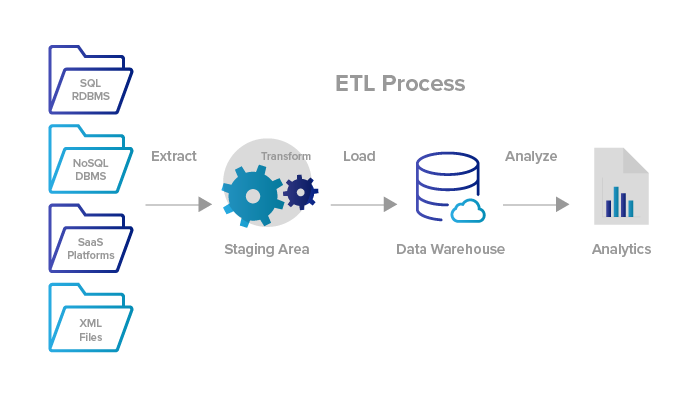
|  |  |  |
| --- | --- | --- |
| REPUBLIQUE DU CAMEROUN  \*\*\*\*\* | Logo université Yaoundé 1 | REPUBLIC OF CAMEROON  \*\*\*\*\* |
| UNIVERSITE DE YAOUNDÉ I  \*\*\*\*\* | UNIVERSITY OF YAOUNDÉ I  \*\*\*\*\* |
| FACULTÉ DES SCIENCES  \*\*\*\*\* | SCIENCES FACULTY  \*\*\*\*\* |
| DÉPARTEMENT D’INFORMATIQUE | COMPUTER SCIENCE DEPARTMENT |



TP INF351 GROUPE 12

Présenté par : AMBOMO TIGA GEDEON

Votre texte ici

# MEMBRES DU GROUPE

# MEMBRES DU GROUPE

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

# SOMMAIRE

# MEMBRES DU GROUPE

[DEFINITION DES BESOINS 4](#_DEFINITION_DES_BESOINS)

[ Analyser le nombre de publications suivant plusieurs axes notamment : 4](#_Toc158573724)

[a. Analyse suivant les dates : 4](#_Toc158573725)

[b. Analyse suivant les localisations géographiques : 4](#_Toc158573726)

[c. Analyse suivant les auteurs : 4](#_Toc158573727)

[d. Analyse suivant le temps et la localisation géographique : 4](#_Toc158573728)

[e. Analyse suivant les auteurs et le temps : 4](#_Toc158573729)

[CONCEPTION DE L’ENTREPÔT 5](#_CONCEPTION_DE_L’ENTREPÔT)

[i. Sujet d’analyse : 5](#_Toc158573731)

[ii. Mesure, description et type : 5](#_Toc158573732)

[iii. Dimensions : 5](#_Toc158573733)

[iv. Choix du modèle : 5](#_Toc158573734)

[v. Définition de l’approche : 6](#_Toc158573737)

[SCHÉMA DE L’ENTREPÔT 7](#_SCHÉMA_DE_L’ENTREPÔT)

[1. Schéma en constellation : 7](#_Toc158573739)

[2. Schéma multidimensionnel : 7](#_Toc158573740)

# 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 ;

## Analyse suivant les dates :

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

## 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.

## 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).

## Analyse suivant le temps et la localisation géographique :

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

## 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

# Sujet d’analyse :

L’analyse porteras essentiellement sur les publications.

# 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. |

# 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. |

# 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 facilité 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**.

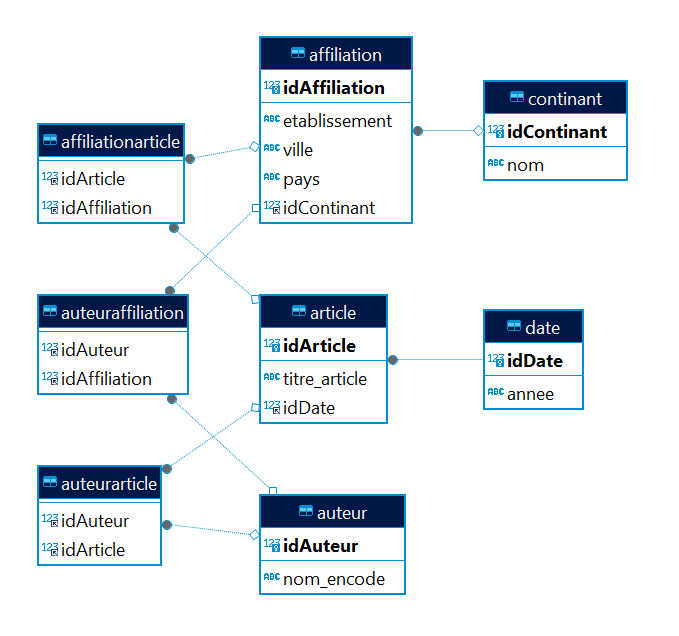
# 

# CONCEPTION DE L’ENTREPÔT

# Définition de l’approche :

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

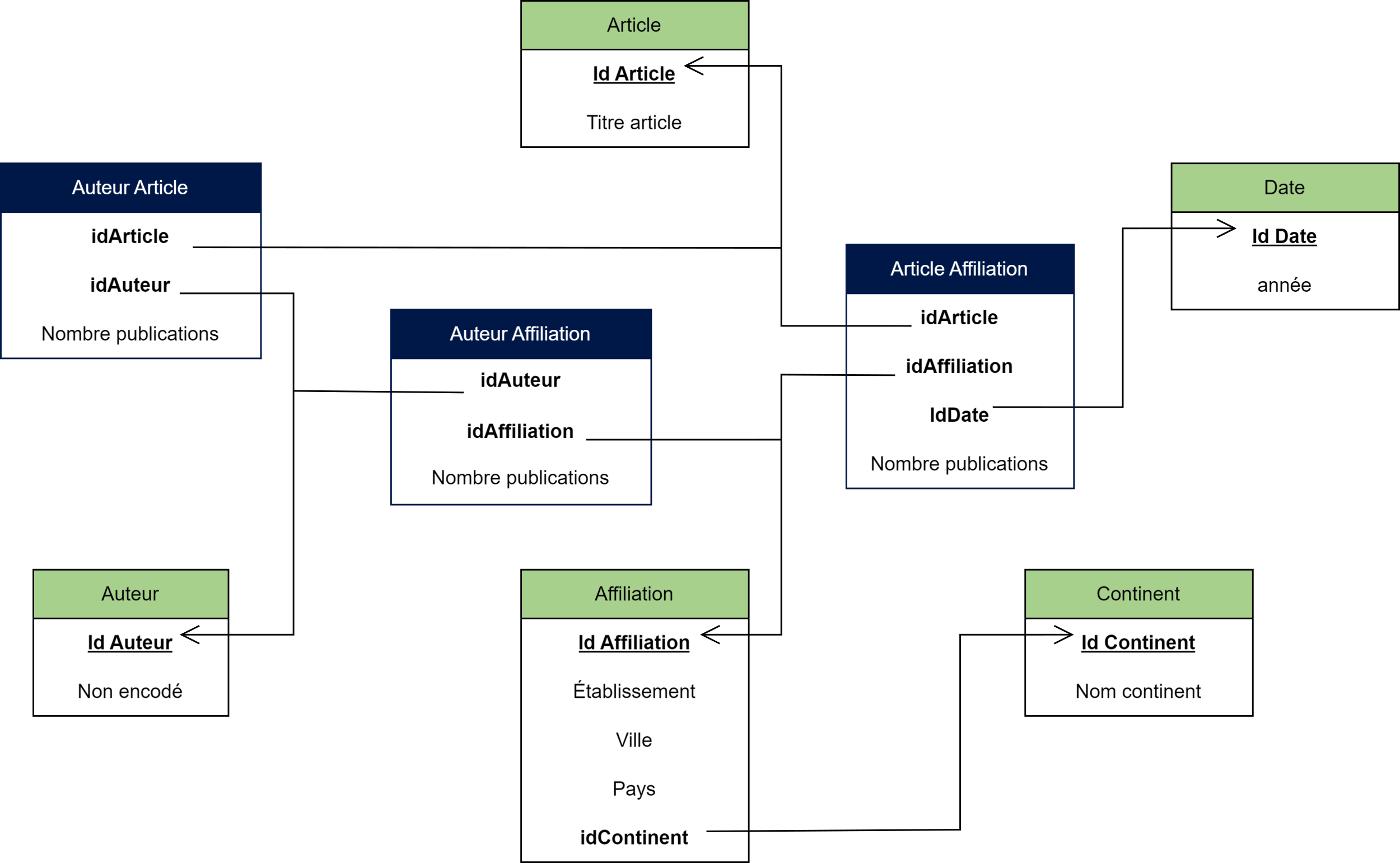
* **Inventaire des sources de données existante (**dans notre cas les fichiers CSV fournis**) ;**
* **Analyse des sources de données (**définition des tables, attributs, clés …**) :**
* **Schéma de la base de données :**

****

* **Identification des besoins métier (**comme défini [**ci-dessus** 👆](#_DEFINITION_DES_BESOINS)**) ;**
* **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

# Schéma en constellation :



# Schéma multidimensionnel :

