



NETFLIX

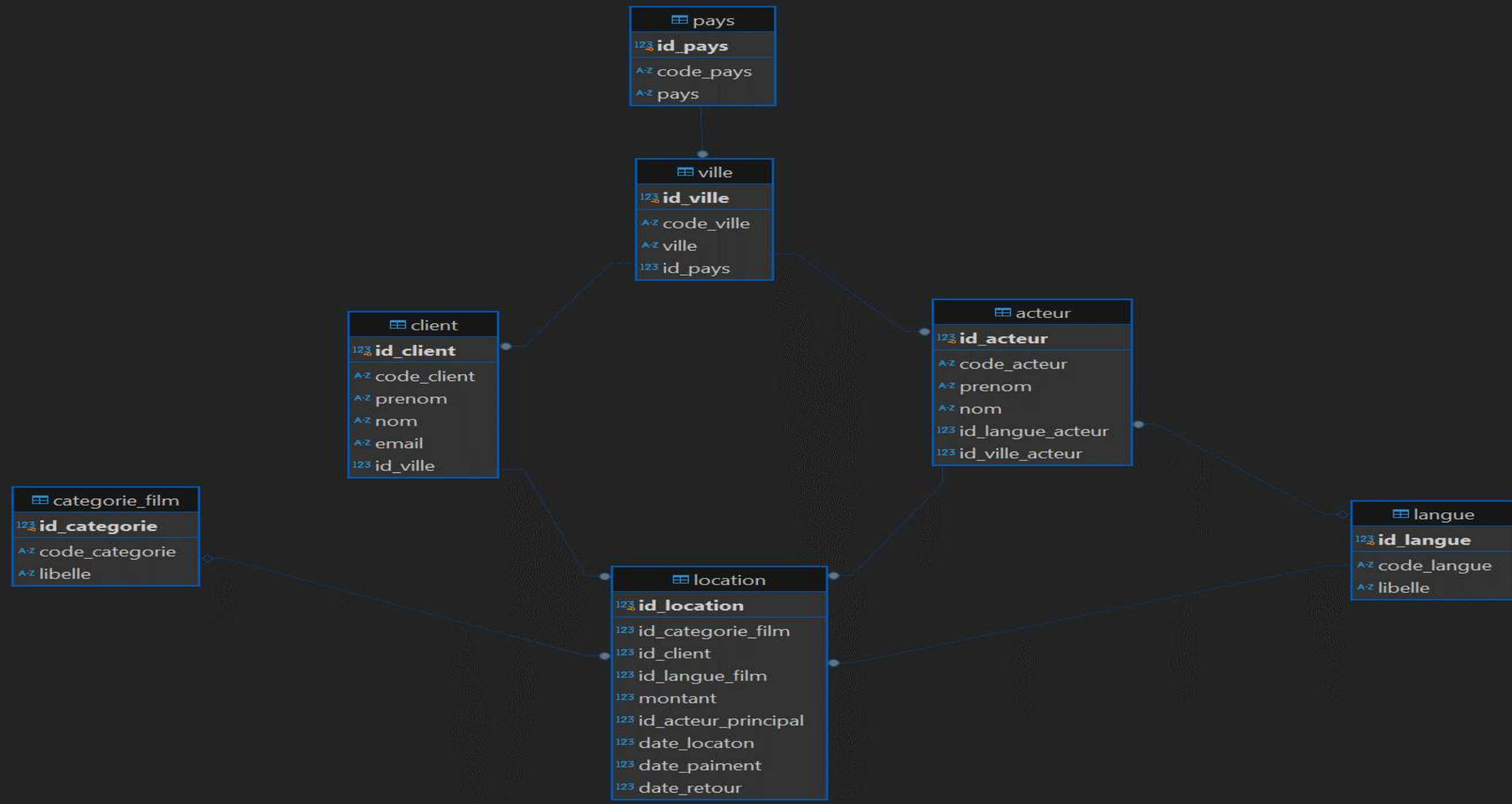
**RAPPORT DU PROJET : MODELISATION ET CONCEPTION DE
BASES DE DONNEES MULTIDIMENSIONNELLES POUR
L'ANALYSE DES DONNEES CHEZ NETFLIX**

PAR MOHAMED FALILOU FALL

INTRODUCTION

- Dans un contexte où les données jouent un rôle central dans les prises de décision stratégiques, ce projet académique vise à concevoir et optimiser une base de données multidimensionnelle dédiée à l'analyse de la performance de la plateforme Netflix. L'objectif est de fournir à l'entreprise un entrepôt de données (Data Warehouse) qui permettra de centraliser et exploiter efficacement les données transactionnelles, comportementales, et géographiques afin de mieux comprendre les préférences des utilisateurs et d'améliorer la stratégie commerciale.
- Ce projet intègre une approche complète, allant de la construction de la base de données opérationnelle à la modélisation avancée pour des analyses multidimensionnelles. En mettant en place un modèle d'entrepôt de données (modèle en étoile), nous visons à faciliter l'accès aux informations clés sur les transactions, utilisateurs, contenus, et périodes. Ces données permettront une analyse approfondie des tendances de consommation et des performances, tant au niveau local (Sénégal) qu'au sein de l'espace UEMOA et au-delà.

CONSTRUCTION DE LA BASE DE DONNÉES OPÉRATIONNELLE NETFLIX



AXES D'ANALYSES

Utilisateurs

Contenus

Periodes

Transactions(paie ment)

Geographie

DÉFINITION DES INDICATEURS CLÉS DE PERFORMANCE (KPI)

MONTANT

NOMBRE DE LOCATIONS

DUREE DES LOCATIONS EN JOURS

REVENUS MENSUELS

TOTAL UTILISATEURS DEBUT ANNEE

UTILISATEURS RETENUS FIN ANNEE

TAUX DE RETENTION

MONTANT TOTAL

NOMBRE TOTAL LOCATIONS

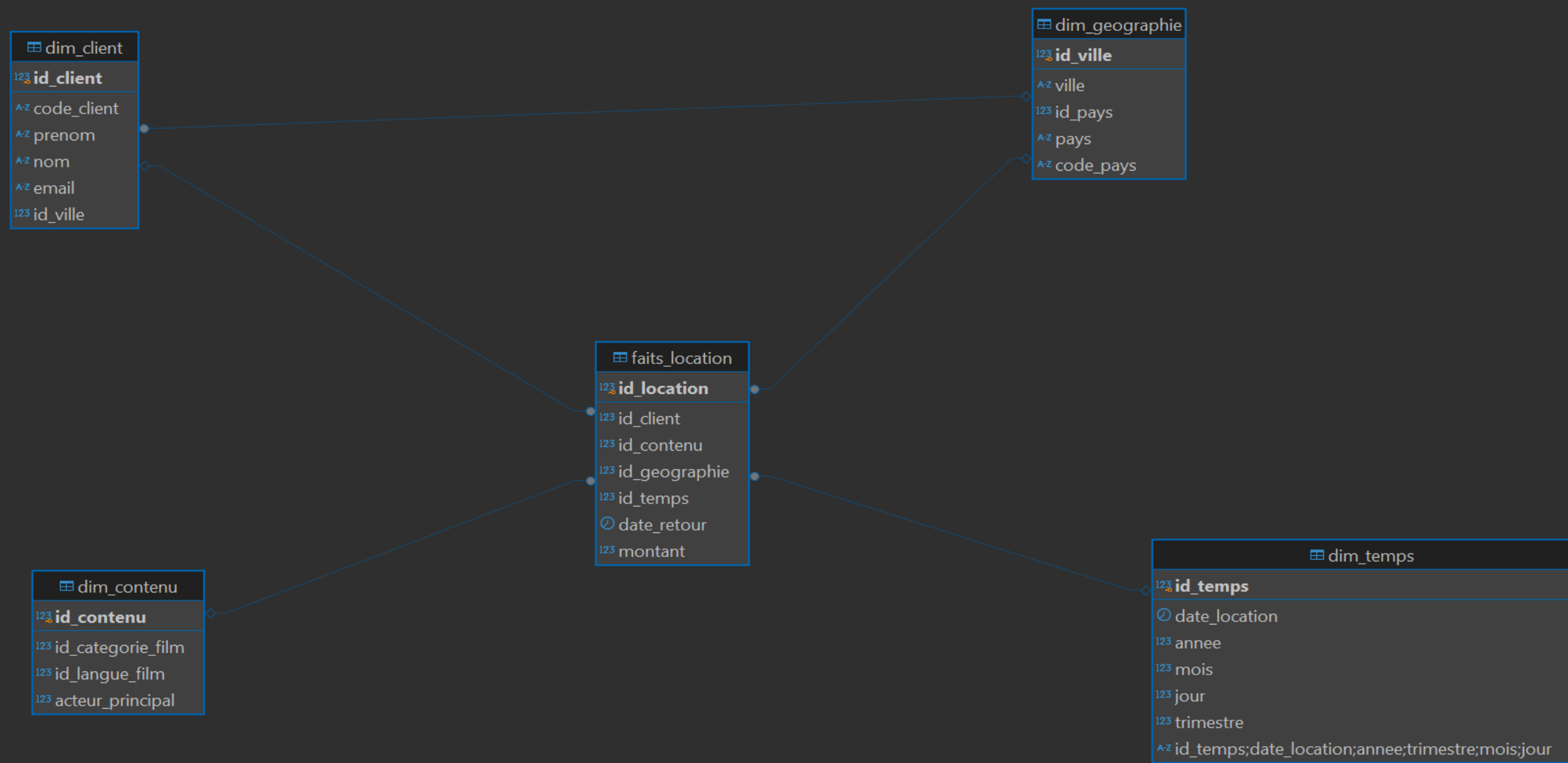
Scenario 1

- suivi des tendances de consommation

Scenario 2

- segmentation des clients

MODELISATION ET IMPLEMENTATION D'UN ENTREPOT DE DONNEES (DWH) DEDIE "DWHNETFLIX "



CONCEPTION DU CUBE OLAP AVEC PENTAHO SCHEMA WORKBENCH

Dimension Client

- id_client
- code_client
- prenom
- nom
- email

Dimension Contenu

- id_contenu
- id_categorie_film
- id_langue_film
- acteur_principal

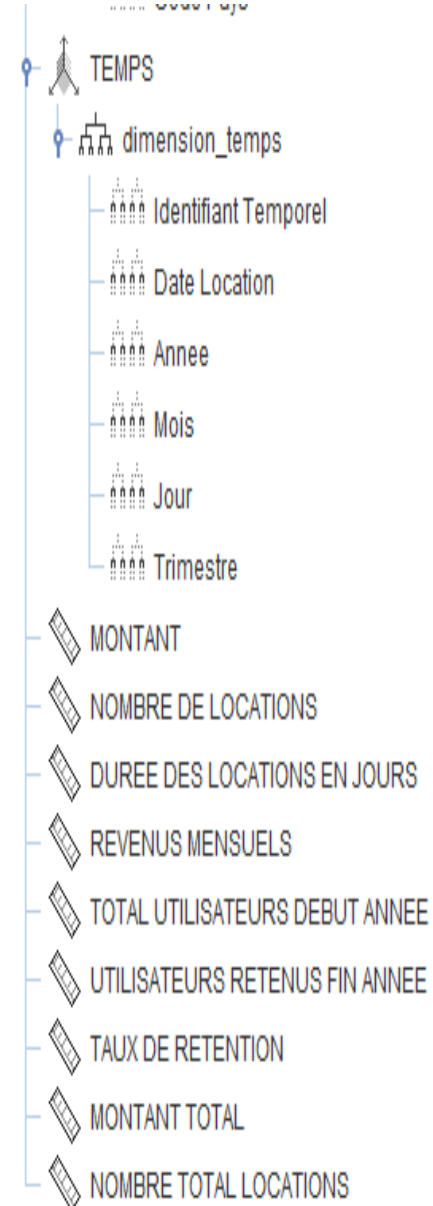
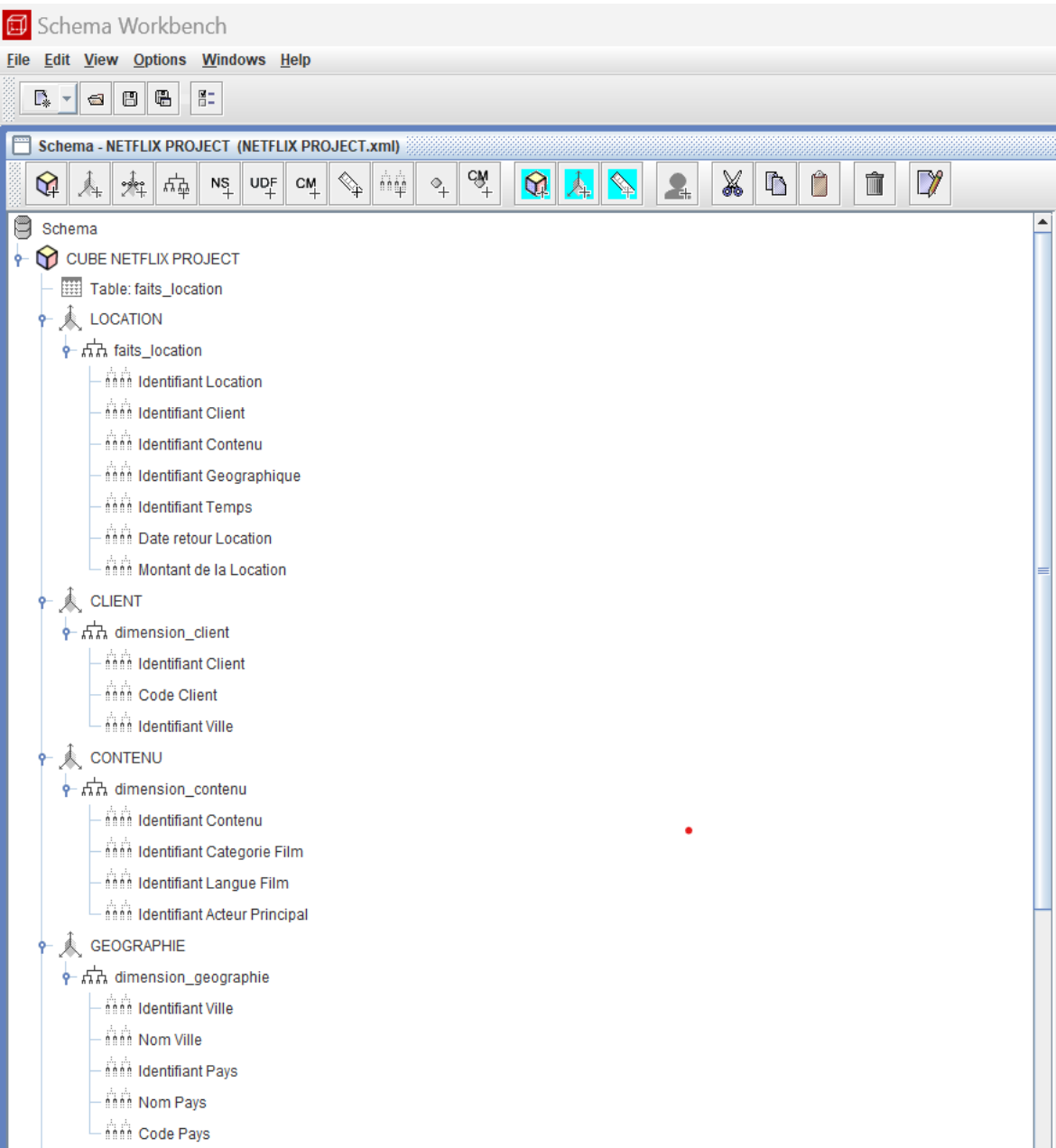
Dimension Temps

- id_temps
- date_location
- annee
- mois
- jour
- trimestre

Dimension Geographie

- id_ville
- ville
- id_pays
- pays
- code_pays

OPTIMISATION DU SCHEMA MONDRIAN POUR LES REQUÊTES MULTIDIMENSIONNELLES



Database - netflix (PostgreSQL)

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

ANALYSE DES REVENUS MENSUELS PAR TYPE DE CONTENU AU SENEGAL

```
WITH
  MEMBER [Measures].[REVENUS MENSUELS] AS
    SUM(
      [Location].[Date retour Location].CurrentMember.Children,
      [Measures].[MONTANT]
    )
SELECT
  {
    [TEMPS].[Mois].Members
  } ON ROWS,
  {
    [CONTENU].[Identifiant Categorie Film].Members
  } ON COLUMNS
FROM
  [CUBE_NETFLIX_PROJECT]
WHERE
  ([GEOGRAPHIE].[Nom Pays].&[Senegal])
```

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

TOP 10 DES FILMS LES PLUS LOUÉS SUR LA PÉRIODE D'UNE ANNÉE AU SENEGAL ET L'ESPACE U.E.M.O.A

```
WITH
  MEMBER [Measures].[NOMBRE DE LOCATIONS] AS
    SUM([LOCATION].[Identifiant Location].MEMBERS, [Measures].[NOMBRE DE LOCATIONS])

SELECT
  TOPCOUNT(
    [CONTENU].[Identifiant Contenu].MEMBERS,
    10,
    [Measures].[NOMBRE DE LOCATIONS]
  ) ON ROWS,

  [TEMPS].[Annee].&[2023] ON COLUMNS

FROM
  [CUBE_NETFLIX_PROJECT]
WHERE
  ([TEMPS].[Mois].MEMBERS, [GEOGRAPHIE].[Nom Pays].&[Sénégal])
```

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

TAUX DE RÉTENTION : POURCENTAGE D'UTILISATEURS QUI RESTENT ABONNÉS SUR UNE PÉRIODE D'UNE ANNÉE PAR RAPPORT AU TOTAL DES UTILISATEURS DU DÉBUT DE LA PÉRIODE (1^{ER} JANVIER) AU SENEGAL

```
WITH
-- Nombre total d'utilisateurs actifs au debut de l'annee (1er janvier)
MEMBER [Measures].[TOTAL UTILISATEURS DEBUT ANNEE] AS
Count(
    Filter(
        [CLIENT].[dimension_client].[Identifiant Client].Members,
        ([TEMPS].[dimension_temps].[Date Location].[1-Jan], [Measures].[NOMBRE DE LOCATIONS]) > 0
    )
)

-- Nombre d'utilisateurs encore actifs a la fin de l'annee (31 décembre)
MEMBER [Measures].[UTILISATEURS RETENUS FIN ANNEE] AS
Count(
    Filter(
        [CLIENT].[dimension_client].[Identifiant Client].Members,
        ([TEMPS].[dimension_temps].[Date Location].[31-Dec], [Measures].[NOMBRE DE LOCATIONS]) > 0
    )
)

-- Calcul du Taux de Retention
MEMBER [Measures].[TAUX DE RETENTION] AS
IIF(
    [Measures].[TOTAL UTILISATEURS DEBUT ANNEE] > 0,
    ([Measures].[ UTILISATEURS RETENUS FIN ANNEE] * 100.0) / [Measures].[ TOTAL UTILISATEURS DEBUT ANNEE],
    NULL
)

SELECT
    {[Measures].[TAUX DE RETENTION]} ON COLUMNS,
    {[GEOGRAPHIE].[dimension_geographie].[Nom Pays].&[Senegal]} ON ROWS
FROM [CUBE_NETFLIX_PROJECT]
WHERE ([TEMPS].[dimension_temps].[Annee].&[2024])
```

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

COMPARAISON DES LOCATIONS ENTRE LE SEMESTRE 1 ET LE SEMESTRE 2 AU SENEGAL

- WITH
- MEMBER [Temps].[Semestre].[S1] AS
- Aggregate(
• {[Temps].[Annee].CurrentMember, [Temps].[Mois].&[1]:[Temps].[Mois].&[6]}
-)
- MEMBER [Temps].[Semestre].[S2] AS
- Aggregate(
• {[Temps].[Annee].CurrentMember, [Temps].[Mois].&[7]:[Temps].[Mois].&[12]}
-)
-
- SELECT
- {[Temps].[Semestre].[S1], [Temps].[Semestre].[S2]} ON COLUMNS,
- {[GEOGRAPHIE].[Nom Pays].&[Senegal]} ON ROWS
- FROM [CUBE_NETFLIX_PROJECT]
- WHERE
- (Measures.[NOMBRE DE LOCATIONS])

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

MONTANT TOTAL DES LOCATIONS AU SENEGAL

```
SELECT
    {[Measures].[MONTANT]} ON COLUMNS,
    {[GEOGRAPHIE].[dimension_geographie].[Nom Pays].[Senegal]} ON ROWS
FROM [CUBE_NETFLIX_PROJECT]
WHERE ([GEOGRAPHIE].[dimension_geographie].[Nom Pays].[Senegal])
```

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

IMPACT DES LANGUES SUR LES LOCATIONS AU SENEGAL

```
SELECT
NON EMPTY
  {[Measures].[NOMBRE DE LOCATIONS], [Measures].[MONTANT]} ON COLUMNS,
NON EMPTY
  (
    [CONTENU].[dimension_contenu].[Identifiant Langue Film].Members
  ) ON ROWS
FROM
  [CUBE_NETFLIX_PROJECT]
WHERE
  ([GEOGRAPHIE].[dimension_geographie].[Nom Pays].&[Senegal])
```

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

ANALYSE DES PICS DE LOCATION PAR JOUR DE LA SEMAINE AU SENEGAL

- SELECT
- NON EMPTY
- [TEMPS].[Jour].Members ON COLUMNS,
- NON EMPTY
- {[Measures].[NOMBRE DE LOCATIONS]} ON ROWS
- FROM
- [CUBE_NETFLIX_PROJECT]
- WHERE
- ([GEOGRAPHIE].[Nom Pays].[Sénégal])

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

EVOLUTION DE LA LOCATION AU FIL DE L'ANNÉE (LES 12 MOIS) AU SENEGAL

```
WITH
MEMBER [Measures].[MONTANT TOTAL] AS
    SUM(
        [GEOGRAPHIE].[Nom Pays].[Senegal],
        [Measures].[MONTANT]
    )
SELECT
    {[Measures].[MONTANT TOTAL]} ON COLUMNS,
    {[TEMPS].[Mois].Members} ON ROWS
FROM [CUBE_NETFLIX_PROJECT]
WHERE
    ([GEOGRAPHIE].[Nom Pays].[Senegal])
```

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

SEGMENTATION DES UTILISATEURS SELON LEURS COMPORTEMENTS DE LOCATION AU SENEGAL

```
SELECT
NON EMPTY
{
    [Measures].[MONTANT],
    [Measures].[NOMBRE DE LOCATIONS]
} ON COLUMNS,

NON EMPTY
(
    [CLIENT].[dimension_client].[Identifiant Client].Members *
    [GEOGRAPHIE].[dimension_geographie].[Nom Pays].&[Senegal] *
    [TEMPS].[dimension_temps].[Annee].Members *
    [TEMPS].[dimension_temps].[Trimestre].Members
) ON ROWS

FROM [CUBE_NETFLIX_PROJECT]

WHERE
(
    [GEOGRAPHIE].[dimension_geographie].[Nom Pays].&[Senegal]
)
```

REQUETES MDX POUR L'ANALYSE STRATEGIQUE

PRÉVISION DE LA DEMANDE BASÉE SUR LES TENDANCES HISTORIQUES AU SENEGAL

```
WITH
MEMBER [Measures].[MONTANT TOTAL] AS
SUM(
    [GEOGRAPHIE].[Nom Pays].&[Senegal],
    [Measures].[MONTANT]
)

MEMBER [Measures].[NOMBRE TOTAL LOCATIONS] AS
SUM(
    [GEOGRAPHIE].[Nom Pays].&[Senegal],
    [Measures].[NOMBRE DE LOCATIONS]
)

SELECT
{
    [Measures].[MONTANT TOTAL],
    [Measures].[Nombre Total Locations]
} ON COLUMNS,

NON EMPTY
(
    [TEMPS].[Annee].[Annee].Members *
    [TEMPS].[Mois].[Mois].Members
) ON ROWS

FROM [CUBE_NETFLIX_PROJECT]

WHERE ([GEOGRAPHIE].[Nom Pays].&[Senegal])
```