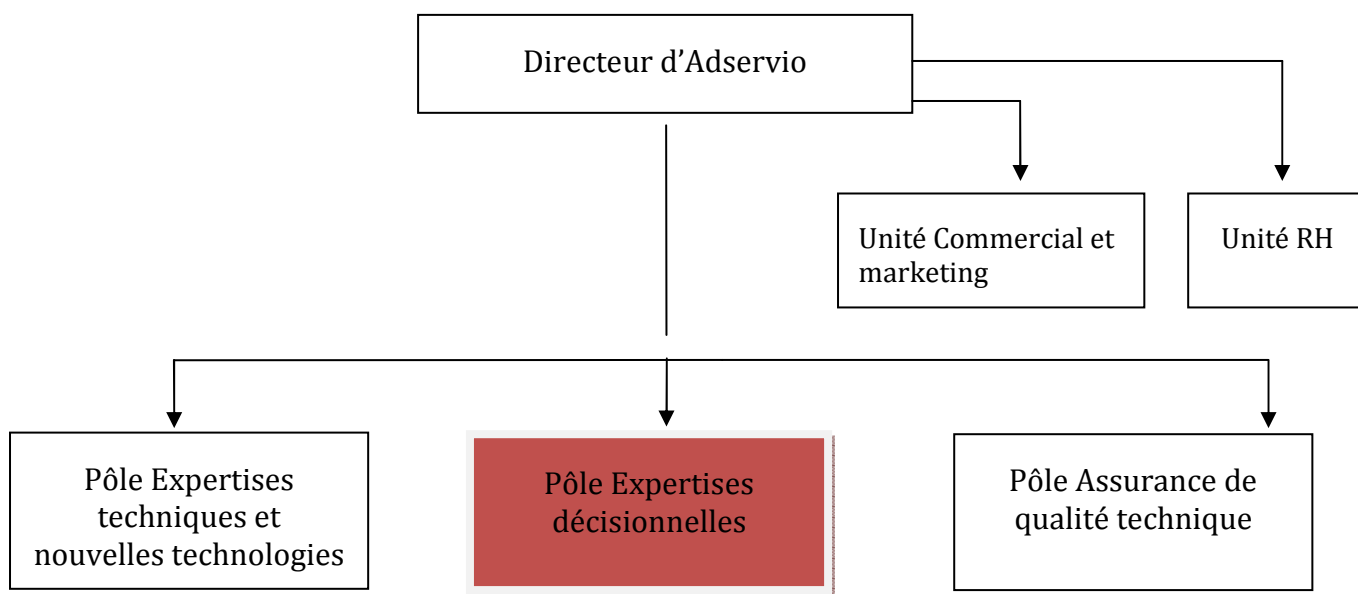


*Figure. Services d'Adservio*

## I.6. Organisation

« Adservio » s'organise de l'unité « Direction », des unités commercial et marketing ainsi que celle des ressources humaines, principalement de trois importants services ; Expertises techniques et nouvelles technologies, Expertises décisionnelles et enfin le pôle assurance de qualité.

Comme suit l'organigramme d'Adservio :



*Figure. Organigramme d'Adservio*

## I.7. Conclusion

Ce chapitre nous a permis de décrire l'organisme d'accueil « Adservio » et aussi d'avoir une idée générale sur ses activités ainsi que sur son organisation.

Le chapitre qui suit, est consacré à la définition du domaine du Business Intelligence, aux technologies et outils de ce dernier.

# **Business Intelligence**

## II. Business Intelligence

---

### II.1. Introduction

Les règles du jeu ont définitivement changées. L'époque des vieilles solutions, compliquées à utiliser, nécessitant des délais de déploiement longs et coûteux, touche à sa fin . Aujourd'hui, les défis macroéconomiques auxquels ils sont confrontés conduisent les clients à exiger des solutions qui peuvent être rentabilisées en quelques semaines, et non en plusieurs mois ou plusieurs années. Les organisations ont besoin d'obtenir des réponses immédiates à leurs questions les plus pressantes sur leur activité, telles que : comment puis-je réduire les coûts de manière sélective ? Optimiser la position de la trésorerie ? Ou répartir plus efficacement les fonds de roulement ? Elles ne peuvent plus se permettre d'attendre des solutions onéreuses, dépendant intégralement du service informatique, lentes ou incapables d'apporter des réponses dans un environnement économique en mutation rapide.

Dans la suite de ce chapitre, nous commencerons par une présentation des fondamentaux du BI, puis une présentation de l'environnement et les outils décisionnels utilisés lors du stage.

### II.2. Définition du Business Intelligence [2]

Aujourd'hui, dans un contexte où les sources d'information sont éclatées, volumineuses et complexes, il y a un réel besoin de consolider et d'analyser ces dernières pour pouvoir avoir une vision globale et optimiser le patrimoine informationnel de l'entreprise. Or, « trop d'information tue l'information » ... l'objectif du BI est de créer, à partir des données de l'entreprise mais aussi externe à celle-ci, l'information et le savoir aidant les membres de l'entreprise, des cadres dirigeants aux opérationnels, dans leur pilotage.

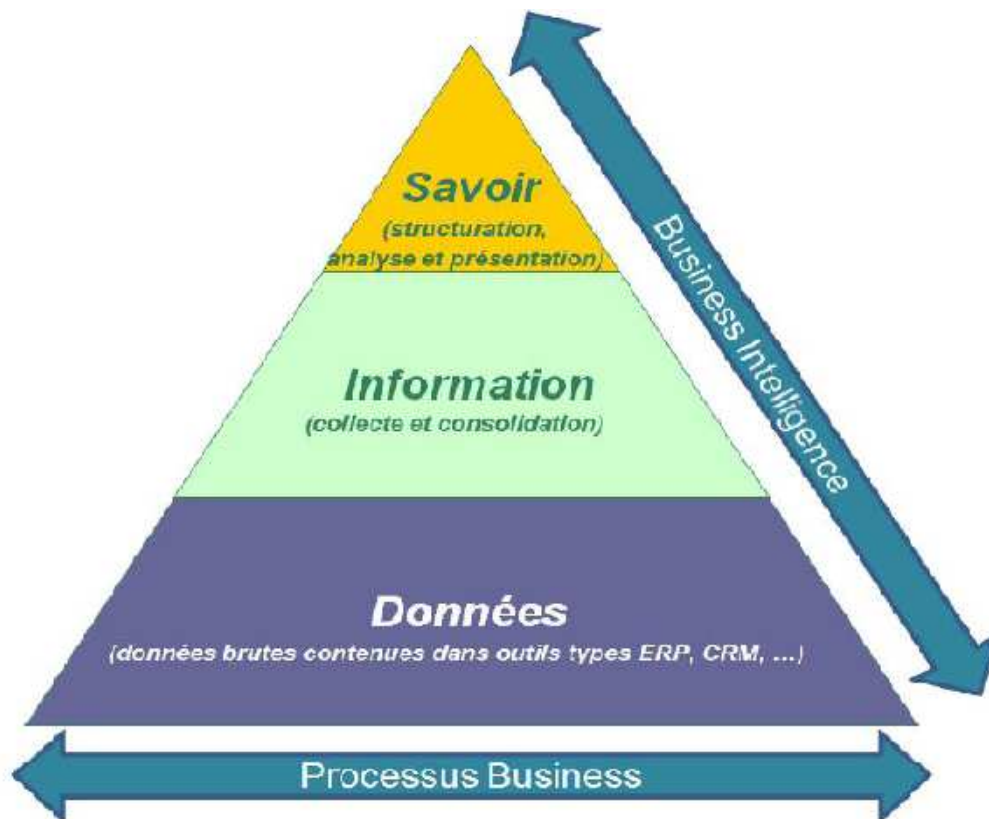
De récentes études montrent que la BI est l'une des préoccupations principales au sein des DSI de grandes entreprises. En effet dans le contexte actuel de crise et d'hyper-concurrence, la BI représente une opportunité pour les entreprises d'optimiser le pilotage de leurs activités, et d'anticiper sur les évolutions du marché, des comportements des clients / consommateurs, ... Les domaines d'utilisation de la BI touchent la plupart des Métiers de l'entreprise :

- Finance, avec les reporting financiers et budgétaires par exemple ;
- Vente et commercial, avec l'analyse des points de ventes, l'analyse de la rentabilité et de l'impact des promotions par exemple ;
- Marketing, avec la segmentation clients, les analyses comportementales par exemple ;
- Logistique, avec l'optimisation de la gestion des stocks, le suivi des livraisons par exemple ;
- Ressources humaines, avec l'optimisation de l'allocation des ressources par exemple ;
- Etc.

La BI est l'ensemble des moyens, outils et méthodes qui supportent le processus de collecte, consolidation, modélisation, analyse et restitution des informations.

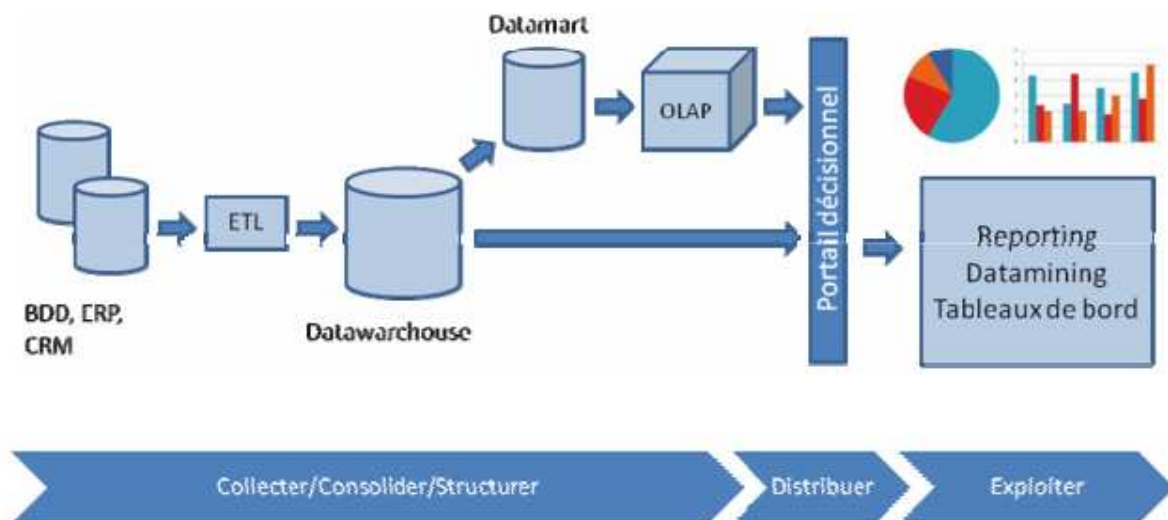
Le processus du BI se voit faire de la récupération des données brutes (contenues dans des outils type ERP, CRM, sources externes provenant des clients / fournisseurs, données de marchés, ...), à les transformer en information et à les diffuser sous forme de tableaux de bord de reporting.

Ci-après le processus du BI :



*Figure. Processus du BI*

Nous pouvons également modéliser le flux informationnel, c'est-à-dire le cheminement depuis la donnée brute provenant de SI sources (ERP, CRM ...), à la production de reportings et autres tableaux de bord de la manière suivante :



*Figure. Flux informationnel lié au processus BI*

Enfin, à ce stade, il est essentiel de définir certains termes, qui seront utilisés dans la suite de ce mémoire :

- **Datawarehouse** (ou entrepôt de données) : base de données utilisée pour collecter et stocker des informations volatiles provenant d'autres bases de données.
- **Datamart** : sous-ensemble logique d'un datawarehouse. Il est généralement exploité en entreprise pour restituer des informations ciblées sur un métier spécifique.
- **Datamining** : les outils de datamining permettent d'extraire des hypothèses à partir de grande quantités de données, par des procédés typiquement statistiques.
- **ETL** (Extract, Transform, Load) : il s'agit d'une technologie informatique intergicielle permettant d'effectuer des synchronisations massives d'informations d'une base de données vers une autre.
- **Cube OLAP** : représentation abstraite d'informations multidimensionnelles. Les données sont rangées selon un principe de dimensions correspondant étroitement aux axes de recherche des utilisateurs (par exemple les ventes de produits dans le temps par zone géographique).



## II.3. Data-Warehouse [3]

Un entrepôt de données se définit par l'ensemble de données historisées variant dans le temps, organisé par sujets, consolidé dans une base de données unique, géré dans un environnement de stockage particulier, aidant à la prise de décision dans l'entreprise.

On trouve trois fonctions essentielles dont, la collecte de données de bases existantes et chargement, la gestion des données dans l'entrepôt et l'analyse de données pour la prise de décision.

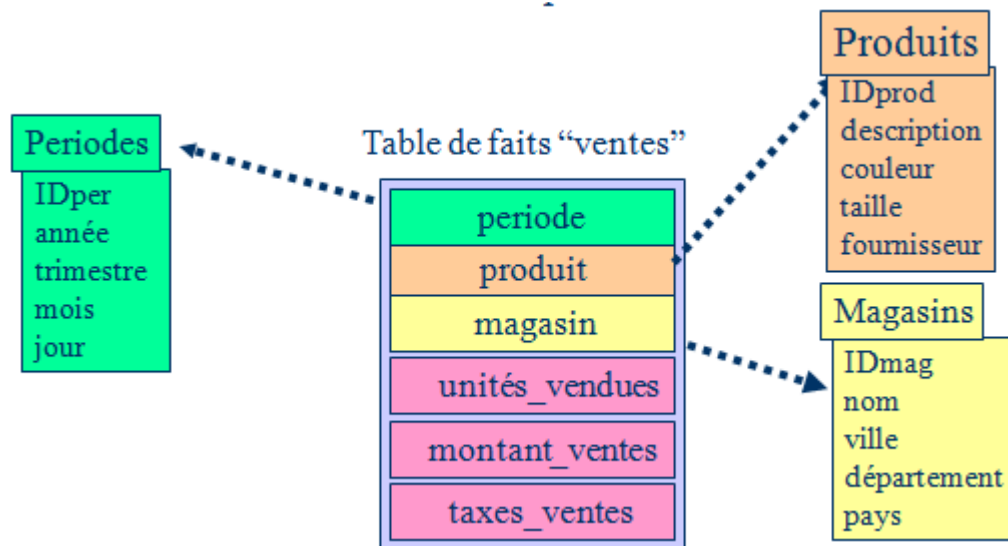
Pour la conception d'un data-warehouse, nous procédons comme suit

- ♦ Export de données des sources
  - Hétérogènes et variées
  - Fichiers, BD patrimoniales, Web, ...
  - Définition des vues exportées
- ♦ Définition d'un schéma global
  - Intègre les données utiles
  - S'appuie sur le modèle relationnel
- ♦ Nécessité d'une gestion de métadonnées
  - Description des sources
  - Description des vues exportées
  - Description du schéma global

Nous pouvons utiliser deux types de schéma, schéma en étoile et schéma en flocons.

Ci-dessous deux exemples des deux types de schéma :

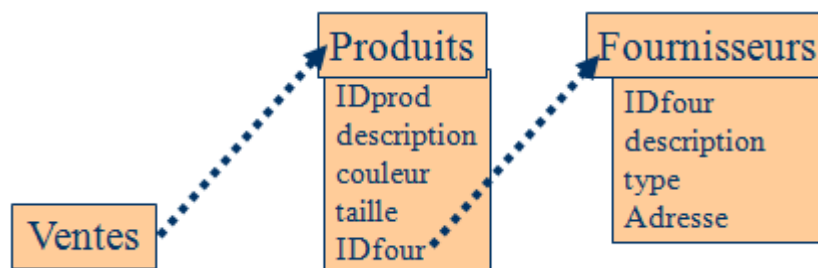
### Schéma en étoile:



*Figure. Schéma en étoile*

### Schéma en flocons :

Raffinement du schéma en étoile avec des tables normalisées par dimensions



*Figure. Schéma en flocons*

Ce schéma évite les redondances et conduit aux constellations (plusieurs tables de faits à dimensions partagées)

## II.4. Datamart

L'entreprise avec ses différents services auront besoin de données qui leurs sont propres. La création d'un datamart est pour regrouper au sein d'une base ces informations propres à un service ou plus généralement à une fonction. Pour toute fonction un datamart propre. Les bases pourront être traitées et préparées plus finement et de manière plus orientée client que sur la totalité du data-warehouse.

## II.5. ETL/ Outils [4]

Extract Transform Load (en anglais) est un terme employé dans le domaine du BI. Les ETL sont les processus de l'Extraction, Transformation et Chargement utilisés pour l'alimentation d'un data-warehouse ou d'un système décisionnel.

1. L'extraction de données de différentes sources de données
2. La transformation de ces données pour les réconcilier entre les différentes sources, pour effectuer des calculs ou du découpage de texte, pour les enrichir avec des données externes et aussi pour respecter le format requis par les systèmes cible.
3. Le chargement des données résultantes dans les différentes applications décisionnelles : data-warehouse, datamart ...

Plusieurs outils d'ETL existent de nos jours, chacun des outils a ses propres particularités mais tous servent à répondre à des besoins client en effectuant souvent la même démarche et pour un même but, qui est de faire de l'extraction, du traitement et du chargement de données.

Les plus connus : Informatica, DataStage, Talend ...



Plus loin dans ce rapport, nous verrons en détail l'outil utilisé dans le stage « Talend ».

## II.6. Reporting/ Outils [5]

« Reporting » désigne une famille d'outils du décisionnel conçus pour assurer la réalisation, la publication et la diffusion de rapports d'activité selon son format prédéfini. Principalement destinés à faciliter la communication de résultats chiffrés ou bien d'un suivi d'avancement.

L'outil de reporting assure l'interrogation des bases de données selon les requêtes SQL préparées lors de l'élaboration du modèle. Par la suite le rapport d'activité peut être publié sur l'Intranet. Il offre des fonctions spécifiques pour l'élaboration du modèle du rapport, des modules de calcul et de présentation (graphiques) afin de concevoir des comptes rendus.

Idem pour les outils de reporting, nous pouvons trouver plusieurs aujourd'hui dans le marché, tels que Qlikview, Microstrategy ...



Plus loin, nous détaillerons l'outil de reporting utilisé durant le stage « Qlikview ».

## II.7. Tableau de bord

Objectifs : piloter l'organisation à travers des indicateurs (regroupés dans un tableau de bord), et informer des résultats à la hiérarchie.

### Principes

#### 1/- Le tableau de bord

Le tableau de bord est un ensemble d'indicateurs de pilotage, construits de façon périodique, à l'intention du responsable, afin de guider ses décisions et ses actions en vue d'atteindre les objectifs de performance.

Le tableau de bord :

- permet le contrôle de gestion en mettant en évidence les performances réelles et potentielles et les dysfonctionnements ;
- est un support de communication entre responsables ;
- favorise la prise de décision, après analyse des valeurs remarquables, et la mise en œuvre des actions correctives ;
- peut être un instrument de veille permettant de déceler les opportunités et risques nouveaux.

La structure et le contenu du tableau de bord dépendent :

- de l'entreprise, son activité, sa taille ;
- de l'environnement : marché, concurrence, conjoncture ;
- du responsable auquel il s'adresse : son champ de responsabilité et d'intervention.

#### 2/- Le reporting

Le reporting est un ensemble d'indicateurs de résultat, construit a posteriori, de façon périodique, afin d'informer la hiérarchie des performances de l'unité.

Dans le cadre d'une gestion décentralisée, le reporting permet de vérifier que les centres de responsabilité respectent leurs engagements contractuels.

## II.8. Conclusion

Dans ce chapitre nous avons pu faire la connaissance du domaine du Business Intelligence, son utilité dans les entreprises ainsi que les outils nous permettant de réaliser une démarche BI. Nous poursuivons avec le chapitre suivant mettant en avant le projet réalisé durant le stage.

# Travaux réalisés

### III. Travaux réalisés

---

#### III.1. Introduction

Pour la réalisation de ce projet, il a été indispensable d'adapter une démarche méthodologique et rigoureuse. Dans notre cas, nous étions un groupe de 4 personnes ; un chef de projet, deux développeurs ainsi que moi-même (stagiaire). Nous étions amenés par exemple à faire des réunions afin de faire des points sur l'avancement du projet, répartition des tâches ... etc.

L'utilisation des outils décisionnels (outil d'ETL et outil de reporting) dans les entreprises aujourd'hui est indispensable pour le gain du temps et d'argents. Ma présence dans le groupe était comme un renforcement pour l'équipe mais aussi une opportunité pour moi de faire des formations.

Cela dit, l'objectif de ce chapitre est d'étudier la problématique du projet APMP (Annual Passport Menthly Payment) Disneyland Paris, ainsi que de présenter les outils utilisés, à savoir « Talend » et « Qlikview ».

#### III.2. Mon arrivé à « Adservio »

Lors de mon arrivé chez « Adservio », j'ai intégré le pôle Expertise décisionnelle comme consultante Business Intelligence stagiaire sous la tutelle du chef du Pôle et directeur associé.

Au sein de l'équipe, j'ai fait des formations en présentation générale du domaine du Business Intelligence et de l'outil d'ETL « Talend », à savoir « Talend Data Integration » et « Talend Data Quality » ainsi que l'outil de reporting « QlikView ». Après la première formation j'ai pu intégrer l'équipe et participer aux développements du projet APMP (Annual Passport Menthly Payment) d'un



des clients d'Adservio, Disneyland Paris, sur l'ensemble de la chaîne décisionnelle sous forme de liste :

- Récupération des données, transformation et alimentation.
- A partir d'une base de données, extraction des données en un seul fichier XML.
- Analyse des données de la base « Galaxy » ainsi que la création de rapport d'analyse.
- Contribution à la conception du tableau de bord en utilisant reporting « QlikView ».

### III.3. Formations effectuées

Avant de commencer ce travail, j'ai fait une formation de trois semaines en interne sur l'outil « Talend Data Integration », comprendre comment l'utiliser pour faire de l'Extraction, de la Transformation et du Chargement des données. D'autres séances de transfert de connaissances par des collègues et particulièrement mon chef de pôle se sont faites au long de mon stage à « Adservio ». La suite a été consacrée à un travail d'intégration de données du projet APMP.

Dans la deuxième période de mon stage, j'ai pu intervenir en poursuivant avec l'extraction de données du projet APMP de Disneyland Paris, j'ai par la suite effectuée une autre formation de deux semaines en « Talend Data Quality » pour réaliser des analyses sur les données Disney, créer des rapports d'analyse ... etc.

Dans la dernière partie de mon stage, j'ai effectué une dernière formation en « Qlikview » pour participer à la conception du tableau de bord de reporting, à savoir des graphes (histogrammes ainsi que des secteurs).

#### III4. Environnement de travail/outils utilisés

Afin de mener à bien ce travail de fin d'études, nous étions amené à utiliser des outils, tels que Talend ainsi que Qlikview, ci-après la définition de ces derniers.

« Talend » est le premier outil utilisé dans le projet.

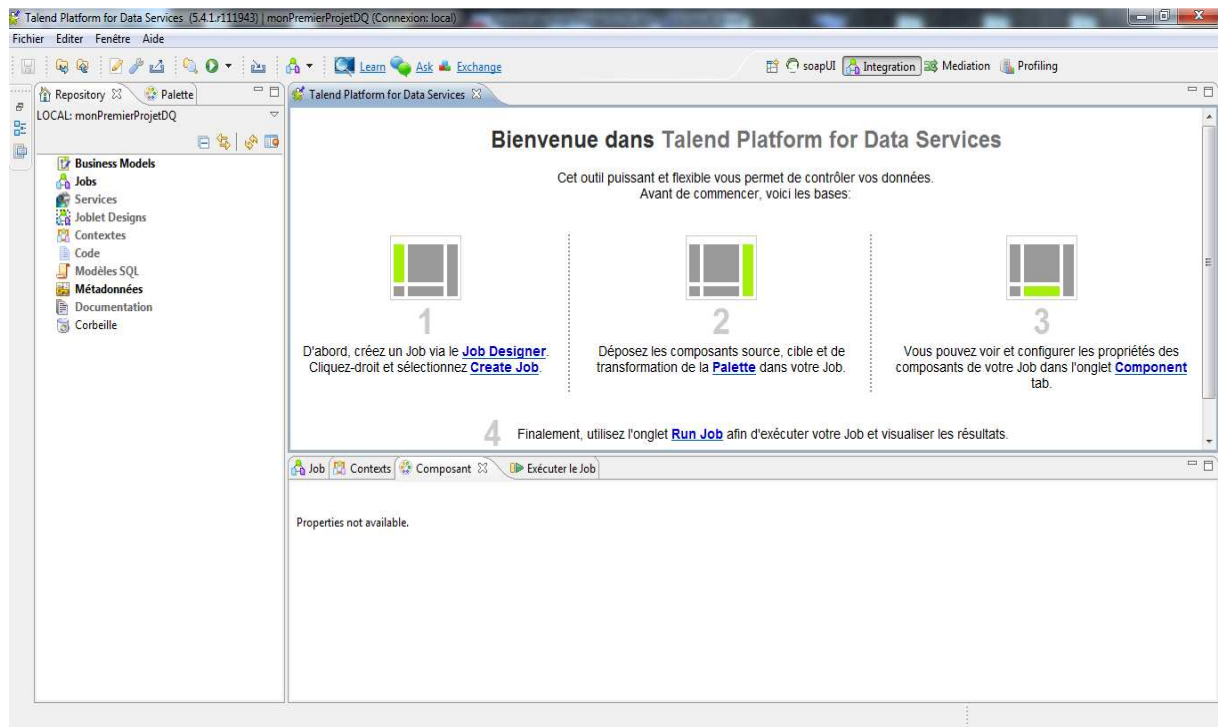
L'outil « Talend » est un ETL open source, développé par la société Française « Talend ». Il permet de créer graphiquement des processus de manipulation et de transformation de données puis de générer l'exécutable correspondant sous forme de programme Java.

« Talend » est notamment l'ETL utilisé ou embarqué par les solutions décisionnelles SpagoBI et JasperSoft. Mais Talend peut être utilisé très largement au-delà des projets décisionnels, dans la synchronisation de données entre applicatifs d'un système d'information.

« Talend » utilise un modèle graphique à base de glisser/déposer pour créer les processus de manipulation de données. De nombreux types d'étapes sont disponibles pour se connecter aux principales bases de données ainsi que pour traiter différents types de fichiers (CSV, Excel, XML). On notera que Talend facilite la construction des requêtes dans les bases de données en détectant les relations entre tables.

Il est également possible d'ajouter simplement de nouvelles fonctions et composants afin de réaliser des processus plus complexes.

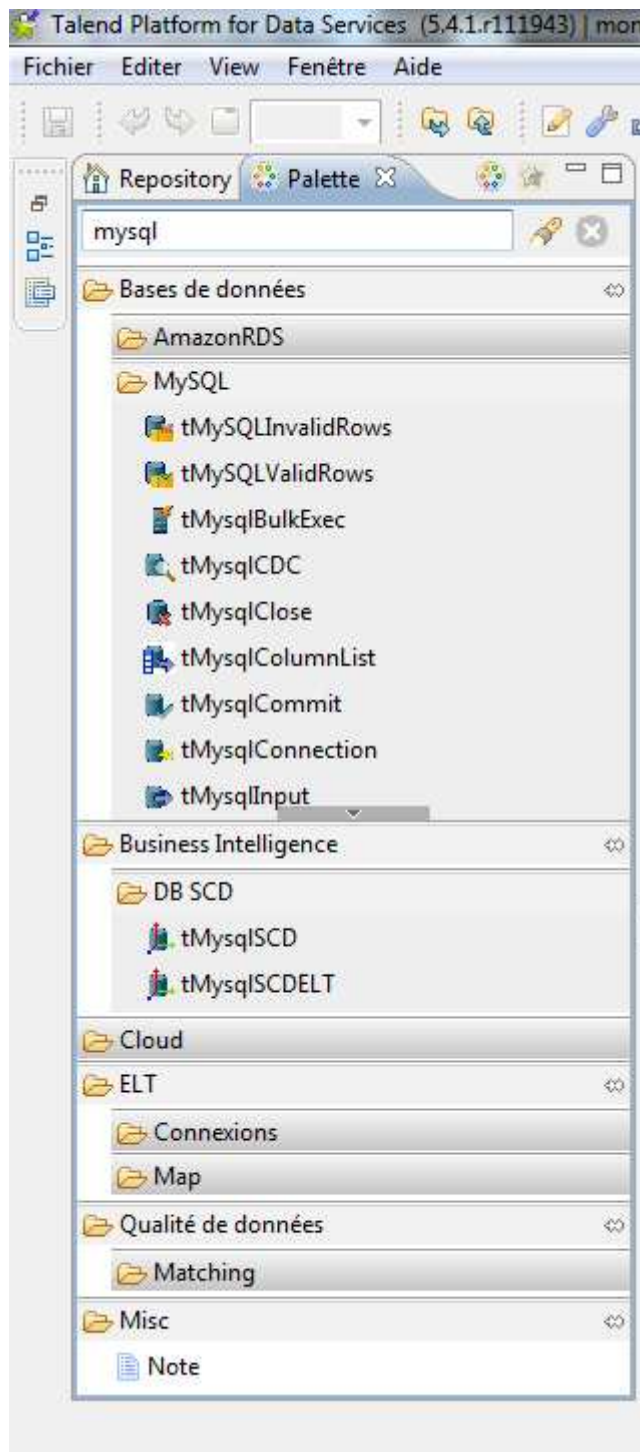
En voici une capture d'écran de la page d'accueil de l'environnement « Talend »



*Figure. Page d'accueil « Talend »*

A gauche de cette capture le répertoire Talend (on peut trouver les connexions aux données, les jobs réalisés, les fichiers stockés ... etc)

La capture d'écran suivante est une capture de la palette contenant les composants que Talend propose afin de réaliser des jobs Talend :



*Figure. Composants « Talend »*

```

1497         row1.DOB = rs_tMysqlInput_1
1498             .getTimestamp(column_index_tMysqlInput_1);
1499     } else {
1500         row1.DOB = (java.util.Date) year0_tMysqlInput_1
1501             .clone();
1502     }
1503 } else {
1504     row1.DOB = null;
1505 }
1506
1507 }
1508
1509 column_index_tMysqlInput_1 = 10;
1510
1511 if (colQtyInRs_tMysqlInput_1 < column_index_tMysqlInput_1) {
1512     row1.Cust_type = null;
1513 } else {
1514

```

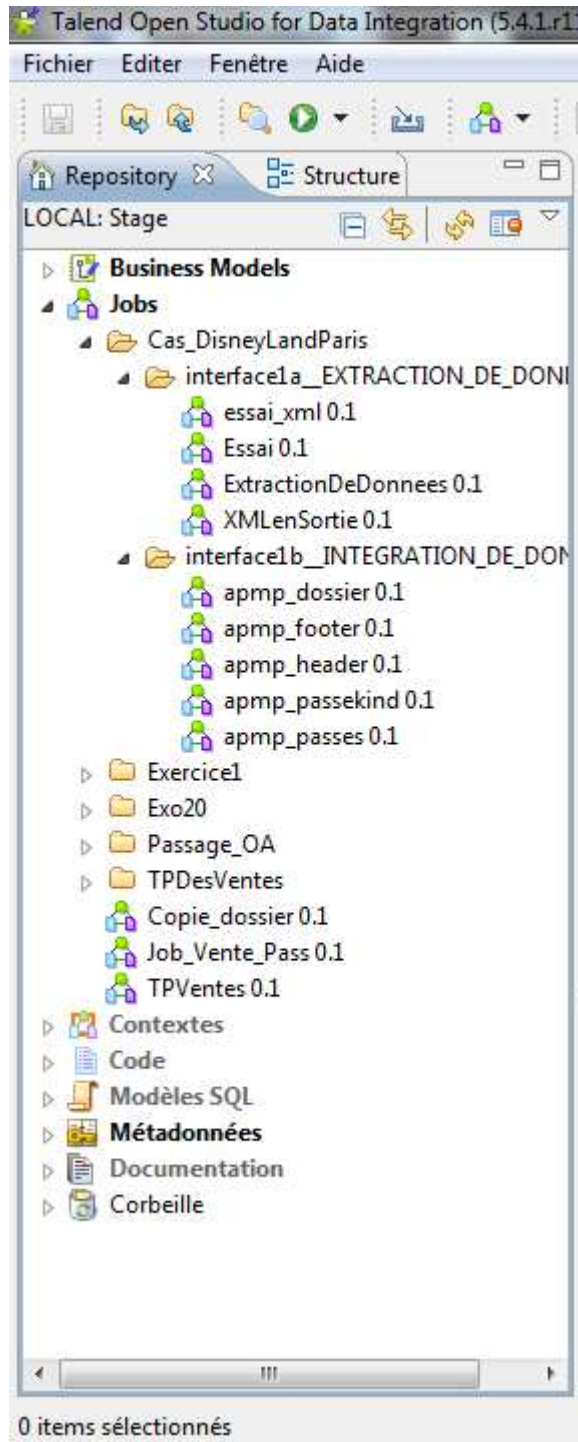
Designer Code JavaScript

Job(crm\_customer\_parsing 0.1) Contexts(Job crm\_customer\_parsing 0.1) Composant Exécuter (Job crm\_customer\_parsing 0.1)

Properties not available.

---

En capture d'écran une liste des jobs réalisée afin de faire le travail d'intégration et d'extraction de données du projet APMP :



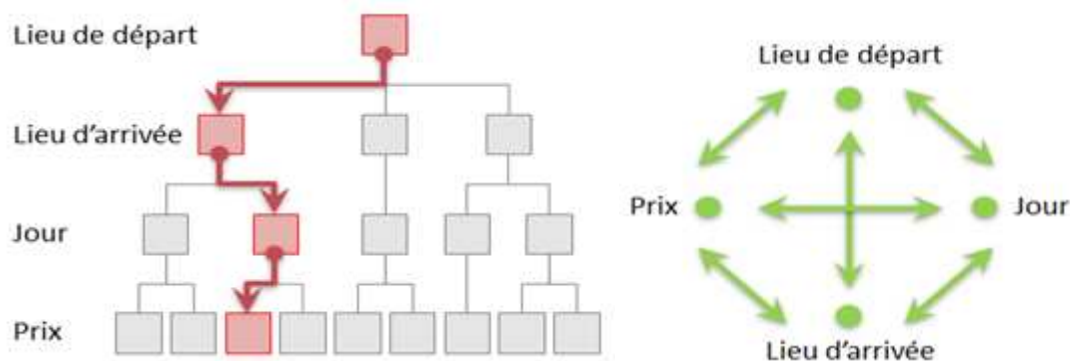
*Figure. Jobs « Talend »*

L'outil de reporting « QlikView », est un outil décisionnel dont les particularités principales sont :

- La facilité d'utilisation.
- Le suivi et l'analyse des indicateurs clés.
- La performance des sélections.
- La capacité de gestion d'importantes volumétries

Contrairement aux outils décisionnels classiques, QlikView n'utilise pas de requêtes prédéfinies ou de base multidimensionnelle mais repose sur une méthode dite de « recherche associative ». Cœur de l'application, la base de données associative est la grande force de QlikView. En effet, grâce à ce système, QlikView met en évidence les relations entre les données sans qu'il soit nécessaire de les définir au préalable. Il devient alors possible d'effectuer une recherche dans les données à partir de n'importe quel point de départ.

Recherche traditionnelle vs recherche associative :



*Figure. Différence entre une recherche traditionnelle et une recherche associative*

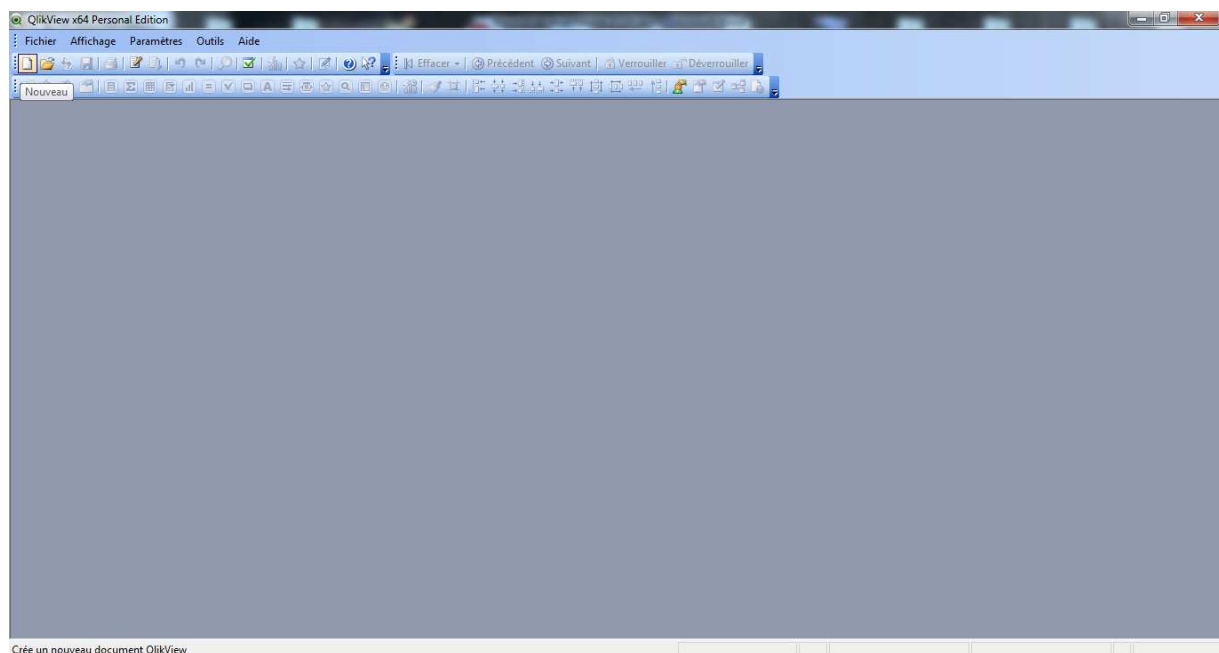
Une application QlikView prend la forme d'un seul fichier dans lequel sont stockées toutes les données et tous les paramètres. Ce fichier est composé:

1. **D'un script:** Extraction, Transformation et Chargement des données
2. **D'une base de données:** Base de données Associative (puissance fonctionnelle) reposant sur une technologie vectorielle (compression des données)
3. **D'une interface graphique utilisateur**



*Figure. Structure d'un fichier Qlikview*

En capture d'écran, la page d'accueil « Qlikview » :



*Figure. Page d'accueil « Qlikview »*



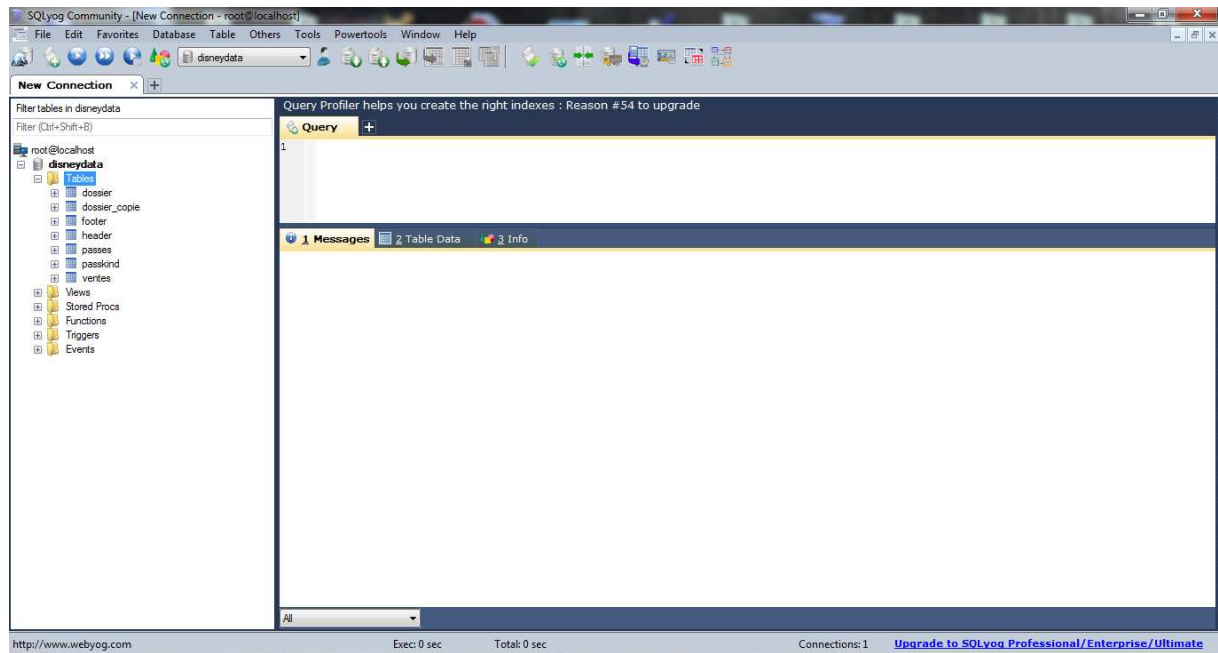
Afin d'alimenter notre base de données ou d'extraire des données depuis la base de données, nous avons eu à utiliser Mysql.

« Mysql » est un serveur de bases de données relationnelles Open Source.

Un serveur de bases de données stocke les données dans des tables séparées plutôt que de tout rassembler dans une seule table. Cela améliore la rapidité et la souplesse de l'ensemble. Les tables sont reliées par des relations définies, qui rendent possible la combinaison de données entre plusieurs tables durant une requête. Le SQL dans "MySQL" signifie "Structured Query Language" : le langage standard pour les traitements de bases de données.

L'utilisation basic passe par l'exploitation de la ligne de commande, cependant pour les projets complexes et pour plus de productivité et de visibilité, des outils graphiques payant ou en open peuvent être utilisés.

En capture d'écran notre base de données « Disneydata » :



*Figure. Base de données Mysql*

### III.5. Projet APMP, Disneyland Paris

#### Présentation du projet « APMP »

#### Le passeport annuel

Le passeport annuel permet à son détenteur d'avoir accès aux deux parcs de Disneyland Paris de 300 à 365 jours par an (en fonction du type de passeport). Il permet également de bénéficier de différents avantages (réduction sur des billets d'entrée aux parcs, dans les hôtels Disney, les boutiques et les restaurants du resort...).

Il existe actuellement trois types de passeports :



##### **Le passeport annuel Francilien**

Il permet un accès aux deux parcs 300 jours par an

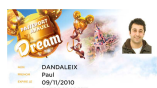
Il offre également à son détenteur la possibilité d'obtenir différents avantages :

- Des réductions sur des billets parc pour les personnes accompagnant le porteur de passeport
- Des réductions pour des soirées spéciales (type Halloween) ou des shows (Buffalo Bill Wild West Show) pour le détenteur
- Des offres spéciales dans les hôtels Disney pour le détenteur...



##### **Le passeport annuel Fantasy**

Ce passeport permet un accès aux deux parcs 335 jours par an et offre globalement le même type de réductions que le Francilien. Il n'est pas limité à la clientèle Francilienne



##### **Le passeport annuel Dream**

Il permet un accès aux parcs tous les jours de l'année. Passeport « premium », il permet d'obtenir des avantages supplémentaires auxquels n'ont pas droit les détenteurs d'autres types de passeport (avec des niveaux de réduction parfois plus importants)

	Fantasia : 99€	Fantasy : 124€	Omnia : 129€
Accès aux 2 Parcs Disney	300 jours / an	333 jours / an	365 jours / an
Billet Soirée Halloween et soirées spéciales	-25%	-25%	-40%
Restaurants	-10% toute l'année	-10% toute l'année	-10% toute l'année
Boutiques	-10% toute l'année	-10% toute l'année	-20% toute l'année
Reduction dans un Hôtel Disney	-20%	-20%	-20%
Reduction famille	-20%	-20%	-20%
Accès aux parkings des 2 Parcs Disney	En option pour 24€	Inclus	Inclus
Billets Parc 1 Jour pour les amis ou vos accompagnants	-10%	-10%	-20%
Dîner Spectacle La Légende de Buffalo Bill	-15%	-15%	-15%

Les différents types de passeports annuels

## Le développement du passeport annuel et les initiatives lancées début 2009

Aujourd'hui, le passeport annuel est vendu via nos principaux canaux et média de distribution directs et indirects :

- Par DLP à l'entrée du parc, sur le web, par téléphone et par courrier, au « bureau des passeports annuels » (BPA)
- Par des partenaires (principalement des acteurs de la billetterie comme la FNAC par exemple)

Jusqu'au premier semestre de l'année 2009, le processus d'obtention d'un passeport annuel était le même, quel que soit le canal/média d'achat :

1. Achat du passeport annuel par le guest chez DLP (CRO, Web, entrée du parc, VPC) ou chez un partenaire : le guest reçoit alors un « billet tempo »
2. Avec son billet tempo, le guest peut entrer dans le parc Disneyland (le tempo est valable pour une entrée unique et le guest a 2 à 6 mois pour l'utiliser)

3. Le jour de son entrée dans le parc, le guest doit se rendre au BPA pour retirer un passeport définitif
4. Au BPA, le guest fournit un certain nombre d'informations (nom, adresse, n° de téléphone...), est pris en photo, signe un contrat d'abonnement et reçoit son passeport définitif

Ces dernières années, les ventes de passeports ont considérablement augmenté pour doubler en FY 09 et atteindre plus de 300 000 passeports.

Un certain nombre d'études de marché ont montré que le prix constitue une barrière importante à l'achat et au renouvellement des passeports annuels, notamment pour les familles. Pour continuer à développer les ventes de passeports annuels Disneyland Paris souhaite lancer un nouveau programme permettant au client d'étaler le paiement des passeports sur 12 mois.

Cette solution de paiement lancée aux Etats-Unis (Disneyland en Californie et Walt Disney World en Floride) a montré l'intérêt des clients pour cette option : augmentation du taux de conversion de tickets en passeports à l'entrée des parcs, augmentation des taux de renouvellement, meilleure fréquence de visite et montée en gamme entre les différents passeports annuels.

Le paiement mensuel est une facilité de paiement proposée en option, pour chaque passeport de la gamme vendu par ailleurs en paiement traditionnel (sauf certains tarifs préférentiels type « actionnaire, partenaire officiel ou famille nombreuse »). Cette option sera proposée sur site (au bureau des passeports annuels) et en vente par correspondance dans une première phase, avant d'élargir la vente au canal Internet.

## **Règles de gestion**

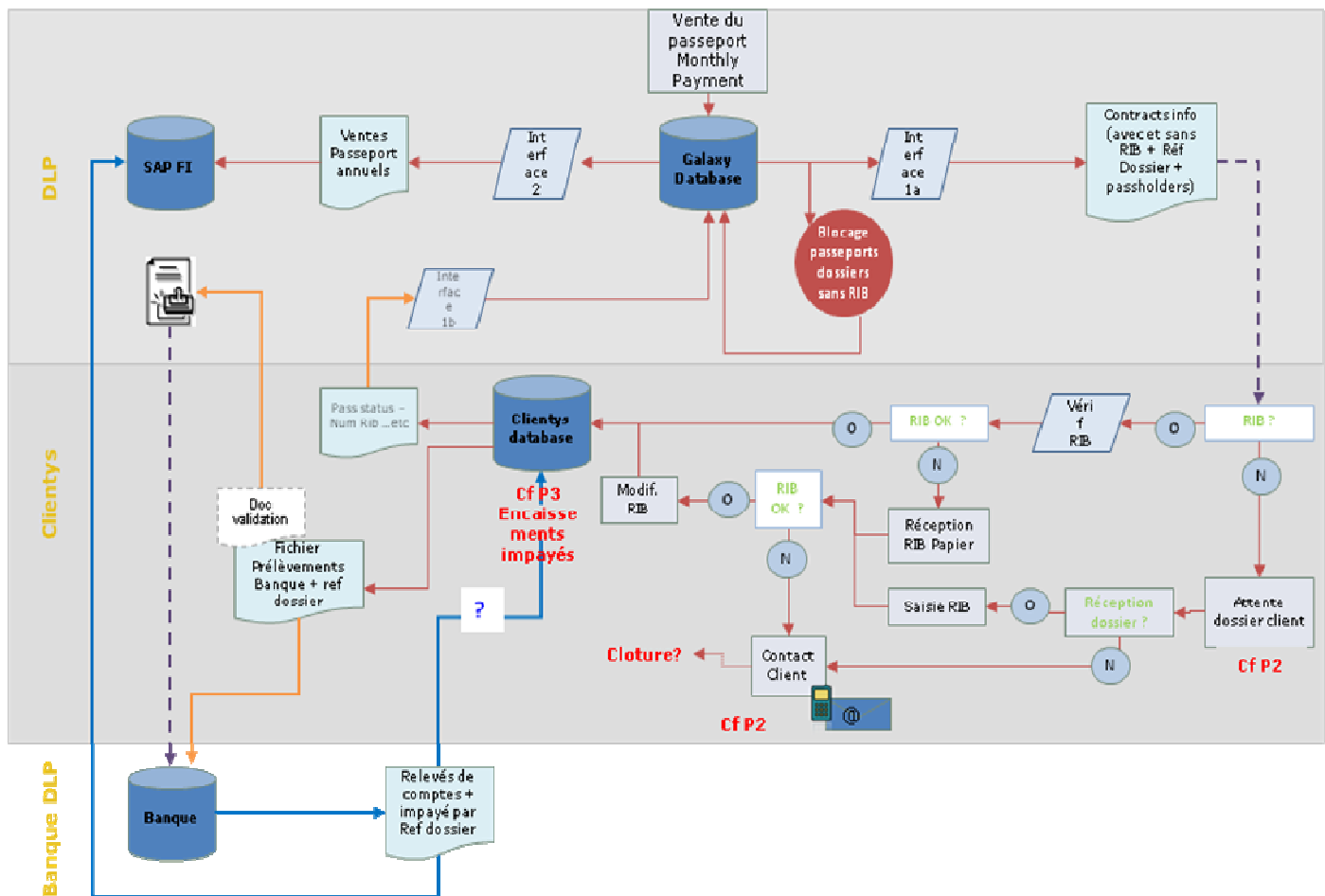
### **Périmètre de l'interface Teotys vers Galaxy**

L'interface 1b devra permettre d'intégrer dans le système de billetterie de Disneyland Paris :

- les informations concernant les RIB des dossiers non finalisés sur site quand le client aura envoyé les éléments manquants
- les blocages temporaires des passeports pour les dossiers en défaut de paiement temporaire

- Les blocages définitifs des passeports annuels pour les dossiers en défaut de paiement ou dossiers non finalisés.
- Les débloques (passeports temporairement bloqués).
- Les modifications de coordonnées (Postales – téléphonique – mail – RIB) pour le payeur du contrat

## Schéma général du flux d'informations



## **Rappel des différents processus fonctionnels du projet APMP**

### **Processus de vente du passeport annuel en Paiement Mensuel**

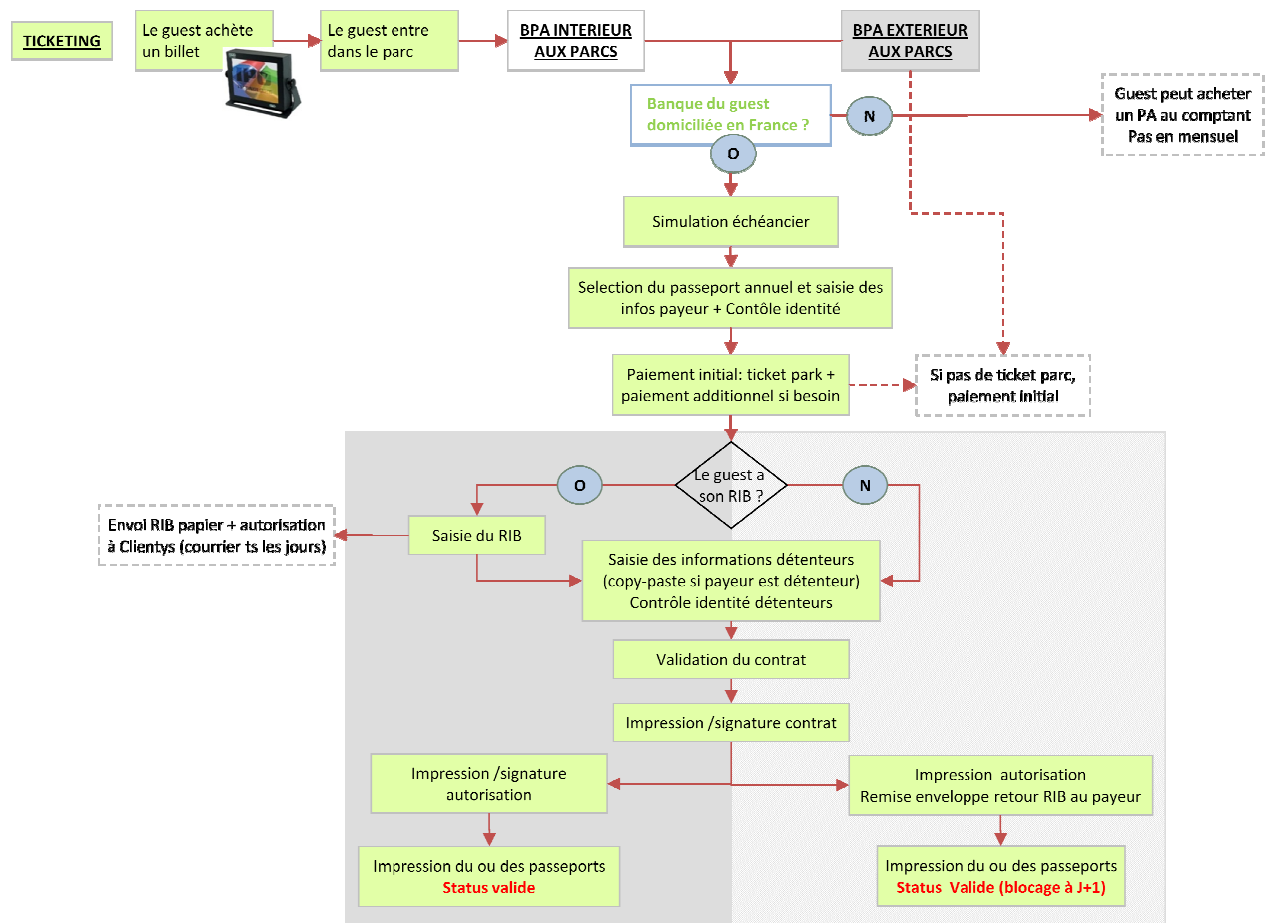
L'obtention d'un passeport annuel peut s'effectuer de deux façons différentes :

- Achat d'un billet parc transformé en Passeport annuel à l'intérieur du parc Disneyland
- Achat d'un passeport annuel à l'entrée des WDS

Concernant le paiement mensuel, un « downpayment » sera demandé pour l'obtention d'un ou plusieurs passeports.

Si l'achat concerne la transformation d'un billet parc, celui-ci fera office de « downpayment ».

Dans le cas d'un achat direct au bureau des Walt Disney Studio, le downpayment pourra être réglé par le visiteur par n'importe quel moyen de paiement à sa convenance.



## Processus de gestion du RIB

Disneyland paris a choisi de ne pas pénaliser les visiteurs qui n'auraient leur RIB finaliser un dossier.

Le payeur devra donc dans un délai imparti retourner le RIB à la société Teotys, et les documents de contractualisation permettant la finalisation de sont dossier.

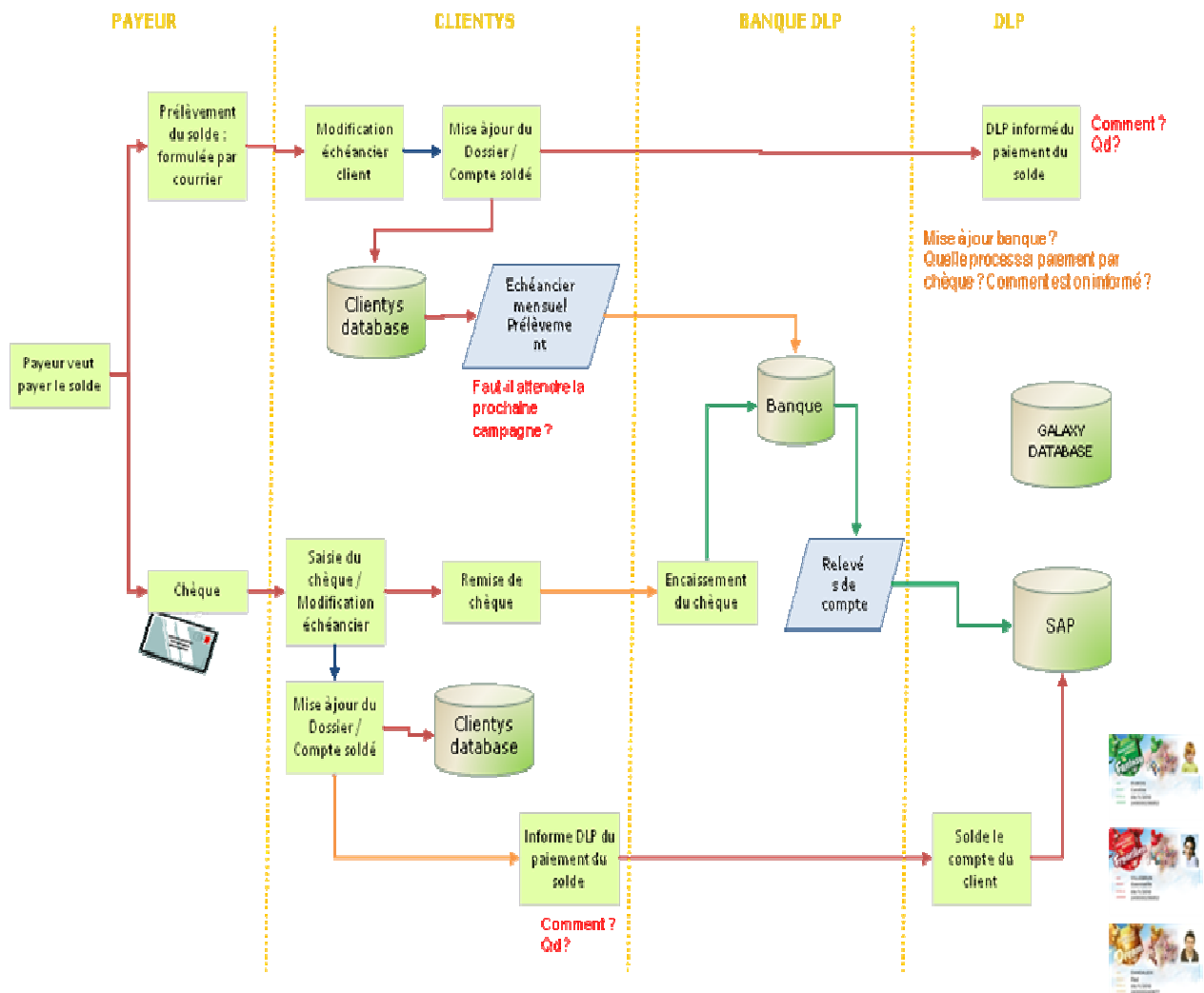
Les détenteurs de passeports annuel rattachés à un dossier non finalisé verront celui-ci bloqué temporairement le lendemain de la production, et ce, jusqu'à l'obtention des éléments manquants par la société Teotys, qui se chargera de les envoyer à Disneyland Paris.



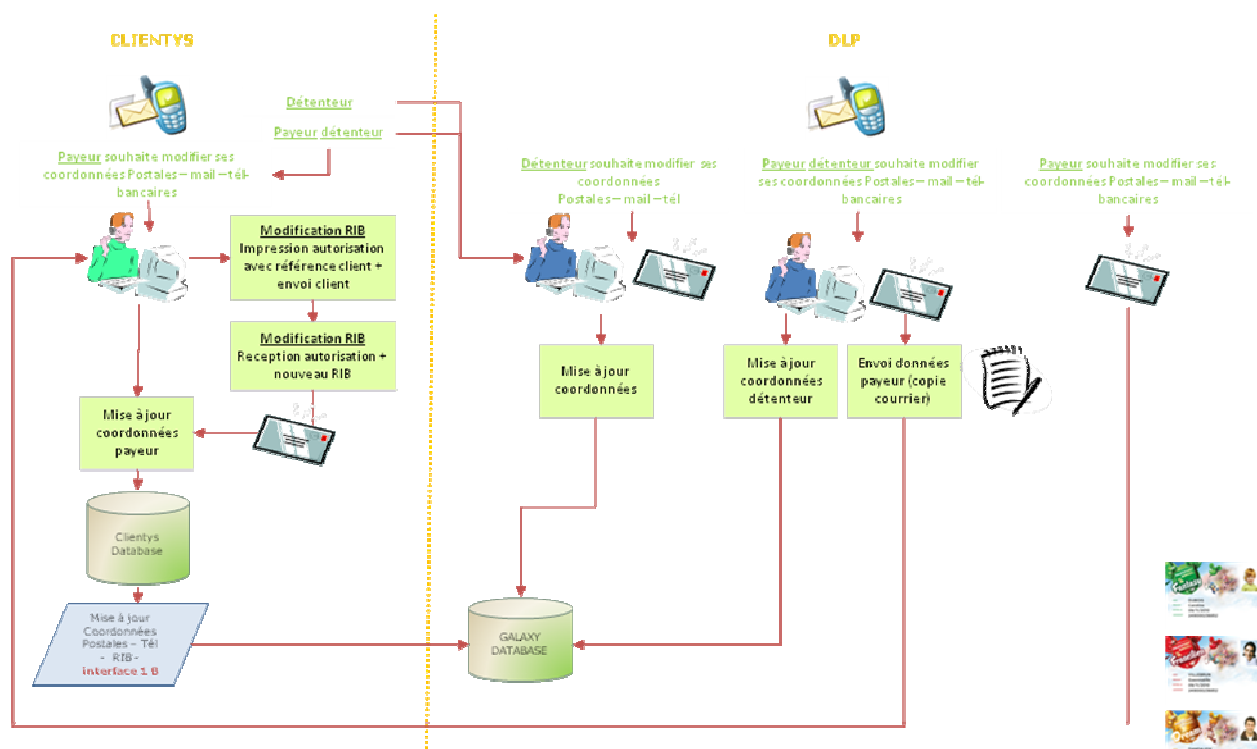


## Paieement du solde

Gestion des paiement mensuel passagère annuel



## Changement de coordonnées (Postales – Mail – Tél)



## Gestion générale des statuts dans la base de données Galaxy

La table Passes de la base de données Galaxy, comprend, outre les informations des détenteurs de passeports annuels, le champ Statut, qui permet de définir l'état du passeport.

Le champ « statut » permettra, dans le cadre du projet APMP, de bloquer ou débloquer un détenteur de passeport, si un défaut de paiement est détecté ou si le dossier du payeur n'est pas complété dans les délais impartis.

Le « userfield10 » de la table passes permettra de préciser les blocages ou déblocage du passeport comme suit :

- 0 = défaut non bloqué
- 1 = tempo dossier incomplet
- 2 = tempo incident de paiement

- 3 = blocage définitif

En complément, le « UserField7 » sera utilisé pour alimenter via interface la date de blocage, c'est-à-dire la date de mise à jour du contenu du « UserField10 » avec les valeurs 1, 2 ou 3. Ce champ est mis à jour par l'interface 1b et sur la base du userfield10 envoyé par Teotys.

#### Cas particulier :

- Un client achète un APMP au BPA mais le dossier n'est pas complet
- Le pass est bloqué temporairement en J+1, mais le client revient quelques jours après directement au BPA pour compléter son dossier
- Le BPA vérifie les documents et débloquent les pass du dossier (A valider)
  - En J+1, le statut des passes ayant été modifié, tous les éléments du dossier sont réextraits en type « Modification détenteurs » dans la 1a. [sous réserve de validation query gateway → to be test asap]
- Les statuts des passeports étant débloqués, le client peut entrer dans le parc librement.
- Les documents du dossier récupérés au BPA sont envoyés à Teotys
- A réception Teotys les vérifie :
  - soit ils sont ok et le dossier est bien complet
    - Teotys modifie le champ qui chez eux correspond au UserField10 à OK et envoie l'info de la mise à jour du statut du dossier dans l'interface 1b
    - DLP intègre la modification du statut du dossier dans Galaxy
  - soit ils sont ko et le dossier n'est toujours pas complet
    - Teotys ne modifie pas le champ qui chez eux correspond au « UserField10 » qui reste à « bloqué » (valeurs 1, 2 ou 3)
    - Teotys envoie quand même systématiquement une mise à jour du champ par l'interface 1b pour que dans Galaxy, le UserField10 des passes du dossier soit repassé en bloqué

### **Élément extraits de la base de données Teotys**

#### **Éléments généraux concernant le fichier généré**

Le fichier généré par Teotys sera au format xml.

Pour chaque modification, tous les éléments de la fiche payeur seront envoyés.

Liste des champs transmis (voir annexe pour format du fichier xml)

### **Payeur**

\*N° de contrat

\*Nom

\*Prénom

\*Adresse1

Adresse2

\*Code postal

\*Ville

\*Pays

N° de Téléphone

Mail

\*Bic

\*Iban

\*Statut

\*Motif (0 – default non bloqué, 1-dossier incomplet, 2- blocage temporaire, 3- blocage définitif)

### **Pied de fichier**

\*Nombre de lignes total du fichier

\*Nombre de dossiers en modification

\*Nom du fichier en sortie d'interface

\*Date de création

\*Numéro séquentiel

Rejet du fichier : si un seul des éléments du pied de page ne correspond pas

Rejet de la ligne (= du dossier) : si un des champs obligatoires (\*) est absent.

## Règle de déclenchement de l'export chez Teotys

L'interface sera déclenchée tous les jours ouvrés à 18 :00, le fichier sera collecté par Disneyland Paris.

En cas d'absence de fichier, un mail sera envoyé automatique à Téotys [Liste destinataires à valider].

## Critères d'extraction

### Mise à jour des contrats

Tous les contrats ayant généré une modification en J devront être extraits pour envoi au système cible excepté les dossiers dont le statut est 3.

## Relance de l'export chez Teotys

### Relance de l'interface en cas d'indisponibilité du système source

L'interface pourra être lancée à la demande, sur les mêmes critères que ceux de l'extraction automatique.

Le fichier généré devra comporter les dernières modifications apportées aux dossiers.

Il ne devra pas comporter l'historique des modifications effectuées sur ces derniers.

## Intégration dans le système cible

### Généralités sur le système cible

Galaxy est le système de billetterie de Disneyland Paris. Il permet la vente de billetterie parc, ainsi que la billetterie de type abonnement, appelée « Passeport Annuel ».

L'acquéreur du passeport annuel obtient un droit d'accès aux Parcs Disney un certain nombre de jours dans l'année selon le type de passeport acquis.

Les billets de type « passeport » requièrent la saisie d'informations sur le détenteur ainsi que sa photo, permettant son identification lors de la séquence d'entrée.

Dans le cadre de la mise en œuvre du paiement mensuel, de nouvelles informations concernant le payeur d'un dossier contenant un ou plusieurs passeports devront être renseignées.

## Traitement du fichier en entrée

### Contrôles de cohérence des éléments envoyés par Teotys

Toute mise à jour de contrat (RIB ou élément du contrat) devra faire l'objet d'une vérification de l'existence de celui-ci, le critère de vérification unique étant le numéro de contrat (RUM).

Si le contrat n'existe pas dans la base de données Galaxy, il sera automatiquement intégré dans un rapport de rejet.

### Contrôle de cohérence entre les différents type de blocage de passeport

	Vérification dans Galaxy			Mise à jour Galaxy		
Mise à jour Teotys	Status initial Galaxy	UserField2 Valorisé ?	Valeur Initiale UserField10	Status Final Galaxy	Mise a jour Userfield 10 et 7	Valeur Userfield10
0	8	Oui	1 ou 2	8	Non	N/A
0	8	Non	1 ou 2	0	Oui	0
1 ou 2	8	Oui	0	8	Oui	2
1ou 2	0	Non	0	8	Oui	2
3	8	Oui	0	8	Oui	3
3	0	Non	0	8	Oui	3

\*Statut Pass  
8 = Blocked  
0 = Actif

### Intégration dans Galaxy

Tout enregistrement (dossier) ne comportant pas les champs obligatoires ou ne correspondant pas à la nomenclature du fichier xml sera rejeté en entrée de l'application Galaxy.

La liste des rejets détaillée sera envoyée par mail à Teotys ainsi qu'aux équipes de support applicatif de Disneyland Paris.

### Traitement des rejets par Teotys

Tout enregistrement contenu dans le fichier de rejet, sera traité par Teotys pour renvoi lors du prochain traitement.

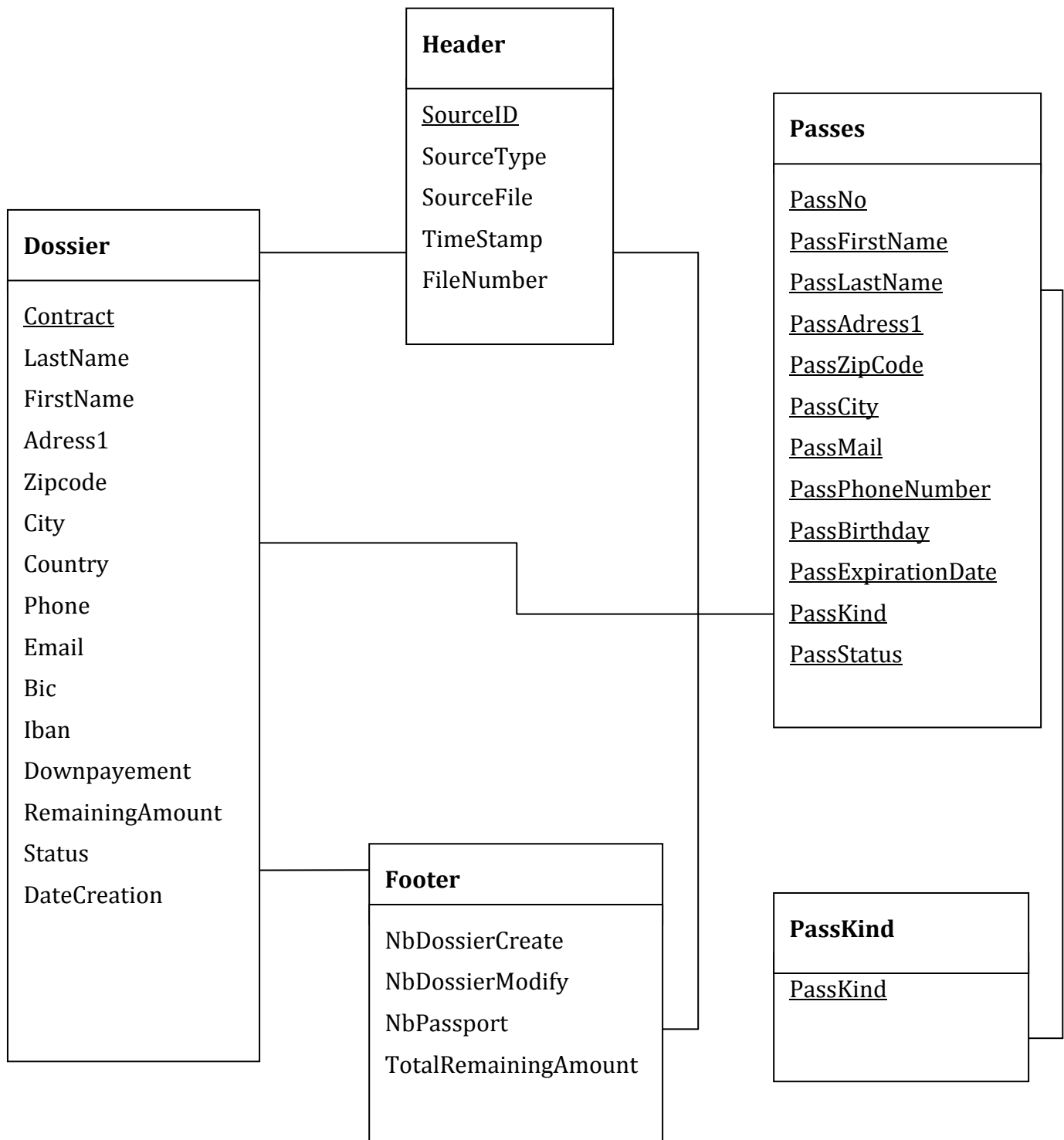
### III.5.1.a/ Intégration des données

Dans le but de mener à bien ce travail, nous avons eu à faire de l'intégration de données avec l'outil « Talend Data Integration », dans l'interface 1b : après l'envoi du fichier par « Teotys » (un partenaire qui s'occupe de la partie paiement et relance client), on s'occupe d'intégrer ce fichier dans « Galaxy » (Base de données Disney) pour mettre à jour les dossiers.

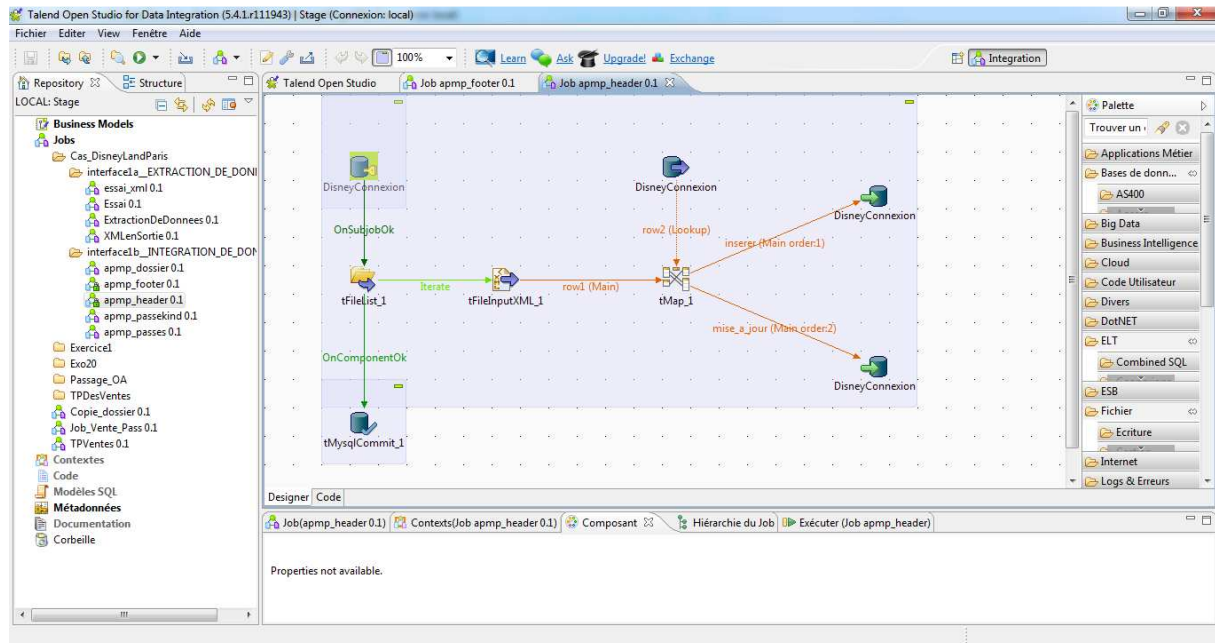
Dans un premier temps on a procédé par la réalisation d'un datamart du projet. Puis par faire de l'alimentation des tables à partir d'une liste de fichiers.



Comme suit le datamart :



En capture d'écran, un job réalisant l'alimentation de la table « header » dans la base de données à partir du fichier xml :



*Figure. Alimentation de la table « header »*

Avec l'utilisation de différents composants Talend, dans ce cas :

- Le composant « tFilelist » permet de parcourir le répertoire ou se trouve le fichier xml (faire des itérations sur les fichiers existant dans le répertoire)
- Le composant « tFileInputXML » permet de mettre en avant le fichier à traiter parmi les fichiers existants dans le répertoire.
- Le composant « tMysqlInput » permet de récupérer la table de la base de données.
- Les composants « tMysqlOutput » permettent la création ou la mise à jour des fichiers.
- Le composant « tMap » permet de faire le traitement sur les données, joindre les données en entrée (dans ce cas le fichier xlm et la table) puis de faire le lien avec les sorties (qui dans ce cas sont les tables créées ou mises à jour).

Ci-dessous l'alimentation de la table « header » dans la base « disneydata » :

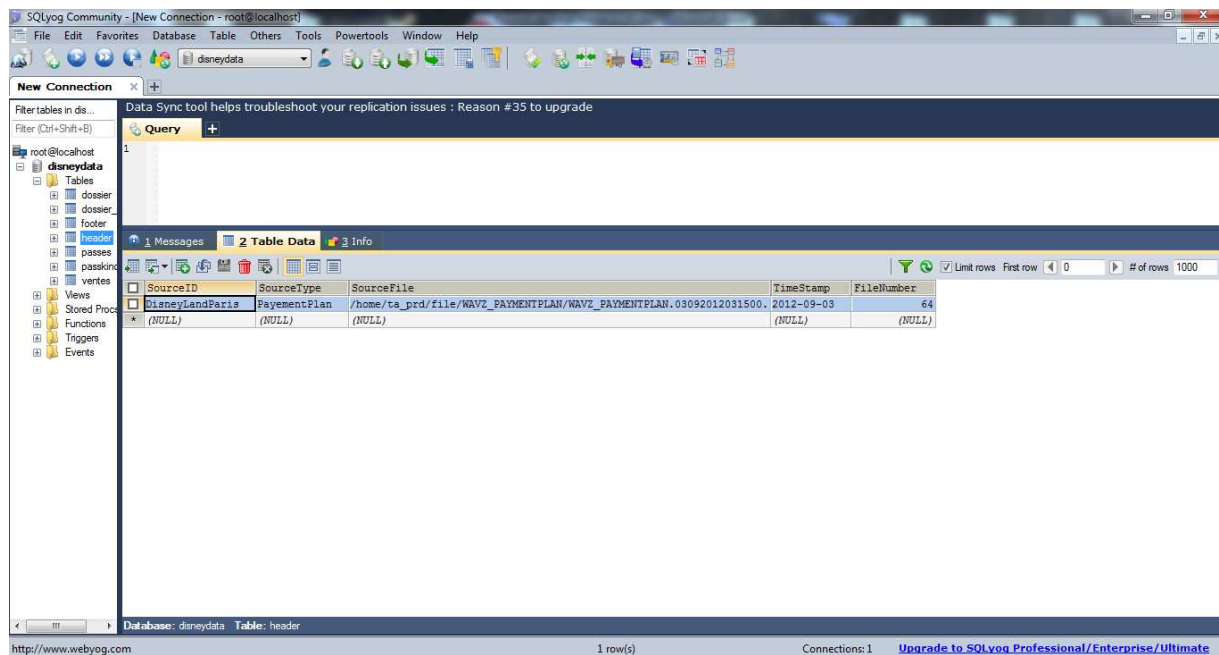


Figure. Table «header »

En capture d'écran, un job réalisant l'alimentation de la table « dossier » dans la base de données à partir du fichier xml :

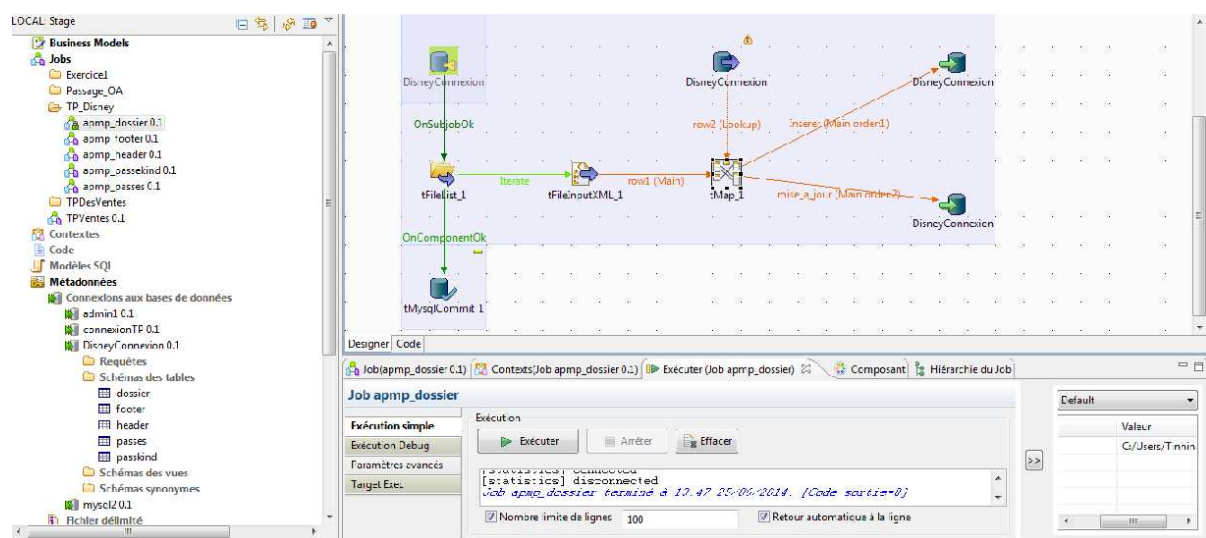


Figure. Alimentation de la table « dossier »

Avec l'utilisation de différents composants Talend, dans ce cas :

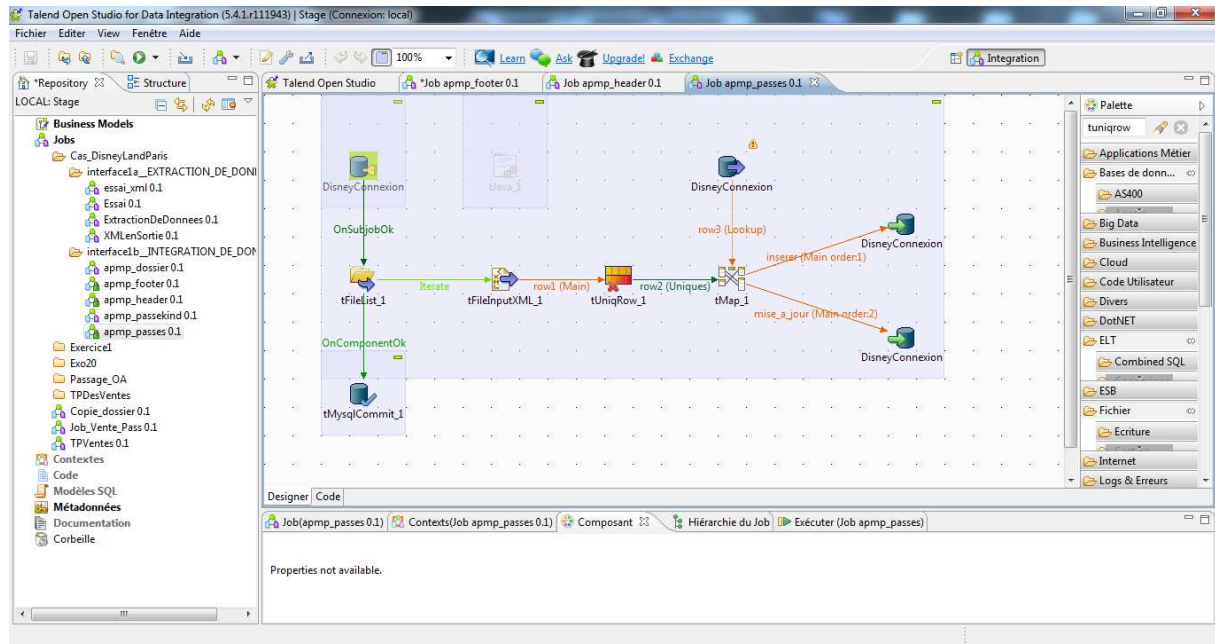
- Le composant « tFilelist » permet de parcourir le répertoire ou se trouve le fichier xml (faire des itérations sur les fichiers existant dans le répertoire)
- Le composant « tFileInputXML » permet de mettre en avant le fichier à traiter parmi les fichiers existants dans le répertoire.
- Le composant « tMysqlInput » permet de récupérer la table de la base de données.
- Les composants « tMysqlOutput » permettent la création ou la mise à jour des fichiers.
- Le composant « tMap » permet de faire le traitement sur les données, joindre les données en entrée (dans ce cas le fichier xml et la table) puis de faire le lien avec les sorties (qui dans ce cas sont les tables créées ou mises à jour).

Ci-dessous l'alimentation de la table « dossier » dans la base « disneydata » :

Contract	LastName	FirstName	Adresse1	Adresse2	Zipcode	City	Country	Phone
6	REINON	GEOFFREY	115 B RUE MARIUS SIDOBRE	(NULL)	94110	ARCUEIL	FR	0632394028
29	BREE	MAXIMILIEN	83 RUE DE LA FERRE	(NULL)	02100	ST QUENTIN	FR	0626803167
37	ZIZERT	DAVID	7 RUE PAUL BELMONDO	(NULL)	93160	NOISY LE GRAND	FR	0663109424
38	BARBOSA	SERGE	5 RUE MOLIERE	(NULL)	94430	CHENNEVIERES	FR	06 63 82 42 3
75	DANICAN	JEREMY	10 RUE PAUL DELOUVRIER	(NULL)	77600	BUSSY ST GEORGES	FR	0611538106
95	BELAZZOUG	PAULINE	1 B RUE DE LA VALLEE	(NULL)	77170	SERVON	FR	0652046574
140	ROBBE	VANESSA	53 RUE DU 6 SEPTEMBRE 1914	(NULL)	59245	RECQIGNIES	FR	0786053640
164	ALVES	SONIA	3 RUE MAURICE RAVEL	(NULL)	60200	COMPIEGNE	FR	0646758935
165	VAN GOOL	NICOLAS	3 RUE MAURICE RAVEL	(NULL)	60200	COMPIEGNE	FR	0642721467
172	DUSART	DORIANE	65 ROUTE DE CURTILLE	(NULL)	38630	LES AVENIERES	FR	0622848387
193	DUBEST	DELPHINE	9 AVENUE JEAN JAURES	APT 63	77000	MELON	FR	0616745182
196	LAMON	ALEXANDRE	APT 301	4 ALLEE BERNARD LETELLIER	77410	CLAYE SOUILLY	FR	0677668673
199	GAGLIANI	ANGE	8 AV ANATOLE FRANCE	TOUR DE LA SEINE	94600	CHOISY LE ROI	FR	0689381037
203	PERRICHON	STEPHANIE	56 ALLEE DU GROS CHENE	(NULL)	94510	LA QUEUE EN BRIE	FR	0620438044
238	TANTY	STEPHANE	2 RUE DE STAINS	(NULL)	77230	VILLENEUVE SOUS DAMMARIN	FR	0660897757
254	LAURET	BRIGITTE	14 RUE DU CHAPERON ROUGE	(NULL)	94400	VITRY SUR SEINE	FR	0682170471
273	POIRAUD	SOPHIE	48 RUE SAINT FARGEAU	(NULL)	75020	PARIS 20EME ARRONDIS	FR	0175504924
304	COLLETTIN	EDDY	2 SQUARE DES TROPHEES	(NULL)	77186	NOISIEL	FR	0660630016
328	LAILLET	SANDRINE	7 B RUE JEAN AMAN	(NULL)	77173	CHEVRY COSSIGNY	FR	0164055263

Figure. Table « dossier »

En capture d'écran, un job réalisant l'alimentation de la table « passes » dans la base de données à partir du fichier xml :



*Figure. Alimentation de la table « passes »*

Avec l'utilisation de différents composants Talend, dans ce cas :

- Le composant « tFilelist » permet de parcourir le répertoire ou se trouve le fichier xml (faire des itérations sur les fichiers existant dans le répertoire)
- Le composant « tFileInputXML » permet de mettre en avant le fichier à traiter parmi les fichiers existants dans le répertoire.
- Le composant « tUniqRow » compare les entrées et supprime les doublons de flux d'entrée
- Le composant « tMysqlInput » permet de récupérer la table de la base de données.
- Les composants « tMysqlOutput » permettent la création ou la mise à jour des fichiers.
- Le composant « tMap » permet de faire le traitement sur les données, joindre les données en entrée (dans ce cas le fichier xlm et la table) puis de faire le lien avec les sorties (qui dans ce cas sont les tables créées ou mises à jour).

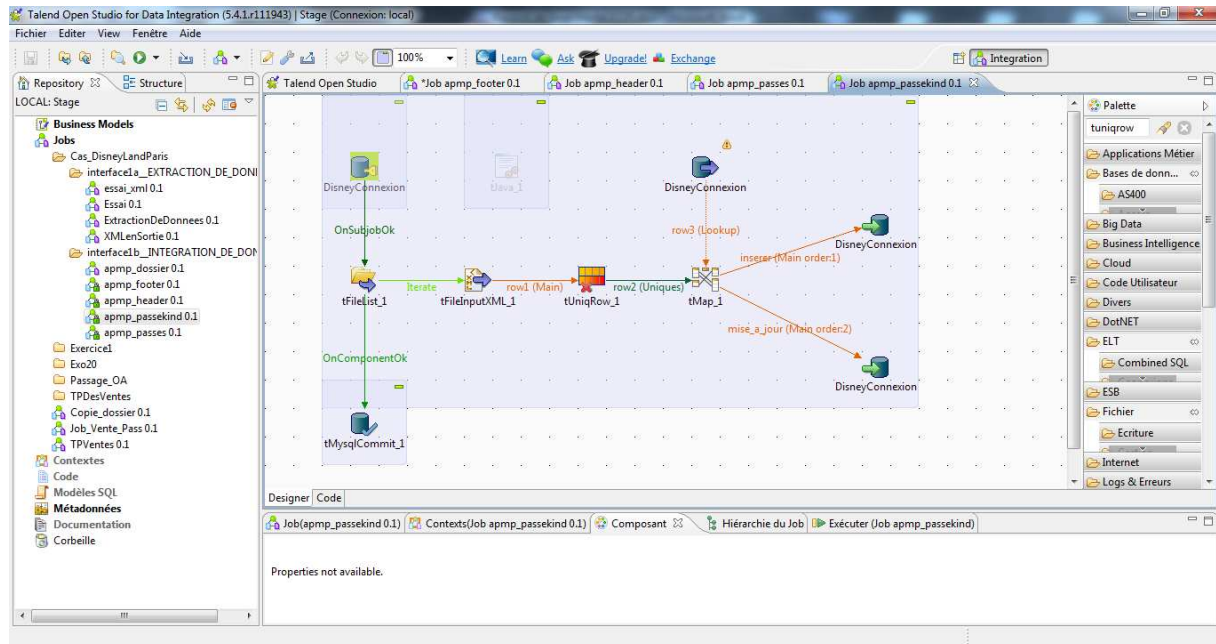
Ci-dessous l'alimentation de la table « passes » dans la base « disneydata » :

PassNo	Contract	PassFirstName	PassLastName	PassAddress1	PassZipcode	PassCity	PassMail
3088379	6	MARIE LAURE	DAHAN	115 B RUE MARIUS SIDOBRE	94110	ARCUEIL	MARIELAURE_DAHAN@YAHOO.COM
3088390	6	GEOFFREY	RENON	115 B RUE MARIUS SIDOBRE	94110	ARCUEIL	BOB_SPACE@YAHOO.FR
3123485	29	MAXIMILIEN	BREE	83 RUE DE LA FERRE	02100	ST QUENTIN	MAXIMBREE@AOL.COM
3133007	37	DAVID	ZIZERT	7 RUE PAUL BELMONDO	93160	NOISY LE GRAND	GOLD093160@YAHOO.FR
3133033	38	SERGE	BARBOSA	5 RUE MOLIERE	94430	CHENNEVIERES	MANDRACKE@VOILA.FR
3150414	75	JEREMY	DANICAN	10 RUE PAUL DELOUVRIER	77600	BUSSY ST GEORGES	DANICAN_JEREMY@HOTMAIL.FR
3155333	95	PAULINE	BELAZZOUG	1 B RUE DE LA VALLEE	77170	SERVON	POPOLIGHT77@GMAIL.COM
3163407	140	JOHAN	ROBBE	53 RUE DU 6 SEPTEMBRE 1914	59245	RECQIGNIES	VANESSA.ROBBE@ORANGE.FR
3163422	140	AURELIEN	ROBBE	53 RUE DU 6 SEPTEMBRE 1914	59245	RECQIGNIES	(NULL)
3163423	140	CASSANDRA	ROBBE	53 RUE DU 6 SEPTEMBRE 1914	59245	RECQIGNIES	(NULL)
3163424	140	ALEXANDRA	ROBBE	53 RUE DU 6 SEPTEMBRE 1914	59245	RECQIGNIES	(NULL)
3163425	140	VANESSA	ROBBE	53 RUE DU 6 SEPTEMBRE 1914	59245	RECQIGNIES	(NULL)
3168115	164	SONIA	ALVES	3 RUE MAURICE RAVEL	60200	COMPIEGNE	SONIANICOLAS27@GMAIL.COM
3168146	165	NICOLAS	VAN GOOL	3 RUE MAURICE RAVEL	60200	COMPIEGNE	NICOLASONIA24@GMAIL.COM
3169701	172	DORIANE	DUSART	65 ROUTE DE CURTILLE	38630	LES AVENIERES	DUSART.LIONEL@ORANGE.FR
3169717	172	LIONEL	DUSART	65 ROUTE DE CURTILLE	38630	LES AVENIERES	Z
3169718	172	JORDAN	DUSART	65 ROUTE DE CURTILLE	38630	LES AVENIERES	Z
3169719	172	JASON	DUSART	65 ROUTE DE CURTILLE	38630	LES AVENIERES	Z
3169720	172	DYLAN	DUSART	65 ROUTE DE CURTILLE	38630	LES AVENIERES	Z

Figure. Table « passes »



En capture d'écran, un job réalisant l'alimentation de la table « passkind » dans la base de données à partir du fichier xml :



*Figure. Alimentation de la table « passkind »*

Avec l'utilisation de différents composants Talend, dans ce cas :

- Le composant « tFilelist » permet de parcourir le répertoire ou se trouve le fichier xml (faire des itérations sur les fichiers existant dans le répertoire)
- Le composant « tFileInputXML » permet de mettre en avant le fichier à traiter parmi les fichiers existants dans le répertoire.
- Le composant « tUniqRow » compare les entrées et supprime les doublons de flux d'entrée
- Le composant « tMysqlInput » permet de récupérer la table de la base de données.
- Les composants « tMysqlOutput » permettent la création ou la mise à jour des fichiers.
- Le composant « tMap » permet de faire le traitement sur les données, joindre les données en entrée (dans ce cas le fichier xlm et la table) puis de faire le lien avec les sorties (qui dans ce cas sont les tables créées ou mises à jour).

Ci-dessous l'alimentation de la table « passkind » dans la base « disneydata » :

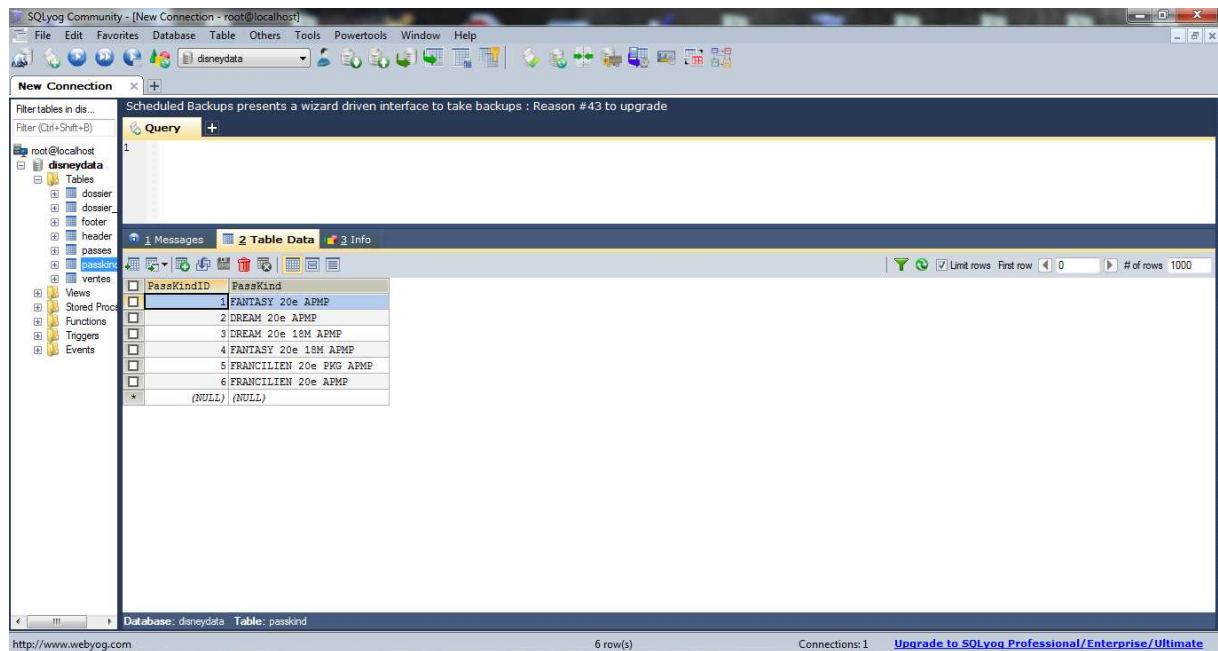


Figure. Table « passkind »

En capture d'écran, un job réalisant l'alimentation de la table « footer » dans la base de données à partir du fichier xml :

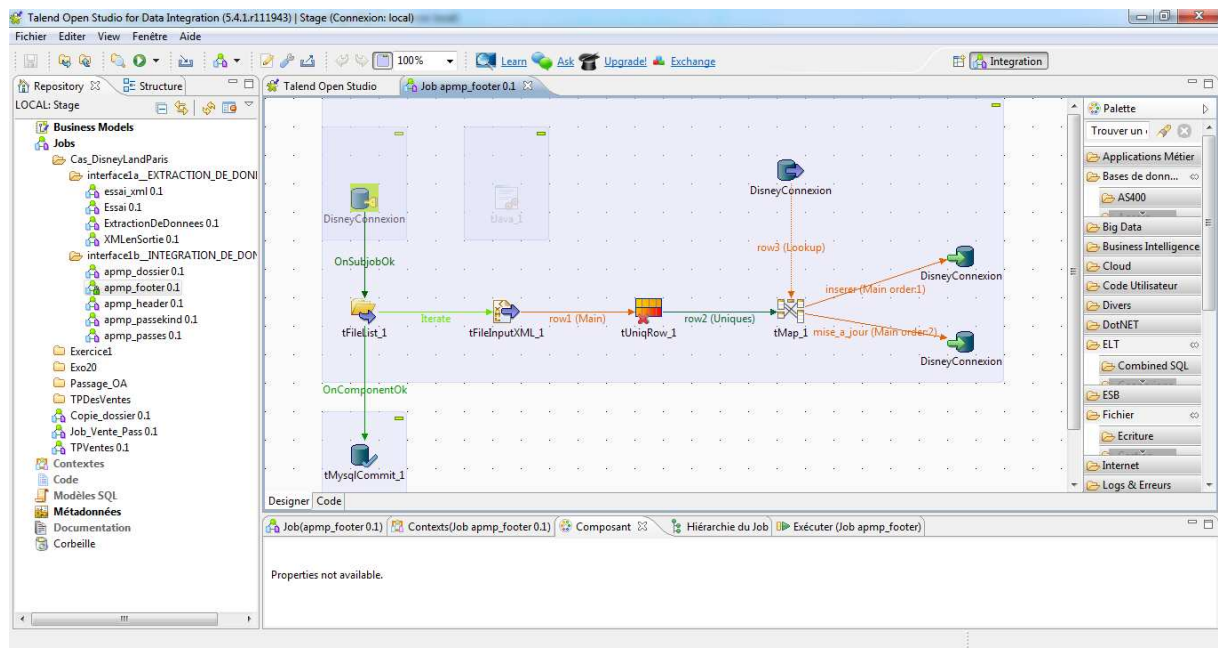


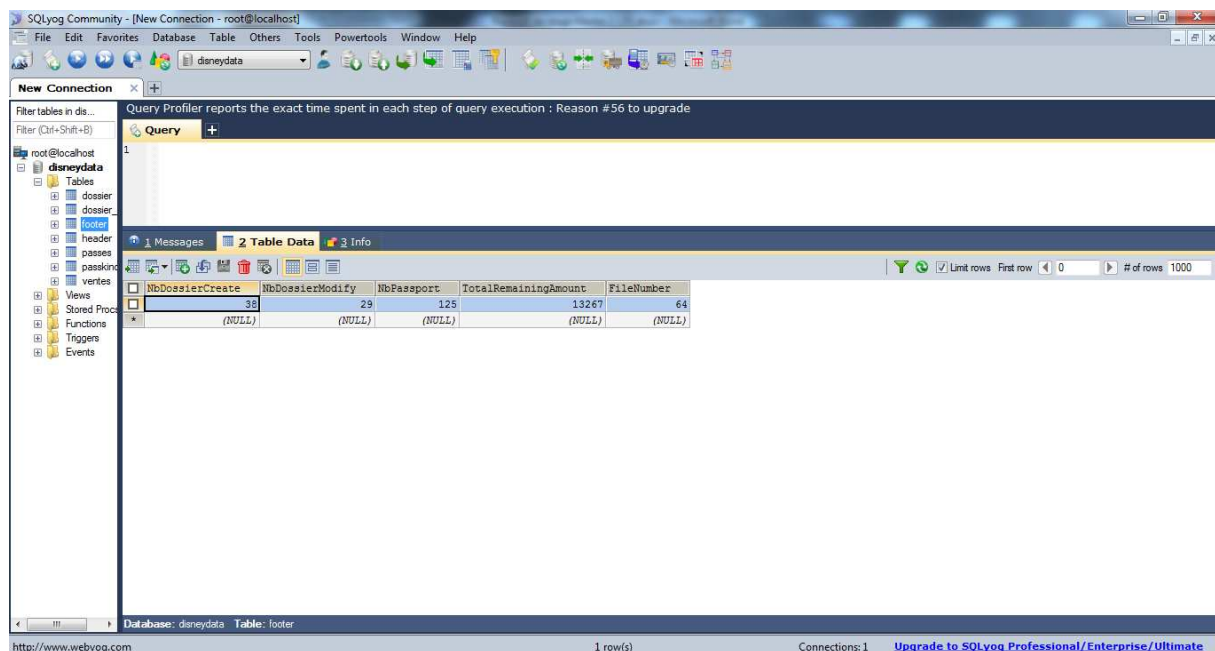
Figure. Alimentation de la table « footer »



Avec l'utilisation de différents composants Talend, dans ce cas :

- Le composant « tFilelist » permet de parcourir le répertoire où se trouve le fichier xml (faire des itérations sur les fichiers existant dans le répertoire)
- Le composant « tFileInputXML » permet de mettre en avant le fichier à traiter parmi les fichiers existants dans le répertoire.
- Le composant « tUniqRow » compare les entrées et supprime les doublons de flux d'entrée
- Le composant « tMysqlInput » permet de récupérer la table de la base de données.
- Les composants « tMysqlOutput » permettent la création ou la mise à jour des fichiers.
- Le composant « tMap » permet de faire le traitement sur les données, joindre les données en entrée (dans ce cas le fichier xml et la table) puis de faire le lien avec les sorties (qui dans ce cas sont les tables créées ou mises à jour).

Ci-dessous l'alimentation de la table « footer » dans la base « disneydata » :



Query Profiler reports the exact time spent in each step of query execution : Reason #56 to upgrade

NbDossierCreate	NbDossierModify	NbPassport	TotalRemainingAmount	FileNumber
38	29	125	13267	64
(NULL)	(NULL)	(NULL)	(NULL)	(NULL)

Database: disneydata Table: footer

Figure. Table «Footer »

Les différentes étapes précédemment définies m'ont permis de mettre en place le travail d'intégration de données demandé, à savoir alimenter les tables et les mettre à jour.

### III.5.1.b/ Extraction des données

Dans l'interface 1a : à partir des tables de notre base de données « Galaxy », nous souhaitons faire de l'extraction de données pour la création d'un fichier xml à envoyer à « Teotys ».

Pour ce faire, nous avons procédé par la réalisation d'un job Talend pour répondre à notre besoin, qui permet d'extraire des données à partir de la base pour les regrouper toutes dans un même fichier xml, prenant compte de la gestion des boucles.

Une capture d'écran illustrant ce dernier :

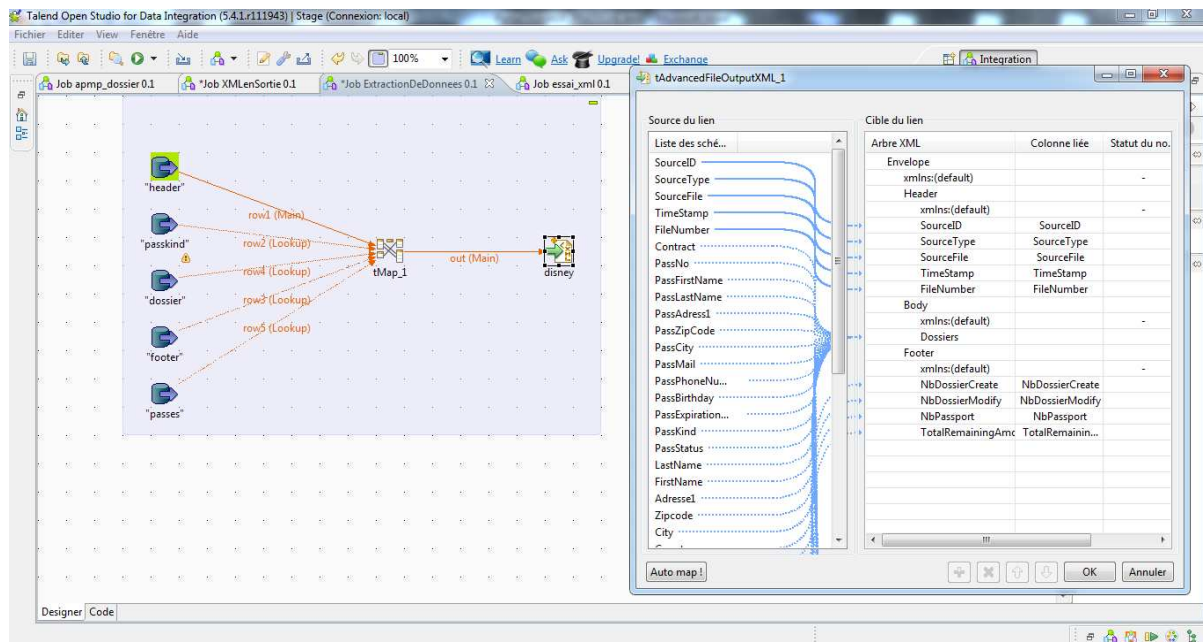
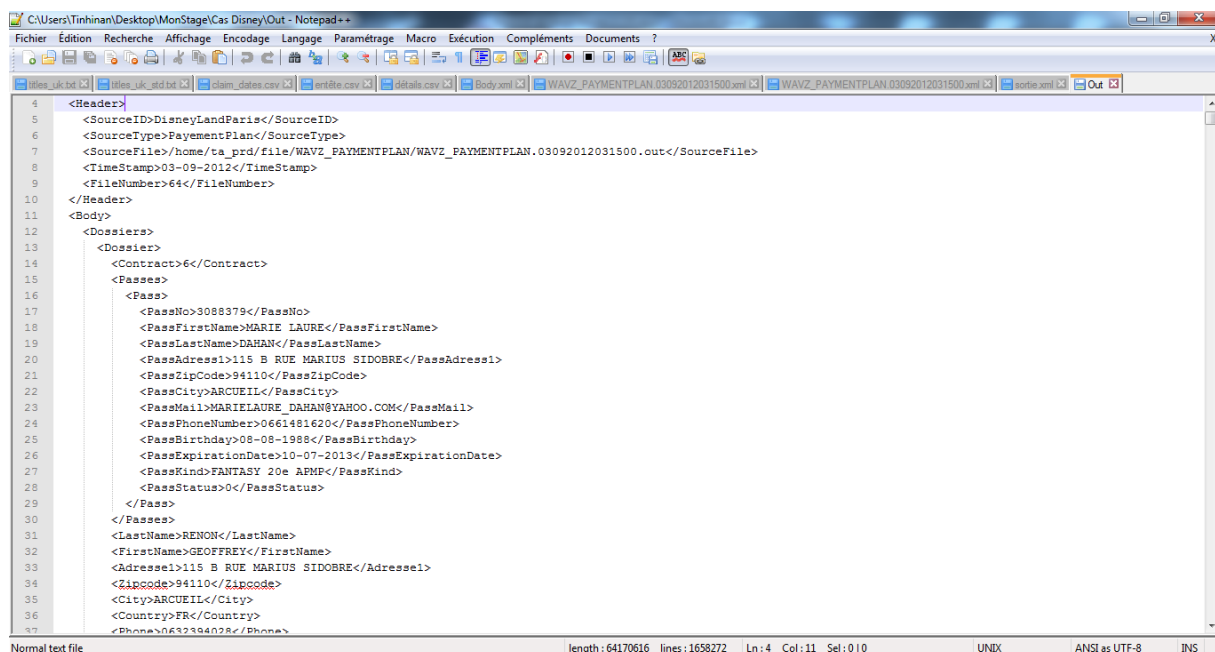


Figure. Extraction de données en un fichier xml

En utilisant différents composants Talend, dans ce cas :

- Les composants en entrée « tMysqlInput » permettent de récupérer nos tables (header, dossier, passes, passkind et footer) depuis notre base de données « Galaxy ».
- Le composant « tMap » permet le traitement sur les données, joindre nos tables en entrée puis les mettre en lien avec la sortie (un fichier xml dans ce cas).
- Le composant « tAdvancedFileOutputXML » transmet mes données depuis ma base vers un fichier de type xml en prenant en compte la gestion des boucles.

Ci-dessous une capture d'écran du fichier résultant :



```
4 <Header>
5 <SourceID>DisneyLandParis</SourceID>
6 <SourceType>PaymentPlan</SourceType>
7 <SourceFile>/home/ta_prd/file/WAVZ_PAYMENTPLAN/WAVZ_PAYMENTPLAN.03092012031500.out</SourceFile>
8 <TimeStamp>03-09-2012</TimeStamp>
9 <FileNumber>64</FileNumber>
10 </Header>
11 <Body>
12 <Dossiers>
13 <Dossier>
14 <Contract>6</Contract>
15 <Passes>
16 <Pass>
17 <PassNo>3088379</PassNo>
18 <PassFirstName>MARIE LAURE</PassFirstName>
19 <PassLastName>DAHAN</PassLastName>
20 <PassAddress1>115 B RUE MARIUS SIDOBRE</PassAddress1>
21 <PassZipCode>94110</PassZipCode>
22 <PassCity>ARCUEIL</PassCity>
23 <PassMail>MARIELAURE_DAHAN@YAHOO.COM</PassMail>
24 <PassPhoneNumber>0661481620</PassPhoneNumber>
25 <PassBirthDay>08-08-1988</PassBirthDay>
26 <PassExpirationDate>10-07-2013</PassExpirationDate>
27 <PassKind>FANTASY 20e AFMP</PassKind>
28 <PassStatus>0</PassStatus>
29 </Pass>
30 </Passes>
31 <LastName>RENON</LastName>
32 <FirstName>GEOFFREY</FirstName>
33 <Address1>115 B RUE MARIUS SIDOBRE</Address1>
34 <Zipcode>94110</Zipcode>
35 <City>ARCUEIL</City>
36 <Country>FR</Country>
37 <Phone>0632394028</Phone>
```

*Figure. Fichier xml réalisé*

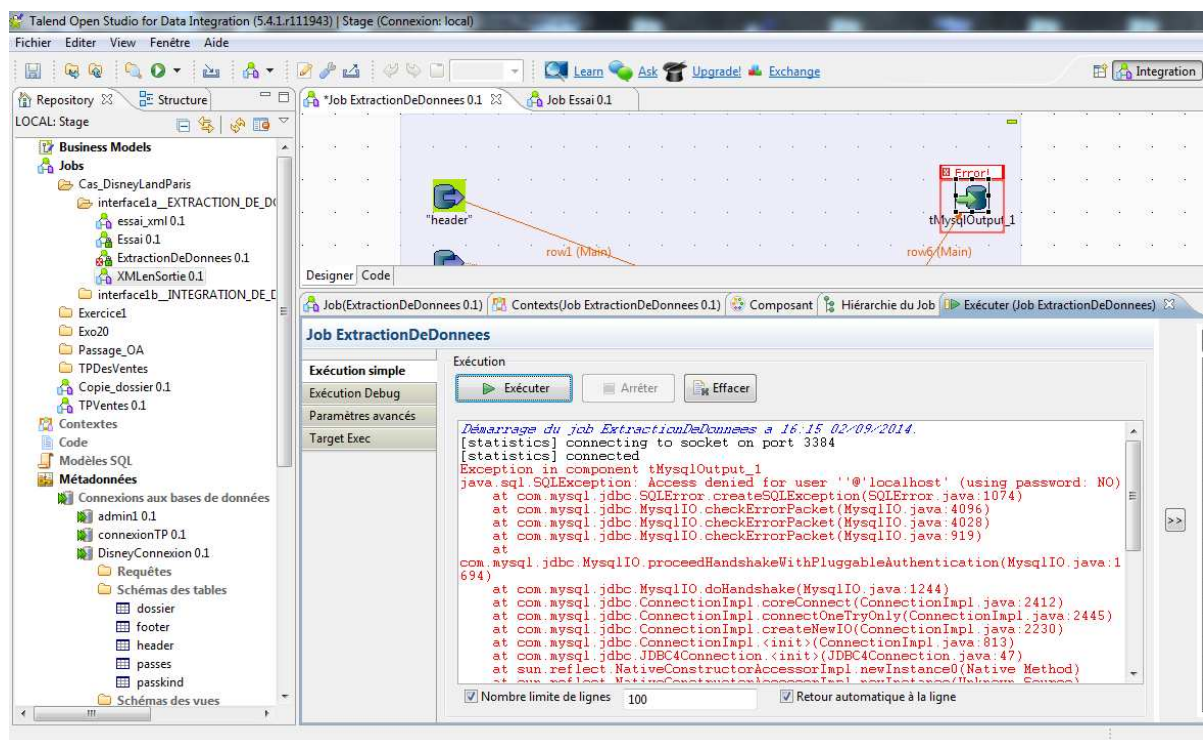
Les étapes ci-dessus m'ont permis de réaliser le module « Extraction de données » du projet APMP Disneyland Paris.

## Tests unitaires:

En ce qui concerne les tests unitaires, ils se sont fait au fur et à mesure avec la création des jobs. Je teste mes composants « Talend » et ces mêmes composants mais liés entre eux ... d'où le job final (job exécutant des tâches bien précises).

Les résultats sont affichés dans la console d'exécution.

Comme suit une capture d'écran de la console d'exécution signalant des erreurs par exemple :



*Figure. Erreur*

Dans cet exemple, dans le job Talend, le composant « tMysqlOutput » n'a pas été correctement paramétré ...

### III.5.2 Qualité de données

Pour mener à bien mon travail, je me suis perfectionné en « Talend », en faisant une autre formation de deux semaines en « Talend Data Quality ».

Si on a une mauvaise qualité de données en entrée on aura une mauvaise qualité de données en sorti, ce qui mène à l'incapacité à prendre des décisions ou alors une prise de décision médiocre.

Avec cette partie, nous avons pu faire l'analyse de la base « Galaxy » ainsi que les données stockées dans cette dernière.

Dans ce travail réalisé, on a commencé par faire l'analyse de connexion à la base « Galaxy », une vue d'ensemble de la structure de la base, à savoir le nombre de tables, le nombre de lignes, le nombre de clés ... etc. Ainsi qu'une exportation d'un rapport permettant le rassemblement des différentes analyses en un document PDF, Excel, XML ou HTML, facile à lire et à partager.

En voici des captures d'écran :

The screenshot shows the Talend Data Quality interface. The left sidebar contains a project tree with folders like 'Analyses (39)', 'Reports (8)', and 'Libraries'. The main workspace displays the 'Analyse\_disneydata 0.1' job. The 'Résumé d'analyse' section shows the following details:

SGBD :	MySQL	Date de création :	18 août 2014 14:58:55
Serveur :	localhost	Date d'exécution :	2 sept. 2014 14:11:17
Port :	3306	Durée de l'exécution :	3.771s
Connecté en tant que :	root	Statut de l'exécution :	Succès
Catalogues :	1	Nombre d'exécutions :	4
Schémas :	0	Dernière exécution réussie :	4

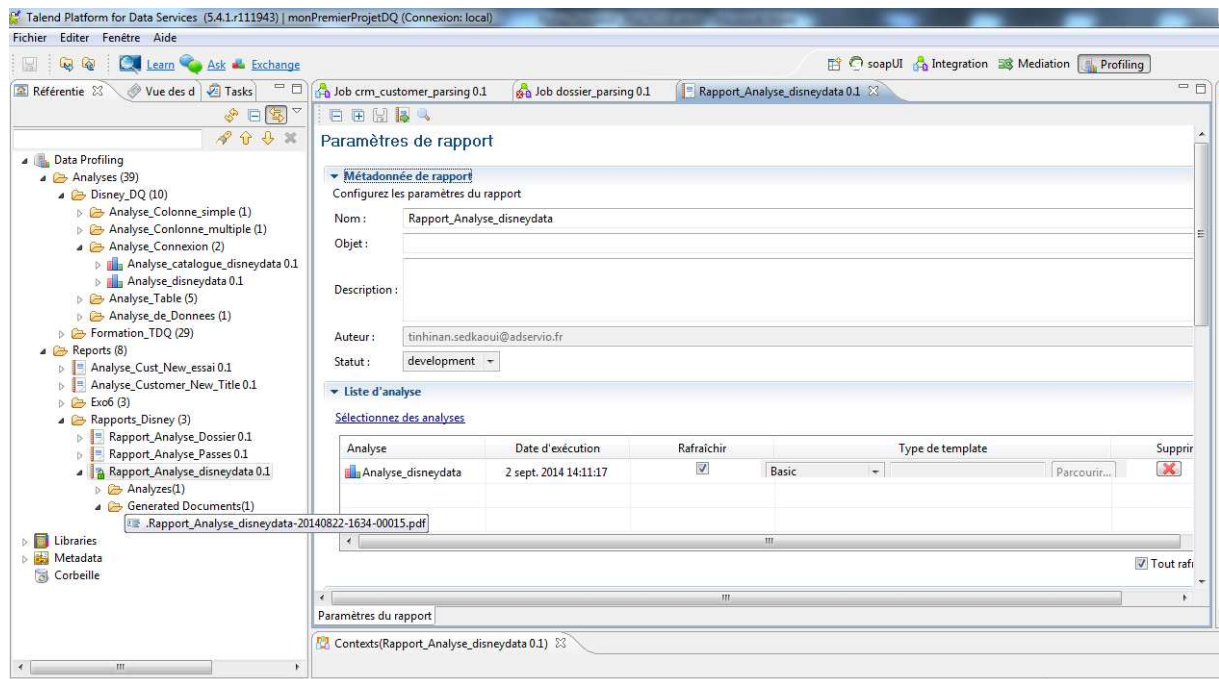
The 'Information statistiques' section displays two tables:

Catalogue	#rows	#tables	#rows/table	#views	#rows/view	#keys	#indexes
disneydata	267	6	44,50	0	NaN	6	7

Table	#rows	#keys	#indexes
dossier	67	1	1
dossier_copie	67	1	1
footer	1	1	1
header	1	1	1
passes	125	1	1
passkind	6	1	2

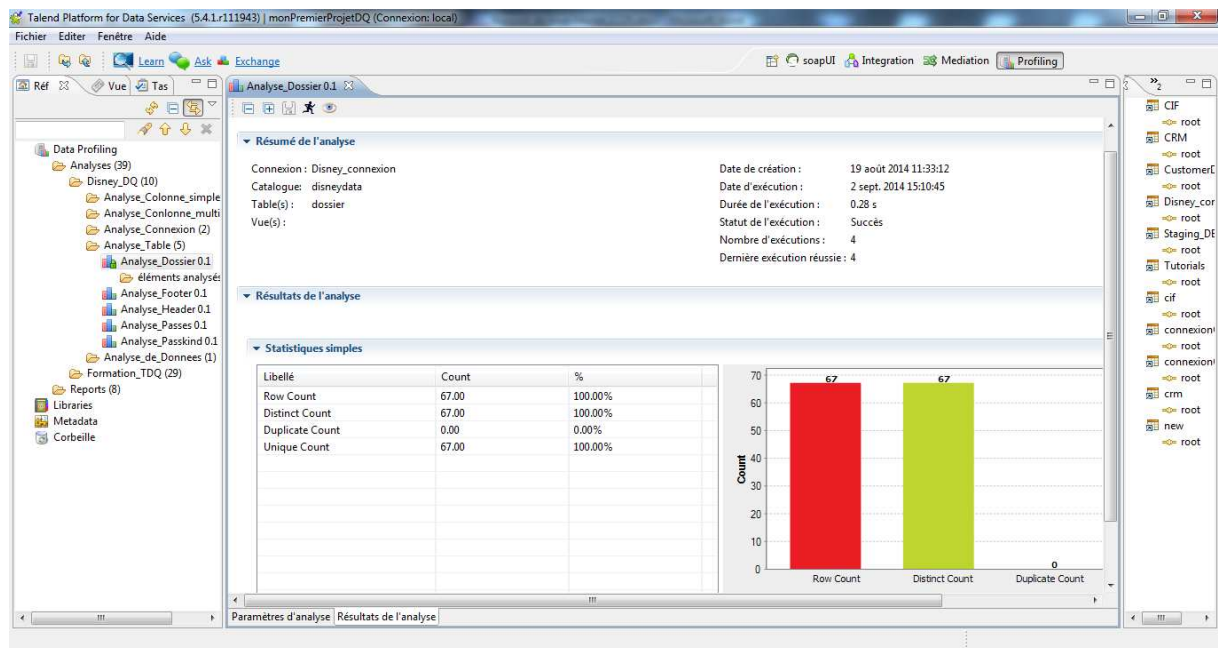
Figure. Analyse de connexion « Disney\_connexion »



*Figure. Rapport de connexion « Disney\_connexion »*

Par la suite, nous avons procédé par une vérification des données dans les tables de notre base, une vue de l'ensemble des colonnes puis en ajoutant des règles métier afin de détecter le contenu des colonnes qui correspond ainsi que celui qui ne correspond pas à notre règle métier afin de trouver des failles et les corriger.

Ci-après une illustration d'analyse de table simple « dossier » :

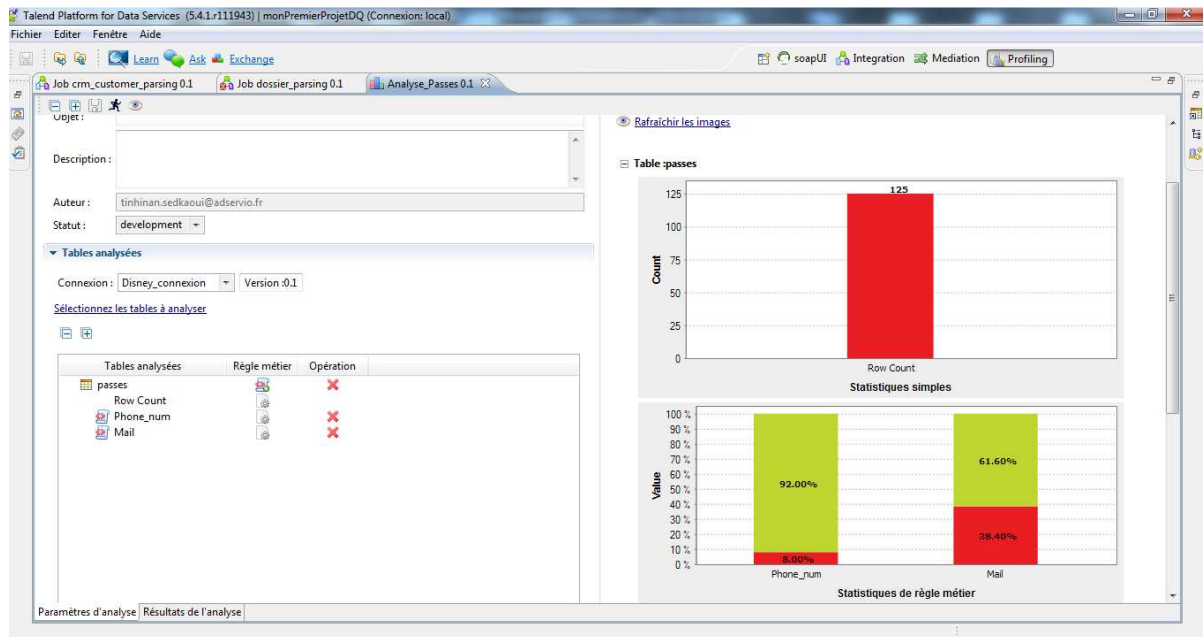


*Figure. Analyse de la table simple « dossier »*

Ceci dit que la table « dossier » a 67 lignes que les 67 lignes sont différentes et qu'il n'y a pas de double.



Une capture d'écran d'analyse de table avec une règle métier :



*Figure. Analyse de la table « passes » avec des règles métiers*

Ce qui veut dire que 92% des numéros de téléphone sont conformes à la règle métier mise comme condition au départ, mais que 08% des numéros ne respectent pas la règle.

Pareil pour les mails 61.60% respectent et 38.40% ne respectent pas avec la règle métier.

Cela nous a permis d'aller plus loin dans les analyses en faisant de l'analyse de colonne d'une ou de deux tables, comme exemple, vérifier la correspondance de la colonne « contract » des tables « dossier » et « passes », résultat trouvé : tous les contrats de la table « dossier » sont retrouvés dans la table « passes » et vis-versa. Ce qui confirme que dans un dossier, on retrouve un ou plusieurs passes.



En voici la capture d'écran :

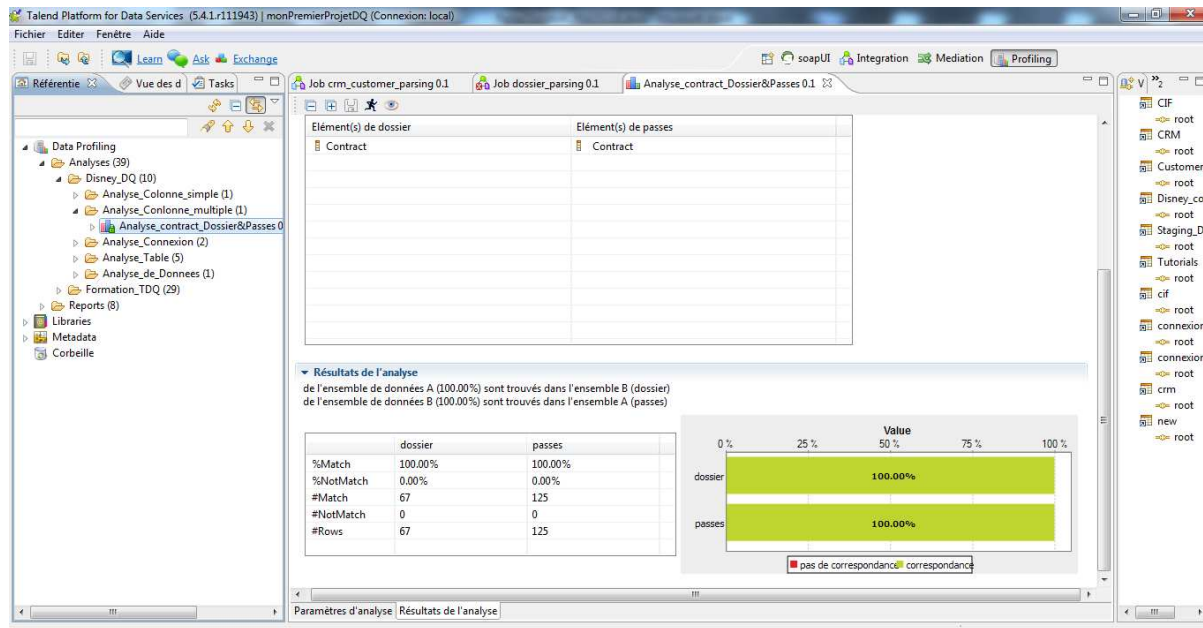
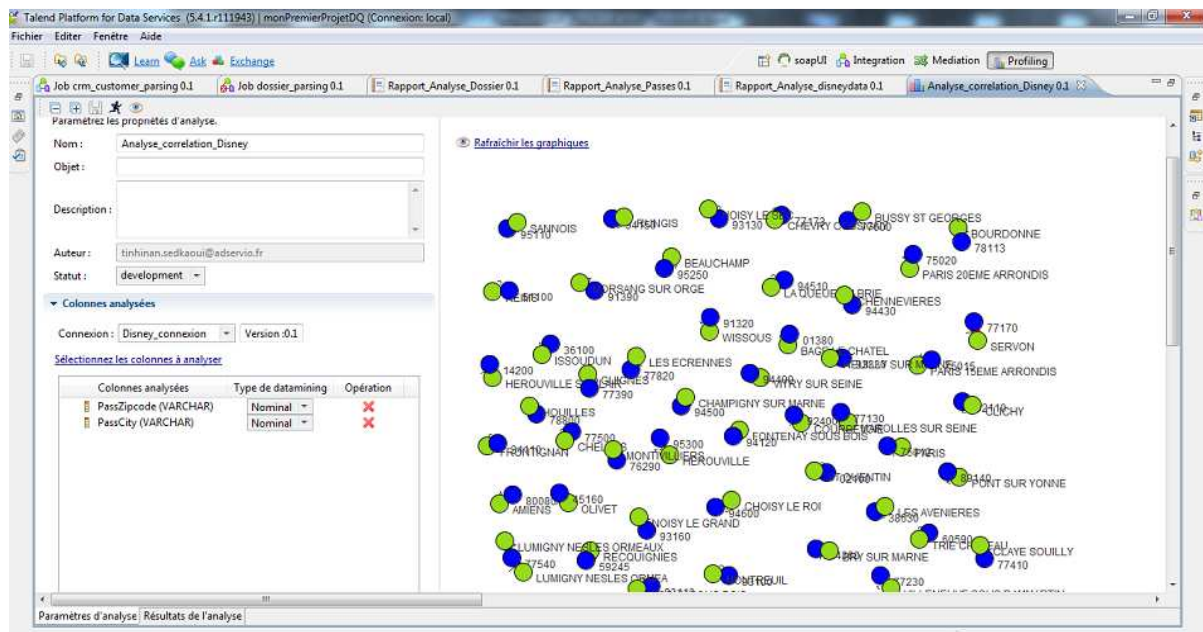


Figure. Analyse de la colonne « contract » dans les tables « dossier » et « passe »

Une analyse de corrélation entre le code postal des clients et leur ville, afin d'afficher les relations entre les valeurs, dans notre cas une analyse de corrélation sur les colonnes « PassZipcode » et « PassCity » de la table « passes ».

En voici la capture d'écran :



*Figure. Analyse de corrélation de données*

Comme exemple, on voit bien que le ZipCode 75020 correspond bien à la ville Paris 20ème arrondissement.

### III.5.3 Qlikview

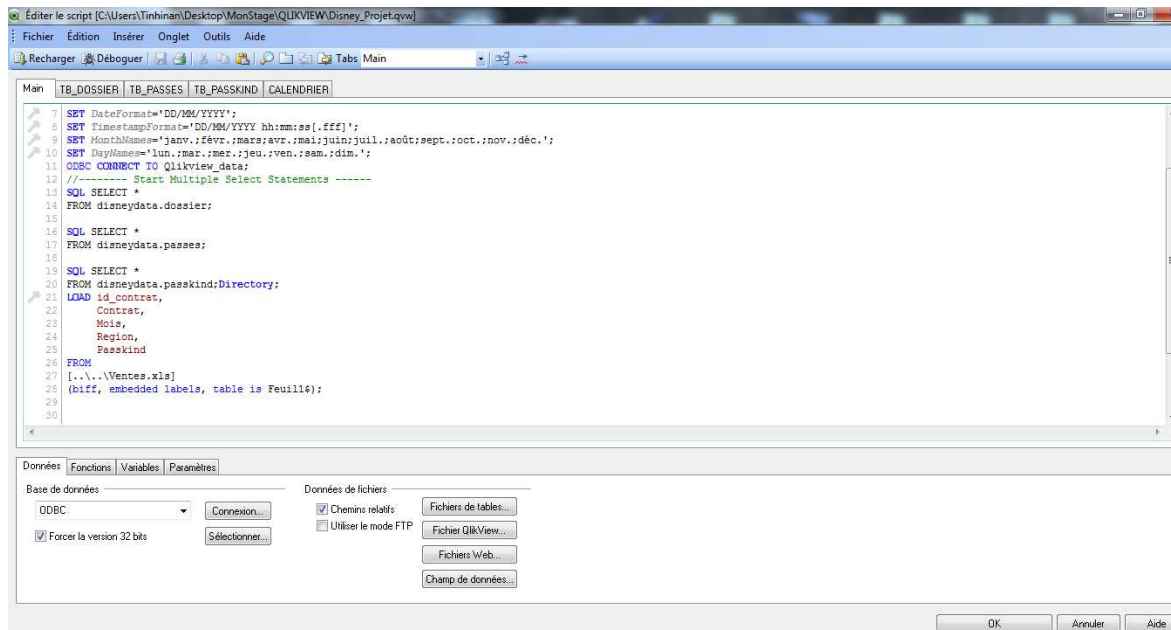
Dans cette partie, j'ai fait une dernière formation de trois semaines en QlikView qui est un outil décisionnel de reporting permettant la réalisation et la diffusion des tableaux de bord de reporting.

J'ai commencé par la découverte de cet outil, son utilisation et ses principes, une application Qlikview/la création d'un tableau de bord de reporting, nécessite des étapes à savoir le chargement de données vers cet outil pour les représenter dans une interface graphique utilisateur.

On a procédé par :

1. Le chargement des données de mes tables de la base « Galaxy » vers « QlikView ».

En voici une capture d'écran :



*Figure. Script chargement de données*

2. La création des listes de sélection contenant les champs ainsi que les indicateurs que je souhaite utiliser dans mon tableau de bord de reporting.
3. Une liste multiple est utilisée dans le cas où j'ai plusieurs listes de sélection.
4. La création des sélections actives mettant en avant la donnée en cours de traitement.
5. Zone de table qui est une nouvelle table regroupant les champs que je souhaite, même si ces derniers proviennent de différentes tables.
6. Graphe d'évolution (histogrammes+secteurs) me permet de faire le calcul d'évolution, dans ce cas, du nombre de contrats par mois, par région ainsi que par type du passe. J'ai aussi mis en avant un graphe d'évolution de la création du contrat par date de création.

Afin de réaliser ce tableau de bord, nous avons eu à utiliser des données (les données chargées) depuis les tables de la BDD « disneydata »,

Ci-dessous des captures d'écran du tableau de bord de reporting :

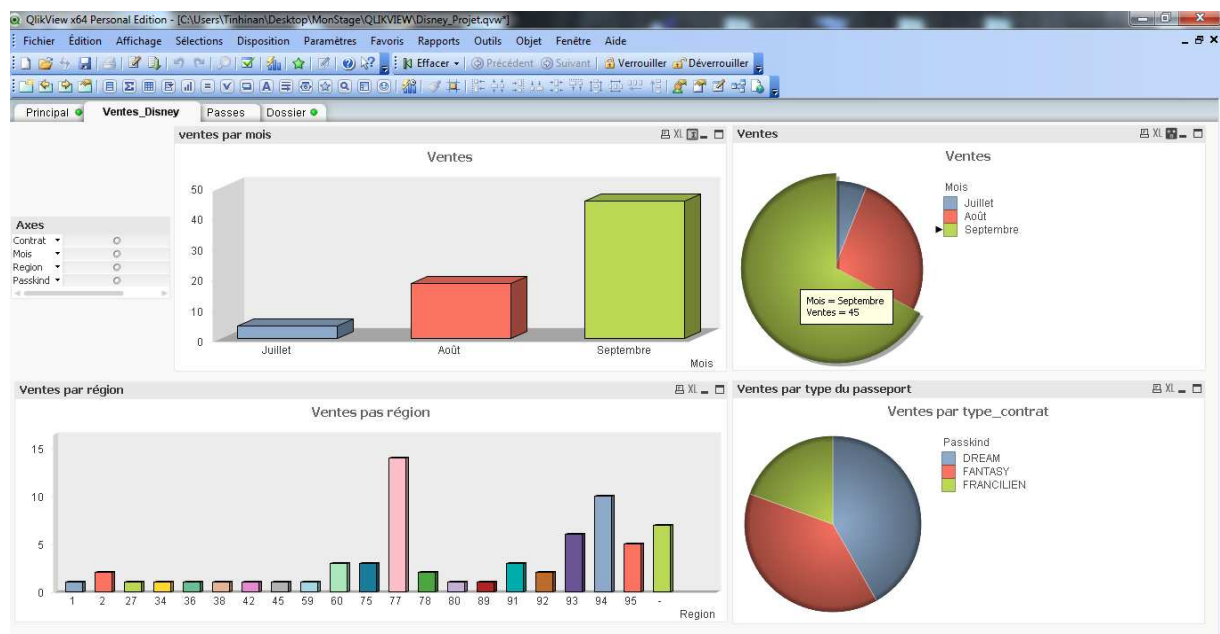


Figure. Tableau de bord d'évolution des ventes par mois, région et type\_contrat

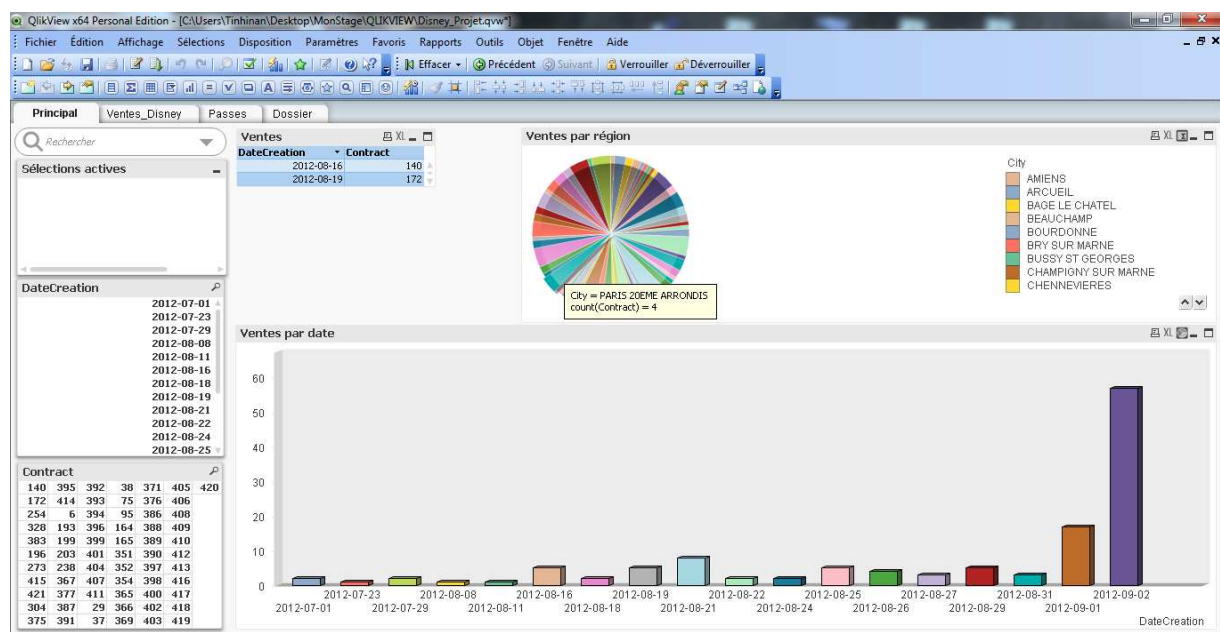


Figure. Vente par date de création

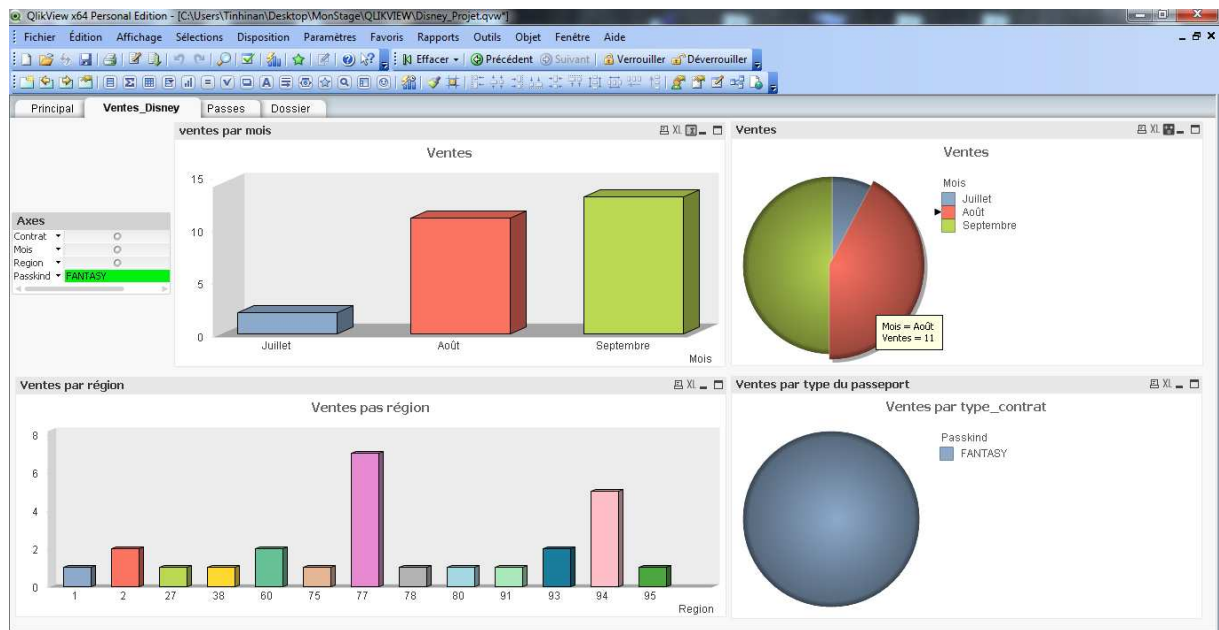


Figure. Evolution par type du contrat

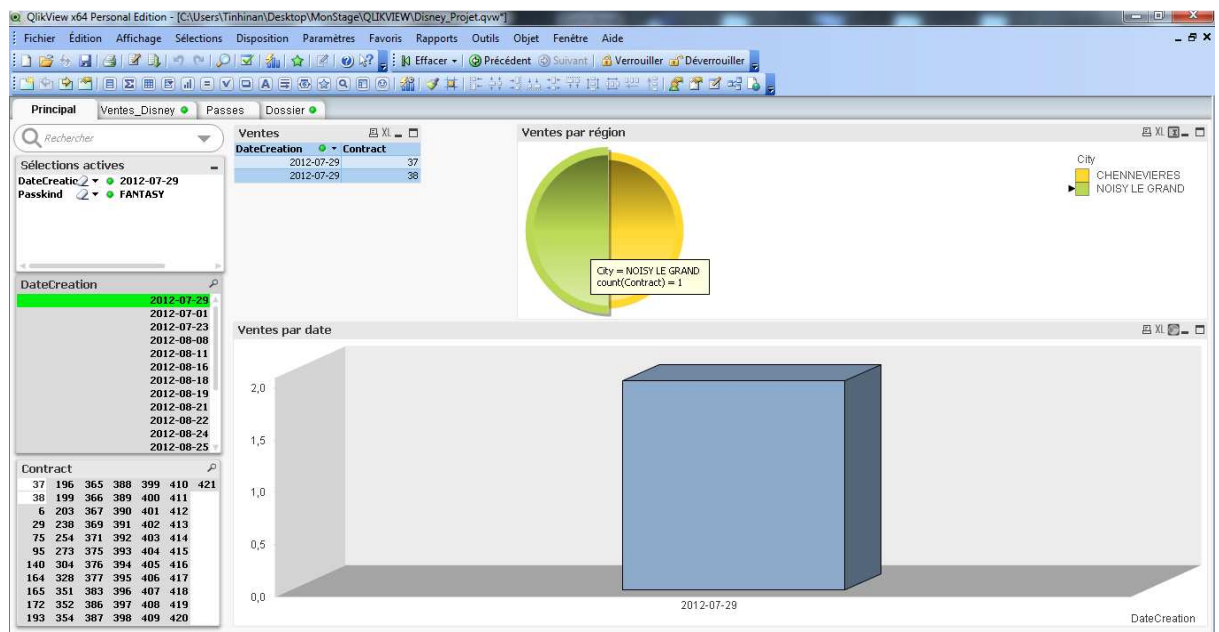


Figure. Evolution par date de création (29/07/2012)

## II.6. Conclusion

Dans le chapitre qui précède nous avons pu mettre en avant les travaux réalisés dans le cadre de ce mémoire de fin d'études, à savoir les formations effectuées, le travail d'intégration, d'extraction ainsi que l'analyse des données disney pour enfin les utiliser en les diffusant dans des tableaux de bord de reporting afin de détecter les périodes, types de passe les meilleurs pour faire des promotions et ainsi garder ou gagner de la clientèle.

Dans la suite de ce mémoire, nous trouverons dans le prochain chapitre les apports du stage.

# Apports du stage

## IV. Les apports du stage

---

### IV.1. Les apports pour l'entreprise

L'entreprise met un accent particulier dans la méthode de travail et la qualité des applications livrées. Ma présence, a été comme un renfort dans l'équipe suite à une croissance d'activité.

### IV.2. Les apports personnels

Ce stage, au sein de la société « Adservio » a été pour moi une occasion de monter en compétences techniques et de travailler sur des sujets réels en prenant en compte les exigences et les contraintes de chaque projet. Il m'a aussi permis un double enrichissement à la fois fonctionnel et technique, de me plonger dans un nouvel univers technologique où les bases de données, ainsi que les nouvelles technologies de développement occupent une place prépondérante.

Cette période en entreprise a donc confirmé les apports techniques que j'ai pu recevoir durant mon cursus universitaire, il m'a également permis de mieux me situer dans l'entreprise. Je me suis totalement intégré dans l'équipe du projet, mais également au sein de la société. Je peux désormais avoir un point de vue global sur la façon de conduire un projet, de la rédaction des spécifications à la mise en recette définitive d'une application en passant par la planification des différentes tâches à effectuer. Ces connaissances me donnent la volonté d'évoluer vers un poste supérieur à moyen terme.



# Conclusion

## Conclusion

---

Ce stage est d'une importance capitale. En effet, j'ai eu l'opportunité de faire des formations en « Talend DI & DQ » ainsi que « QlikView », d'apprendre à utiliser ces outils pour répondre à des besoins clients et de travailler sur un projet de cas réel (Projet APMP Disneyland Paris). Il m'a permis, de perfectionner mon niveau technique, d'approfondir mes connaissances, théoriques et pratiques.

Par ailleurs, travailler dans un environnement professionnel permet une approche rationnelle des résolutions des problèmes en environnement de développement : identification et résolution de problèmes sont intégrées au quotidien et l'entreprise permet de bénéficier de l'expérience d'autrui pour s'améliorer. Finalement, ce stage est important pour les aspects humains et relationnels : apprendre à écouter, discuter des problèmes et solutions, défendre des idées.

En d'autre terme ce stage m'a permis de mettre en application mes connaissances théoriques acquises durant mon cursus universitaire et d'acquérir des compétences professionnelles et fonctionnelles me permettant ainsi de bien définir mon projet professionnel ce qui me facilitera par la suite le passage à la vie active.

# Bibliographie



## BIBLIOGRAPHIE

---

- [1]: [adservio.fr](http://adservio.fr)
- [2]: [cigref.fr/big-data-vision-grandes-entreprises](http://cigref.fr/big-data-vision-grandes-entreprises)
- [3]: [georges.gardarin.free.fr/Cours\\_Total/15-Datawarehouse](http://georges.gardarin.free.fr/Cours_Total/15-Datawarehouse)
- [4]: [talend.fr](http://talend.fr)
- [5]: <http://www.piloter.org/business-intelligence/reporting.htm>