

République Algérienne Démocratique Et Populaire
Ministère de l'Enseignement Supérieur Et De La Recherche
Scientifique

Université Saad Dahleb Blida

Faculté des sciences

Département d'informatique



Mémoire de fin d'études pour l'obtention du

Diplôme de **Master en informatique**

Spécialité : Ingénierie des Logiciels

Thème

**Automatisation des flux de Reporting Techniques &
Comptables pour le bouclage du chiffre d'affaires d'un
opérateur de télécom mobile**

Réalisé par : Mme. YASMINE BAHAMID

Organisme d'accueil : KPMG Spa, Algérie

Devant le jury composé de :

Président : M.FERFERA Sofiane

Examinatrice : Mme .BACHA Sihem

Promoteur : M .CHIKHI Nacim Fateh

Encadrant : M. BELRAMOUL Amine

Promotion : 2020/2021

Remerciements & dédicaces

Je remercie tout d'abord le tout puissant qui, par sa grâce m'a permis d'arriver au bout de mes efforts en me donnant la santé, la force, le courage et en me faisant entourer des merveilleuses personnes dont je tiens à remercier. Je remercie :

Mon Superviseur, M.BELRAMOUL Amine Senior Manager IT Advisory à KPMG Algérie, pour son encadrement, sa rigueur au travail, ses orientations et sa disponibilité.

Mon encadrant, M.CHIKHI Nacim Fateh, pour ses contrôles, ses orientations, ses relectures, ses corrections et son attention bienveillante dans ce travail.

Je tiens à remercier sincèrement KPMG Algérie en particulier l'équipe IT Advisory de m'avoir accueilli chaleureusement parmi eux et pour leurs soutien et encouragement.

Je tiens à remercier également les membres du jury qui me font le grand honneur d'évaluer ce travail.

Ainsi que tous ceux qui de près ou de loin ont contribué à l'accomplissement de ce travail.

Je dédie ce travail à mes parents, mes sœurs Lydia et Sara, sans oublié mes grands-parents Jedi et Zizou, mes tantes et mes oncles qui n'ont cessés de m'encourager et m'ont soutenus jusqu'aux derniers moments.

BAHAMID.Y

Résumé

Le processus de bouclage de chiffre d'affaires des opérateurs télécom est un processus complexe, qui inclut la comptabilisation et la réconciliation des revenus et revenus différés et qui malgré sa criticité se fait jusqu'à présent manuellement.

L'objectif de notre travail consiste à concevoir une solution logicielle pour l'automatisation des flux et traitements de données entrants dans ce processus, afin de diminuer le risque d'erreur.

Notre solution, est une plateforme de Data Analytics basée sur la programmation visuelle, permettant de créer des workflows de traitement.

Chaque workflow sera capable de se connecter à différentes sources de données afin d'importer les données entrantes dans le processus de contrôle, à effectuer différents traitements de nettoyage, transformation et extraction sur les données importées et enfin visualiser et exporter les résultats vers d'autres destinations.

Le workflow généré pourra, par la suite, être réutilisé et adapté.

Ce travail a été réalisé dans le cadre d'un stage de fin d'étude, au sein de KPMG spa Algé

Abstract

The revenue closure process of telecom operators is a complex process, which includes the accounting and reconciliation of revenues and deferred revenues and which, despite its criticality, is still done manually.

The objective of our work is to design a software solution for the automation of the incoming data flows and treatments in this process, in order to decrease the risk of errors.

Our solution is a Data Analytics platform based on visual programming, allowing to create processing workflows.

Each workflow will be able to connect to different data sources in order to import the incoming data into the control process, to perform different cleaning, transformation and extraction processes on the imported data and finally to visualize and export the results to other destinations.

The generated workflow can then be reused and adapted.

This work was carried out within KPMG spa Algeria.

ملخص

تعتبر عملية إغلاق رقم الأعمال لمشغلي الاتصالات عملية معقدة، والتي تشمل الاعتراف وتسوية الدخل والدخل المؤجل، والتي على الرغم من أهميتها تم تنفيذها يدويًا حتى الآن.

الهدف من عملنا هو تصميم حل برمجي لأتمتة تدفق ومعالجة البيانات التي تدخل هذه العملية، من أجل تقليل مخاطر الخطأ.

الحل الذي نقدمه هو نظام أساسي لتحليلات البيانات يعتمد على البرمجة المرئية، مما يسمح بإنشاء عمليات سير عمل المعالجة.

سيكون كل سير عمل قادرًا على الاتصال بمصادر بيانات مختلفة من أجل استيراد البيانات الواردة إلى عملية التحكم، لإجراء معالجات مختلفة: تنظيف وتحويل واستخراج على البيانات المستوردة وأخيراً لتصور النتائج وتصديرها إلى جهات أخرى.

يمكن بعد ذلك إعادة استخدام وتكييف سير العمل الذي تم إنشاؤه.

تم تنفيذ هذا العمل كجزء من تدريب نهاية الدراسة في KPMG spa Algeria .

Sommaire

Remerciements & dédicaces	2
Résumé	3
Abstract	4
ملخص	5
Sommaire	6
Table des figures.....	8
Liste des tableaux	10
Liste des abréviations	11
Introduction générale	12
Première partie : État de l'art	14
Chapitre I : Revenue assurance	15
1 Introduction.....	15
2 Présentation de l'organisme d'accueil.....	16
3 Marché télécom en Algérie.....	19
4 Notions en comptabilité	20
5 Revenue assurance	22
6 Conclusion.....	30
Chapitre II : Data analytics.....	31
1 Introduction.....	31
2 Data analytics	32
3 Plateforme de Data analytics	37
4 Conclusion.....	39
Chapitre III : Programmation basée sur le flux	40
1 Introduction.....	40
2 Diagramme de flux de données	41
3 Programmation basée sur le flux :	43
4 Workflow.....	48
5 Conclusion.....	50
Deuxième partie : Solution.....	51
Chapitre I : Conception de la solution.....	52
1 Introduction.....	52
2 Description de l'approche.....	53

3	Modélisation <i>de l'approche</i>	54
4	Conclusion :	64
	Chapitre II : Implémentation de la solution	65
1	Introduction	65
2	Outils de Développement	66
3	Développement du Système :	69
4	Test de l'application	84
5	Conclusion	90
	Conclusion Générale et Perspectives	91
	Références bibliographiques	93

Table des figures

Figure 1 : KPMG dans le monde.

Figure 1.1 : Organigramme de KPMG Algérie.

Figure 1.2 : Répartition des abonnés par opérateur et par type d'abonnement pour le 4ème trimestre de 2020.

Figure 1.3 : Processus de génération de revenu Télécom.

Figure 1.4 : Norme IFRS 15.

Figure 2.1 : Les types de la DA.

Figure 2.2 : Processus DA.

Figure 2.3 : Transformation de données.

Figure 2.4 : exemple de tableau de bord.

Figure 3.1 : DFD d'un système de livraison [1].

Figure 3.2 : nœud PBF.

Figure 3.3 : Exemple de Composant-Composite.

Figure 3.4 : Connexion FIFO.

Figure 3.5 : Paquet d'Information.

Figure 3.6 : Processus de transformation de lait.

Figure 4.1 : Diagramme des Cas d'utilisation.

Figure 4.2 : Diagramme de séquence Système « Importer Données ».

Figure 4.3 : Diagramme de séquence Système « Traiter Données ».

Figure 4.4 : Diagramme de séquence Système « Visualiser Données ».

Figure 4.5 : Diagramme de séquence Système « Exporter Données ».

Figure 4.6 : Diagramme de Classes de Conception.

Figure 5.1 : Logo Python.

Figure 5.2 : Logo PyCharm.

Figure 5.3 : Logo Pandas

Figure 5.4 : Logo PyQt5

Figure 5.5 : Logo NumPy.

Figure 5.6 : Logo Matplotlib.

Figure 5.7 : interface graphique Système.

Figure 5.8 : Ajout de Nœuds.

Figure 5.9 : Insertion des arcs.

Figure 5.10 : Enregistrement d'un WF vide.

Figure 5.11 : Exécution du WF.

Figure 5.12 : Structure de Data frame.

Figure 5.13 : Fenêtre de dialogue du nœud Tableau.

Figure 5.14 : Fenêtre de dialogue du nœud Base de données.

Figure 5.15 : Fenêtre de dialogue du nœud Sélection.

Figure 5.16 : Fenêtre de dialogue du nœud projection.

Figure 5.17 : fenêtre de dialogue du Nœud Nettoyage.

Figure 5.18 : fenêtre de dialogue de Renommer Colonnes.

Figure 5.19 : Fenêtre de dialogue du Nœud Remplacer.

Figure 5.20 : fenêtre de dialogue du nœud Join.

Figure 5.21 : fenêtre de dialogue du nœud Union.

Figure 5.22 : fenêtre de dialogue du nœud Opération.

Figure 5.23 : Fenêtre de dialogue du Nœud Graphe.

Figure 5.24 : Fenêtre de dialogue du Nœud Graphe.

Figure 5 : Workflow de Test.

Figure 5.25 : résultat de l'importation de données.

Figure 5.26 : Résultat du nettoyage de données.

Figure 5.27 : Résultat obtenu après extraction de la date.

Figure 5.28 : Résultat de la requête 4.1.

Figure 5.29 : Résultat de la requête 4.2.

Figure 5.30 : Résultat du nœud 5.1.

Figure 5.31 : Résultat du nœud 5.2.

Figure 5.32 : Résultat du de la requête 6.1.

Figure 5.32 : Résultat du de la requête 6.2.

Figure 5.33 : Résultat du tri.

Figure 5.34 : Résultat du de l'union.

Figure 5.35 : Résultat de la différence.

Figure 5.37 : Résultat du de la jointure.

Figure 5.38 : Résultat du nœud 12.

Liste des tableaux

Tableau 1.1 : Nombre d'abonnés par opérateur pour les quatre trimestres de 2020.

Tableau 1.2 : Journal Comptable.

Tableau 1.3 : Exemple d'outils RA.

Tableau 2.1 : Exemple de plateformes de DA.

Tableau 3.1 : Exemple Comparatif Algorithmique Vs FBP.

Tableau 3.2 : Systèmes Basé sur le Flux.

Tableau 4.1 : Description textuelle « Importer Données ».

Tableau 4.2 : Description textuelle « Traiter Données ».

Tableau 4.3 : Description textuelle « Visualiser Données ».

Tableau 4.4 : Description textuelle « Exporter Données ».

Tableau 4.5 : Description textuelle « Gérer Workflow ».

Tableau 5.1 : Caractéristiques de la machine.

Liste des abréviations

Chapitre I : Revenu assurance

PwC : PricewaterhouseCoopers, un des Big *four* cabinet d'audit mondial.

EY : Ernst & Young, un des Big four cabinet d'audit mondial.

ARPCÉ : AUTORITÉ DE RÉGULATION DE LA POSTE ET DES COMMUNICATIONS ÉLECTRONIQUES

RA : Revenu Assurance.

Air Time : Durée d'utilisation des services télécoms (monnaie, temps, data ou unité

DFC : Directeur des Finances et de la Comptabilité

TPE : Terminal de paiement électronique

DPST : Direction Produit Service & Taxation

Voucher : Voucher Physique (Carte de Recharge)

e-Voucher : Voucher Electronique

PDV : Point De Vente

PCA : Produit Constaté d'Avance

Chapitre II : Data analytics

DA : Data Analytics

XML: Extensible Markup Language

NoSQL: Not Only SQL

JSON : JavaScript Objet Notation

Chapitre III : Programmation basée sur le flux

DFD : diagramme de flux de données

PBF : programmation/ programme basée sur le flux

Introduction générale

Contexte

L'introduction des smartphones a entraîné une augmentation soudaine de l'utilisation des données (DATA), cependant les opérateurs télécom n'ont pas mis en place les plans de tarification et les mécanismes adéquats pour en tirer profit. De plus, ces opérateurs sont confrontés à la perte de revenus, ce qui n'est pas surprenant étant donné les très gros volumes de données à traiter, la complexité de l'évolution des éléments de réseau, des multiples plateformes d'applications et des offres de produits et de services.

Cependant, les opérateurs de télécommunications utilisent une multitude de techniques et de méthodes pour identifier et recouvrir les pertes de revenus et pour détecter et prévenir les erreurs qui entraînent des revenus non facturés ou non recouverts.

Problématique

Revenue assurance est l'utilisation des données et méthodes d'amélioration des processus pour améliorer les bénéfices, les revenus et les flux de trésorerie sans influencer la demande du marché. Aussi, les états financiers doivent également réconcilier ce qui a été comptabilisé en tant que revenu avec ce qui a été reconnu comme chiffre d'affaire et cela se fait le plus souvent *manuellement*, avec un risque d'erreur qui peut avoir un impact sur le bilan de l'opérateur de manière très significative avec d'importantes implications fiscales.

Objectifs

Les outils et plateformes de Revenue assurance disponibles sur le marché offrent pratiquement les mêmes services en termes de rapprochements de flux et d'abonnés, et fournissent une visibilité détaillée sur les différents écarts entre tous les nœuds du réseau de télécommunication d'un opérateur, et tout cela de manière automatique. Cependant, ces outils ne traitent pas le rapprochement entre les

revenus générés et ceux reconnus en chiffre d'affaires, qui est une étape critique dans le processus de revenue assurance et qui se fait manuellement.

Le but de ce travail est de concevoir une plateforme de Data analytics qui se connectera aux différentes sources d'information ou d'importation permettant ainsi de produire des rapports indispensables pour la validation des comptes revenus et revenus différés d'un opérateur télécom mobile et ce de manière **automatique**. Avec cet outil, il sera possible de créer des workflows complets pour le contrôle et traitement de fichiers financiers entrant dans le processus de bouclage du chiffre d'affaires de l'opérateur télécom.

Organisation du mémoire

Ce document est organisé essentiellement en deux parties distinctes. La première partie intitulée **État de l'art** est composée de trois chapitres.

Dans le premier chapitre intitulé **Revenue assurance** nous présenterons l'organisme d'accueil, nous aborderons aussi le marché télécom en Algérie, et la procédure de bouclage du chiffre d'affaires afin de mieux cerner le besoin. Dans le second chapitre intitulé **Data analytics** nous aborderons l'analytique de données en tant que révolution dans le monde de la science de données, son processus et son impact sur les organisations. Le troisième chapitre dénommé **Programmation basée sur le flux** traitera quant à lui ce nouveau paradigme de programmation, son principe, ainsi que son lien avec le principe d'automatisation.

La deuxième partie de ce mémoire intitulée **Solution** est déclinée en deux chapitres. Dans le premier intitulé **Conception de la solution** nous exposons une description suivie d'une modélisation de notre solution pour l'automatisation des flux de reporting.

Le deuxième et dernier chapitre dénommé **Implémentation de la solution** sera une présentation de l'aspect technique et de l'implémentation de l'outil en citant les différents modules et technologies utilisés. Puis nous clôturons ce chapitre ainsi que le travail avec une démonstration de l'outil.

Première partie

État de l'art

- Chapitre I : Revenue assurance
- Chapitre II : Data analytics
- Chapitre III : Programmation basée sur le flux

Chapitre I : Revenu assurance

1 Introduction

De nos jours, le terme revenu assurance est fréquemment utilisé notamment dans le domaine de l'industrie télécom. Le revenu assurance a toujours été un problème pour les entreprises de télécommunications notamment depuis la vulgarisation des smartphones, le suivi des impulsions, des minutes, des comptes, des octets, etc. n'a jamais été aussi difficile.

Dans la hâte de mettre les nouvelles technologies sur le marché, les systèmes de revenu assurance sont toujours à la traîne. L'assurance des revenus dans un environnement de télécommunications couvre un large éventail d'aspects techniques et commerciaux [8].

Dans ce chapitre, nous allons présenter brièvement l'organisme d'accueil ainsi que l'environnement télécom en Algérie, suivi de quelques notions en comptabilité qui nous seront utiles pour comprendre le revenu assurance et particulièrement le processus de comptabilisation du chiffre d'affaires des opérateurs.

2 Présentation de l'organisme d'accueil

Ce projet de fin d'étude a été réalisé au sein de la Structure « IT Advisory » du cabinet « KPMG Algérie Spa ». Nous donnons ci-après quelques chiffres le concernant ainsi que les services et activités qu'il propose à ses clients [7].

2.1 KPMG International

KPMG (Klynveld, Peat, Marwick, Goerdeler) est l'un des réseaux globaux d'audit et de conseil les plus étendus dans le monde, KPMG est l'un des quatre plus grands réseaux de cabinets internationaux d'audit et de conseil, faisant partie de ce qu'on appelle le « Big Four » avec Deloitte, PwC, et EY [7].

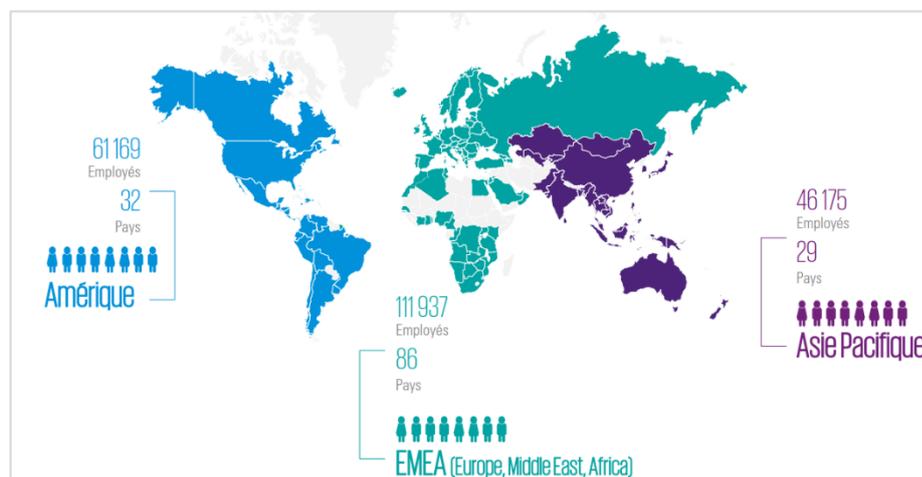


Figure 1 : KPMG dans le monde

KPMG est considéré aujourd'hui comme une référence mondiale dans le cercle des affaires, elle est présente dans 147 pays avec près de 220 000 collaborateurs, tous affiliés à KPMG International, société de droit anglais. Ses cabinets membres sont légalement distincts et indépendants les uns des autres.

2.2 KPMG Algérie

KPMG Algérie Spa, membre algérien du réseau KPMG International Coopérative, a été l'un des premiers cabinets de conseil et d'audit à exercer en Algérie. Elle existe sous la forme d'une société de droit Algérien depuis 2002 [7].

KPMG Algérie Spa, résume le savoir-faire international conjugué à une connaissance approfondie de l'environnement économique algérien.

En mai 2009, et pour se rapprocher de ses clients de l'ouest, KPMG Algérie a ouvert un bureau à Oran, puis en 2014, elle a inauguré son nouveau siège au centre des affaires à Bab Ezzouar, qui représente une concrétisation de son engagement.

Le cabinet KPMG Algérie compte plus de cent cinquante experts formés aux normes universelles du label KPMG dans les disciplines de la finance, du conseil, de l'audit, des systèmes d'information et du droit des entreprises.

2.3 Organigramme de l'entreprise

Le cabinet KPMG Algérie est organisé en cinq départements [7] :

- i. Le Département Audit.
- ii. Le Département Conseil (Advisory).
- iii. Le Département Fiscal et Juridique (Tax & Legal).
- iv. Le Département d'assistance en comptabilité (AAS : Accountancy Assistance Services).
- v. Le Département Administration et finance : Organisé en trois services : Comptabilité et Finances, Ressources Humaines et Moyens Généraux-Informatique, en plus du secrétariat.

La direction de chaque département est confiée à un Associé (Partner). Le Partner Président Directeur Général est à la tête des Départements d'administration et d'audit. Ce dernier se charge aussi du bureau d'Oran dont la principale mission est de promouvoir les services de KPMG dans la région.

L'organigramme suivant permet de mieux visualiser ce découpage.

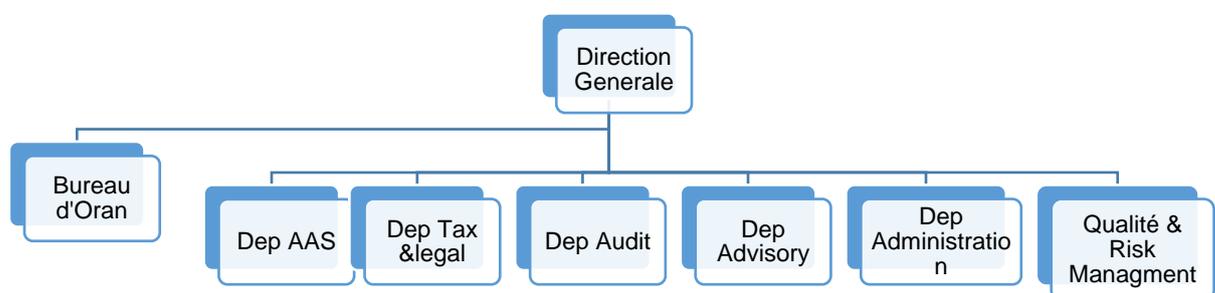


Figure 1.1 : Organigramme de KPMG Algérie

2.4 Activités de KPMG Algérie

Parmi les activités de KPMG Algérie on trouve [7] :

- i. **Audit contractuel** : KPMG Algérie apporte à ses clients une révision intègre et objective de la situation financière de leurs entreprises, ce qui leurs permet d'améliorer l'efficacité de leurs activités.
- ii. **Prestations fiscales** : KPMG assure pour ses clients :
 - Revue des déclarations fiscales et Audit fiscal.
 - Revue fiscale dans le cadre de l'exécution des contrats.
 - Assistance à la mise en place de procédures de conformité.
- iii. **Prestations juridiques** : KPMG accompagne ses clients dans le cadre de
 - Création de sociétés et Revue juridique de conformité
 - Études sur les formes d'implantation et la réalisation d'une opération
 - Assistance en matière contractuelle et Secrétariat juridique
- iv. **Management Consulting et Risk consulting** :
 - La définition et la conduite de projets de transformation et d'amélioration de la performance.
 - Mise en place de dispositif de contrôle interne et de management des risques.
- v. **Conseil informatique (IT Advisory)** :
 - Audit des systèmes d'informations.
 - Cahier des charges et choix de solutions informatiques.
 - Le conseil en management des risques.
 - Accompagnement dans la mise en œuvre.
 - Cyber sécurité.
 - Contrôle interne IT et attestation.
 - Accompagnement de la DSI.

Parmi les départements cités, la problématique que nous traitons apparaît au niveau du département Advisory, plus précisément dans les services d'IT Advisory, dans le cadre de l'audit et contrôle.

3 Marché télécom en Algérie

3.1 Acteurs

Ci-dessous le Tableau 1.1 montrant le nombre d'abonnés actifs par opérateur télécom pendant les quatre trimestres de l'année 2020 [18] :

	1er trimestre 2020	2eme trimestre 2020	3eme trimestre 2020	4eme trimestre 2020
Algérie Télécom Mobile (Mobilis)	18 874 336	18 654 330	18 757 780	18 974 678
Optimum Télécom Algérie (Djezzy)	14 224 144	13 952 347	14 473 544	14 363 102
Wataniya Télécom Algérie (Ooredoo)	12 044 478	11 805 053	11 990 227	12 217 893
Total	45 142 958	44 411 730	45 221 551	45 555 673

Tableau 1.1 : Nombre d'abonnés par opérateur pour les quatre trimestres de 2020

3.2 Services

On distingue deux services télécom qui sont les plus utilisés en Algérie :

- i. GSM : Système cellulaire numérique de communication avec des mobiles ou entre mobiles, destiné principalement aux communications téléphoniques (voix) aussi appelé 2G [9].
- ii. 3G/4G : La troisième et quatrième génération de systèmes cellulaires sont des générations de systèmes mobiles qui permettent des services de communication plus rapides, notamment pour la voix, la télécopie et l'Internet de n'importe quel endroit et à tout moment [9].

D'après l'ARPCE, sur les 45,556 millions d'abonnés actifs en 2020, 6,783 millions sont des abonnés au réseau GSM soit 15% contre 38,773 millions abonnés au réseau 3G/4G, soit 85% [4].

3.3 Types d'abonnement

- i. Abonnement Prépayé : le paiement est effectué à l'avance avant l'utilisation du service, autrement dit payer puis consommer.
- ii. Abonnement Postpayé : utilisation des services proposés tout au long d'une période définie (un mois par exemple), le client reçoit la facture à la fin de la période, c'est consommer puis payer.

D'après l'ARPCÉ, la répartition des abonnés par opérateur et type d'abonnement, où on remarque que 95,05% des abonnés ont adopté le prépayé, contre 4,95% pour le postpayé [4].

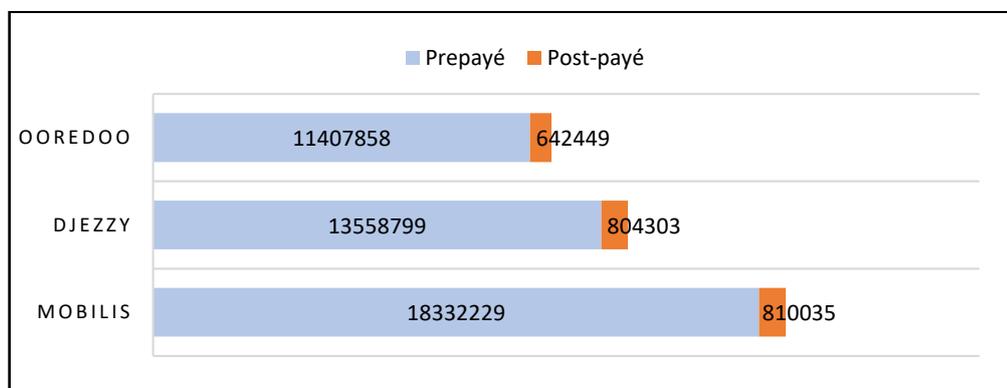


Figure 1.2 : Répartition des abonnés par opérateur et par type d'abonnement pour le 4ème trimestre de 2020

4 Notions en comptabilité

4.1 Définition d'un revenu

C'est le flux entrant brut de profits économiques (espèces, créances, autres actifs) résultant des activités opérationnelles ordinaires d'une entité (les ventes de biens, les ventes de services, les intérêts, les redevances et les dividendes) [8].

4.2 Définition d'un chiffre d'affaires

Le chiffre d'affaires (ou CA) d'une entreprise est la somme des ventes des biens ou des services facturés sur un exercice, autrement dit la somme des revenus.

Il peut être exprimé en hors taxes (HT) ou en toutes taxes comprises (TTC) s'il inclut la TVA.

Selon le besoin, il pourra être intéressant d'observer plusieurs indicateurs, basés sur le chiffre d'affaires, qui permettront d'interpréter la performance de l'entreprise (exemple : sa part du marché, son évolution,...etc.)

Le calcul du chiffre d'affaires peut être réalisé à la période souhaitée : le jour, la semaine, le mois, le trimestre, le semestre ou l'année [14] .

Voici la formule permettant de calculer le chiffre d'affaires :

$$CA = \sum (\text{Quantités vendues} \times \text{prix de vente unitaire})$$

4.3 Définition d'un revenu différé

Les revenus différés, également appelés revenus non gagnés ou produits constatés d'avance (PCA), font référence aux paiements anticipés (prépaiement) qu'une entreprise reçoit pour des produits ou des services qui doivent être livrés ou rendus à l'avenir.

L'entreprise qui reçoit le paiement anticipé enregistre le montant en tant que revenu différé, un passif sur son bilan, car ils reflètent des revenus qui n'ont pas été gagnés mais qui sont dus à un client [15].

Une fois que l'entreprise a fourni le service ou le produit, l'écriture du journal consiste à débiter ou réduire le compte de passif et à créditer ou augmenter le compte revenu. Et c'est là que les revenus différés se transforment en « revenus réels ». (Voir *Tableau 1.2*)

	Cash	Revenu	Revenu Différé
Paiement + Service non rendu	1500 DZD	0 DZD	1500 DZD
Paiement + Service rendu		1500 DZD	0 DZD

Tableau 1.2 : Journal Comptable

5 Revenu assurance

5.1 Définition

Revenu Assurance (RA) est un processus permettant à un fournisseur de service de communication (opérateur mobile) de vérifier et de contrôler l'exhaustivité, l'exactitude et l'intégrité des données du processus opérationnel de bout en bout de la capture, à l'enregistrement, à la facturation et du rapport de tous les événements ou transactions génératrices de revenus, depuis les ventes et le marketing jusqu'au recouvrement, en passant par les comptes clients [1] .

Un processus de revenu assurance peut ainsi contribuer à augmenter les bénéfices, à réduire les gaspillages et à améliorer l'efficacité opérationnelle de l'entreprise.

5.2 RA et les organisations Télécom

Bien que le revenu assurance ait toujours fait partie du jargon des télécommunications, elle est récemment passée au premier plan des préoccupations des directions télécoms. Cela est dû à plusieurs facteurs, notamment [2] :

- i. Profit : pression croissante sur les coûts et diminution des marges. L'époque des profits élevés pour la plupart des opérateurs de télécommunications est révolue. Ils doivent tous trouver d'autres moyens d'augmenter leurs marges en assurant un suivi efficace de leurs revenus.
- ii. La réglementation : La nouvelle structure réglementaire et les exigences de conformité qui obligent les opérateurs de télécommunications à déclarer leurs revenus avec précision.
- iii. Innovation technologique : S'assurer que les nouvelles technologies et les nouveaux produits fonctionnent selon les plans perçus. Suivre les nouvelles technologies et coexister avec les systèmes existants.
- iv. Fusions et acquisitions : Avec l'augmentation des fusions et acquisitions dans le secteur des télécommunications, les organisations ont du mal à gérer plusieurs systèmes, notamment la facturation, la médiation, etc.

5.3 Processus de génération de revenu Télécom

La facturation des télécommunications est un processus qui consiste à collecter l'utilisation, à appliquer les frais d'utilisation, et enfin à générer les factures. Ce processus comprend également la réception et l'enregistrement des paiements des clients [10].

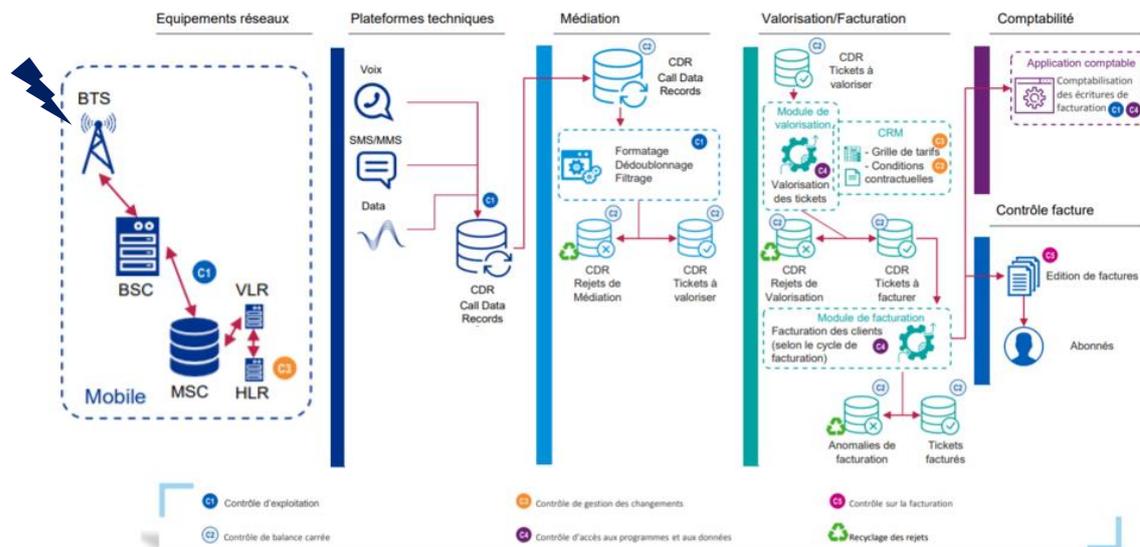


Figure 1.3 : Processus de génération de revenu Télécom

En se référant à l'architecture de la *Figure 1.3* ci-dessus, on distingue que le processus de génération des revenus télécom passe par les étapes suivantes :

Étape 1 : Après qu'une utilisation a été générée par le client final (appel passé par exemple), les données d'utilisation sont capturées auprès du commutateur réseau (couche d'Équipements Réseaux).

Étape 2 : Les plateformes techniques établissent un enregistrement des détails de l'appel dans un fichier nommé CDR. Ce CDR doit contenir le numéro de la partie "A" et le numéro de la partie "B", ainsi que les dates et heures de début et de fin de l'utilisation et le type du service utilisé.

Étape 3 : le processus de médiation, stocke par la suite les CDRs afin qu'ils puissent être évalués. Pour évaluer un appel par exemple, le CDR est examiné pour voir si l'appel est, un numéro vert, un appel local couvert par un plan d'appel local, un appel international ou un appel payant. Des informations telles que l'heure de l'appel et le code de la ville ou du pays d'où il a été passé sont utilisées pour calculer le tarif de l'appel.

Étape 4 : Une fois que chaque service est évalué, ces informations sont stockées jusqu'à ce que la facture soit établie, généralement une fois par mois (pour les abonnements). Lorsque la facture est établie, d'autres frais de non-utilisation, tels que des remises ou des frais mensuels, peuvent être appliqués à la facture.

Il peut y avoir un tarif réduit ou une remise sur la durée de facturation, différents paiements effectués par les clients, différents ajustements donnés, toutes ces informations contribuent à la génération de la facture finale, entre autres le revenu généré.

5.4 Reconnaissance des revenus

Le principe général de la reconnaissance est de comptabiliser le revenu sur la base du transfert des biens ou services promis aux clients. Le montant du revenu doit refléter la contrepartie qu'ils s'attendent à recevoir en échange de ces biens ou services. Le revenu est comptabilisé [8]:

- i. De manière continue, d'une façon qui représente la performance de la société.
- ii. A une date donnée dans le temps, quand le contrôle des biens ou services est transféré au client.

5.4.1 Vérification de La norme IFRS 15

IFRS 15 est une norme internationale sur la comptabilisation du chiffre d'affaires qui a été publiée en Mai 2014. Elle concerne les secteurs de l'économie, industrie et services, financiers et non financiers [16].

La *Figure 1.4* illustre les principales étapes requises par IFRS 15 pour comptabiliser le chiffre d'affaires résultant de la signature d'un contrat avec un client.

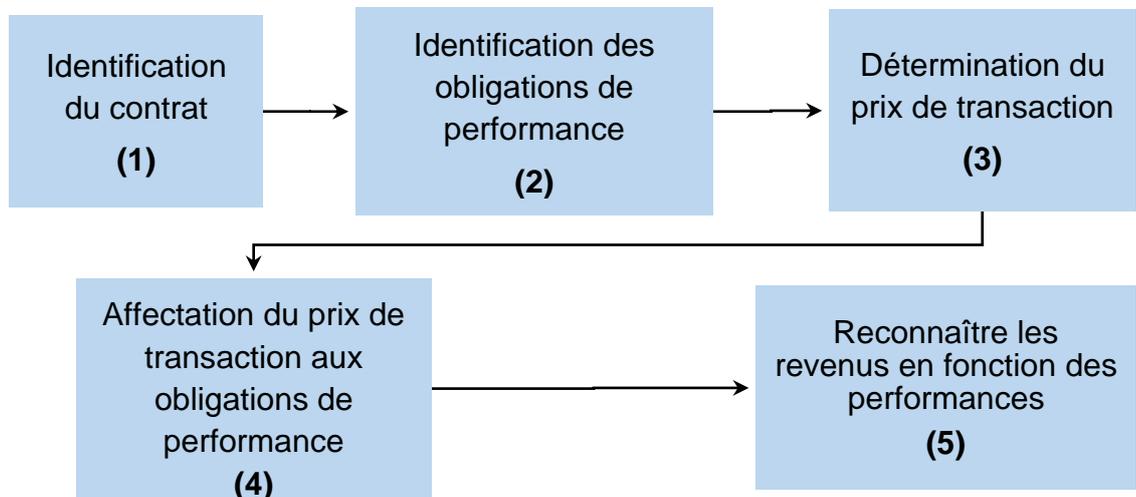


Figure 1.4 : Norme IFRS 15

i. Identification des contrats

Les contrats peuvent être écrits, oraux ou implicites, et nécessitent également de remplir tous les critères suivants :

- Le recouvrement du prix est probable.
- Les droits aux biens ou services et les conditions de règlement peuvent être identifiés.
- Il a une substance commerciale.
- Le contrat est approuvé et les parties sont engagées à respecter leurs obligations.

ii. Les identifications des obligations de prestation

L'opérateur identifie chaque promesse de fourniture d'un service contenue dans un contrat conclu avec ses clients. Les biens et services de télécommunication offerts par l'opérateur X sont identifiables et séparés lors de leurs fournitures, et sont dénommés comme suit [8] :

- Voix : service d'appels téléphoniques commutés.
- SMS : service de messagerie textuelle SMS.
- MMS : service de messagerie multimédia.
- DATA : Service d'accès à Internet.
- VAS : Service à Valeurs Ajoutées (contenus multimédia, information, transfert de balance...etc.).
- Terminaux : équipement de télécommunication compatible avec réseau GSM/3G/4G.
- Prépaie Air Time : Volume de communication prépayé transférable sous forme physique en carte de recharge et ticket imprimé et/ou forme électronique en service de transfert Top Up.

iii. Détermination du prix de la transaction

Le prix de chaque service doit être défini dans le catalogue des offres principales et offres secondaires que l'opérateur propose à ses clients [8].

iv. Allocation du prix de la transaction

Le système d'information de l'opérateur X enregistre le prix des prestations et rabais pour chaque événement réalisé sur son réseau de télécommunication [8].

v. Comptabilisation du revenu

Le revenu est comptabilisé lorsque l'opérateur a rempli une obligation de prestation en transférant le contrôle d'un bien ou service au client final (abonnés prépayés et/ou post payés) [16].

5.4.2 Description du processus

Nous allons nous intéresser qu'à la dernière étape du processus de vérification de la norme IFRS 15, à savoir la reconnaissance du CA. [8].

i. Comptabilisation des PCA & Avances Clients :

La sous-direction de comptabilité de l'opérateur reçoit des directions commerciales les factures de ventes de service suivant [8] :

- a) Ventes Indirectes : (via les distributeurs)
 - Factures des Ventes de Vouchers Physiques (Cartes de recharge et TPE) aux distributeurs et Points de Ventes (PDV) agréés par l'opérateur.
 - Factures des Ventes des e-Vouchers (aux distributeurs et Points de Ventes (PDV) agréés par l'opérateur.

- b) Ventes Directes (VD)
 - Etats des Ventes des Vouchers Physiques & TPE au niveau des points de ventes de l'opérateur.
 - Etats des Ventes e-Vouchers au niveau des points de ventes de l'opérateur.

On distingue deux types d'écritures comptables sur les produits comptabilisés d'avance :

a) Vente au niveau régional

Pour chaque direction régionale, un compte PCA dédié pour comptabiliser les ventes AirTime régionales (voucher physique).

b) Vente au niveau central

Le département de la comptabilité établit une fiche d'imputation à laquelle il attribue des numéros dans la séquence du journal des ventes directes et indirectes. Ces comptes sont dédiés pour comptabiliser les ventes des e-Vouchers, des Vouchers Masters, les usages des Vouchers et les Avances de Paiement de Factures Post payés.

ii. *Réconciliation des revenus* :

La réconciliation est un processus comptable qui cherche à vérifier deux ensembles d'enregistrements, souvent internes et externes, pour s'assurer que les chiffres sont corrects et concordants.

Le rapprochement des comptes est important car il permet de détecter les erreurs, écarts ou fraudes dans les livres comptables qui pourraient gravement affecter la santé financière de l'entreprise. La réconciliation est une bonne pratique commerciale qui peut contribuer au succès d'une entreprise [17].

Il n'y a pas de méthode ni d'outils spécifique pour effectuer le rapprochement, mais en général, il se fait manuellement en comparant les transactions du registre interne et le compte externe (relevé bancaire par exemple).

Dans notre cas, en fin de mois, les rapports suivants devront être édités afin de procéder au contrôle des écritures comptables enregistrées [8] :

- Etat des ventes indirectes.
- Etat des ventes directes.
- Etat d'usage des Cartes de Recharge & TPE.
- Etat des ventes Directes Autres Channel.
- Etat des Recharges Post-Payés (Avances)
- Etat des Activations d'offres avec balances initiales principales.

Les écarts trouvés doivent être analysés et documentés par des structures compétentes. Le produit constaté d'avance est éclaté par produit, seuls les transferts et rechargements établis sur les balances Distributeurs et Prépayés sont à comptabiliser en PCA [8].

iii. *Reconnaissance du chiffre d'affaires*

La sous-direction de comptabilité reçoit des directions commerciales les états de consommation des services mobiles suivants [8] :

- Voix
- SMS et MMS
- Data
- Souscriptions aux Plans supplémentaires
- Souscriptions aux Service de Contenues
- Frais de Transferts de solde entre Abonnés

Le rapport des consommations réelles éclatées par produit et catégorie de service est donc édité.

Le département de la comptabilité établit une fiche d'imputation pour les consommations prépayées.

Pour les consommations Post-Payées avant facturation finale, le département de la comptabilité établit une fiche d'imputation en provision pour les consommations post-payées.

iv. Solutions RA existantes

Le *Tableau 1.3* suivant présente quelques outils dédiés à l'assurance des revenus ainsi qu'une description de leurs fonctionnalités :

Outils	Description
 <p>MOBILEUM Revenu Assurance [11]</p>	<p>Cet outil garantit une :</p> <ul style="list-style-type: none"> - Assurance du provisionnement des services mobiles et filaires. - Assurance de l'utilisation du réseau filaire. - Validation du solde prépayé. - Validation de la tarification et de la facturation.
 <p>SUBEX Revenu Assurance [12]</p>	<p>Cet outil est doté de :</p> <ul style="list-style-type: none"> - Plate-forme ETL avancée. - Analyseur virtuel. - Gestionnaire robuste des cas. - Moteur d'audit. - Tarification et facturation avancées.
<p>PANAMAX Revenu Assurance [13]</p> 	<p>Cet outil garantit :</p> <ul style="list-style-type: none"> - Évaluation du risque de fuite de revenus pour chaque source. - Assurance d'abonnement, d'utilisation, de tarification, de médiation, d'interconnexion, de facturation et d'itinérance. - Validation de ce qui est vendu/commandé exactement avec ce qui est fourni aux clients. - Analyse des écarts d'incohérence et des tendances. - Planification des actions pour le recouvrement des revenus

Tableau 1.3 : Exemple d'outils RA

v. *Problématique*

Comme l'a montré le *Tableau 1.3*, les outils de revenu assurance existants n'automatisent que la partie vérification de l'infrastructure qui consiste à vérifier que les données sont bien collectées, qu'il n'y a pas eu de fuite et que les normes sont bien respectées (IFRS 15) en donnant une visibilité détaillée des différents écarts entre l'ensemble des nœuds du réseau télécom de l'opérateur.

En revanche, la partie réconciliation des revenus se fait encore manuellement.

L'objet de notre travail est d'automatiser le processus de réconciliation des revenus et ce avec la création d'un workflow qui sera paramétrable et réutilisable selon le besoin afin de réduire le risque d'erreur et faciliter la tâche de validation des comptes Revenus et Revenus différés des Opérateurs Télécom Mobiles.

6 Conclusion

Dans ce chapitre, nous avons exposé le marché télécom en Algérie suivi des notions de chiffre d'affaires, de revenu et de revenu différé qui nous sont utiles pour comprendre le revenu assurance qui est le processus global qui résume tout le parcours suivi par un revenu en commençant par le processus de génération à la reconnaissance de chiffre d'affaires en passant par la réconciliation.

Les exemples d'outils existants nous ont aidés à cerner l'objectif de notre travail consistant à automatiser la procédure de comptabilisation et réconciliation.

Chapitre II : Data analytics

1 Introduction

Lorsqu'on parle de données dans le monde des affaires d'aujourd'hui, un certain nombre de termes apparaissent généralement. Parmi les termes les plus courants en science des données, on retrouve Data Analytics (DA).

Data analytics est aujourd'hui au cœur des préoccupations des entreprises. Il s'agit d'une pratique essentielle pour augmenter significativement son chiffre d'affaires, mais également pour rester compétitif dans la plupart des industries. Toutefois, avant de s'engager dans ce nouveau domaine, il est nécessaire de bien comprendre ce dont il s'agit, et c'est ce que va traiter ce chapitre.

2 Data analytics

2.1. Définition

Data analytics (DA) est la science, discipline, qui consiste à analyser, examiner et explorer des données brutes à l'aide de nombreuses techniques et processus d'analyse de données afin de tirer des conclusions et trouver les tendances sur ces informations [19].

2.2. Types de Data analytics

Le processus de Data analytics est réparti en fonction de l'objectif de l'analyse des données. On distingue quatre types [6] :

2.2.1. L'analytique Descriptive

Elle fournit des informations pertinentes sur ce qui a été réalisé et permet donc de comprendre ce qui s'est produit sur une période passée. Un simple rapport sur les ventes au cours du dernier trimestre est un exemple d'analyse descriptive. Bien qu'elle soit utile, elle ne traite pas le "pourquoi" des données, et c'est là que l'analyse diagnostique entre en jeu.

2.2.2. L'analytique Diagnostique

Elle se focalise davantage sur les raisons pour lesquelles quelque chose s'est produit. Cela implique des entrées de données plus diverses et un peu de formulation d'hypothèses. En s'appuyant sur l'exemple précédent : un rapport sur les ventes du trimestre précédent pourrait être un exemple d'analyse diagnostique s'il essayait également d'utiliser les données pour expliquer pourquoi les ventes ont chuté, par exemple, d'un pourcentage étonnamment élevé.

2.2.3. L'Analytique Prédictive

Elle utilise les résultats de l'analyse descriptive et diagnostique pour prévoir les tendances futures. Tout exercice d'analyse de données bien développé commence par des informations qui peuvent être extraites de sources internes et externes. Les données transitent ensuite par un moteur de prévision qui peut utiliser des modèles statistiques et de la machine learning pour émettre

des prédictions. Celles-ci peuvent aider à générer des rapports de prévision et permettre à la direction de prendre des décisions plus avisées.

2.2.4. L'Analytique Prescriptive

Cette analytique s'appuie sur les tendances favorables découvertes au cours des trois types précédents en donnant des recommandations d'action. Un exemple serait un moteur d'intelligence artificielle interprétant les données pour suggérer à une entreprise de dépenser davantage de son budget pour un emplacement particulier via un canal spécifique afin de maximiser son taux de croissance.

La *Figure 2.1* suivante montre les quatre types de Data analytics.

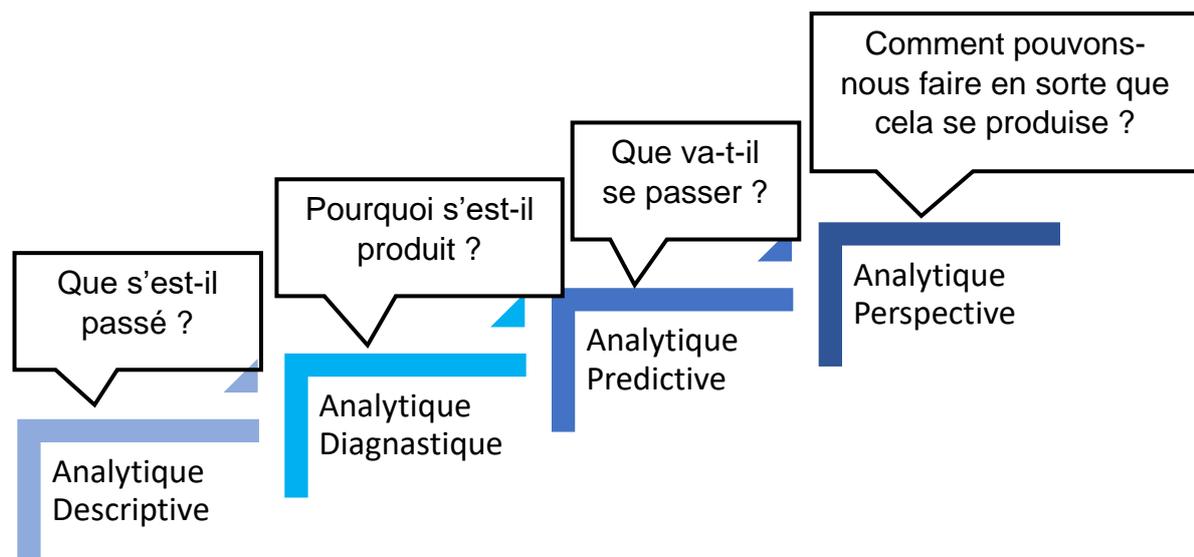


Figure 2.1 : Les types de la DA

2.3. Processus de Data analytics

Le processus de Data analytics comporte plusieurs phases de traitement des données et de visualisation, comme le montre la *Figure 2.2* suivante [6] :

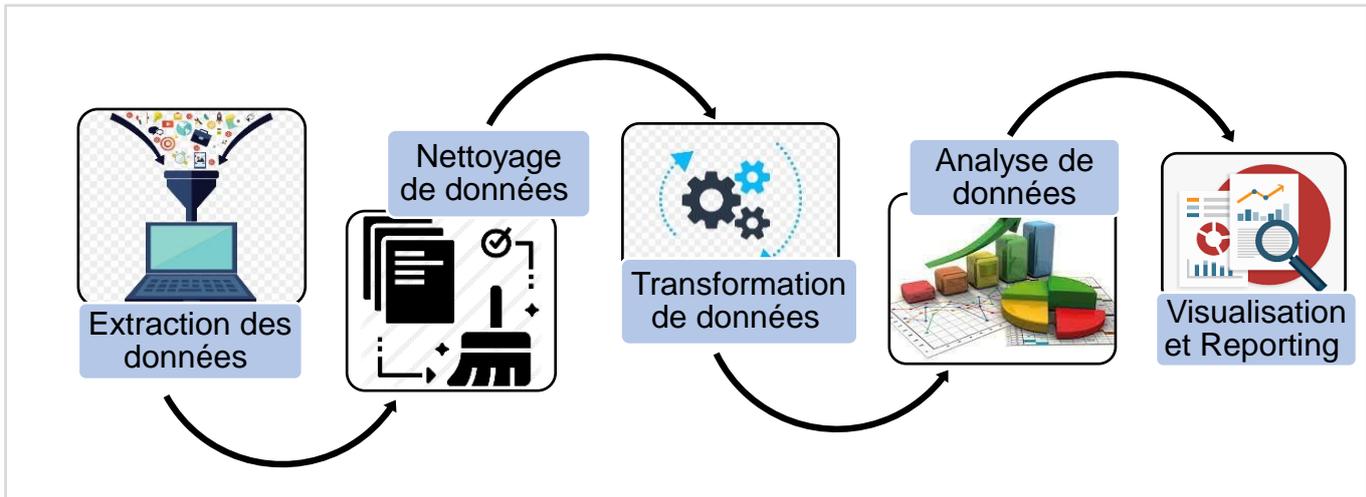


Figure 2.2 : Processus DA

2.3.1. Extraction de données

C'est La collecte de données de différents types à partir de plusieurs sources. Les formats des données introduites dans le flux d'analyse des données peuvent être groupés comme suit :

- *Données structurées* : sont des données qui ont un format et une structure bien précise, et qui sont placées dans une base de données relationnelle, car ces données ont été formatées dans des champs précisément définis (type, taille, ...) pour être facilement interrogées.
- *Données semi-structurées* : qui n'ont pas de définition précise ou fixe du format, mais elles peuvent être identifiées, séparés et regroupés sur la base d'un schéma standard ou d'autres règles de métadonnées. Par exemple, un fichier XML utilise le balisage pour contenir les données, tandis que le fichier JSON contient les données par paires clé-valeur. Les bases de données NoSQL comme MongoDB sont également utilisées pour stocker des données semi-structurées.
- Des données non structurées : sont des données stockées dans leur format d'origine et non traitées avant leur utilisation, tels que les conversations sur les médias sociaux, les images, les clips audio, etc.

2.3.2. Nettoyage de données

Le nettoyage de données ou Data Cleaning est effectué pour garantir la cohérence et la disponibilité de données pertinentes pour les étapes ultérieures d'un processus.

Les principales opérations de nettoyage dans l'analytique des données sont les suivantes :

- Détection et élimination des valeurs aberrantes dans les volumes de données.
- Suppression des doublons dans l'ensemble des données.
- Traitement des entrées manquantes dans les enregistrements de données en tenant compte de la fonctionnalité ou des cas d'utilisation.
- Validation des valeurs de champ autorisées dans les enregistrements de données, par exemple "31 février" ne peut être une valeur valide dans aucun des champs de date.

2.3.3. Transformation de données

Les données nettoyées sont ensuite transformées dans un format approprié pour l'analyse de données (*Voir Figure 2.3*).

Les transformations de données comprennent :

- Le filtrage des enregistrements de données non désirés.
- L'assemblage des données provenant de différentes sources.
- L'agrégation ou le regroupement des données.
- La catégorisation des données.

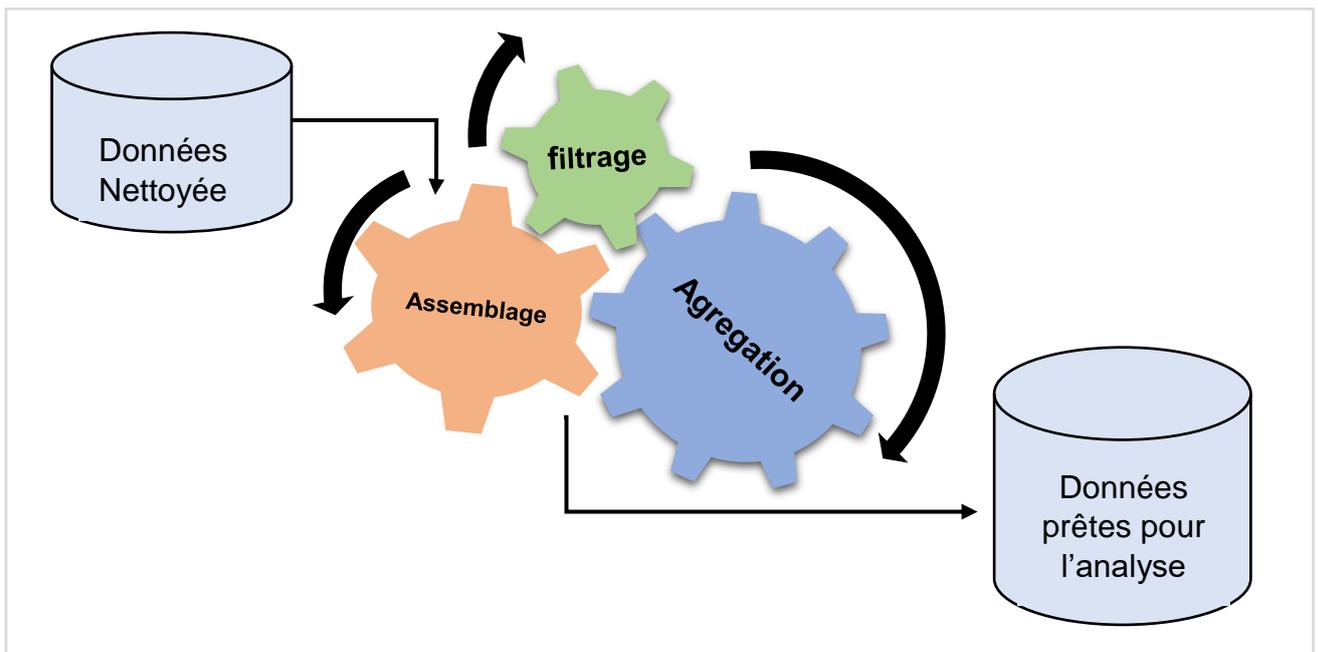


Figure 2.3 : Transformation de données

2.3.4. Analyse des données :

En fonction de l'objectif de l'analyse, l'analyse des données est effectuée à l'aide de diverses techniques et méthodes d'exploration de données (Data Mining) et d'apprentissage profond (Deep Learning) pour évaluer les indicateurs clés de performance (KPI) ou pour obtenir des informations précieuses à partir des données nettoyées et transformées.

2.3.5. Visualisation des données et Reporting :

La visualisation des données est le processus de présentation claire, transparent et efficace des modèles découverts, des conclusions dérivées depuis les données en utilisant des graphiques, des tracés et des tableaux de bord (*Voir Figure 2.4*).

Les outils de reporting de données tels que QlikView, Tableau, etc., affichent les indicateurs clés de performance et autres mesures dérivées à différents niveaux de granularité. De plus, ils permettent aux utilisateurs finaux de créer des rapports personnalisés.

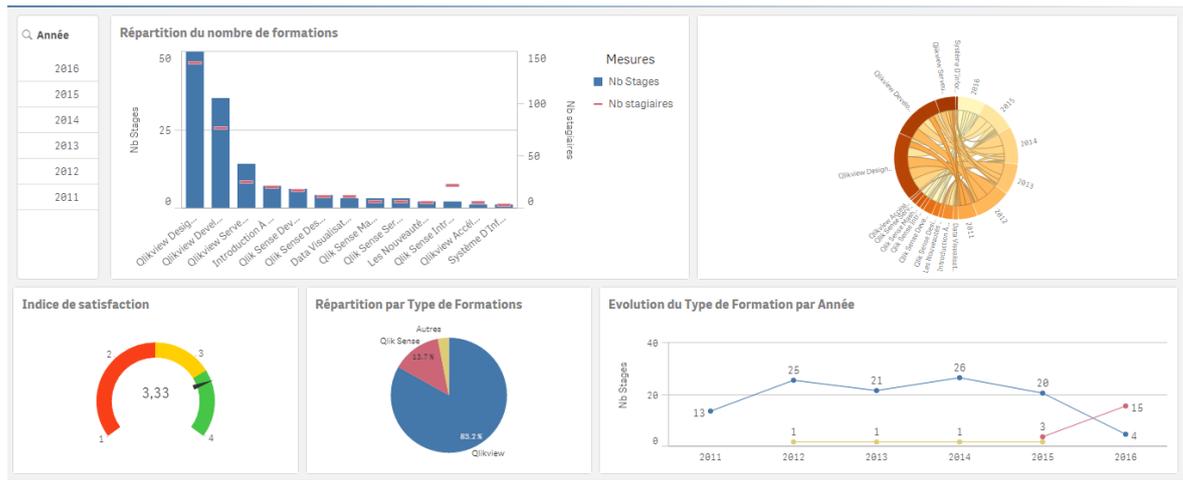


Figure 2.4 : exemple de tableau de bord

3 Plateforme de Data analytics

3.1. Définition

Une plateforme de DA est une solution unifiée, un écosystème qui combine des technologies pour répondre aux besoins des entreprises tout au long du cycle de vie de l'analyse de bout en bout, de la sélection et le stockage à la gestion et la préparation des données vers d'autres processus d'analyse des données [20] .

3.2. Avantages d'une plateforme DA

L'investissement dans une plateforme d'analyse offre de précieux avantages sur le plan stratégique [20] :

- i. Une plateforme analytique peut contribuer à réduire le temps consacré à la préparation des données.
- ii. Permettre aux utilisateurs professionnels de découvrir des informations fiables en libre-service et intégrer les données dans l'ensemble de l'entreprise afin d'accélérer l'obtention de ces informations.
- iii. Les entreprises modernes constatent un élargissement des cas d'utilisation de l'analytique, ce qui nécessite une plateforme analytique complète pour rendre l'analytique pleinement opérationnelle dans plusieurs domaines d'activité.

- iv. En mettant en œuvre une plateforme analytique, les entreprises sont en mesure de donner aux utilisateurs les moyens de prendre des décisions fondées sur les données au sein des équipes et de favoriser l'innovation.

3.3. Exemples de plateformes de DA

Des exemples d'outils data analytique sont présentés dans le tableau suivant :

	Description	License	Critique
 [22]	<p><i>Alteryx Analytics</i> est un éditeur de logiciels d'analyse de données en libre-service, spécialisé dans la préparation et le mélange de données, permettant aux utilisateurs d'organiser, de nettoyer et d'analyser des données dans un flux de travail réutilisable.</p>	<p>5,195USD /utilisateur par année</p>	<ul style="list-style-type: none"> - Manque de fonctions de visualisation de données. - Logiciel payant.
 [23]	<p><i>Altair</i> offre une plateforme d'analyse de données ouverte, évolutive, unifiée et extensible avec des outils intégrés de préparation, transformation de données, d'analyse prédictive, apprentissage automatique.</p>	<p>/</p>	<ul style="list-style-type: none"> - Pas de version gratuite - Pas de version d'essai
 [24]	<p><i>SAS Visual Analytics</i> permet d'explorer visuellement les données et intégrer des données provenant de diverses sources et gérer des modèles complexes. SAS propose des outils de gestion de données, d'IoT, de protection de données personnelles.</p>	<p>Tarif abonnement : 8 000,00 \$US /Année</p>	<ul style="list-style-type: none"> - Pas de version gratuite - Ne fait pas de traitement de données (data processing)

Tableau 2.1 : Exemple de plateforme D&A

4 Conclusion

Data analytics aide à répondre à des questions, à tester des hypothèses, à réfuter des théories pour prendre de meilleures décisions, cela via des plateformes qui collectent, nettoient, transforment et analysent les données pour obtenir et visualiser des informations significatives. Ce chapitre a permis de prendre connaissance de cette discipline, pour comprendre l'approche de notre solution.

Chapitre III : Programmation basée sur le flux

1 Introduction

La programmation ou le codage consiste à rédiger un code informatique pour la résolution d'un problème. Lors du développement, les développeurs sont confrontés à deux problèmes :

- i. La résolution du problème lui-même.
- ii. La complexité de l'implémentation de la solution :
 - Choix du Framework, langage...
 - Choix de l'architecture ...

Pour mieux se focaliser sur la résolution du vrai problème, Jean Paul Morrison a proposé une nouvelle approche qui consiste à changer de paradigme. Plusieurs besoins notamment celui de modification (maintenance) sans affecter le système complet, l'ont poussé à penser à une programmation dite visuelle ou graphique, où la structure du système est claire, évidente et simple et ce en s'inspirant du diagramme de flux de données.

Cette approche est connue sous le nom de Programmation basée sur le flux (PBF).

2 Diagramme de flux de données

2.1 Définition

Un diagramme de flux de données (DFD) est un diagramme qui schématise le flux d'informations de n'importe quel processus ou système. Il utilise des symboles définis, ainsi que des étiquettes aux libellés courts, pour représenter les entrées de données, les sorties, les points de stockage et les trajets entre chaque destination.

Il permet d'avoir une vue sur les processus, et ainsi modéliser des flux de données mieux conçus et plus performants à travers les processus métier [26].

2.2 Composants d'un DFD :

Dans un diagramme de flux de données on trouve [26] :

- i. **Processus** : est une entité qui transforme des données et produit un résultat, par exemple : des calculs, tri de données. Une courte description est utilisée pour décrire le processus. Dans l'exemple ci-dessous (Figure 3.1), on trouve : Commander le repas, vérifier la commande, créer le rapport.
- ii. **Entité externe** : qui émet ou reçoit des données, et communique avec le système représenté. Elles sont également appelées sources et puits, ou acteurs. Ils sont souvent dessinés sur les bords du diagramme. Dans l'exemple ci-dessous on trouve Client, Fournisseur, Manager, Cuisine).
- iii. **Magasins de données** : fichiers ou dépôts qui contiennent des informations pour une utilisation ultérieure, comme une table de base de données ou un formulaire d'adhésion. Ils sont dotés d'une étiquette simple, telle que "Liste des commandes" dans l'exemple ci-dessous.
- iv. **Flux de données** : le chemin emprunté par les données entre les entités externes, les processus et les magasins de données. Il est représenté par des flèches, généralement étiquetées avec un nom de données court, tel que "Commande" et "Facture".

La Figure 3.1 suivante représente un DFD de système de livraison d'un restaurant.

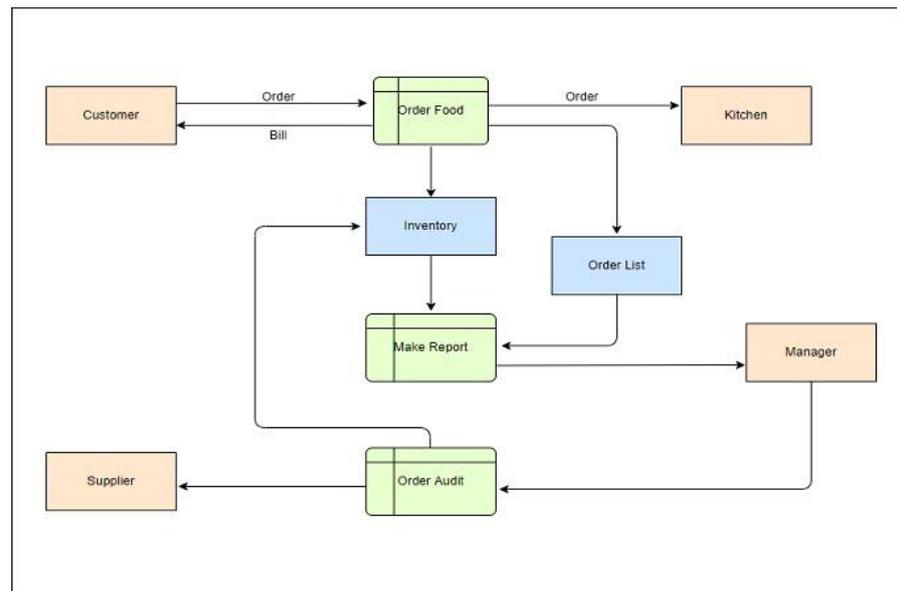


Figure 3.1 : DFD d'un système de livraison [1]

2.3 Domaines d'utilisation des diagrammes de flux de données

Les DFD sont utiles pour l'analyse et la modélisation de différents types de systèmes dans un grand nombre de domaines [26].

- i. Le Génie Logiciel : Qui est le principal point de départ des DFD dans les années 1970. Les diagrammes peuvent fournir une approche centrée sur le développement technique, dans laquelle une recherche plus poussée est effectuée en avant le codage.
- ii. L'analyse d'Entreprise : Les analystes financiers utilisent les DFD pour analyser les systèmes en vigueur et identifier les lacunes et inefficacités. La cartographie des processus peut dévoiler des étapes qui pourraient autrement être oubliées ou mal comprises.
- iii. La Réingénierie des processus d'affaires : Les DFD peuvent être utilisés pour modéliser des flux de données meilleurs et plus efficaces à travers les

processus métier afin d'aider les compagnies à réduire les coûts opérationnels, à renforcer le service client et à être plus compétitives.

- vi. Développement Agile : Les DFD peuvent être un outil de communication et de collaboration simple, mais puissant, pour un développement rapide.

3 Programmation basée sur le flux :

3.1 Définition

Inspirée du diagramme de flux de données, la programmation basée sur le flux est une manière graphique de créer des programmes informatiques, et ce en créant des flux au lieu de lignes de code.

Ces flux sont des réseaux de boîtes noires (nœuds), qui sont en réalité des séries de fonctions dotées d'entrées et / ou sorties, et qui exécutent une tâche spécifique.

Ce principe consiste à bâtir un système informatique à partir d'un diagramme (réseau ou workflow), en créant des connexions entre composants, en connectant la sortie de l'un avec l'entrée d'un autre composant, grâce à un arc, et ainsi chaque combinaison permettra d'avoir un processus [3].

Le *Tableau 3.1* ci-dessous contient un exemple comparatif entre un algorithme et un réseau PBF d'une simple opération arithmétique dans lequel :

- Addition et Multiplication sont des fonctions arithmétiques (+ et *).
- A, B, C des entrées.
- R est une sortie.
- Les flèches représentent les connexions.

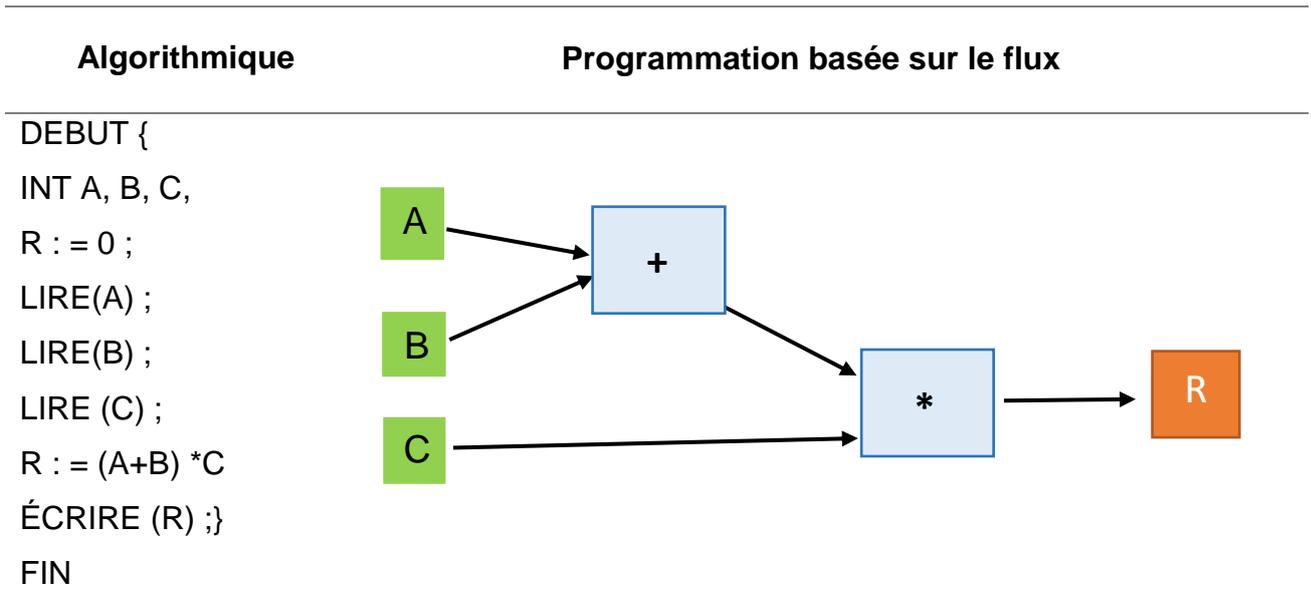


Tableau 3.1 : Exemple Comparatif Algorithmique Vs FBP

3.2 Concepts de la programmation basée sur le flux :

3.2.1. Nœud

Ce concept est "atomique", au sens qu'il ne peut pas être décomposé en notion plus fondamentale, il est solidement fondé sur l'idée de la fonction mathématique. Cette fonction peut être standard (lecture, écriture, recherche...), comme elle peut être créée pour un besoin bien précis. Chaque nœud est caractérisé par un nom, description, entrées et/ou sorties, conduits par des déclencheurs avec le principe d'enchaînement ou chainage où toute fonction se déroule après sa précédente, en plus de la présence de tous ses paramètres d'entrées (voir Figure 3.2).



Figure 3.2 : nœud PBF

De plus, Il est possible de regrouper les nœuds de fonctions, qui composent un système basé sur le flux, en des packages qui peuvent être utilisés comme s'il s'agissait de composants à part entière. Ce type de composant est appelé « **composant composite** ». Il est construit avec des composants de niveau inférieur (voir *Figure 3.3*), mais à l'extérieur, il ressemble à n'importe quel autre composant.

Les composants qui ne sont pas constitués de composants de niveau inférieur sont appelés « **élémentaires** » (exemple du *Tableau 3.1*), et peuvent être écrits dans n'importe quel langage de programmation appropriée. Pour faire en sorte qu'un composant composite ressemble à d'autres composants de l'extérieur, il doit évidemment avoir ses propres ports, les points où les connexions se rencontrent. Les composants ne peuvent pas nommer directement leurs voisins ni leurs connexions, elles s'y réfèrent en nommant les « ports » [3].

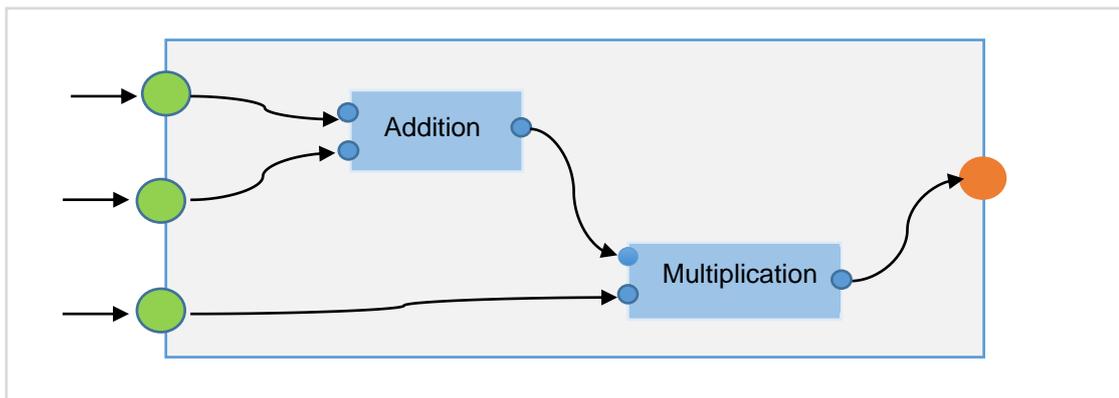


Figure 3.3 : Exemple de Composant-Composite

3.2.2. Connexion

Les connexions sont de types FIFO ou files d'attente (premier entré, premier sorti) pouvant contenir jusqu'à un certain nombre maximum de paquets d'information appelé « capacité d'une file d'attente » (voir *Figure 3.4*).



Figure 3.4 : Connexion FIFO

Comme tout réseau, un service de "réception" peut être bloqué quand il n'y a aucune donnée dans la Connexion, et un service « envoi » peut être bloqué quand la connexion est complète et ne peut plus accepter de données pendant un certain temps. Il est nécessaire de donner un maximum de capacité de connexion, non seulement pour que les applications s'intègrent dans le stockage disponible, mais également pour que toutes les données soient finalement traitées (sinon les données pourraient simplement s'accumuler dans les connexions et ne seront jamais traitées) [3].

3.2.1 Paquet d'Information

Le terme 'flux' dans PBF remonte au flux de données qui circulent dans le réseau. Ces paquets de données (voir *Figure 3.5*) doivent être dans un format que la fonction en amont peut générer et la fonction en aval peut gérer.

La méthodologie basée sur le flux nécessite un concepteur pour décrire d'abord les données, puis définir les informations sur ces données (émetteur, numéro du port...) et les transformations qui leurs sont applicables [3].

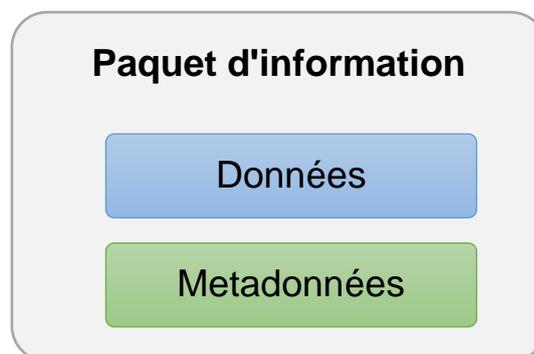


Figure 3.5 : Paquet d'Information

3.3 Les Composants d'un système basé sur le flux

D'après ce qui a été cité précédemment, on déduit que les systèmes basés sur le flux se composent essentiellement des composants suivants [2] :

- i. Un certain ensemble de fonctions pré testées, fournies sous forme de boîtes noires, et non sous forme de code source, avec des ports où les connexions entre les processus sont établies.

- ii. Un pilote : logiciel qui coordonne les modules comprenant une application et implémente l'interface de programmation d'application qu'ils utilisent pour demander des services.
- iii. Une notation : Description pour spécifier comment les composants sont connectés pour construire l'application.

3.4 La Réutilisabilité, un avantage de la PBF

En génie logiciel, la réutilisation s'appuie fréquemment sur le concept de modularité, en faisant projection sur la PFB, les modules seront dans ce cas les fonctions qui composent le système basé sur le flux. On distingue deux types :

- i. Boîte Noire

L'utilisateur ne peut pas voir l'intérieur du composant (code de la fonction), donc n'a jamais besoin d'être modifié, une fois fonctionnel, il peut être fiable pour fonctionner correctement dans n'importe quel contexte et où chaque instantiation de la fonction sera donc une copie du code de la fonction.

- ii. Boîte Claire

L'utilisateur aura accès au code ce qui peut réduire le coût de développement, mais la quantité de nouveau code qui doit être maintenue augmente. De plus, si un bug est trouvé dans le code réutilisé, il n'y a pas moyen de savoir s'il est sûr de réparer toutes les instances de celui-ci.

3.5 Exemples d'outils conçus avec PBF :

Le *tableau 3.2* présente quelques systèmes conçus avec la programmation basée sur le flux :

Nom	Auteur	URL	Licence	Date	Domaine d'application
Visuino	Boian Mitov	www.visuino.com	Commercial	12/16/2016	Microcontrôleurs de type Arduino
Pure Data	Miller Puckette	http://puredata.info/	Open Source	1/8/2010	MultiMedia
Apache NiFi	Apache	https://nifi.apache.org	Apache 2.0	1/6/2017	Data Distribution, logging

Tableau 3.2 : Systèmes Basé sur le Flux

4 Workflow

4.1 Définition

On appelle Workflow la modélisation et la gestion informatique de l'ensemble des tâches à accomplir et des différents acteurs impliqués dans la réalisation d'un processus métier [27].

C'est donc l'automatisation de tout ou d'une partie d'un processus d'entreprise. Il contribue à l'automatisation du travail humain, il vise à assister l'homme dans ses interactions.

4.2 L'automatisation un des principes du workflow

Avec le Workflow, les tâches, les données, les échanges sont automatisés de façon à se succéder dans un ordre bien défini pour obtenir des résultats attendus, dans une optique de :

- i. Renforcement et Maximisation de la productivité, la fiabilité, la disponibilité et des performances.

- ii. Réduction des coûts d'exploitation.
- iii. Diminution des interventions manuelles.
- vii. De meilleurs flux de données sans erreurs et des rapports plus rapides.

4.3 Workflow et PFB

En prenant l'exemple du principe de fonctionnement des usines (voir *Figure 3.6*), où de nombreux processus sont synchrones et se déroulent en même temps, simultanément et parallèlement, ce qui n'incite pas vraiment de savoir si une machine fonctionne avant ou après une autre, tant que les processus sont appliqués à un travail donné dans le bon ordre.

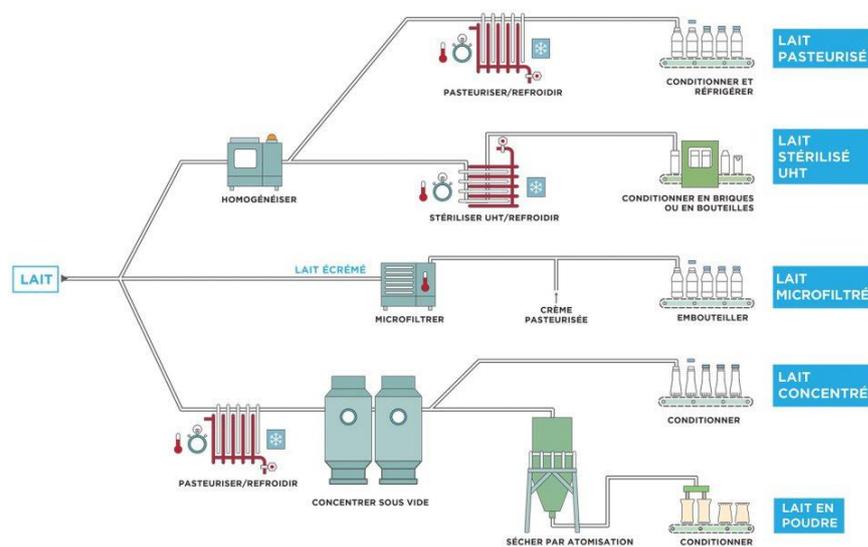


Figure 3.6 : Processus de transformation de lait

Contrairement à la programmation traditionnelle où ce principe de simultanété et de parallélisme n'existe pas, l'exécution du code se fait par ligne et par bloc selon les événements et les conditions.

Dans une véritable usine, des contraintes inutiles de ce type signifieraient que certaines machines ne seraient pas utilisées efficacement.

On constate que le principe de la Programmation basée sur le flux, par rapport à la programmation conventionnelle, est l'approche la plus adéquate pour répondre au besoin d'automatisation, vu la ressemblance du principe de l'enchaînement de processus métiers avec celui de chaînage de fonction [3].

4.4 Inconvénients de l'approche conventionnelle

On peut résumer les problèmes rencontrés avec l'approche conventionnelle comme suit :

- i. Le code est procédural, séquentiel.
- ii. Une vue du traitement une étape à la fois.
- iii. La structure hiérarchique du code.
- iv. Le programmeur doit contrôler la séquence exacte des événements.

5 Conclusion

Ce chapitre a permis de découvrir la programmation basée sur le flux comme une approche de programmation basée sur l'aspect visuel inspirée du diagramme de flux de données, son principe de création d'application ressemble beaucoup à celui du workflow. Le principe d'automatisation des opérations et des processus avec les workflows a permis d'adopter cette approche pour résoudre notre problématique.

Deuxième partie

Solution

- Chapitre I : Conception de la solution
- Chapitre II : Implémentation de la solution

Chapitre I : Conception de la solution

1 Introduction

Ce chapitre va exposer la solution à notre problématique liée à la revenue assurance, en commençant par la description de notre approche, où on va montrer les liens entre les notions précédemment exposées et l'objectif à atteindre puis on va exposer les différents diagrammes qui ont servi à la modélisation de la solution.

2 Description de l'approche

Comme vu précédemment, dans le premier chapitre dédié à la **Revenue Assurance**, la RA (dans le cadre d'un opérateur télécom mobile) consistait à réconcilier et à rapprocher le revenu reconnu comme chiffre d'affaires avec ce qui a été comptabilisé. Ce processus est enfin de compte, d'un point de vue informatique, un ensemble de traitement et de manipulations à savoir :

- Des opérations de Nettoyage.
- Des Sélections : Projections et restrictions.
- Des Jointures et des Unions.
- Des Agrégations.
- Des Operations arithmétiques (additions, soustraction, divisions,...)

Ces manipulations sont effectuées sur des masses de fichiers clients, des CDRs, ...etc. Permettant ainsi de vérifier l'intégrité des comptes, leur régularité, leur conformité et leur aptitude à refléter une image fidèle de l'état des finances et actions comptables de l'entité auditée (Opérateur télécom).

Ces opérations sont résumées par le terme **Data Analytics**, qui est la discipline consistant à analyser, examiner et explorer des données brutes à l'aide de nombreuses techniques et processus d'analyse de données.

Notre approche a pour but de créer une plateforme de Data Analytics avec interface interactive pour automatiser le processus de réconciliation des revenus.

Afin d'obtenir l'aspect automatique et dynamique du traitement, notre approche se base sur **la programmation visuelle, basée sur le flux**, où l'utilisateur peut suivre visuellement l'enchaînement et l'exécution de ses traitements en créant un workflow complet de traitement, grâce à une palette de fonctions analytiques et statistiques mises à sa disposition. Les Workflows créés permettront d'automatiser les traitements récurrents et répétitifs de vérification, manuellement fait généralement, ce qui diminuera le risque d'erreur et facilitera la tâche d'audit.

3 Modélisation de l'approche

Le langage **UML** fournit plusieurs modèles servant de manuel de construction et de documentation du système à développer, où sont spécifiés les structures et le comportement du système.

3.1 Diagramme des Cas d'utilisation

C'est un diagramme comportemental d'UML qui décrit la portée du système entre autres les besoins fonctionnels seulement. Il identifie les interactions entre le système et ses acteurs.

La *Figure 4.1* résume les fonctionnalités principales de notre système avec un diagramme des cas d'utilisation.

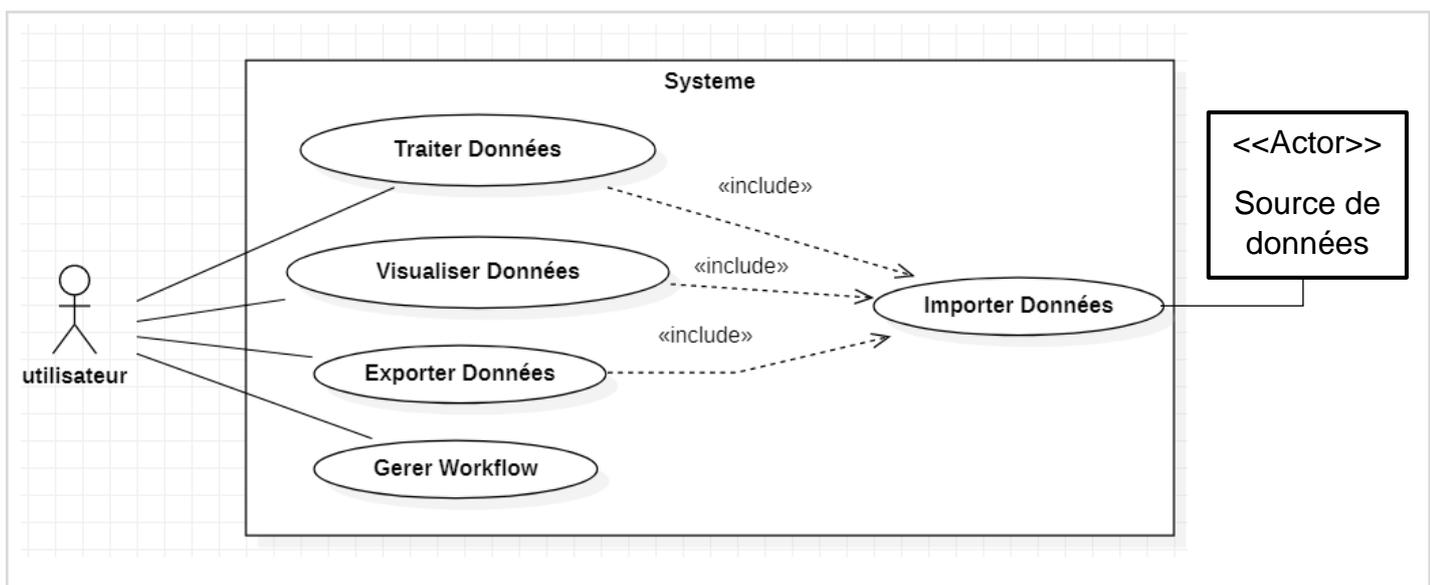


Figure 4.1 : Diagramme des Cas d'utilisation

Ci-dessous la description Textuelle des fonctionnalités principales de notre système :

i. Description Textuelle « Importer Données » (Voir *Tableau 4.1*)

Cas d'utilisation	Importer Données
<i>Description</i>	Récupérer des données qui sont dans des sources diverses via la plateforme.
<i>Acteurs</i>	<ul style="list-style-type: none"> • Utilisateur. • Source de données.
<i>Evènement Déclencheur</i>	L'utilisateur souhaite effectuer des traitements, ou visualiser des données.
<i>Scenario de Base</i>	<ol style="list-style-type: none"> 1. Utilisateur ajoute un nœud de type <i>Input</i>. 2. Une fenêtre de dialogue est affichée. 3. Il saisit les renseignements sur la source de données (nom du fichier ou de la base,.. selon le type de la source fichier ou BDD). 4. Le système vérifie les paramètres. 5. Le système affiche les métadonnées de l'input. 6. L'utilisateur clique sur le bouton exécuter. 7. Le système importe les données.
<i>Précondition</i>	Source de donnée déjà existante
<i>Post-Condition</i>	Données importées

Tableau 4.1 : Description textuelle « Importer Données »

ii. Description Textuelle « Traiter Données » (Voir *Tableau 4.2*)

Cas d'utilisation	Traiter Données
<i>Description</i>	Effectuer un traitement sur des données.
<i>Acteurs</i>	<ul style="list-style-type: none"> Utilisateur.
<i>Evènement Déclencheur</i>	L'utilisateur souhaite effectuer un traitement.
<i>Scenario de Base</i>	<ol style="list-style-type: none"> Utilisateur ajoute un nœud de type <i>opération</i> selon le traitement souhaité. L'utilisateur relie le nœud de traitement <i>X</i> avec un nœud <i>Input</i> contenant les données brut (ou avec un autre nœud d'opération <i>Y</i> retournant à son tour des données), sur lesquelles le traitement sera effectué. Le nœud opération récupère les métadonnées. Une fenêtre de dialogue est affichée. L'utilisateur saisit les paramètres nécessaires au traitement. Le système enregistre les paramètres. L'utilisateur click sur le bouton exécuter. Le workflow s'exécute et le système effectue le(s) traitement.
<i>Précondition</i>	Données importées soit par importation direct (relier directement a un nœud input) ou indirecte (données résultantes d'un traitement)
<i>Post-Condition</i>	Données Traitées

Tableau 4.2 : Description textuelle « Traiter Données »

iii. Description Textuelle « Visualiser Données » (Voir *Tableau 4.3*)

Cas d'utilisation	Visualiser Données
<i>Description</i>	Afficher des données sous forme de graphe
<i>Acteurs</i>	<ul style="list-style-type: none"> • Utilisateur.
<i>Evènement Déclencheur</i>	L'utilisateur souhaite visualiser des données ou résultat de traitement.
<i>Scenario de Base</i>	<ol style="list-style-type: none"> 1. Utilisateur ajoute un nœud nommé <i>graphe</i>. 2. L'utilisateur relie le nœud <i>graphe</i> avec un nœud <i>Input</i> ou un nœud <i>opération</i> contenant des données. 3. Le nœud <i>graphe</i> récupère les métadonnées. 4. Une fenêtre de dialogue est affichée. 5. L'utilisateur saisit les paramètres (titre, axes,..). 6. Le système enregistre les paramètres. 7. L'utilisateur demande l'exécution du workflow 8. Une fois le Workflow exécuté, Le système crée le graphe. 9. L'utilisateur a le choix d'enregistrer le graphe en image.
<i>Précondition</i>	Données importées soit par importation direct (relier directement a un nœud input) ou indirecte (données résultantes d'un traitement)
<i>Post-Condition</i>	Graphe créé.

Tableau 4.3 : Description textuelle « Visualiser Données »

iv. Description Textuelle « Exporter Données » (Voir *Tableau 4.4*)

Cas d'utilisation	Exporter Données
<i>Description</i>	Charger et sauvegarder des données dans un fichier Excel/csv
<i>Acteurs</i>	<ul style="list-style-type: none"> • Utilisateur.
<i>Evènement Déclencheur</i>	L'utilisateur souhaite sauvegarder des données résultantes d'un ou plusieurs traitements.
<i>Scenarion de Base</i>	<ol style="list-style-type: none"> 1. Utilisateur ajoute un nœud de type output nommé export. 2. L'utilisateur relie le nœud avec un nœud <i>Input</i> ou un nœud <i>opération</i> contenant les données à exporter. 3. Le nœud <i>export</i> récupère les métadonnées. 4. Une fenêtre de dialogue est affichée. 5. L'utilisateur saisit les paramètres (type du fichier, chemin, nom,...) 6. Le système enregistre les paramètres. 7. L'utilisateur click sur le bouton d'exécution. 8. Le workflow s'exécute et les données sont chargées dans un fichier.
<i>Précondition</i>	Données importées soit par importation direct (relier directement a un nœud input) ou indirecte (données résultantes d'un traitement)
<i>Post-Condition</i>	Données chargées dans un fichier.

Tableau 4.4 : Description textuelle « Exporter Données »

v. Description Textuelle « Gérer workflow » (Voir *Tableau 4.5*)

Cas d'utilisation	Gérer Workflow
<i>Description</i>	Permet de créer, modifier, supprimer, enregistrer et exécuter un flux de travail.
<i>Acteurs</i>	<ul style="list-style-type: none"> • Utilisateur.
<i>Evènement Déclencheur</i>	L'utilisateur souhaite créer un nouveau flux de travail afin d'automatiser un ou plusieurs traitements.
<i>Scenario de Base</i>	<ol style="list-style-type: none"> 1. Utilisateur crée un nouveau workflow (vide). 2. L'utilisateur fait glisser un ou plusieurs nœuds depuis la palette afin de les rajouter au workflow. 3. L'utilisateur relie les nœuds ajoutés. 4. Il peut supprimer/déplacer un ou plusieurs nœuds. 5. Il paramètre chaque nœud à travers sa fenêtre de dialogue. 6. L'utilisateur déclenche l'exécution du workflow. 7. Il peut enregistrer le workflow.
<i>Précondition</i>	/
<i>Post-Condition</i>	Un traitement automatisé et un workflow enregistré prêt à être réutilisé

Tableau 4.5 : Description textuelle « Gérer Workflow »

3.2 Diagrammes de Séquence

C'est un diagramme comportemental d'UML qui exprime l'aspect dynamique ou les interactions entre les objets. Ces interactions sont présentées en messages échangés dans un ordre chronologique.

Pour chaque cas d'utilisation, nous avons dressé un diagramme de Séquence :

i. Diagramme de séquence Système « Importer Données » (Voir Figure 4.2)

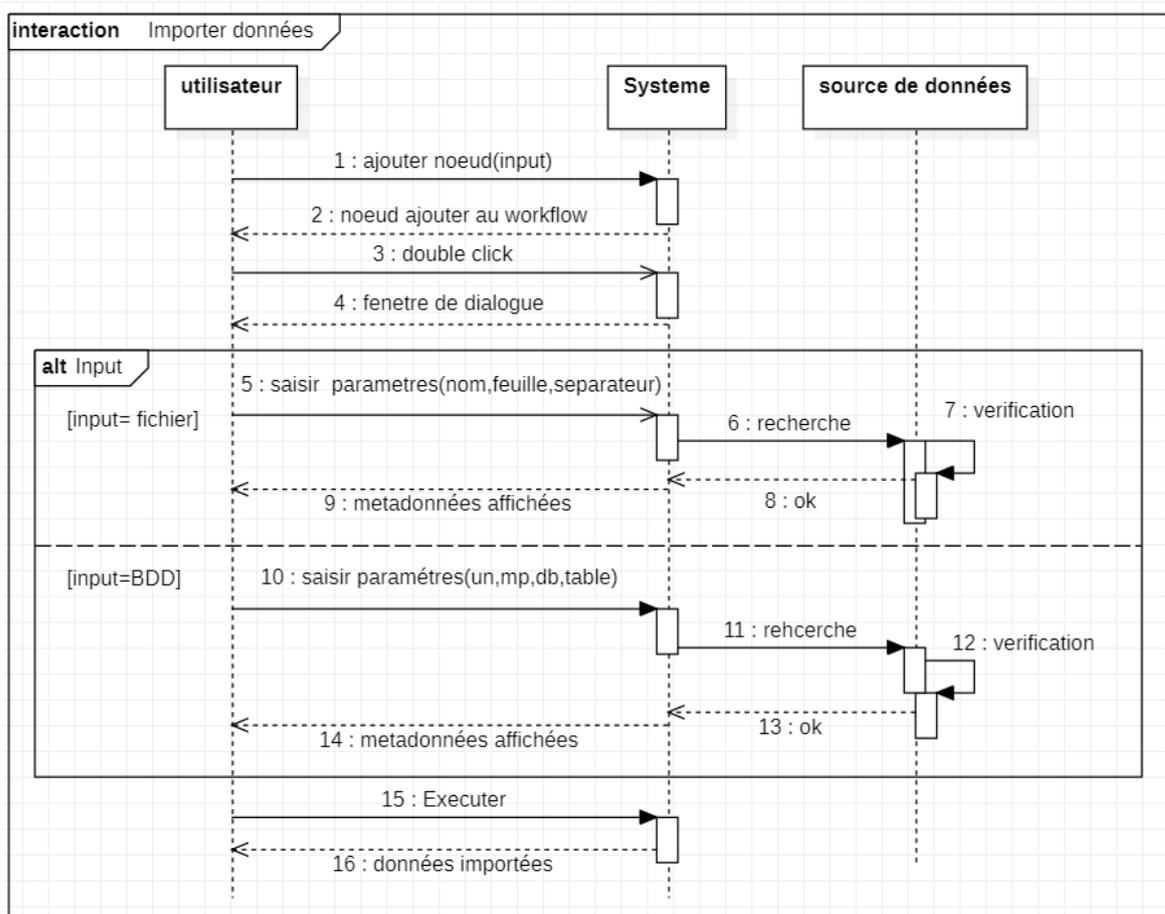


Figure 4.2 : Diagramme de séquence Système « Importer Données »

ii. Diagramme de séquence Système « Traiter Données »

supposant que les données sont récupérées d'une importation directe (Voir Figure 4.3)

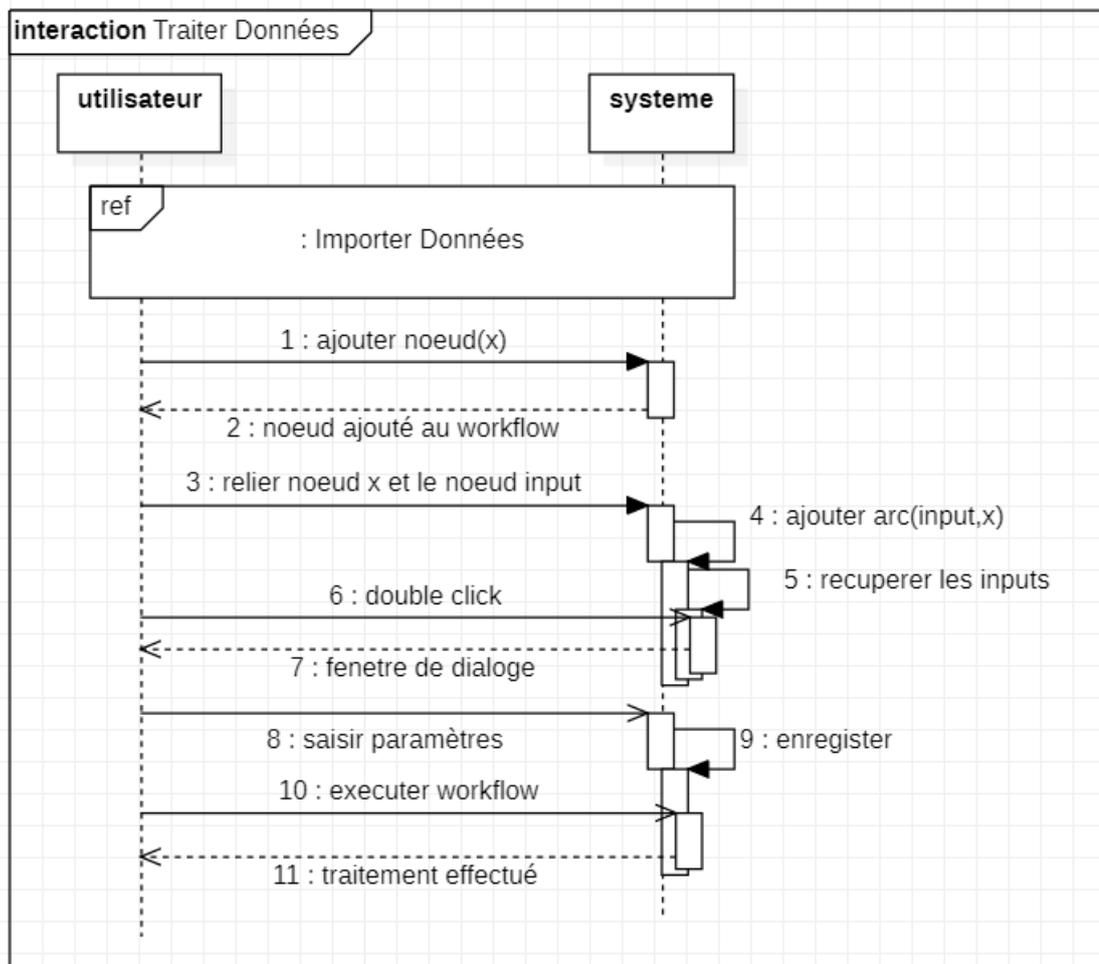


Figure 4.3 : Diagramme de séquence Système « Traiter Données »

iii. Diagramme de séquence Système « Visualiser Données »

supposant que les données sont récupérées d'une importation directe (Voir Figure 4.4)

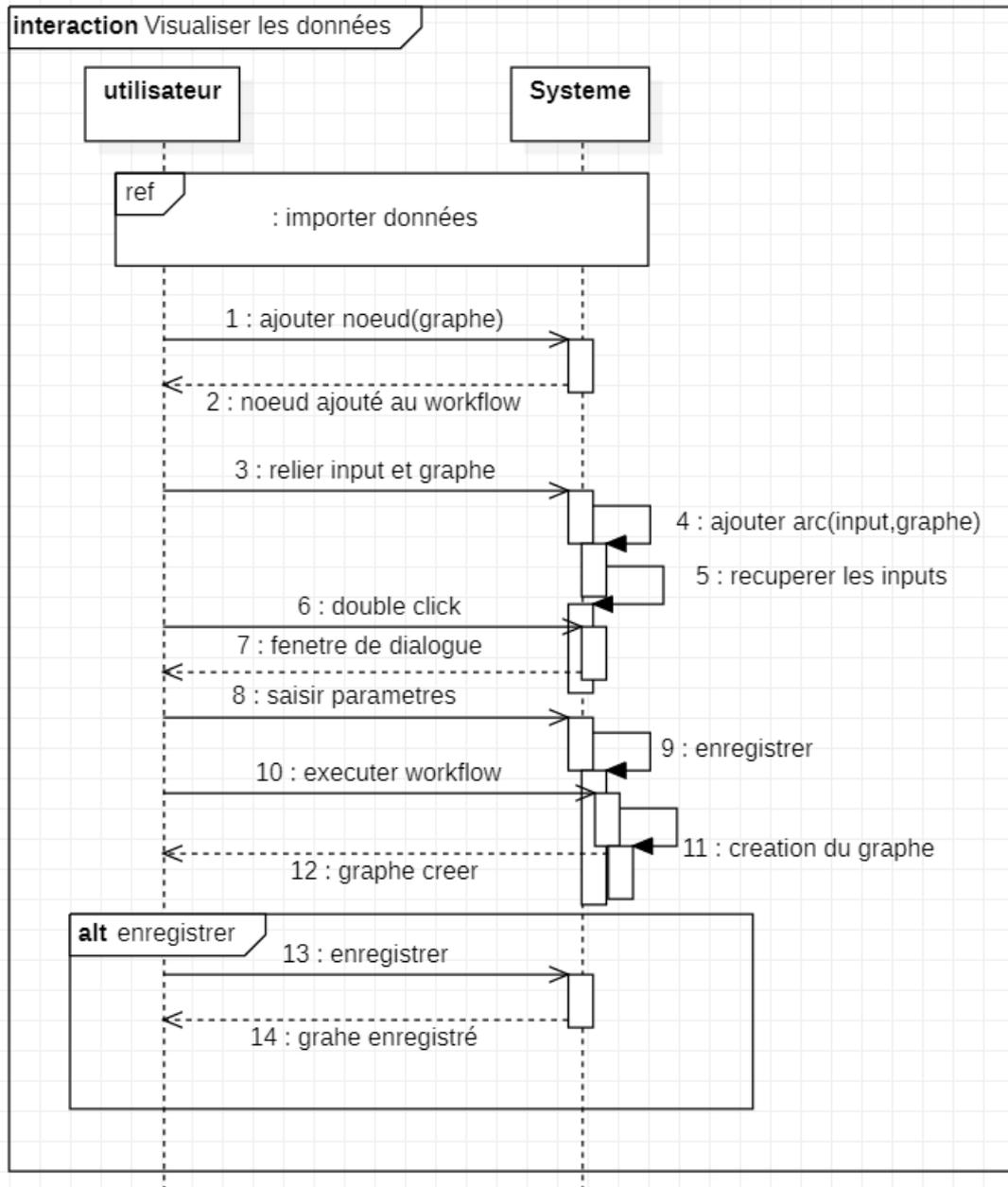


Figure 4.4 : Diagramme de séquence Système « Visualiser Données »

iv. Diagramme de séquence Système « Exporter Données » (Voir Figure 4.5)

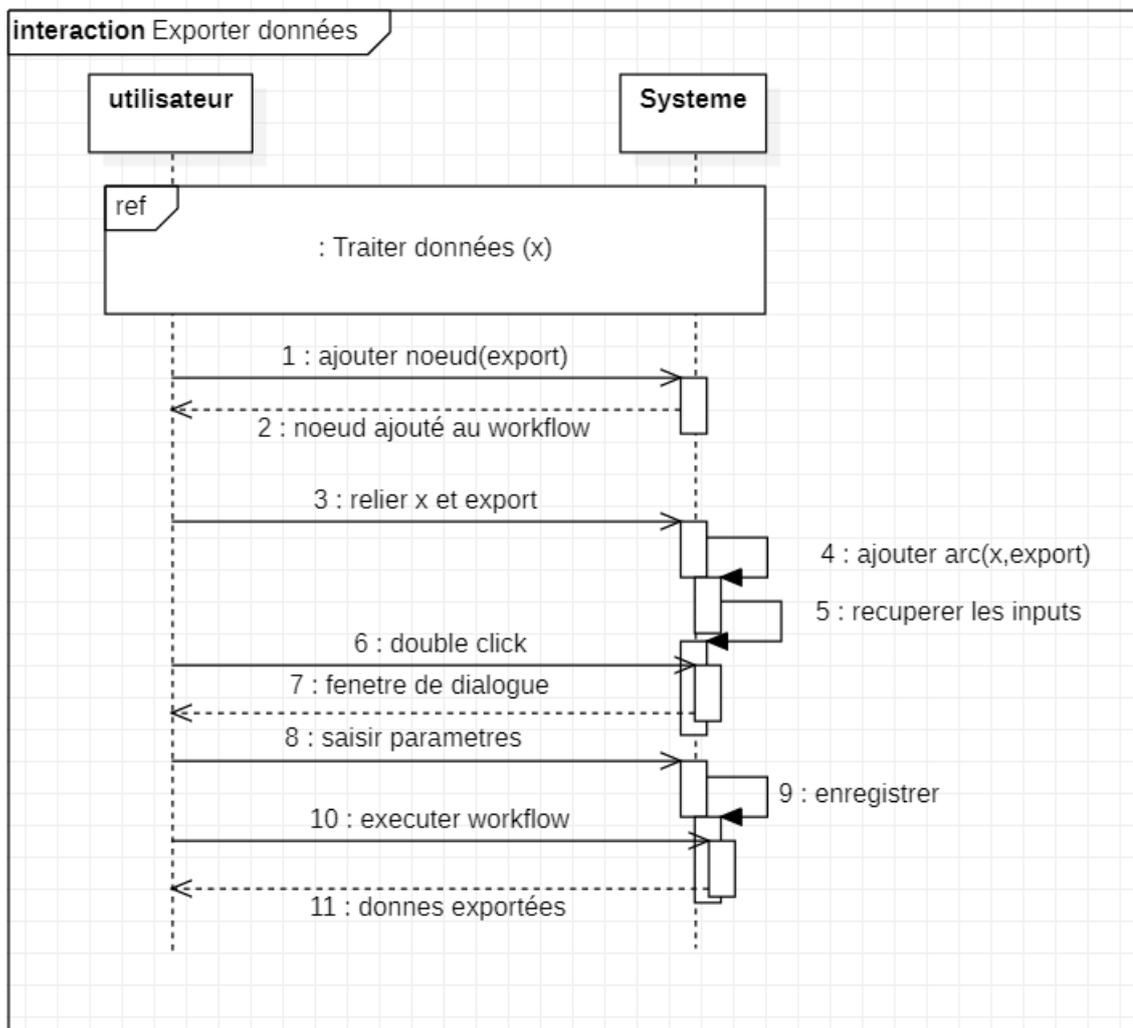


Figure 4.5 : Diagramme de séquence Système « Exporter Données »

3.3 Diagramme de classes :

Le diagramme de classes est un diagramme structurel qui modélise la structure statique du système, en modélisant les objets qui le constituent, et les relations entre ces objets.

La figure 4.6 représente le diagramme de classes UML, contenant les principales classes utilisées dans l'implémentation de notre système.

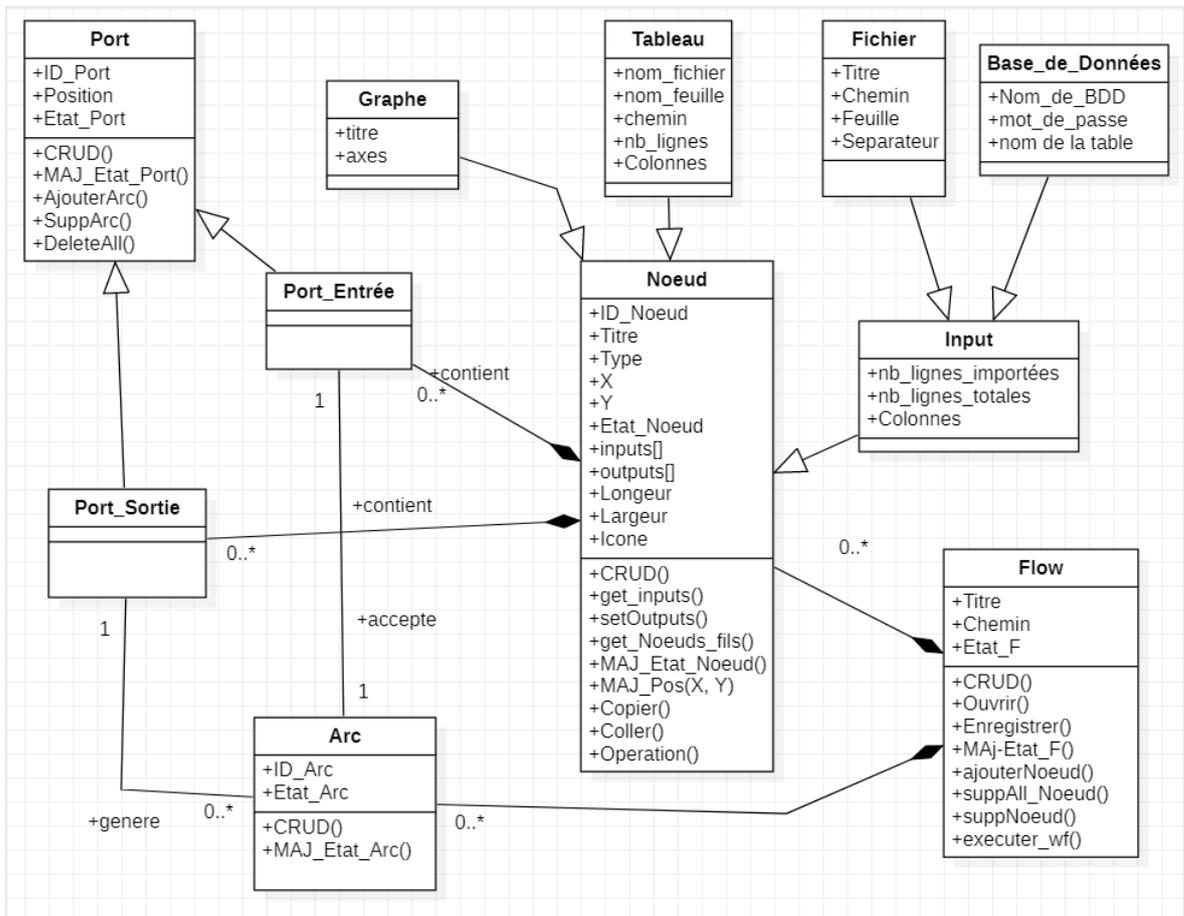


Figure 4.6 : Diagramme de Classes de Conception

4 Conclusion :

Ce chapitre a décrit la conception de notre solution, de la description aux différents diagrammes UML pour donner une vision plus claire sur le fonctionnement et la structure du système. Le prochain chapitre exposera des détails sur l'implémentation et le développement de cette solution.

Chapitre II : Implémentation de la solution

1 Introduction

Ce chapitre sera organisé en trois parties : dans une première partie nous aborderons les outils, ressources et technologies utilisés pour le développement et l'implémentation de notre système. La deuxième partie sera consacrée à l'exposition des différents modules développés et mis en œuvre pour répondre à notre problématique. La troisième partie quant à elle, sera consacrée à la phase de test et résultats où nous y présenterons nos données de test suivis d'une démonstration des résultats obtenus.

2 Outils de Développement

2.1 Ressources matérielles

Notre système a été déployé sur une machine ayant les caractéristiques exposées dans le *Tableau 5.1* suivant :

Système d'exploitation	Windows 10 Professionnel, 64 bits
Processeur	Intel(R) Core(TM) i7-8550U CPU @ 1.80GHz 1.99 GHz
Mémoire RAM	8,00 Go

Tableau 5.1 : Caractéristiques de la machine

2.2 Ressources logicielles

2.2.1 Langage de Programmation



Figure 5.1 : Logo Python

Python est un langage de programmation open source créé par La Python Software Foundation. C'est le langage le plus utilisé pour la Data Science, Pour cause, ce langage est simple, lisible, propre, flexible et compatible avec de nombreuses plateformes, ses nombreuses bibliothèques, telles que pandas et Numpy permettent d'effectuer une large variété de tâches [28].

Pour le développement de notre projet nous avons utilisé la version 3.9.5 de python.

2.2.2 IDE



Figure 5.2 : Logo PyCharm

PyCharm est un environnement de développement intégré développé par la société JetBrains pour programmer en langage Python. C'est un logiciel multiplateforme décliné en deux versions communautaire et professionnelle. La version communautaire a été utilisée dans le cadre de ce projet, étant gratuite et riche en contenant tous les outils nécessaires à l'exécution de notre projet [29].

2.2.3 Bibliothèques

i. Module Pandas



Figure 5.3 : Logo Pandas

Pandas est l'une des bibliothèques de science de données les plus populaires. Elle est aujourd'hui utilisée par un grand nombre de scientifiques et d'analystes. Elle offre de nombreuses fonctionnalités natives très utiles. De la lecture de données en provenance de nombreuses sources, à la création de data frames à partir de ces sources, et permet d'effectuer des analyses agrégées. On peut aussi s'en servir pour la manipulation de tableaux numériques et de séries temporelles [33].

ii. Module PyQt5



Figure 5.4 : Logo PyQt5

PyQt5 est un ensemble complet de liaisons Python pour Qtv5 qui est un ensemble de bibliothèques C++ multiplateformes qui implémentent des API de haut niveau. *PyQt5* est implémenté sous forme de plus de 35 modules permettant ainsi de créer des interfaces graphiques en Python. Elle est publiée sous licence GPL v3 et sous licence commerciale qui permet le développement d'applications propriétaires [30].

iii. Module Numpy



Figure 5.5 : Logo NumPy

NumPy est un package fondamental, multiplateformes et open Source pour le calcul scientifique en *Python*. Il s'agit d'une bibliothèque qui fournit un objet tableau multidimensionnel et un assortiment de routines pour des opérations rapides sur ces tableaux, notamment mathématiques, logiques, manipulation de forme, tri...etc. [31].

iv. Module Matplotlib



Figure 5.6 : Logo Matplotlib

Matplotlib est une bibliothèque complète permettant de créer des visualisations statiques, animées et interactives en *Python*. Elle peut être combinée avec les bibliothèques python de calcul scientifique comme *NumPy*. La licence Matplotlib est basée sur la licence Python Software Foundation (PSF) [32].

3 Développement du Système :

Pour répondre à notre problématique principale de bouclage de chiffre d'affaire et de réconciliation des revenus d'un opérateur télécom mobile, nous avons mis en œuvre, dans le cadre de ce projet une solution logicielle, conçue pour réaliser des traitements et des analyses de données, autrement dit de Data Analytics. Ce système se base sur un nouveau paradigme de programmation qui est la programmation visuelle ou basée sur le flux qui se mettra en évidence à travers l'interface graphique du système. Nous pouvons donc reconnaître deux éléments constitutifs de ce système : L'interface graphique et les nœuds.

3.1 Interface Graphique (GUI)

Notre système s'inspire de la programmation visuelle basée sur le flux, de ce fait le déploiement de l'interface graphique ou GUI a une grande importance dans le développement de l'intégralité du projet. Comme précédemment mentionné, nous avons utilisé le module PyQt5 pour le développement de cette interface.

3.1.1 Présentation de l'interface graphique :

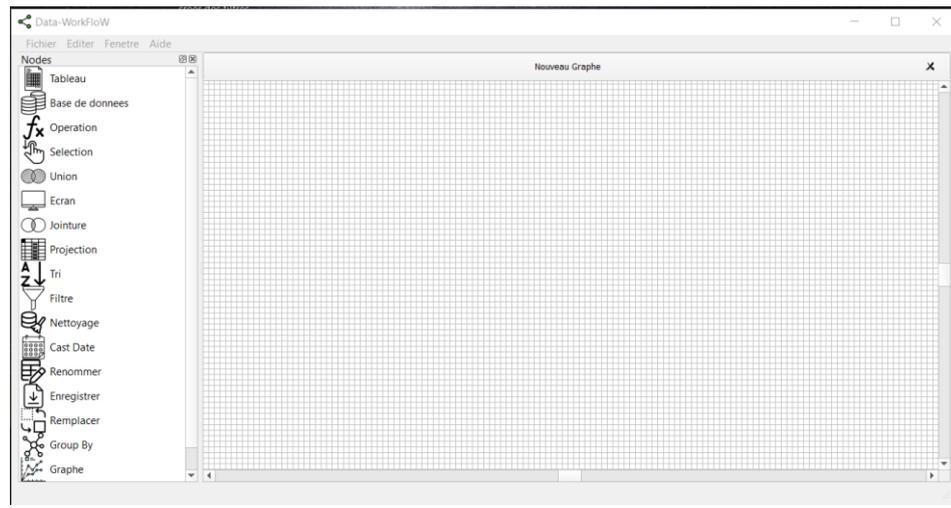


Figure 5.7 : interface graphique Système

Notre GUI se présente, comme le montre la Figure 5.7 ci-dessus, en une fenêtre contenant :

i. Un menu Principal

Il se trouve en haut de la fenêtre, qui se compose de quatre éléments

a) *Aide* : affiche une description sur le projet.

b) *Fenêtre* : permet de :

- Afficher/ masquer la palette de fonctions
- Gérer les différentes feuilles ou zones de travail actives : réordonner, changer de feuille, fermer une feuille, fermer toutes les feuilles.

c) *Editer* : ce sous menu est chargé de la gestion et édition des workflows :

- Copier /Coller : pour copier/coller tout ou une partie du workflow, l'utilisateur peut avoir le même résultat en sélectionnant la partie souhaitée et cliquer sur *Ctrl +c* pour copier et sur *Ctrl+v* pour coller.
- Undo/Redo : pour annuler/refaire la dernière action faite par l'utilisateur. Sinon il suffit de sélectionner la partie souhaitée et cliquer sur *Ctrl +Z* pour Undo et *Ctrl+shift+Z* pour Redo.

- Couper : pour couper ou supprimer un arc, sinon il suffit de sélectionner l'arc et cliquer *Ctrl +x*.
 - Supprimer : permet de supprimer tout ou une partie du workflow. L'utilisateur peut aussi sélectionner la partie souhaitée et cliquer sur *Del*.
- d) *Fichier* : ce sous menu est chargé de la gestion de la zone graphique de travail :
- Nouveau : créer et ouvrir une nouvelle feuille de travail vide
 - Ouvrir : ouvrir une feuille de travail déjà existante (enregistrée)
 - Enregistrer / Enregistrer sous :
 - Fermer : permet de fermer l'application
- ii. Une palette / boite à outils :
- Qui est une liste de fonctions DA mise à disposition par le système, et qui se trouve au volet gauche de la fenêtre.
- iii. Un espace de travail :
- C'est une zone graphique, où seront créés les Workflows et qui se situent au centre de la fenêtre principale.

3.1.2 Workflow :

Le terme workflow dans le cadre de notre projet désigne d'un point de vue visuel une suite de nœuds connectés ou reliés avec des arcs. D'un point de vue fonctionnel c'est une succession/enchaînement d'Operations Data Analytics qui doit impérativement commencer par une importation de données puis suivie d'un ensemble de fonctions ajoutées selon les besoins et traitements souhaités par l'utilisateur afin d'automatiser ces traitements.

- i. Composants d'un Workflow :
- a. Nœud : c'est un composant graphique caractérisé par un titre, une icône, ainsi qu'un ou plusieurs ports permettant de le relier à d'autres

nœuds via des arcs, et par une fonction précise pour le traitement de données.

- b. Arc : une ligne graphique créée à partir d'un un port A indiquant le début vers un port B qui est la fin.

ii. Création d'un Workflow :

- a. L'Ajout d'un nœud :

Avec l'option glisser /déposer offerte par PyQt5, l'utilisateur fait glisser le nœud souhaité depuis la palette et le dépose dans l'espace de travail. (Voir *Figure5.8*)

L'utilisateur doit paramétrer chaque nœud à travers sa fenêtre de dialogue qui s'affiche suite à un double clic sur le nœud.

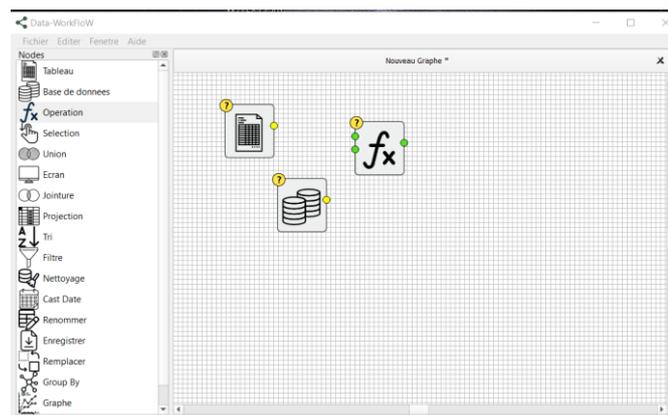


Figure 5.8 : Ajout de Nœuds

- b. L'Ajout d'un arc :

Afin d'effectuer cette opération, il est nécessaire d'avoir au moins deux nœuds. Il suffit de cliquer sur le port de sortie du nœud A, et ainsi sans relâcher, jusqu'au port d'entrée du nœud B (*Figure5.9*).

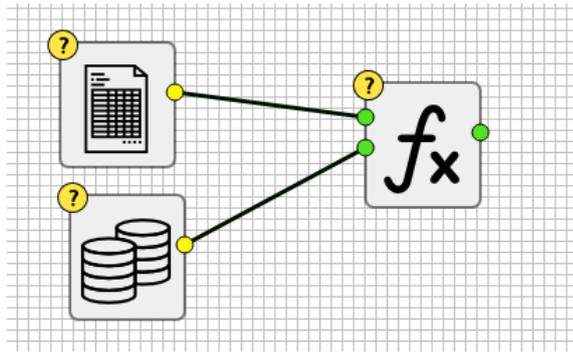


Figure 5.9 : Insertion des arcs

iii. Enregistrement d'un Workflow :

Nous avons conçu une fonction qui enregistre le workflow en un fichier structuré en format JSON qui comportera les informations sur les nœuds et les arcs composants le workflow.

La *Figure 5.10* représente l'enregistrement d'un workflow vide (n'ayant pas de nœuds ni d'arcs) qui ne contient que l'id du workflow, dimension de la zone graphique et les listes de nœuds et arcs sont vides. Au fur et à mesure de l'ajout des nœuds et des arcs, les listes « nodes » et « edges » seront alimentées.

```
{  
  "id": 238445400,  
  "scene_width": 64000,  
  "scene_height": 64000,  
  "nodes": [],  
  "edges": []  
}
```

Figure 5.10 : Enregistrement d'un WF vide

v. Exécution d'un workflow :

Après la création du workflow, les métadonnées sont transmises à partir des arcs pour effectuer le paramétrage des nœuds. L'utilisateur enchaîne ensuite avec

l'exécution du workflow, qui est une suite d'exécution partielle des nœuds composants le workflow selon une file par ordre FIFO.

Cette exécution comprend l'exécution de la fonction principale du nœud ainsi qu'une évaluation de l'état du nœud :

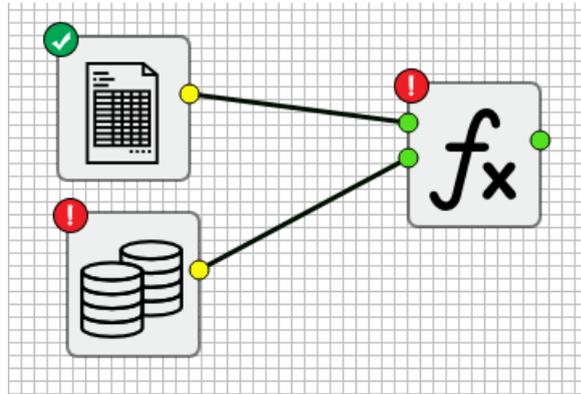


Figure 5.11 : Exécution du WF

Le Vert signifie que :

- Les entrées sont bien connectées.
- les Données récupérées des entrées ont un format valide, ne sont pas vides et correspondent aux paramètres saisis.
- le résultat a un format valide.

Le Rouge signifie que soit :

- Les entrées ne sont pas connectées.
- les Données récupérées des entrées sont vides.
- les paramètres sont invalides et ne correspondent pas aux données.
- problème lors de l'exécution de la fonction, un message contenant l'exception sera généré.

3.2 Fonctions implémentées

Pour ce projet, nous avons mis à la disposition de l'utilisateur une palette contenant de fonctions de data analytics intégrées dans des nœuds, caractérisés par une icône spécifique, un nom et une fonction bien précise. Ses nœuds sont classés en trois types : Inputs, Operations et Outputs.

3.2.1 Les inputs

Sont les nœuds chargés de l'importation des données, ils retournent les données importées dans un format nommé data frame (*Figure 5.12*) qui est une structure identique à celle de table SQL. Dans cette catégorie on trouve deux nœuds :

	Name	Team	Number	Position	Age
0	Avery Bradley	Boston Celtics	0.0	PG	25.0
1	John Holland	Boston Celtics	30.0	SG	27.0
2	Jonas Jerebko	Boston Celtics	8.0	PF	29.0
3	Jordan Mickey	Boston Celtics	NaN	PF	21.0
4	Terry Rozier	Boston Celtics	12.0	PG	22.0
5	Jared Sullinger	Boston Celtics	7.0	C	NaN
6	Evan Turner	Boston Celtics	11.0	SG	27.0

Figure 5.12 : Structure de Data frame

i. Tableau

Ce nœud permet d'importer des données en provenance de fichier *Excel* ou *Csv*

Pour le *paramétrage*, l'utilisateur utilise la fenêtre de dialogue suivante (Voir *Figure 5.13*) afin de saisir les paramètres suivants nécessaires pour importer les données :

- Le chemin et nom du fichier à importer.
- le séparateur dans le cas de fichier *csv* ou nom de la feuille dans le cas de fichier *Excel*.
- Nombre de lignes à importer.

En plus des métadonnées, la fenêtre va donc afficher à l'utilisateur l'information suivante :

- le temps de chargement.
- nombre totale de lignes.
- nom et types des colonnes du fichier.

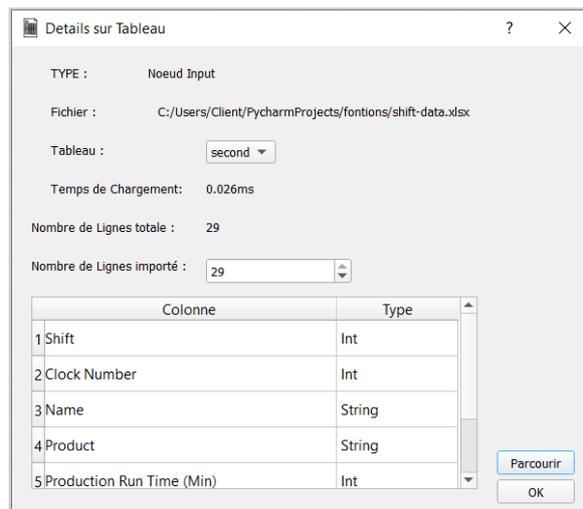


Figure 5.13 : Fenêtre de dialogue du nœud Tableau

Le nœud Tableau, utilise principalement les fonctions `read_csv` ou `read_excel` de la bibliothèque *Pandas*, qui servent à lire les données depuis les paramètres récupérés précédemment.

Les données importé sont donc mises dans un data frame.

ii. Base de données

C'est un nœud qui importe des données depuis des bases de données relationnelles (MySQL par exemple).

Afin de se connecter à la base de données l'utilisateur doit nécessairement indiquer :

- Nom de l'utilisateur et Mot de passe
- Nom de la base de données.
- Nom de la table.
- Nombre de lignes à Importer.

La fenêtre va donc afficher à l'utilisateur l'information suivante : (voir *Figure 5.14*)

- le temps de chargement.
- nombre total de lignes.
- nom et types des colonnes du fichier.

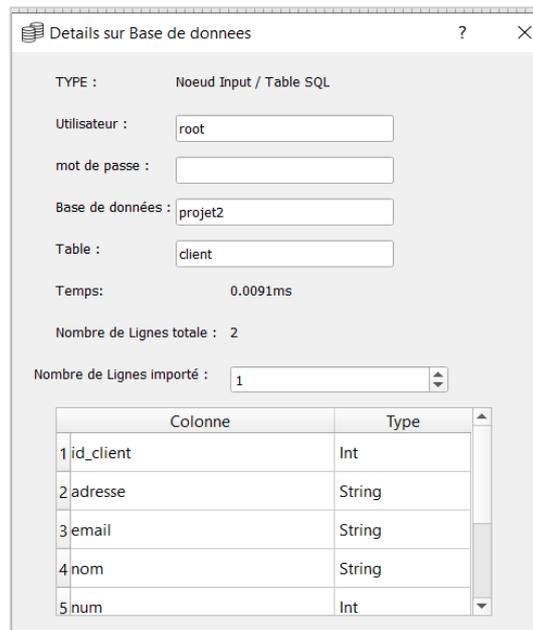


Figure 5.14 : Fenêtre de dialogue du nœud Base de données

Le nœud Base de données commence d'abord par établir la connexion avec la BDD en utilisant le module *mysqlConnector* puis importe les données depuis la table indiquée avec la fonction *read_sql* de pandas.

3.2.2 Les nœuds d'opération :

Sont des nœuds qui effectuent des traitements sur les données introduites en entrée. Ils retournent un data frame de données traitées ; on peut les classer en quatre classes : Nœuds d'Extraction, Nœud de Nettoyage, Nœuds de Transformation et Nœuds de Traitement.

a) Nœuds d'Extraction

Retournent un sous ensemble de données à partir du data frame initial. On y trouve deux nœuds :

i. Sélection

Ce nœud permet de faire une sélection horizontale d'un ensemble de lignes à partir d'un index de début, index de fin et le nombre de lignes (pas) pour chaque index Qui est similaire au *Limit* et *Offset* en SQL (Voir *Figure 5.15*).

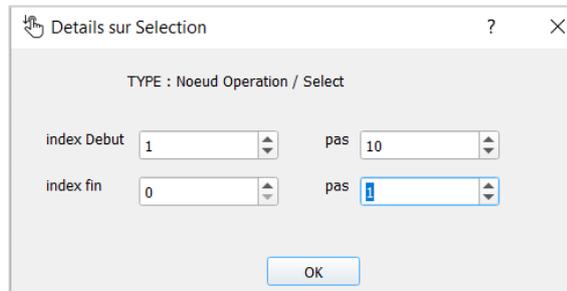


Figure 5.15 : Fenêtre de dialogue du nœud Sélection

ii. projection

Il Permet de sélectionner des colonnes (sélection verticale) depuis un data frame introduit en entrée, d'où créer un sous data frame contenant que les colonnes sélectionnées par l'utilisateur (*Figure 5.16*).

En SQL, c'est l'équivalent de : ' Select c1, c2, c3... from table'.

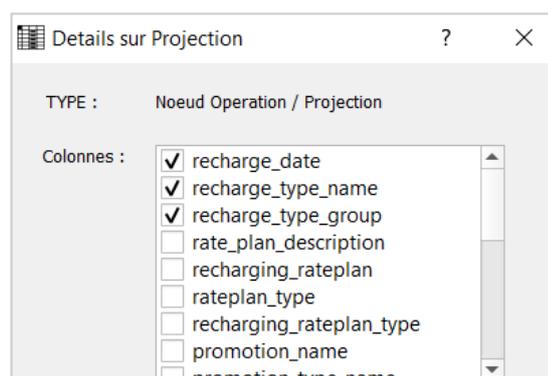


Figure 5.16 : Fenêtre de dialogue du nœud projection

b) Nœud de Nettoyage

Ce nœud Retourne un data frame de données nettoyées, on y trouve trois fonctions intégrées (*Figure 5.17*) dans un même nœud à savoir :

- Suppression des lignes vides.
- Suppression des colonnes vides.
- Suppression des lignes en double.

Pour ce faire, l'utilisateur doit d'abord cocher des cases afin de choisir le traitement à effectuer (voir *Figure 5.21*).

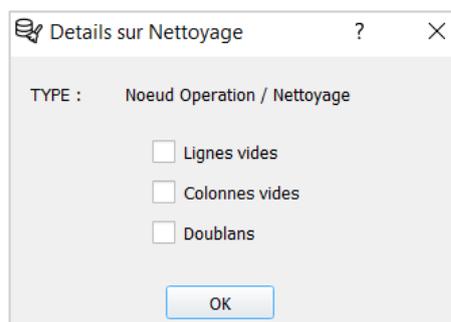


Figure 5.17 : fenêtre de dialogue du Nœud Nettoyage

c) Nœuds de Transformation

Ces nœuds permettent d'effectuer des modifications sur les données importées.

i. Renommer

Depuis la fenêtre de dialogue suivante l'utilisateur peut changer et modifier les noms des colonnes d'un ensemble de valeur. Ça ressemble à la notion d'alias en SQL (Voir *Figure 5.18*).

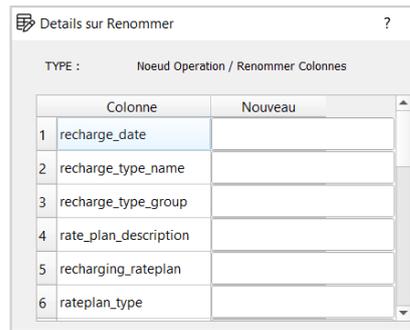


Figure 5.18 : fenêtre de dialogue de Renommer Colonnes

ii. Remplacer

La Figure 5.19 est une la fenêtre de dialogue qui permet de remplacer une valeur X existante dans une colonne C par une nouvelle valeur Y, toutes récupérées depuis un tableau.

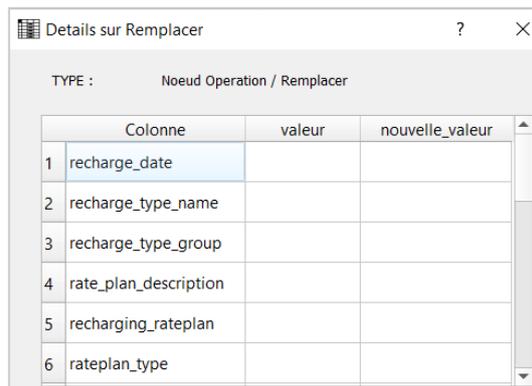


Figure 5.19 : Fenêtre de dialogue du Nœud Remplacer

iii. Extract Date

Cette fonction va trouver et extraire la date existante dans une chaîne de caractères puis remplacer cette chaîne par la date trouvée. Ainsi le type de la colonne changera de String à Date.

d) Nœuds de Traitement

i. Filtre

Ce nœud ne retourne que les lignes où les conditions introduites par l'utilisateur sont satisfaites, c'est l'équivalent du *WHERE* en SQL.

L'utilisateur choisit pour chaque colonne un parmi les opérateurs : égale, inférieur, supérieur, inférieur ou égale, supérieur ou égale, différent ainsi que la valeur et ce via la fenêtre de dialogue.

ii. Join

Ce nœud permet d'assembler les informations d'un tableau A avec celles d'un autre tableau B selon des critères choisis. Ces conditions sont faites sur une ou plusieurs colonnes communes à A et B qui effectuent une correspondance entre les 2 tableaux. Les conditions de jointure sont récupérées depuis le tableau de la fenêtre montré dans la *Figure 5.20*.

Details sur Jointure

TYPE : Noeud Operation / Join

	Colonne A	Operateur	Colonne B
1	recharge_date	Egale	recharge_date
2	recharge_type_...	Different	recharge_date
3	recharge_type_...	Aucun	recharge_date
4	rate_plan_descri...	Different	recharge_date
5	recharging_rate...	Egale	recharge_date
6	rateplan_type	Egale	recharge_date

Figure 5.20 : fenêtre de dialogue du nœud Join

iii. Union

Il permet de créer un tableau (data frame) qui est le résultat de la concaténation du tableau A et B. Ce nouveau tableau va donc contenir toutes les lignes du premier tableau A et toutes les lignes du second. C'est l'équivalent d'une *UNION* en SQL (Voir *Figure 5.21*).

A partir de la fenêtre de dialogue suivante l'utilisateur pourra choisir les colonnes du résultat de concaténation.

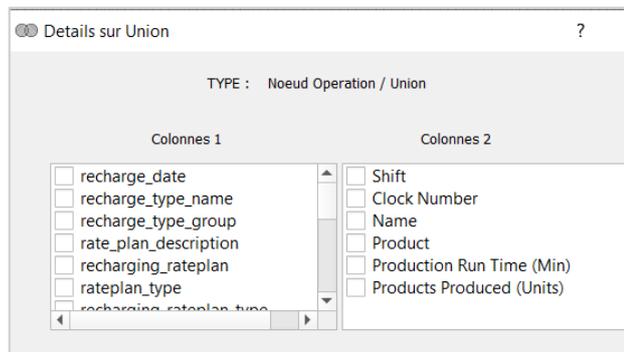


Figure 5.21 : fenêtre de dialogue du nœud Union

iv. Group By

Cette fonction permet d'accomplir une agrégation par groupe avec les opérateurs Max, Min, SUM, AVG, Count. C'est semblable à ce qu'on peut obtenir en utilisant le mot clé GROUP BY en SQL.

v. Order By

Ce nœud permet d'ordonner l'ensemble de données introduites en entrée selon un ordre croissant ou décroissant et selon le choix de l'utilisateur sur une colonne, c'est l'équivalent d'un Order By en SQL.

vi. Opération

Cette fonction permet d'effectuer des opérations arithmétiques sur les données, à savoir des additions, soustractions, multiplications et divisions, entre deux colonnes, une colonne et un entier ou entres deux entier, puis le résultat sera mis dans une nouvelle colonne nommée par l'utilisateur via la fenêtre de dialogue (Voir Figure 5.22).

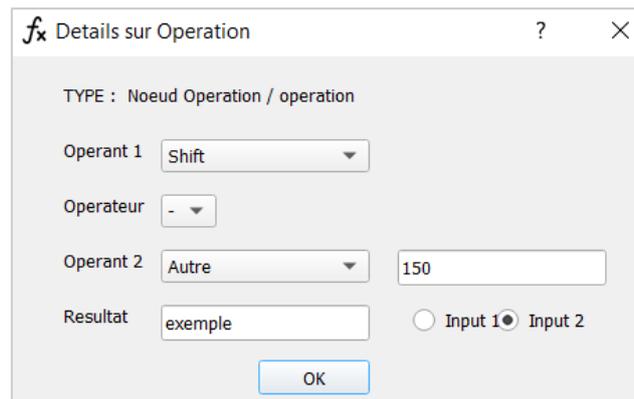


Figure 5.22 : fenêtre de dialogue du nœud Opération

3.2.3 Les Outputs

Sont les nœuds chargés de la présentation des résultats obtenus après traitements. On y trouve trois nœuds

i. Ecran

Permet d'avoir une vue sur le contenu d'un data frame, en affichant les données introduites à travers le port d'entrée dans un tableau.

ii. Graphe

Cette fonction crée des graphes à partir des données mises en entrée. L'utilisateur va donc choisir les axes, ainsi que leurs noms et le nom du graphe via la fenêtre de dialogue (Voir Figure 5.23).

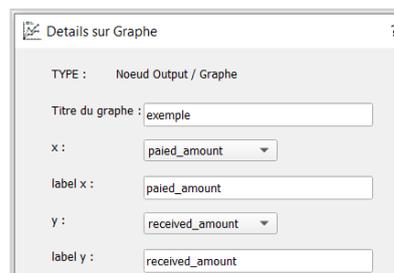


Figure 5.23 : Fenêtre de dialogue du Nœud Graphe

iii. Export

Ce nœud récupère le chemin de destination, nom et caractéristiques du fichier : nom de la feuille, séparateur, type du fichier (Voir *Figure 5.24*).

Puis il enregistre les données introduites en entrées sous format data frame dans un fichier.

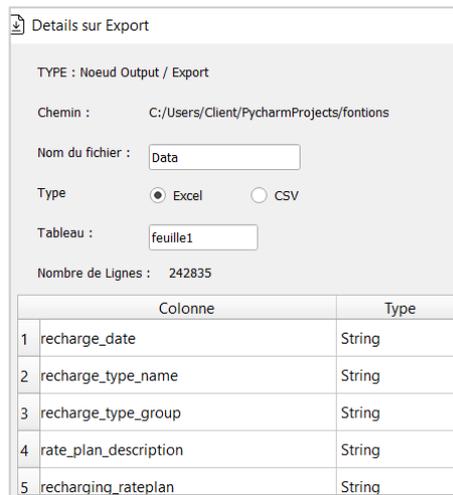


Figure 5.24 : Fenêtre de dialogue du Nœud Graphe

4 Test de l'application

4.1 Workflow de test

La *Figure 5* suivante montre le workflow qui va permettre de tester les nœuds de l'application.

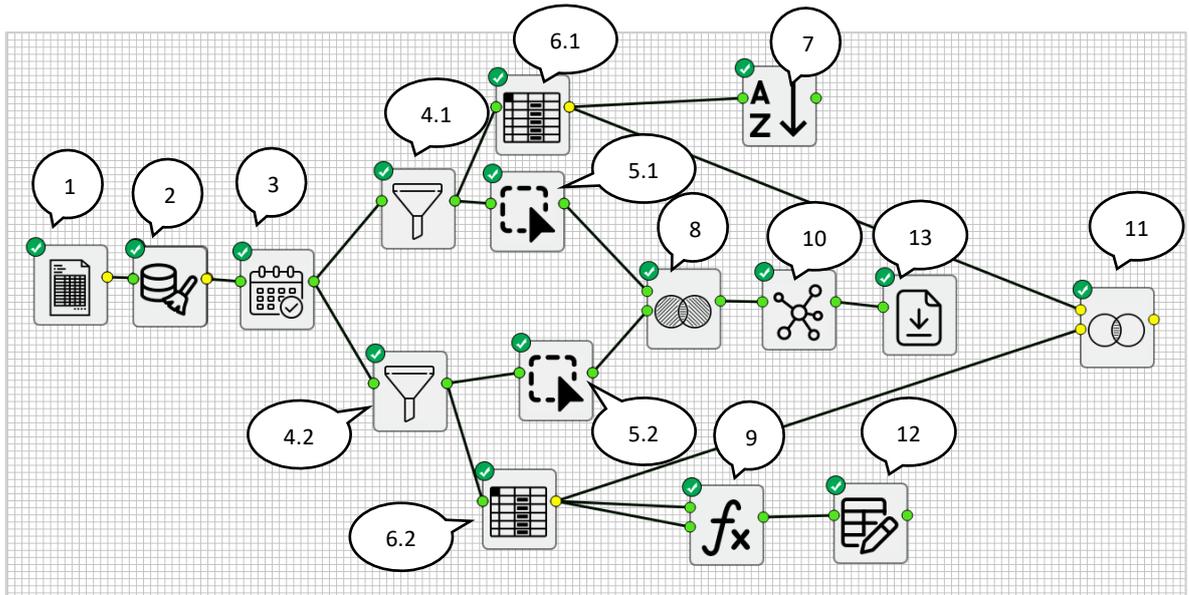


Figure 5 : Workflow de Test

4.2 Résultat de l'exécution

1. Tableau

La figure 5.26 est une vue des données importer avec le nœud tableau

	call_date	service	rvive_grou	destination_nam	destination_type_name	or_intern	arge_amo	ainaccour	ed_durati	ne_gro	one_nam	ator_sms
0	2020-01-08T00:00:00.00...	Texto (SMS)SMS-...	SMS	FRANCE	International	I	10.0	10.0	0...	nan	nan	nan
1	2020-01-28T00:00:00.00...	Communications	CALL	Luxembourg	International	I	35.0	0.0	0.0	nan	nan	nan
2	2020-01-07T00:00:00.00...	Communications	CALL	Djezzy	DJEZZY	L	111.67	21.16	4.45	nan	nan	nan
3	2020-01-16T00:00:00.00...	GPRS Basic Servic...	GPRS	nan	nan	nan	0.0	0.0	0.0	nan	nan	nan
4	2020-01-07T00:00:00.00...	GPRS Basic Servic...	GPRS	nan	nan	nan	1.21	1.21	0.0	nan	nan	nan

Figure 5.25 : résultat de l'importation de données

2. Nettoyage

La figure 5.26 représente le résultat retourné par le nœud de nettoyage après la suppression des lignes et colonnes vides.

Ecran								
	call_date	service	service_group	destination_name	stination_type_nar	charge_amount	mainaccount	rated_duration
0	2020-01-08T00:00:00.00000000	Texto (SMS)SM...	SMS	FRANCE	International	10.0	10.0	0.0333333333333333
1	2020-01-28T00:00:00.00000000	Communications	CALL	Luxembourg	International	35.0	0.0	0.0
2	2020-01-07T00:00:00.00000000	Communications	CALL	Djezzy	DJEZZY	111.67	21.16	4.45
3	2020-01-16T00:00:00.00000000	GPRS Basic ...	GPRS	nan	nan	0.0	0.0	0.0
4	2020-01-07T00:00:00.00000000	GPRS Basic ...	GPRS	nan	nan	1.21	1.21	0.0
5	2020-01-23T00:00:00.00000000	GPRS Web pag...	GPRS	nan	nan	0.0	0.0	0.0
6	2020-01-15T00:00:00.00000000	Communications	CALL	callcentre	CALL CENTER	12.0	0.0	0.0
7	2020-01-15T00:00:00.00000000	Communications	CALL	Wataniya ...	OOREDOO	0.0	0.0	0.0

Figure 5.26 : Résultat du nettoyage de données

3. Extraction de la Date

La figure 5.27 suivante montre le résultat obtenu après reformatage du champ call_date.

Ecran								
	call_date	service	service_group	destination_name	stination_type_nar	charge_amount	mainaccount	rated_duration
	2020-01-08	Texto (SMS)SM...	SMS	FRANCE	International	10.0	10.0	0...
	2020-01-28	Communications	CALL	Luxembourg	International	35.0	0.0	0.0
	2020-01-07	Communications	CALL	Djezzy	DJEZZY	111.67	21.16	4.45
	2020-01-16	GPRS Basic ...	GPRS	nan	nan	0.0	0.0	0.0
	2020-01-07	GPRS Basic ...	GPRS	nan	nan	1.21	1.21	0.0
	2020-01-23	GPRS Web pag...	GPRS	nan	nan	0.0	0.0	0.0
	2020-01-15	Communications	CALL	callcentre	CALL CENTER	12.0	0.0	0.0

Figure 5.27 : Résultat obtenu après extraction de la date

4. Filtrage par service_group

4.1. La figure 5.28 suivante représente le sous ensemble de données satisfaisant la condition : service_group = 'CALL'

Ecran									
	call_date	service	service_group	destination_name	stination_type_nar	indicator_internation	charge_amount	mainaccount	rated_duration
	2020-01-28	Communications	CALL	Luxembourg	International	I	35.0	0.0	0.0
	2020-01-07	Communications	CALL	Djezzy	DJEZZY	L	111.67	21.16	4.45
	2020-01-15	Communications	CALL	callcentre	CALL CENTER	L	12.0	0.0	0.0
	2020-01-15	Communications	CALL	Wataniya ...	OOREDOO	L	0.0	0.0	0.0
	2020-01-14	Communications	CALL	Voicemail	VOICE MAIL	L	0.0	0.0	0.0

Figure 5.28 : Résultat de la requête 4.1

4.2. La *figure 5.29* suivante représente le sous ensemble de données satisfaisant la condition : `service_group <>'CALL'`



call_date	service	service_group	destination_name	stination_type_nar	dicator_internation	charge_amount	mainaccount	rated_duration
2020-01-08	Texto (SMS)SM...	SMS	FRANCE	International	I	10.0	10.0	0...
2020-01-16	GPRS Basic ...	GPRS	nan	nan	nan	0.0	0.0	0.0
2020-01-07	GPRS Basic ...	GPRS	nan	nan	nan	1.21	1.21	0.0
2020-01-23	GPRS Web pag...	GPRS	nan	nan	nan	0.0	0.0	0.0
2020-01-14	Texto (SMS)SM...	SMS	Djezzy	DJEZZY	L	2.0	0.0	0.0
2020-01-31	GPRS Basic ...	GPRS	nan	nan	nan	0.0	0.0	0.0

Figure 5.29 : Résultat de la requête 4.2

5. Sélection de ligne à partir de leurs index

5.1. Sélection de trois premières lignes du résultat de la requête 4.1. (Voir *Figure 5.30*)

	call_date	service	service_group	destination_name	stination_type_nar	dicator_internation	charge_amount	mainaccount	rated_duration
1	2020-01-28	Communications	CALL	Luxembourg	International	I	35.0	0.0	0.0
2	2020-01-07	Communications	CALL	Djezzy	DJEZZY	L	111.67	21.16	4.45
6	2020-01-15	Communications	CALL	callcentre	CALL CENTER	L	12.0	0.0	0.0

Figure 5.30 : Résultat du nœud 5.1

5.2. Sélection de 3 dernières lignes du résultat de la requête 4.2. (Voir *Figure 5.31*)

	call_date	service	service_group	destination_name	stination_type_nar	dicator_internation	charge_amount	mainaccount	rated_duration
196	2020-01-20	Texto (SMS)SM...	SMS	Wataniya ...	OOREDOO	L	5.0	5.0	0...
197	2020-01-24	GPRS Basic ...	GPRS	nan	nan	nan	0.0	0.0	0.0
198	2020-01-24	GPRS Basic ...	GPRS	nan	nan	nan	0.0	0.0	0.0

Figure 5.31 : Résultat du nœud 5.2

6. Projection

6.1. Sur les colonnes : 'call_date', 'service_group' et 'service' du résultat de la requête du nœud 4.1 (Voir Figure 5.32).

	call_date	service_group	service
1	2020-01-28	CALL	Communications
2	2020-01-07	CALL	Communications
6	2020-01-15	CALL	Communications

Figure 5.32 : Résultat du de la requête 6.1

6.2. Sur les colonnes 'call_date', 'charge_amount' et 'mainaccount' du résultat de la requête du nœud 4.2 (Voir Figure 5.32).

	call_date	charge_amount	mainaccount
0	2020-01-08	10.0	10.0
3	2020-01-16	0.0	0.0
4	2020-01-07	1.21	1.21

Figure 5.32 : Résultat du de la requête 6.2

7. Tri de données retournées par le nœud 6.1 par ordre croissant de la colonne 'call_date' (Voir Figure 5.33).

2020-01-05	CALL	Communications
2020-01-06	CALL	Communications
2020-01-07	CALL	CALL_FORWARD
2020-01-07	CALL	Communications

Figure 5.33 : Résultat du tri

8. Le nœud *Union* concatène les résultats de 5.1 et 5.2 pour obtenir le résultat suivant (Voir *Figure 5.34*)

call_date	service	charge_amount
2020-01-28	Communications	35.0
2020-01-07	Communications	111.67
2020-01-15	Communications	12.0
2020-01-20	Texto (SMS)SM...	5.0
2020-01-24	GPRS Basic ...	0.0
2020-01-24	GPRS Basic ...	0.0

Figure 5.34 : Résultat du de l'union

9. Fonction arithmétique

La colonne 'Résultat' (voir *Figure 5.35*), est la différence entre les colonnes 'charge_amount' et 'mainaccount' .

charge_amount	mainaccount	Resultat
10.0	10.0	0.0
0.0	0.0	0.0

Figure 5.35 : Résultat de la différence

10. Le nœud Group By groupe les données par la colonne call_date et type de service
11. Jointure naturelle du résultat retourné du nœud 6.1 et celui du nœud 6.2. (voir *Figure 5.37*)

call_date	service_group	service	charge_amount	mainaccount
2020-01-28	CALL	Communications	0.0	0.0
2020-01-28	CALL	Communications	0.0	0.0
2020-01-28	CALL	Communications	0.0	0.0
2020-01-28	CALL	Communications	0.0	0.0

Figure 5.37 : Résultat du de la jointure

12. Renommer les colonnes 'call_date', 'charge_amount', 'mainaccount' par Date, Montant-Chargé et compte respectivement. (Voir la *Figure 5.38*)

	Date	Montant-chargé	compte	Resultat
0	2020-01-08	10.0	10.0	0.0

Figure 5.38 : Résultat du nœud 12

5 Conclusion

Dans ce chapitre nous avons présenté les ressources matérielles et technologies utilisées pour la réalisation de ce travail. Puis nous avons énuméré toutes les fonctionnalités de notre application, depuis l'importation de données, jusqu'à l'exportation en passant par l'extraction, le nettoyage et la transformation.

Nous avons également testé notre application, à travers un workflow de test. D'après les résultats obtenus on constate que notre plateforme répond correctement au besoin de la problématique et au objectif fixés à savoir une plateforme de data analytics qui automatise les traitements de données financières.

Conclusion Générale et Perspectives

Le bouclage du chiffre d'affaires d'un opérateur télécom mobile est un processus complexe qui se résume par une étape de vérification de l'infrastructure suivie de la comptabilisation et réconciliation des revenus et revenus différés pour enfin reconnaître le Chiffre d'Affaires.

Les outils dédiés au revenue assurance existants ne prennent en charge que l'automatisation de la vérification de l'infrastructure, d'où est apparu le besoin d'automatisation de la suite du processus à savoir la comptabilisation et la réconciliation des revenus et revenus différés qui fait l'objet de ce travail.

Notre solution à cette problématique est une plateforme de Data Analytics facilitant le contrôle et comptabilisation du Chiffre d'Affaires.

Via cette plateforme, l'utilisateur pourra importer des données depuis différentes sources en mode offline (fichier et/ou Bases de données). Ces fichiers représenteront les inputs ou la matière sur laquelle seront effectués les contrôles.

Le contrôle du chiffre d'affaires exige des filtrages par type de revenus, services et/ou par date, des regroupements par date et/ou type de services, des agrégats par somme, moyenne ou autres pour comptabiliser les revenus, en passant par des étapes de nettoyage, de tri et de transformation de données. Les résultats obtenus seront concaténés par des jointures et/ou unions pour enfin être visualisés ou exportés vers d'autres sources.

Pour satisfaire le besoin d'automatisation, nous avons opté pour l'utilisation de la programmation visuelle qui apporte un meilleur contrôle sur les procédures de tests de bouclage des données financières sans qu'il y ait besoin d'avoir des connaissances pré-requises en programmation avancée.

Cet outil est adaptable et peut sauvegarder des versions de séries de traitements de données qui peuvent être réutilisées ultérieurement. Ainsi il garantit que le traitement soit automatique et uniforme à travers un workflow complet qui pourra faire objet d'une preuve d'audit.

Bien que notre outil soit fonctionnel et répond correctement à notre besoin d'automatisation de flux de reporting pour le bouclage de chiffres d'affaire d'un opérateur télécom, il est toujours amené à être amélioré. Ainsi, en guise de mise en perspective, nous proposons :

- Se connecter a des sources de données NoSQL tels que Mongo DB et Hadoop base ou HDFS par exemple.
- Effectuer des extractions de données à partir de fichiers non structurés.
- Exporter les données dans des bases NoSQL.
- Enrichir l'outil en rajoutant d'autres fonctions spécifiques au type de données (traitement avancé des dates par exemple).
- Généraliser l'outil pour qu'il soit utilisable dans n'importe quel domaine.
- Développer une version Web de l'outil pour en faciliter l'installation et la mise à jour.

Références bibliographiques

- [1] H. Alexander *et al.*, "The importance of telco's revenue assurance function study," *WCC 2000 - ICCT 2000. 2000 International Conference on Communication Technology Proceedings (Cat. No.00EX420)*, 2000, pp. 826-833 vol.1, doi: 10.1109/ICCT.2000.889322.
- [2] R.Mattison, "The Telco's revenue assurance handbook" , Edited and illustrated by Chrysanthemum Mattison Brigitte Mattison XiT Press Oakwood Hills, Illinois ; Published by XiT Press, Oakwood Hills, Illinois, USA - 2005 Copyright © Rob Mattison, 2005, Internatinonal Standard Book Number: 1-4116-2801-2
- [3] J.Paul Morrison. (2009). FLOW-BASED PROGRAMMING (van Nostrand Reinhold, 1994 éd.). <http://www.jpaulmorrison.com/fbp/book.pdf>
- [4] Observatoire du marché de la téléphonie mobile en Algérie 4ème trimestre 2020. (2020). AUTORITÉ DE RÉGULATION DE LA POSTE ET DES COMMUNICATIONS ELECTRONIQUES. <https://www.arpce.dz/fr/file/k3x9f7>
- [5] IIBA. (2019). INTRODUCTION TO BUSINESS DATA ANALYTICS _ A PRACTITIONER View. International Institute of Business Analysis, Toronto, Ontario, Canada. <https://www.iiba.org/globalassets/standards-and-resources/business-data-analytics/files/introduction-to-bda-practitioner-2.pdf>
- [6] KPMG Algérie S.P.A. (2021). Data Analytics : Un must have pour les prochaines décennies. KPMG ACTU'. Published.
- [7] KPMG Algerie. (2021). *Livret d'accueil KPMG Algérie*.
- [8] KPMG Algerie. (2017). *Document interne*.
- [9] 1G 2G 3G 4G *généralisation de systéme cellulaire définitive*. (s. d.). marche-public.fr. Consulté le 17 juin 2021, à l'adresse <http://www.marche-public.fr/Terminologie/Entrees/1G.htm>
- [10] *Telecom Billing - Quick Guide - Tutorialspoint*. (s. d.). Tutorialspoint. Consulté le 23 juin 2021, à l'adresse <https://www.tutorialspoint.com/telecom-billing/quick-billing-guide.htm>
- [11] *Mobileum*. (s. d.). Risk & Assurance Group. Consulté le 19 juin 2021, à l'adresse <https://riskandassurancegroup.org/mobileum/>

- [12] Subex Limited. (s. d.). *Telecom Business Assurance | Revenue Assurance | Risk | 5G*. Consulté le 19 juin 2021, à l'adresse <https://www.subex.com/business-assurance/>
- [13] Panamax Inc. (s. d.). *Telecom Revenue Assurance Solution | Revenue Assurance System*. Panamax, Inc. Consulté le 19 juin 2021, à l'adresse <https://www.panamaxil.com/revenue-assurance>
- [14] RÉ ; Daction, L. (2019, 3 février). *Chiffre d'affaires (CA) : définition, calcul simple, en anglais*. JDN. <https://www.journaldunet.fr/business/dictionnaire-comptable-et-fiscal/1198401-chiffre-d-affaires-ca-definition-traduction/>
- [15] *Deferred Revenue Definition*. (s. d.). Investopedia. Consulté le 16 septembre 2021, à l'adresse <https://www.investopedia.com/terms/d/deferredrevenue.asp>
- [16] Sion, M. (2020, 14 février). *Qu'est-ce que la norme IFRS 15 ? 1/2*. Le blog des Fonctions financières. <https://www.leblogdesfinanciers.fr/2019/03/11/quest-ce-que-la-norme-ifrs-15-1-2/>
- [17] *What Is Reconciliation in Accounting?* (s. d.). Investopedia. Consulté le 16 septembre 2021, à l'adresse <https://www.investopedia.com/terms/r/reconciliation.asp>
- [18] *Impôt sur le revenu - Comment sont imposés les revenus différés ?* (s. d.). service-public.fr. Consulté le 17 juin 2021, à l'adresse <https://www.service-public.fr/particuliers/vosdroits/F34064#:~:text=Les%20revenus%20diff%C3%A9r%C3%A9s%20sont%20des,de%20salaire%20ou%20d'avancement>
- [19] Stedman, C. (s. d.). data analytics (DA). SearchDataManagement. Consulté le 20 mai 2021, à l'adresse <https://searchdatamanagement.techtarget.com/definition/data-analytics>
- [20] What is an Analytics Platform? (s. d.). TIBCO Software. Consulté le 22 mai 2021, à l'adresse <https://www.tibco.com/reference-center/what-is-an-analytics-platform>
- [21] Salunke, R. (s. d.). Data Analytics Technique. EDUCBA. Consulté le 25 mai 2021, à l'adresse <https://www.educba.com/data-analytics-technique/>
- [22] What is Data Analytics. (s. d.). EDUCBA. Consulté le 15 mai 2021, à l'adresse <https://www.educba.com/what-is-data-analytics/>
- [23] Self-Service Analytics, Data Science & Process Automation | Alteryx. (s. d.). alteryx. Consulté le 25 mai 2021, à l'adresse <https://www.alteryx.com/fr>
- [24] Data Analytics and Artificial Intelligence Applications | Altair. (s. d.). ALTAIR. Consulté le 25 juin 2021, à l'adresse <https://www.altair.com/data-analytics/>
- [25] Analytique avancée : Data Mining et analyse prédictive. (s. d.). SAS FR. Consulté le 25 juin 2021, à l'adresse https://www.sas.com/fr_fr/solutions/analytics.html

- [26] *Qu'est-ce qu'un diagramme de flux de données ?* (s. d.). Lucidchart. Consulté le 12 avril 2021, à l'adresse <https://www.lucidchart.com/pages/fr/diagramme-de-flux-de-donnees>
- [27] *Workflow [Veille technologique]*. (s. d.). ICS, Amexio. Consulté le 26 avril 2021, à l'adresse <https://ics.utc.fr/c2m/DOCS/L2e/html/co/Workflow.html>
- [28] *Welcome to*. (s. d.). Python.Org. Consulté le 14 août 2021, à l'adresse <https://www.python.org/>
- [29] KooR.fr. (s. d.). *KooR.fr - Présentation de l'IDE PyCharm - Le tutoriel/cours sur le langage de programmation Python*. Consulté le 14 août 2021, à l'adresse https://kooR.fr/Python/Tutorial/python_ide_pycharm.wp
- [30] *PyQt5*. (2021, 10 mars). PyPI. <https://pypi.org/project/PyQt5/>
- [31] *Overview — NumPy v1.21 Manual*. (s. d.). OverView. Consulté le 14 août 2021, à l'adresse <https://numpy.org/doc/stable/>
- [31] *Matplotlib : Python plotting — Matplotlib 3.4.3 documentation*. (s. d.). matplotlib. Consulté le 14 août 2021, à l'adresse <https://matplotlib.org/>
- [32] *pandas documentation — pandas 1.3.3 documentation*. (s. d.). Pandas. Consulté le 12 août 2021, à l'adresse <https://pandas.pydata.org/docs/>