

Université de Yaoundé I
Faculté des Sciences
Département d'Informatique
Master II Pro
Matière : INF561



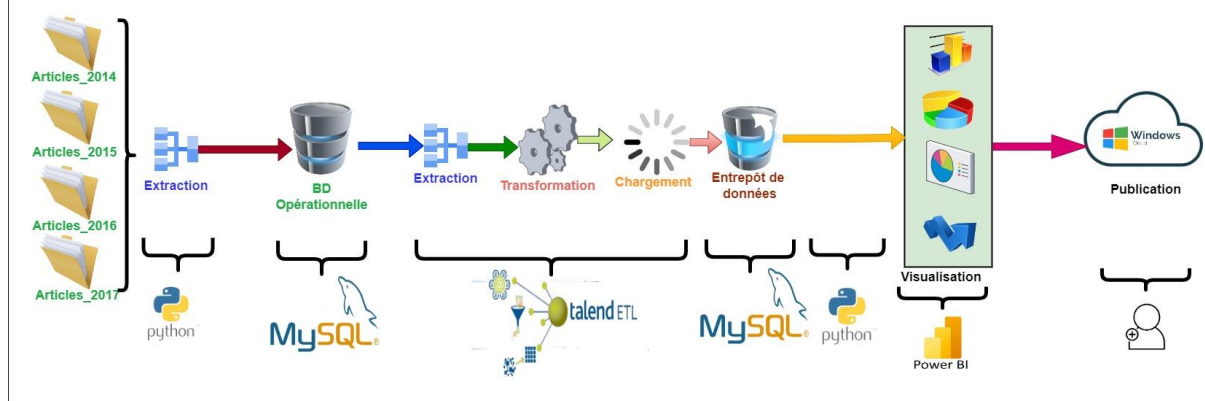
University of Yaounde I
Faculty of Sciences
Department Of Computer Science
Level : Master II

PROJET DE SYNTHÈSE



Mise en œuvre d'un Système d'Aide à la Décision

Processus général pour la réalisation du Projet



Noms et Prénoms des Etudiants :

❖ ABDELAZIZ MAHAMAT LOUKY
❖ MOUSSA ABAKAR ABAZENE
❖ SEMDI HONORE

Matricule

18T2916
23V2834
14L1992

Enseignant : Dr. Valery MONTHE

Année Académique : 2024 - 2025

Table des matières

Liste des figures	ii
Liste des tableaux	iii
Introduction	1
Partie 1 : Conception, implémentation et alimentation de la base de données opérationnelle (lot1)	2
1. Analyse et conception	2
a) Proposons un modèle conceptuel de données.....	2
b) Produire le modèle logique de données.....	2
2. Implémentation	2
3. Alimentation	3
a) Analysons la structure (contenu) des différents fichiers d'articles et compréhension	3
b) A l'aide du langage Python, mettre en place le processus ETL permettant d'extraire les données de ces fichiers, de les transformer (traiter et formater) et de les stocker dans la base de données.....	3
Partie 2 : Conception, implémentation et alimentation de la base de données opérationnelle (lot1)	6
1) Analyse	6
2) Conception	10
Hierarchies des Dimensions	10
- Faire le schéma du processus à mettre en place.	10
- Utiliser l'outil Talent OS-DI pour réaliser le processus ETL et alimenter l'entrepôt de données à partir de la BD opérationnelle	11
- Décrire les composants utilisés dans votre/vos job(s).....	12
3) Alimentation	13
Partie 3 : Conception et implémentation de tableaux de bord et reporting (lot3)	14
1. Définition des indicateurs pertinents	14
2. Concevoir plusieurs types de tableaux de bord	14
3. Utiliser le langage Python et l'outil Power BI pour mettre en œuvre les tableaux de bord	14
4. Créer des comptes sur Power BI service et publier les rapports.	17
Conclusion	18

Liste des figures

Figure 1: Modèle Conceptuel.....	2
Figure 2: Répertoires des données brutes à analyser.....	3
Figure 3: Processus de l'API-Python.....	4
Figure 4: Sous processus (Parcourir/Extraction/Insertion)	4
Figure 5: Résultat de l'extraction des données	4
Figure 6: Description des tables dans la BD	5
Figure 7: Les résultats des données après extraction dans les différentes tables	5
Figure 8: Schéma de l'Entrepôt de données	10
Figure 9: Processus ETL avec Talend et visualisation avec PowerBI	10
Figure 10: ETL avec Talend de la BD Opérationnelle à l'Entrepot (Unique).....	11
Figure 11: ETL avec Talend de la BD Opérationnelle à l'Entrepot (via les tables).....	12
Figure 12: Résultats après ETL des données dans les différentes tables	13
Figure 13: Tableau de bord 1	15
Figure 14: Tableau de Bord 2.....	15
Figure 15: Résultat du drill-down	16
Figure 16: Résultat du drill-down	16
Figure 17: Résultat du roll-up	16
Figure 18: Résultat de la publication des rapports sur PowerBI Service	17
Figure 19: Affichage des rapports sur PowerBI Service.....	17

Liste des tableaux

Tableau 1: Description de la table de fait	6
Tableau 2: Description des différentes dimensions	6
Tableau 3: Structure de la table de Fait	7
Tableau 4: Description des mesures	8
Tableau 5: Description de la Dimension Article	9
Tableau 6: : Description de la Dimension Auteur	9
Tableau 7: : Description de la Dimension Instituion	9
Tableau 8: : Description de la Dimension Pays	9
Tableau 9: Description de la dimension Ville	9
Tableau 10; Description de la Dimension Temps (annee)	9

Introduction

Dans le cadre de ce travail pratique, nous allons nous plonger dans le domaine passionnant des systèmes d'aide à la décision, en nous concentrant sur un cas concret : l'analyse et l'optimisation de la conférence scientifique RecSys. Notre objectif est de développer un système d'information décisionnel (SID) robuste et efficace pour aider les organisateurs de RecSys à mieux comprendre l'évolution de leur conférence et à prendre des décisions éclairées pour sa gestion future.

Ce TP nous permettra de mettre en pratique des compétences essentielles en analyse de données, conception de bases de données, et business intelligence. Nous allons :

1. Analyser les besoins spécifiques des organisateurs de RecSys et définir les fonctionnalités clés du SID.
2. Concevoir et implémenter une architecture de système décisionnel complète, incluant la collecte de données à partir de fichiers d'articles, leur stockage dans une base de données opérationnelle, et la création d'un entrepôt de données.
3. Identifier et développer des indicateurs pertinents pour l'évaluation de divers aspects de la conférence, tels que le taux de publication par institution, ville, pays et continent.
4. Créer des tableaux de bord intuitifs et informatifs pour présenter ces indicateurs de manière claire et actionnable.
5. Documenter minutieusement chaque étape de notre processus, assurant ainsi la transparence et la reproductibilité de notre travail.

Ce projet nous permettra non seulement d'approfondir nos connaissances techniques en gestion et analyse de données, mais aussi de comprendre les enjeux et les opportunités liés à l'utilisation de l'intelligence décisionnelle dans un contexte réel de gestion d'événements scientifiques. À travers ce processus, nous développons une compréhension pratique des étapes nécessaires à la création d'un système d'aide à la décision efficace, depuis l'analyse des besoins, la modélisation, conception de base de données Opérationnelles, l'Extraction des données via une API (en python), l'alimentation de cette base, le traitement des ces données (Transformation) et analyse de données (visualisation) afin de prendre une décision jusqu'à la présentation des résultats, en passant par la conception de l'architecture et l'implémentation des fonctionnalités.

Partie 1 : Conception, implémentation et alimentation de la base de données opérationnelle (lot1)

1. Analyse et conception

a) Proposons un modèle conceptuel de données

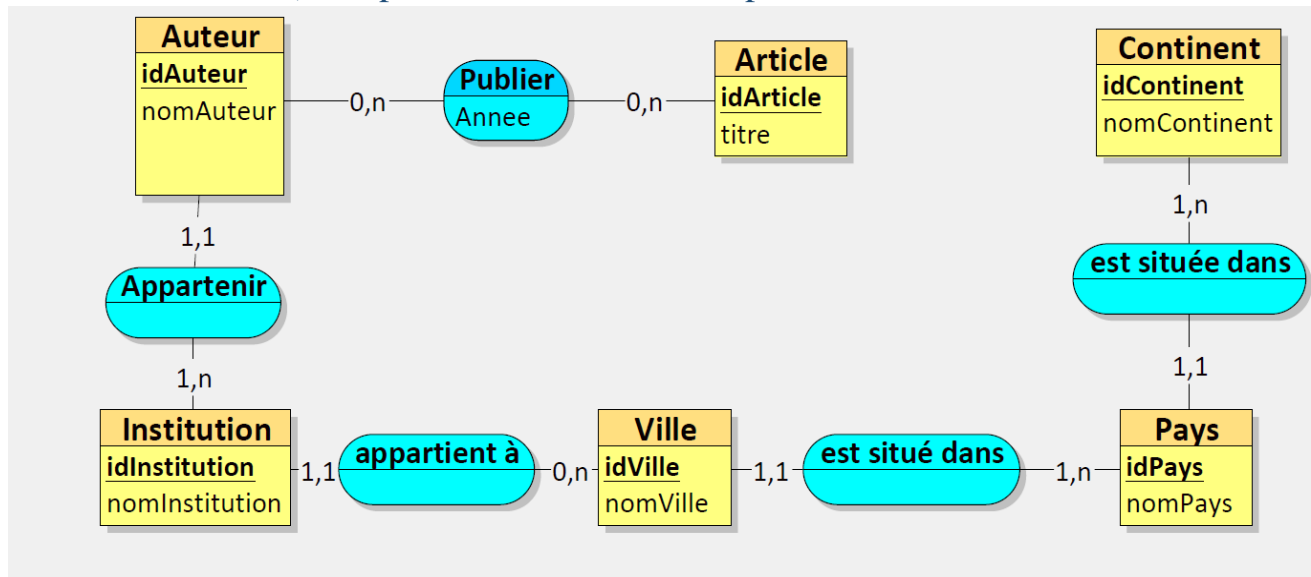


Figure 1: Modèle Conceptuel

b) Produire le modèle logique de données.

- Auteurs (idAuteur, nomAuteur)
- Article(idArticle, titre)
- Publication (idAuteur, idArticle, annee)
- Institution (idInstitution, nomInstitution, #idVille)
- Ville(idVille, nomVille, #idPays)
- Pays(idPays, nomPays, #idContinent)
- Continent(idContinent, nomContinent)

NB : la table continent est présente dans la description mais n'a pas des données dans les fichiers articles.

2. Implémentation

Choisissons un SGBD, l'installons et implémentons la base de données.

Le choix de SGBD s'est porté sur MySQL pour plusieurs raisons clés :

1. **Open-source et gratuit** : MySQL est un logiciel open-source, ce qui signifie qu'il est gratuit et bénéficie d'une communauté active qui contribue à son développement et à sa maintenance.
2. **Performant et multi-plateforme** : MySQL est reconnu pour ses performances élevées, particulièrement pour les requêtes
3. **Convivialité et documentation** : MySQL est bien documenté et facile à apprendre, ce qui facilite son utilisation pour les développeurs.

4. **Sécurité robuste** : MySQL offre une sécurité robuste avec des systèmes de gestion des privilèges d'accès et des comptes utilisateurs, ce qui est essentiel pour les environnements multi-utilisateurs où le contrôle des accès est crucial
5. **Adaptabilité et évolutivité** : Grâce à sa communauté active, MySQL bénéficie de mises à jour régulières qui améliorent ses performances et ses fonctionnalités, ce qui permet de l'adapter facilement à l'évolution des besoins du projet.

Une fois que MySQL est téléchargé et installé, il est accessible sur invite de commande via :

```
➤ mysql -u root -p
```

3. Alimentation

- a) Analysons la structure (contenu) des différents fichiers d'articles et compréhension

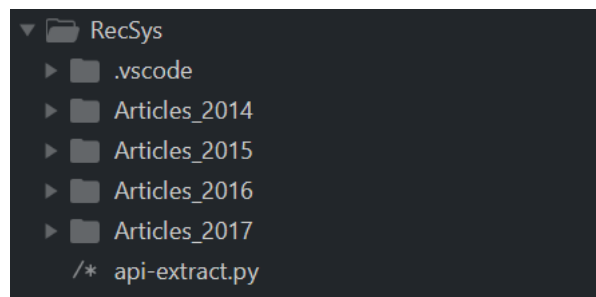


Figure 2: Répertoires des données brutes à analyser

Nous avons créé un répertoire RecSys dont nous avons les 4 dossiers qui sont : Articles_2014, Articles_2015, Articles_2016, Articles_2017. Dans le répertoire **Articles_2014**, il contient :

- La liste supérieure est le titre de l'article ;
- Les lignes suivantes sont les lignes qui contiennent les informations suivantes : Noms des auteurs, les noms des Institutions, les Villes, les Pays et chaque auteur (et ses affiliations) est séparé de l'autre par le symbole pipe « | » ;

Articles_2015 à Articles_2017, il contient :

- La liste supérieure est le titre de l'article ;
- Les lignes suivantes sont les lignes qui contiennent les informations suivantes : Noms des auteurs, les noms des Institutions, les Villes, les Pays et chaque auteur (et ses affiliations) est séparé de l'autre par le symbole pipe « | » ;
- Les résumés de l'article, liste des références bibliographiques, etc.

NB : ses dernières lignes ne sont pas utiles pour notre analyse.

- b) A l'aide du langage Python, mettre en place le processus ETL permettant d'extraire les données de ces fichiers, de les transformer (traiter et formater) et de les stocker dans la base de données.

Le code python sera en zip avec ce présent rapport mais au mois voici une description qui donne l'idée sur comment ce code fonctionne.

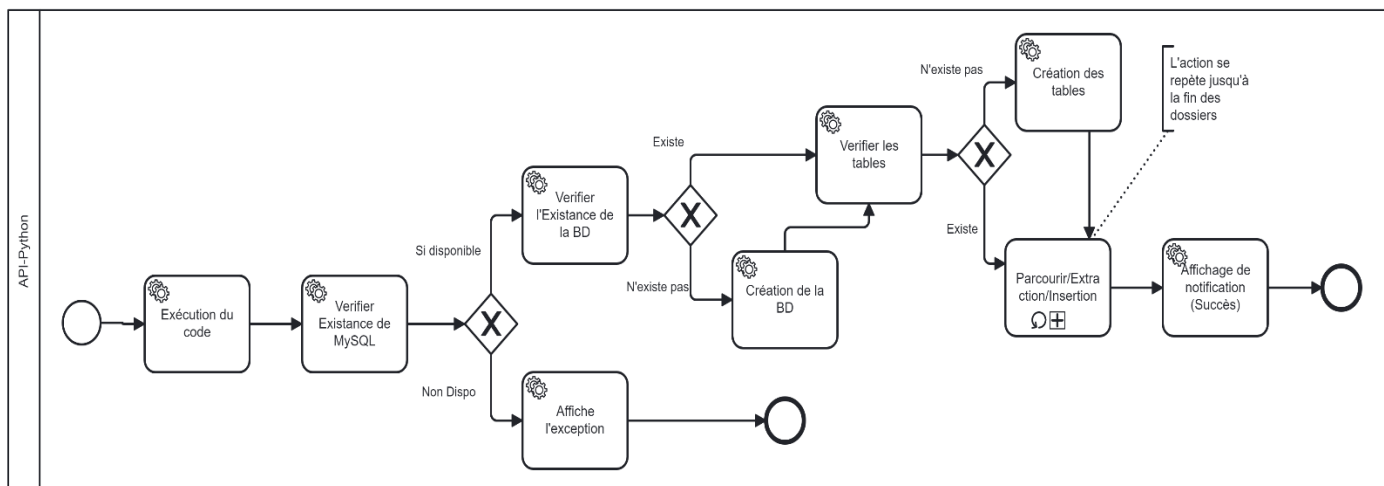


Figure 3: Processus de l'API-Python

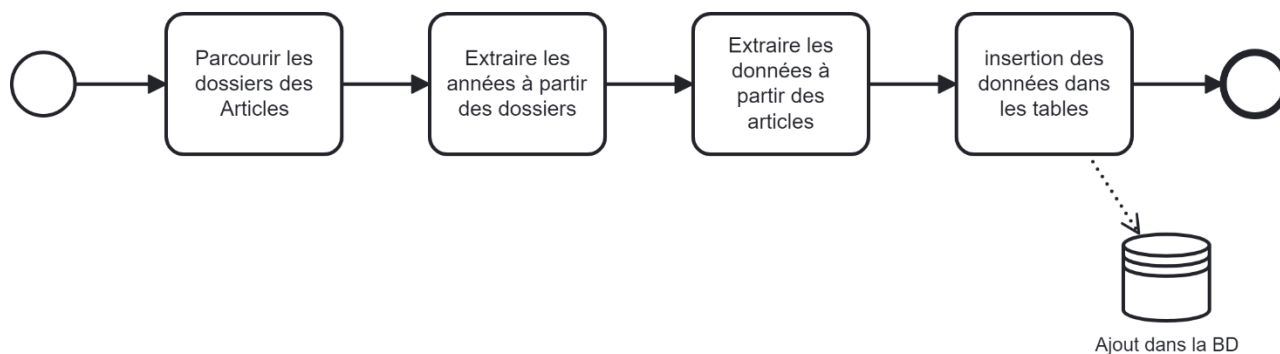


Figure 4: Sous processus (Parcourir/Extraction/Insertion)

NB : L'API codé en python permet de faire tout le travail possible, il suffit juste d'installer MySQL sur la machine, installer PyMySQL (à travers pip install pymysql) et indiqué le chemin des dossiers des articles puis lancer le code (à travers python api-extract.py).

Résultats obtenus

Une fois le code lancé via python api-extract.py

Voici les résultats sous la console :

```
PS C:\Users\User\Desktop\RecSys> python test5.py
⚠ La base de données 'Articles_DB' n'existe pas. Création en cours...
✅ Base de données 'Articles_DB' créée avec succès !
✅ Connexion réussie à la base de données 'Articles_DB' !
✅ Connexion réussie à la base de données !
✅ Importation terminée avec succès.
```

Figure 5: Résultat de l'extraction des données

Dans le SGBD, on constate la création des tables suivantes :


```

MariaDB [Articles_DB]> show tables;
+-----+
| Tables_in_articles_db |
+-----+
| article                |
| auteur                 |
| continent              |
| ecrire                 |
| institution            |
| pays                   |
| ville                  |
+-----+
7 rows in set (0.001 sec)

```

Figure 6: Description des tables dans la BD

Nous allons faire un count sur les données extraites sur les différentes tables

```

MariaDB [Articles_db]> select count(*) from auteur;
+-----+
| count(*) |
+-----+
|      308 |
+-----+
1 row in set (0.001 sec)

MariaDB [Articles_db]> select count(*) from article;
+-----+
| count(*) |
+-----+
|      111 |
+-----+
1 row in set (0.001 sec)

MariaDB [Articles_db]> select count(*) from Institution;
+-----+
| count(*) |
+-----+
|      326 |
+-----+
1 row in set (0.001 sec)

MariaDB [Articles_db]> select count(*) from Publication;
+-----+
| count(*) |
+-----+
|      364 |
+-----+
1 row in set (0.001 sec)

```

Figure 7: Les résultats des données après extraction dans les différentes tables

Partie 2 : Conception, implémentation et alimentation de la base de données opérationnelle (lot1)

1) Analyse

- Identifier et décrire les différents éléments qui vont constituer votre entrepôt de données : faits, mesures, dimensions et leurs niveaux de granularités appropriés:

Faits (Table centrale)

Le fait principal dans notre cas est la publication d'un article dans une conférence.

Tableau 1: Description de la table de fait

Table des faits	Description
Publication	Enregistrement de chaque article publié dans une conférence, avec son auteur et son institution

Mesures (Indicateurs d'analyse)

Les mesures sont les valeurs numériques que nous allons analyser. Voici quelques mesures clés :

- Nombre total d'articles publiés
- Nombre d'auteurs uniques par année
- Nombre d'institutions ayant publié des articles
- Répartition des publications par pays et continent

Dimensions (Axes d'analyse)

Les dimensions permettent de décrire et de structurer les faits selon différents axes d'analyse. La granularité définit le niveau de détail des données stockées.

Tableau 2: Description des différentes dimensions

Dimension	Description	Granularite
Article	Identifiant, titre	Article de niveau individuel
Auteur	Identifiant, nom	Niveau auteur individuel
Institution	Identifiant, nom, ville	Niveau institution individuelle
Ville	Identifiant, nom,	Niveau de rémunération individuel
Temps	Annee,	Niveau

Pourquoi cette granularité ?

Elle permet :

- d'analyser chaque auteur et chaque institution individuellement .
- d'analyser les tendances des publications par jour, mois et année .

- d'agréger les données si nécessaire (ex : total des publications par pays ou par année).

Définir en détail les faits, en produisant le dictionnaire des mesures:

Définition des Faits

Un fait représente un événement mesurable dans le système.

Dans notre projet, l'événement principal est la publication d'un article scientifique par un auteur affilié à une institution dans une conférence.

Table des Faits :Publication

- La table des faitsPublication contient les informations concernant un article, un auteur, son institution, la conférence et la date de publication .
- Chaque enregistrement dans ce tableau représente une publication d'un article par un auteur donné .

Structure de la Table Publication

Tableau 3: Structure de la table de Fait

Nom de l'attribut	Description
idPublication	Identifiant unique de la publication dans la table des faits. Chaque ligne représente une publication unique.
idArticle	Référence vers l'article publié. Plusieurs auteurs peuvent être associés à un même article.
idAuteur	Référence vers l'auteur qui a contribué à l'article. Si un article a plusieurs auteurs, il sera dupliqué avec chaque auteur dans la table
idInstitution	Référence vers l'institution de l'auteur au moment de la publication.
idPays	Référence vers le pays de l'institution affiliée. Cela permet d'effectuer des analyses géographiques.
idTemps	Référence vers la dimension temps (date de soumission de l'article). Permet d'analyser les tendances de publication par période (jour, mois, année).
nbpublication	Nombre total de fois où l'article a été publié dans différentes conférences (généralement = 1 sauf cas spécifiques).

Granularité de la table Publication

- Chaque ligne représente une contribution individuelle d'un auteur à un article.

- Si un article est écrit par 3 auteurs, il apparaîtra 3 fois dans la table des faits .
- Cela permet d'effectuer des analyses précises sur la contribution des auteurs et des institutions.

1. Dictionnaire des Mesures:

Le **dictionnaire des mesures** est une liste d'indicateurs qui seront utilisés pour effectuer des analyses dans l'entrepôt de données. Les mesures sont généralement des **comptages** ou des **agrégations** des données stockées dans la table des faits.

Tableau 4: Description des mesures

Nom de la mesure	Description
Nombre total d'articles publiés	Nombre distinct d'articles ayant été publiés dans une conférence. Permet d'analyser la production scientifique globale.
Nombre total d'auteurs	Nombre d'auteurs distincts ayant participé à au moins une publication. Permet de voir combien de chercheurs ont contribué.
Nombre total de publications	Nombre total d'articles en prenant en compte chaque auteur individuellement (si un article à 3 auteurs, il sera compté 3 fois).
Nombre de publications par institution	Nombre d'articles produits par chaque institution. Permet d'analyser les institutions les plus productives.
Nombre de publications par pays	Nombre d'articles produits par pays. Permet de voir ce qui paie publier le plus.
Nombre de publications par conférence	Nombre d'articles présentés dans chaque conférence. Permet d'identifier les conférences les plus actives.
Nombre de publications par année	Nombre total d'articles publiés par année. Permet de voir l'évolution du nombre de publications au fil du temps.
Nombre moyen de publications par auteur	Nombre moyen d'articles publiés par auteur. Indiquer la productivité moyenne des chercheurs.

-Définir en détail les dimensions : dictionnaire des attributs.

Définition Détaillée des Dimensions : Dictionnaire des Attributs

Les dimensions sont les axes d'analyse qui permettent d'examiner les faits sous différents angles.

Elles sont utilisées pour décrire et structurer les événements stockés dans la table des faitsPublication .

1. **Dimension Article:** La dimension Article contient les informations sur les articles scientifiques publiés .

Tableau 5: Description de la Dimension Article

Nom de l'attribut	Description
idArticle	Identifiant unique de l'article
titre	Titre de l'article publié

2. Dimension Auteur: La dimension Auteur contient **les informations sur les auteurs des articles**

Tableau 6: : Description de la Dimension Auteur

Nom de l'attribut	Description
idAuteur	Identifiant unique de l'auteur
nom	Nom complet de l'auteur

3. Dimension Institution: La dimension Institution contient **les informations sur les universités, laboratoires et entreprises** où travaillent les auteurs.

Tableau 7: : Description de la Dimension Instituion

Nom de l'attribut	Description
idInstitution	Identifiant unique de l'institution
nom	Nom de l'institution

4. Dimension Pays: La dimension Pays contient **les informations géographiques des institutions** .

Tableau 8: : Description de la Dimension Pays

Nom de l'attribut	Description
idPays	Identifiant unique du pays
nom	Nom du pays

Tableau 9: Description de la dimension Ville

Nom de l'attribut	Description
idVille	Identifiant unique de la ville
nomVille	Nom de la ville

6. Dimension Temps: La dimension Tempscontient **les informations sur la date de soumission des articles** .

Tableau 10; Description de la Dimension Temps (annee)

Nom de l'attribut	Description
idTemps	Identifiant unique de la date
annee	Année de publication

2) Conception

Choix du modèle approprié pour l'entrepôt: **Le Modèle en Flocon de neige**

Hiérarchies des Dimensions: Les hiérarchies nous permettent de structurer les dimensions avec des niveaux différents de détail.

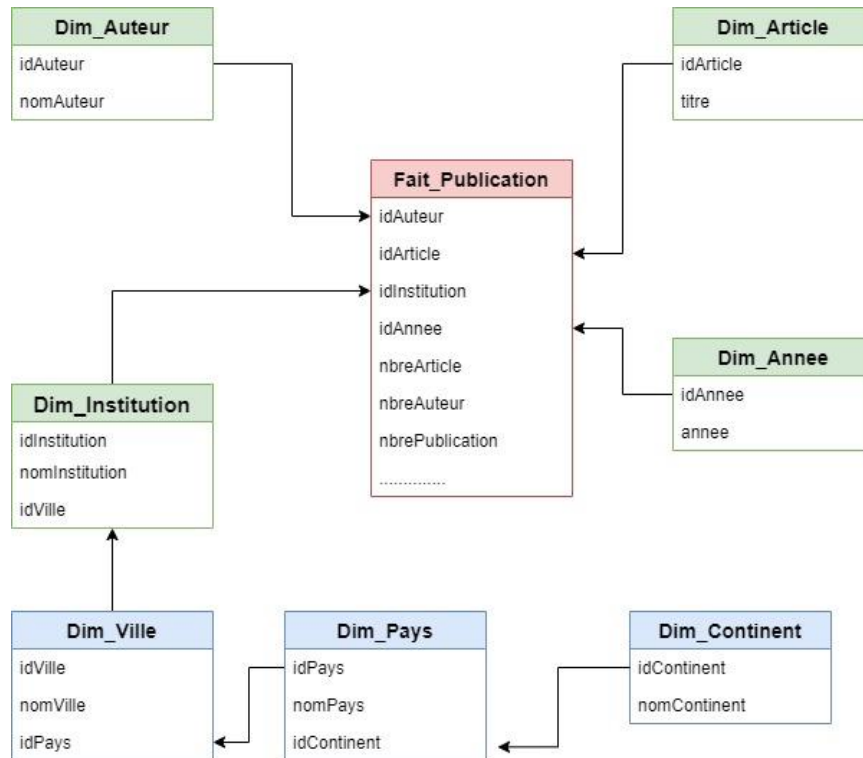


Figure 8: Schéma de l'Entrepôt de données

- Faire le schéma du processus à mettre en place.

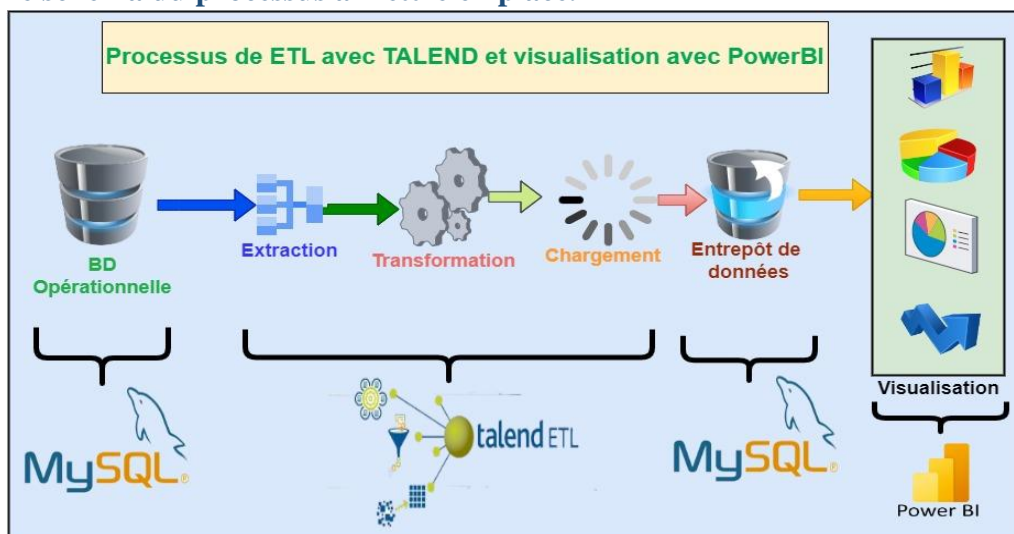


Figure 9: Processus ETL avec Talend et visualisation avec PowerBI

- Utiliser l'outil Talend OS-DI pour réaliser le processus ETL et alimenter l'entrepôt de données à partir de la BD opérationnelle

Pour extraire les données de la base de données Opérationnelle vers l'entrepôt des données avec Talend, on procède comme suit :

- Installer et configurer Talend ;
- Créer un ou plusieurs jobs ;
- Configurer la connexion à la base de données Opérationnelle (ici on a utilisé les contextes pour faciliter/simplifier la configuration et surtout sauvegarder pour une utilisation ultérieure) ;
- Ajouter les composants (tDBInputs, tDBOutput, tMap) et appliqué les filtrages possibles pour les doublons, ...
- Puis relier les composants de l'entrée (tDBInput) au composant tMap puis tMap à la sortie. Ici on peut avoir une seule sortie (si l'entrepôt de données n'est pas créer) tout comme on peut avoir plusieurs sortie (si l'entrepôt est créer) puis configurer les variables et les Id ;
- Exécuter le programme (code ou design talend)

Voici les figures Talend

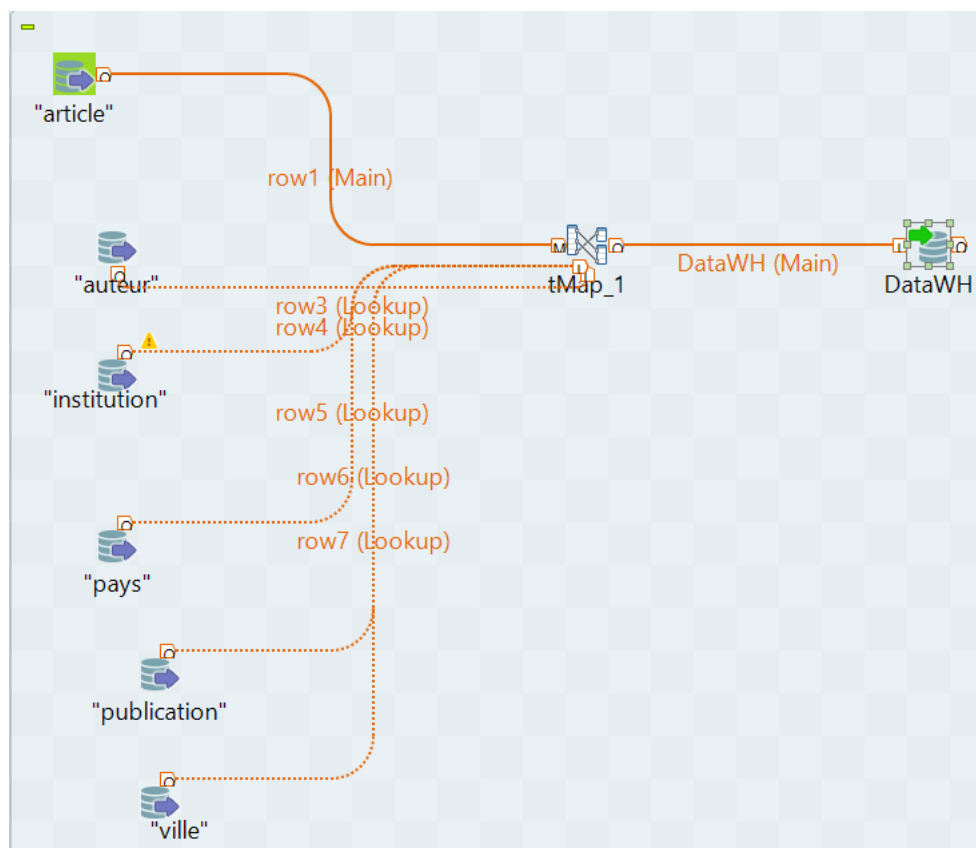


Figure 10: ETL avec Talend de la BD Opérationnelle à l'Entrepot (Unique)

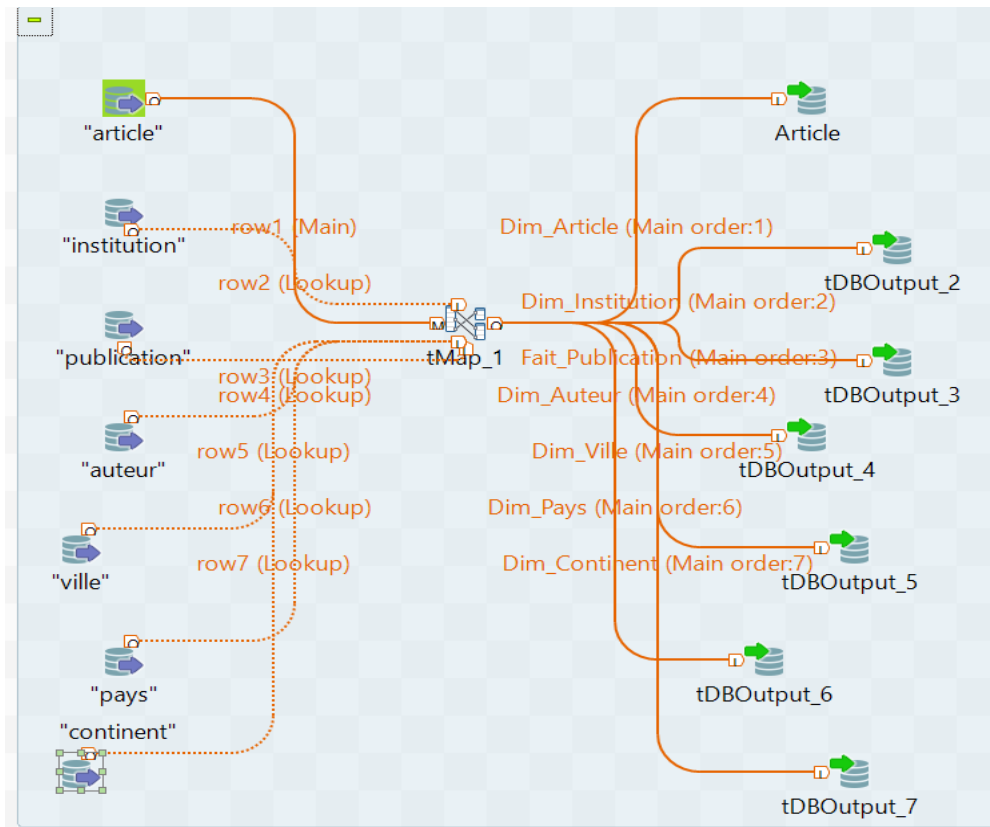


Figure 11: ETL avec Talend de la BD Opérationnelle à l'Entrepot (via les tables)

- Décrire les composants utilisés dans votre/vos job(s).

Les composants :

- **tDBInput** : est utilisé pour extraire des données d'une base de données. Il permet de se connecter à diverses bases de données (comme MySQL, Oracle, PostgreSQL, etc.) et d'exécuter des requêtes SQL pour récupérer des données. Ces données peuvent ensuite être traitées ou transformées dans le flux de données Talend.
- **tDBOutput** : est utilisé pour insérer, mettre à jour ou supprimer des données dans une base de données. Il permet de se connecter à une base de données cible et d'exécuter des requêtes SQL pour y écrire les données transformées ou traitées dans le flux de données Talend.
- **tMap** : est l'un des plus puissants et polyvalents de Talend. Il permet de réaliser des opérations complexes sur les données, telles que des jointures, des transformations, des filtrages, et des rejets. Il peut gérer plusieurs entrées et sorties, ce qui le rend très flexible pour manipuler les données de manière complexe.

Les Jobs :

Dans notre TP, nous avons utilisé qu'un seul Job, Extraction puisque nos données viennent d'une seule source.

Mais cependant, nous avons utilisé Deux variables de contextes pour la connexion à nos deux Bases de données car ceci nous facilité la tâche au lieu de configurer pour chaque table en Output ses variables comme le login, le server (localhost), le mot de passe, le nom de la Base de données, le port (3306) et le paramètre additionnel (AdditionnalParams).

3) Alimentation

Une fois le programme Talend exécuté, les données de la base de données seront extraites (via les filtres mis en place) et envoyées vers l'entrepôt. Voici les résultats :

```
+-----+
| count(*) |
+-----+
|      111 |
+-----+
1 row in set (0.001 sec)

MariaDB [articles_datawh]> select count(*) from Dim_Institution;
+-----+
| count(*) |
+-----+
|      326 |
+-----+
1 row in set (0.001 sec)

MariaDB [articles_datawh]> select count(*) from Dim_Auteur;
+-----+
| count(*) |
+-----+
|      308 |
+-----+
1 row in set (0.002 sec)

MariaDB [articles_datawh]> select count(*) from Dim_Publication;
+-----+
| count(*) |
+-----+
|      364 |
+-----+
1 row in set (0.002 sec)
```

Figure 12: Résultats après ETL des données dans les différentes tables

Partie 3 : Conception et implémentation de tableaux de bord et reporting (lot3)

1. Définition des indicateurs pertinents

Pour aider les organisateurs de RecSys à prendre des décisions, voici des indicateurs clés :

Indicateurs de publication

- Nombre total d'articles publiés (par année, par pays)
- Taux de croissance des publications (%)
- Nombre moyen d'auteurs par article

Indicateurs par institution et localisation

- Nombre d'articles par institution
- Classement des institutions les plus contributrices
- Nombre d'articles par ville, pays et continent
- Répartition géographique des publications (%)

Indicateurs de collaboration

- Nombre d'articles coécrits par plusieurs institutions
- Nombre d'auteurs ayant publié plusieurs articles
- Réseau de collaboration entre auteurs et institutions

Indicateurs d'évolution et tendances

- Évolution du nombre de publications par année
- Progression des contributions par ville et pays
- Évolution des collaborations internationales

2. Concevoir plusieurs types de tableaux de bord

Les tableaux de bord sont conçus dans le logiciel PowerBI. Ils sont ensemble avec la question 3.

3. Utiliser le langage Python et l'outil Power BI pour mettre en œuvre les tableaux de bord

On a utilisé les données transformées avec le logiciel Talend en posant des filtres pour la suppression des doublons et d'autres données brutes dont voici les résultats de ce deux données.

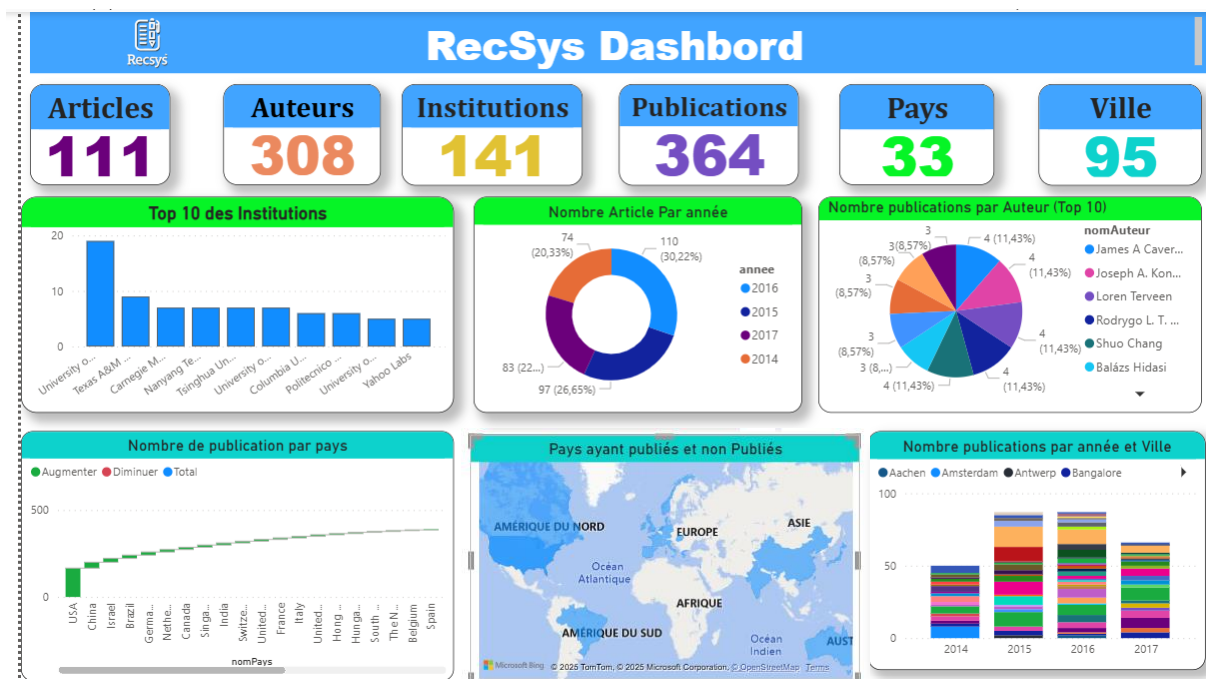


Figure 13: Tableau de bord 1

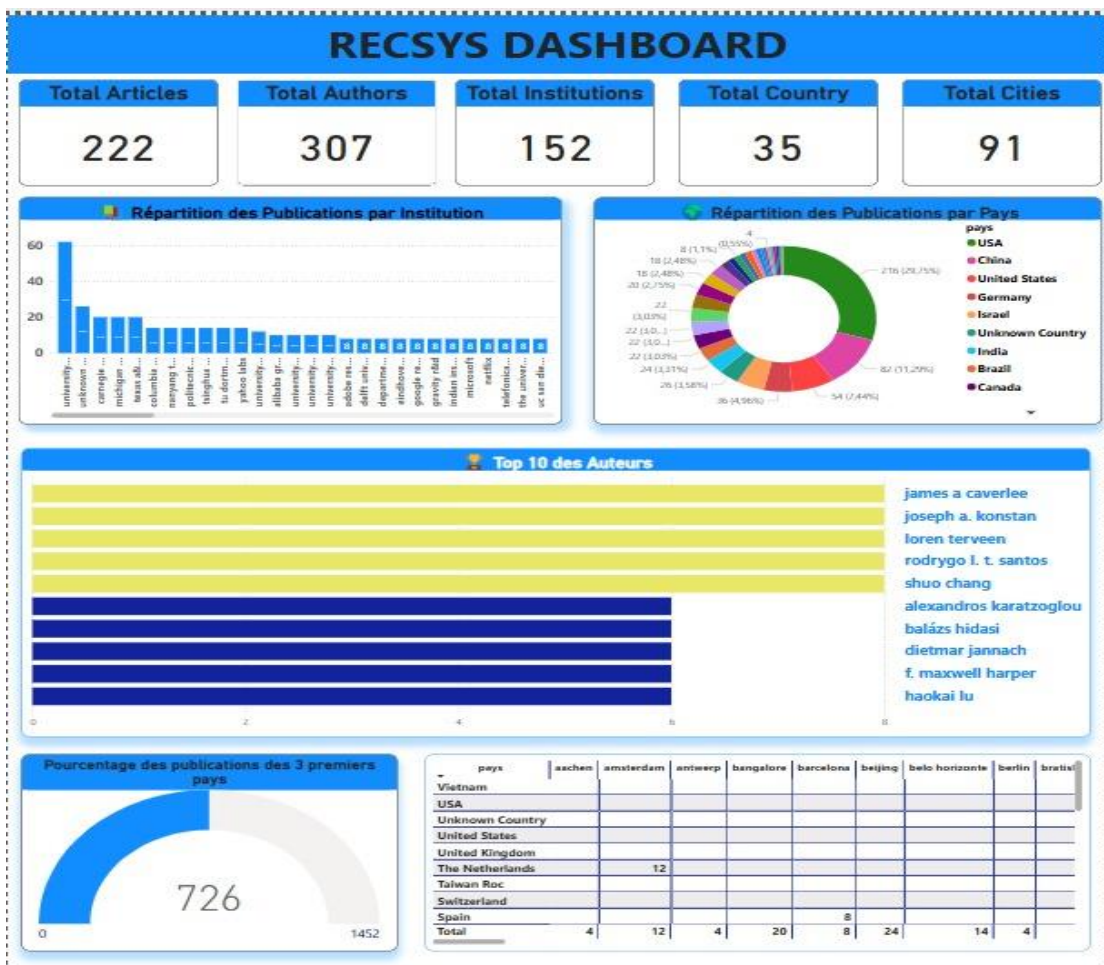


Figure 14: Tableau de Bord 2

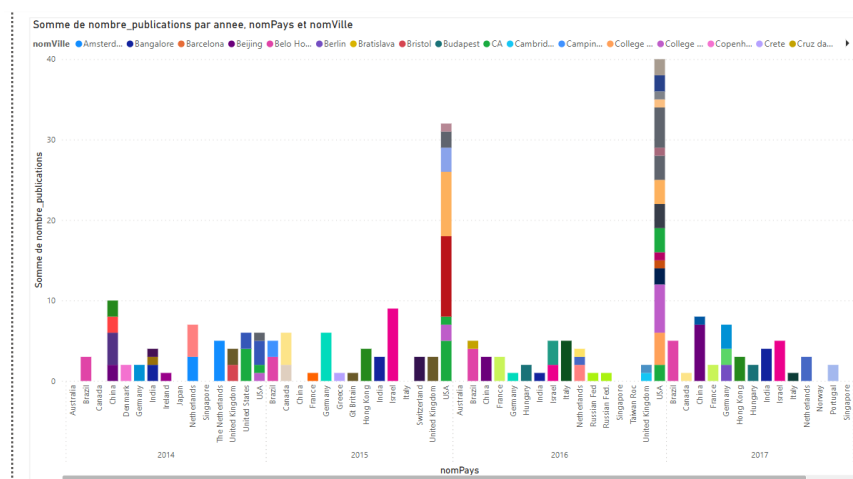


Figure 15: Résultat du drill-down

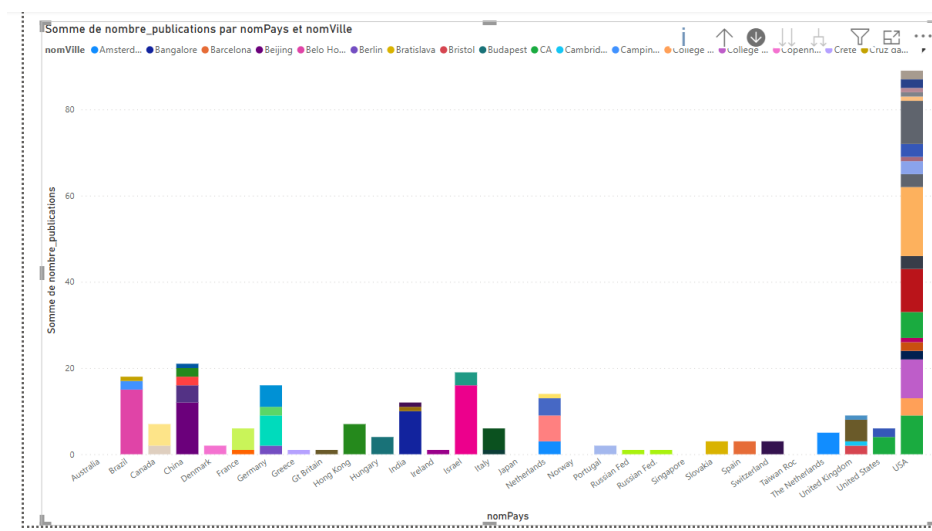


Figure 16: Résultat du drill-down

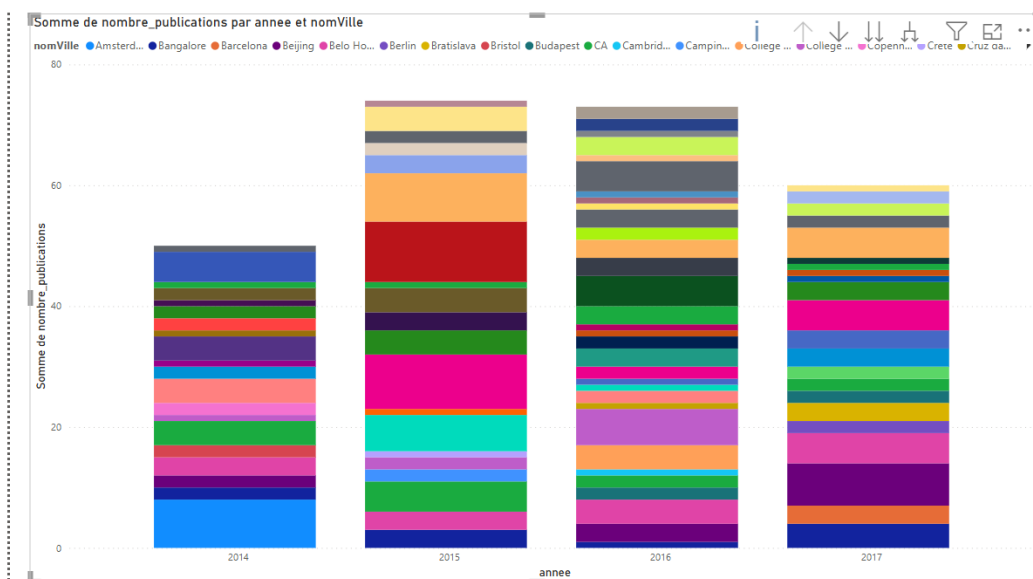


Figure 17: Résultat du roll-up

4. Créer des comptes sur Power BI service et publier les rapports.

Pour publier un rapport sur PowerBI, il faut au préalable avoir un compte professionnel ou un compte étudiant. Puis créer un compte sur PowerBI service et publier les rapports. Donc voici nos résultats.



Figure 18: Résultat de la publication des rapports sur PowerBI Service

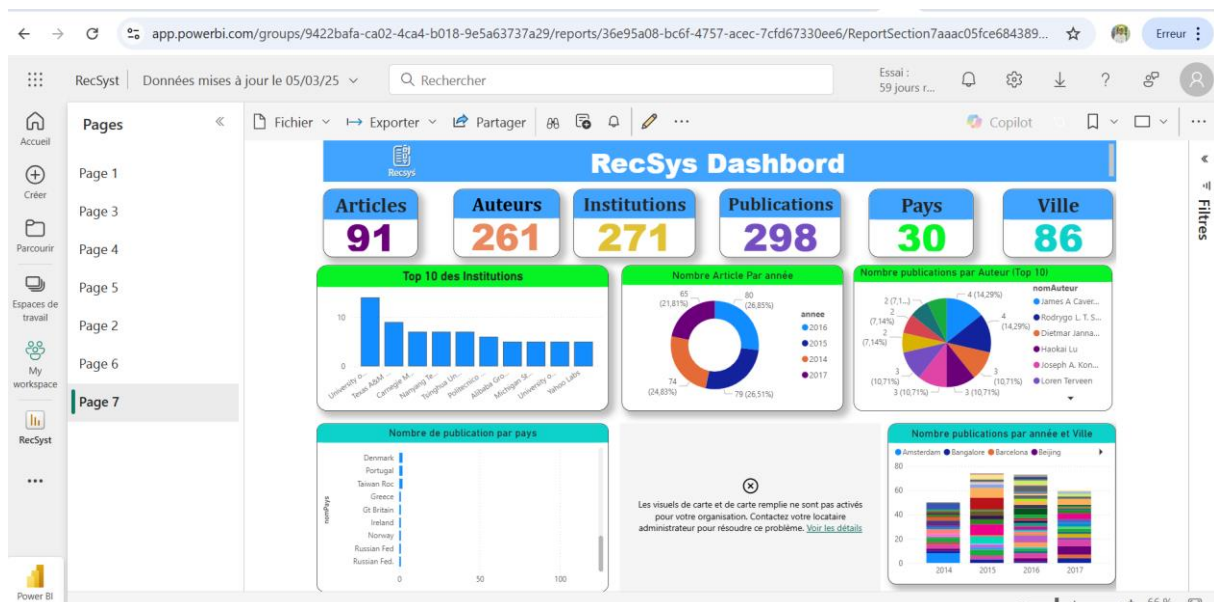


Figure 19: Affichage des rapports sur PowerBI Service

Remarque : Dans l’affichage du rapport, la carte ne s’affiche pas car la visualisation est désactivée. Il faut être un administrateur pour l’organisme qui a

fourni le mail professionnel pour activer cette fonction. Comme on avait fourni le mail étudiant, on n'a pas la permission de modifier les paramètres.

Ci-joint le lien des rapports :

<https://app.powerbi.com/groups/9422bafa-ca02-4ca4-b018-9e5a63737a29/reports/36e95a08-bc6f-4757-acec-7cfd67330ee6/ReportSection7aaac05fce684389981f?experience=power-bi>

Conclusion

Ce projet a permis de mettre en œuvre un processus complet de gestion et d'analyse des données, en passant par plusieurs étapes essentielles. Tout d'abord, une **Base de Données Opérationnelle** a été conçue et implémentée sous **MySQL** pour stocker les données brutes. Ensuite, un script **Python** a été développé pour **extraire et charger** ces données depuis différents fichiers répartis dans quatre dossiers distincts vers la base de données.

Après cette phase d'intégration, un **entrepôt de données** a été conçu et implémenté sous **MySQL**, structuré de manière à optimiser l'analyse des données. Pour le processus **ETL (Extract, Transform, Load)**, l'outil **Talend OS-DI** a été utilisé afin d'extraire les données de la base opérationnelle, de les transformer via le composant **tMap**, puis de les charger dans l'entrepôt à l'aide des composants **tDBInput** et **tDBOutput**.

Une fois les données correctement chargées dans l'entrepôt, **Power BI** a été utilisé pour se connecter à cette source et concevoir des **tableaux de bord interactifs**. Ces tableaux de bord ont permis de visualiser et d'analyser les données de manière efficace, offrant ainsi une meilleure prise de décision. Enfin, les rapports ont été **publiés sur Power BI Service**, facilitant ainsi leur partage et leur accessibilité en ligne.

Ce projet a donc couvert l'ensemble du **cycle de vie de la donnée**, depuis son extraction jusqu'à sa visualisation, en mettant en pratique des concepts clés de **l'intégration, de la transformation et de l'analyse des données**.