

REPUBLICUE DU
CAMEROUN

UNIVERSITE DE
YAOUNDÉ I

FACULTÉ DES SCIENCES

DÉPARTEMENT
D'INFORMATIQUE

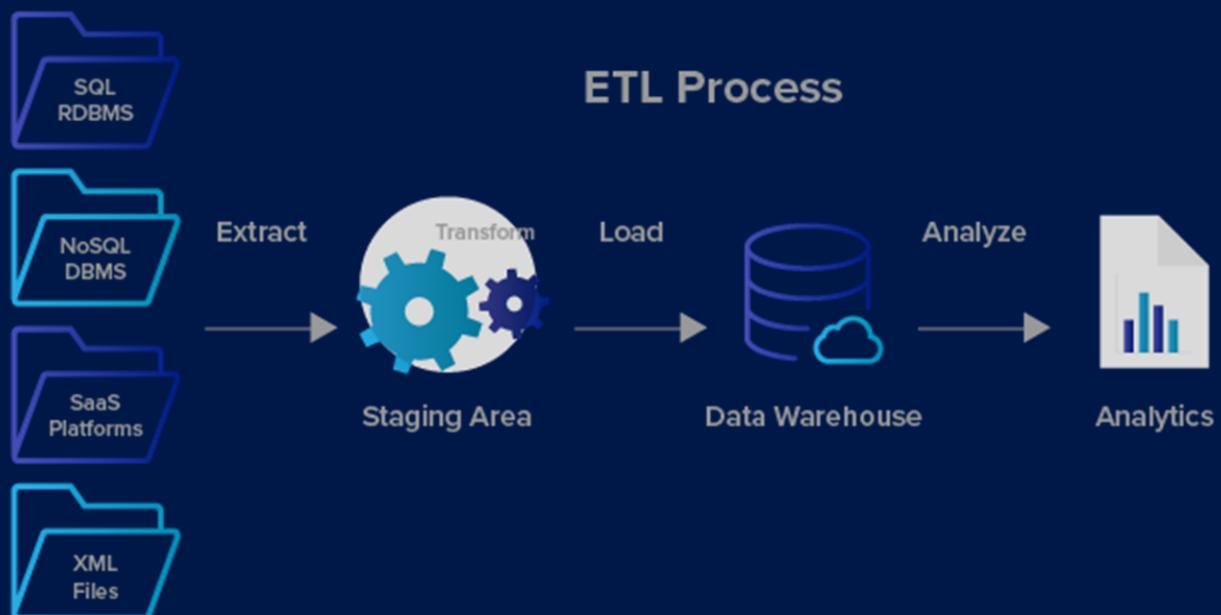


REPUBLIC OF CAMEROON

UNIVERSITY OF YAOUNDÉ I

SCIENCES FACULTY

COMPUTER SCIENCE
DEPARTMENT



TP INF351 GROUPE 12

Présenté par : AMBOMO TIGA GEDEON

MEMBRES DU GROUPE

NOMS ET PRENOM	MATRICULES	PARTICIPATIONS
AMBOMO TIGA GEDEON	21T2496	100%

SOMMAIRE

DEFINITION DES BESOINS	4
✓ Analyser le nombre de publications suivant plusieurs axes notamment :	4
a. Analyse suivant les dates :	4
b. Analyse suivant les localisations géographiques :	4
c. Analyse suivant les auteurs :	4
d. Analyse suivant le temps et la localisation géographique :	4
e. Analyse suivant les auteurs et le temps :	4
CONCEPTION DE L'ENTREPÔT	5
i. Sujet d'analyse :	5
ii. Mesure, description et type :	5
iii. Dimensions :	5
iv. Choix du modèle :	5
v. Définition de l'approche :	6
SCHÉMA DE L'ENTREPÔT	7
1. Schéma en constellation :	7
2. Schéma multidimensionnel :	7

DEFINITION DES BESOINS

✓ Analyser le nombre de publications suivant plusieurs axes notamment :

- Le temps (date) ;
- La localisation géographique ;
- Les auteurs ;
- Le temps et la localisation géographique ;
- Les auteurs et le temps ;

a. Analyse suivant les dates :

Ici nous réalisons cette analyse afin de déterminer les éventuelles années les plus productives et déterminer si possible les fondements.

b. Analyse suivant les localisations géographiques :

Ici, cette analyse est réalisée dans le cas où nous souhaitons connaître les institutions, villes, pays et continents les plus dynamiques pour un éventuel classement lors d'une conférence.

c. Analyse suivant les auteurs :

Ici, cela est réalisée dans l'éventualité l'on souhaite primer le meilleur auteur (dans ce cas celui qui a le plus grand nombre de publications).

d. Analyse suivant le temps et la localisation géographique :

Cette analyse est réalisée afin de connaître l'évolution des publications par établissement sur le temps.

e. Analyse suivant les auteurs et le temps :

Elle est réalisée afin de connaître l'évolution des publications par établissement sur le temps.

CONCEPTION DE L'ENTREPÔT

i. Sujet d'analyse :

L'analyse portera essentiellement sur les publications.

ii. Mesure, description et type :

Mesure	Type	Description
Nombre de publication	Entier	Renvoie au nombre total de publication en fonction des auteurs, institutions, villes, pays, continents et années.

iii. Dimensions :

Dimensions	Description
Auteur	Obtenir les publications suivant les auteurs.
Temps (Date)	Obtenir les publications suivant leurs années de publication.
Géographique	Obtenir les publications suivant leurs institutions, villes, pays et continents.

iv. Choix du modèle :

Le modèle choisi est celui en **constellation** car on a plusieurs tables de faits et des liens entre les dimensions.

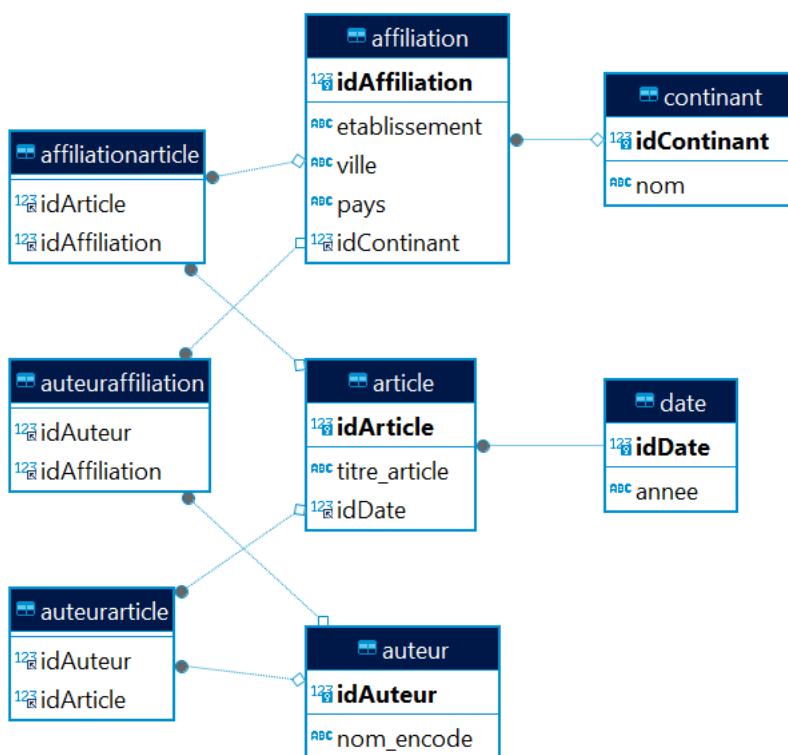
NB : Afin d'améliorer les performances de l'application et de faciliter l'extraction de données des fichiers Excel fournis, les hiérarchies suivant la dimension géographique telles que : institution - ville – pays, seront représentées comme attributs dans la table de dimension **Affiliation**.

CONCEPTION DE L'ENTREPÔT

v. Définition de l'approche :

L'approche utilisée ici est **BOTTOM-UP**, car la conception de notre entrepôt s'est faite suivant les étapes traduites dans cette approche qui sont entre autres :

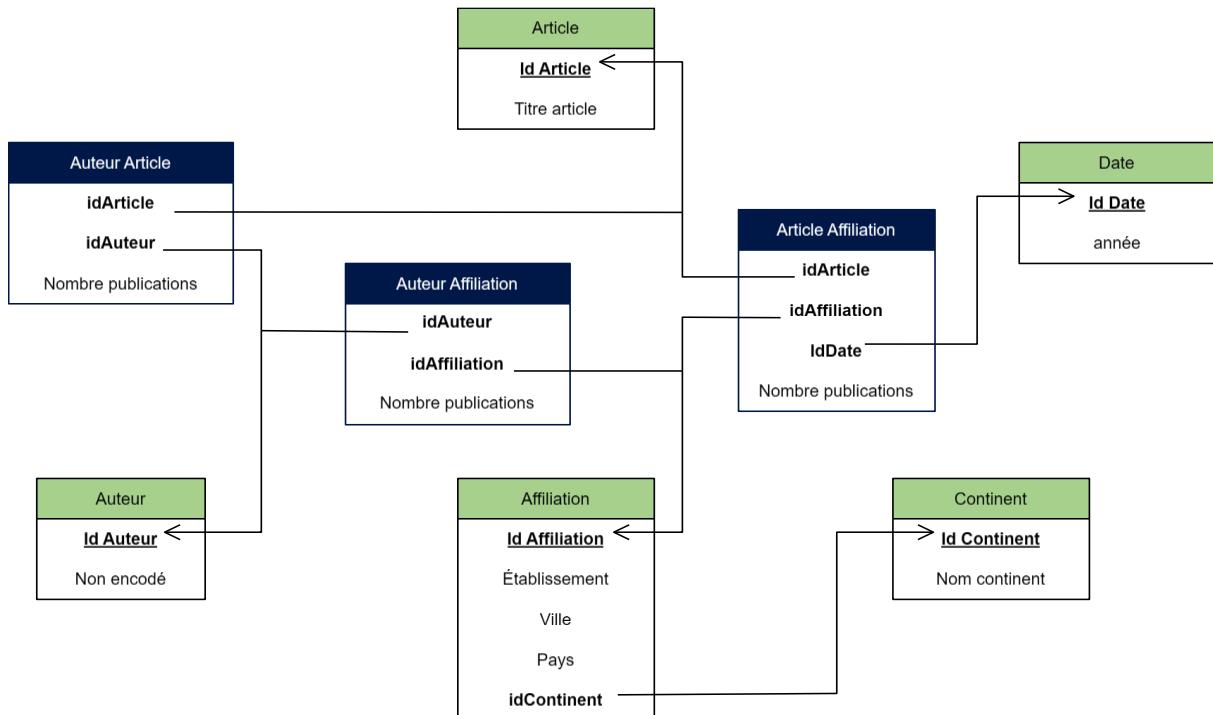
- **Inventaire des sources de données existante** (dans notre cas les fichiers CSV fournis) ;
- **Analyse des sources de données** (définition des tables, attributs, clés ...) :
 - **Schéma de la base de données :**



- **Identification des besoins métier** (comme défini ci-dessus ↗) ;
- **Conception du modèle de données** ;
- **Développement du ou des processus ETL** (à travers un script python) ;
 - **Requêtes SQL utilisées dans le script python** :
- **Conception du Dashboard** (conception des différents visuels dans power BI).

SCHÉMA DE L'ENTREPÔT

1. Schéma en constellation :



2. Schéma multidimensionnel :

