# Créer la data base chu

Create database ChuData ;

Use ChuData

## Lister les permissions actuelle dans CLOUDERA

hdfs dfs -ls -R /user/cloudera/ChuData

activer les permissions partt :

hdfs dfs -chmod -R 777 /user/cloudera/ChuData

# DIMENSIONS

PATIENT

CREATE TABLE IF NOT EXISTS Dimension\_Patient (

IdPatient INT,

Age INT,

Sexe STRING,

Departement STRING,

EstDeces BOOLEAN,

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ';'

STORED AS TEXTFILE

LOCATION '/user/cloudera/ChuData/Patient';

DATE

CREATE TABLE IF NOT EXISTS Dimension\_Date (

IdDate INT,

Jour INT,

Mois INT,

Annee INT

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ';'

STORED AS TEXTFILE

LOCATION '/user/cloudera/ChuData/Date';

PROFESSIONNEL :

CREATE TABLE IF NOT EXISTS Dimension\_Professionnel (

Nom STRING,

Prenom STRING,

Specialite STRING,

IdProfessionnel INT,

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ';'

STORED AS TEXTFILE

LOCATION '/user/cloudera/ChuData/Professionnel';

Diagnostic :

CREATE TABLE IF NOT EXISTS Dimension\_Diagnostic (

Description\_Diagnostic STRING,

Code\_Diagnostic STRING,

Jours\_Hospitalisation INT,

Identifiant\_patient INT,

IdConsultation INT

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ';'

STORED AS TEXTFILE

LOCATION '/user/cloudera/ChuData/Diagnostic';

Localisation :

CREATE TABLE IF NOT EXISTS Dimension\_Localisation (

IdLocalisation INT,

Departement STRING,

Ville STRING,

Pays STRING,

Etablissement STRING,

Satisfaction\_Etablissement DECIMAL

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ';'

STORED AS TEXTFILE

LOCATION '/user/cloudera/ChuData/Localisation';

Table des faits :

CREATE TABLE IF NOT EXISTS Faits\_Consultation (

FK\_IdPatient INT,

FK\_IdDate INT,

FK\_IdProfessionnel INT,

FK\_IdConsultation INT,

Nombre\_Consultation INT,

Nombre\_Hospitalisation INT,

Nombre\_Deces INT,

Taux\_Satisfaction DECIMAL

)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY '';

STORED AS TEXTFILE

LOCATION '/user/cloudera/ChuData/Faits\_consultation';

# Peuplement des dimensions ( vérifier si les fichiers sont stockés dans user/cloudera ou user/hive)

LOAD DATA INPATH '/user/cloudera/ChuData/Patient/Patient.txt' INTO TABLE Dimension\_Patient;

LOAD DATA INPATH '/user/cloudera/ChuData/Date/Date.txt' INTO TABLE Dimension\_Date;

LOAD DATA INPATH '/user/cloudera/ChuData/Professionnel/Professionnel.txt' INTO TABLE Dimension\_Professionnel;

LOAD DATA INPATH '/user/cloudera/ChuData/Diagnostic/Diagnostic.txt' INTO TABLE Dimension\_Diagnostic;

LOAD DATA INPATH '/user/cloudera/ChuData/Localisation/Localisation.txt' INTO TABLE Dimension\_Localisation;

# Créer la table des faits

## CREATE TABLE Faits\_Consultation (

## IdPatient INT,

## IdDate INT,

## IdProfessionnel INT,

## IdConsultation INT,

## IdLocalisation INT,

## Nombre\_Consultation INT,

## Nombre\_Hospitalisation INT,

## Nombre\_Deces INT,

## Taux\_Satisfaction DECIMAL(10, 2)

## )

## ROW FORMAT DELIMITED

## FIELDS TERMINATED BY ','

## STORED AS TEXTFILE;

## Loader la table des faits

INSERT INTO TABLE chudata.faits\_consultation

SELECT

d.id\_patient AS IdPatient,

da.id\_date AS IdDate,

p.idprofessionnel AS IdProfessionnel,

d.id\_consultation AS IdConsultation,

l.idlocalisation AS IdLocalisation,

COUNT(d.id\_consultation) AS Nombre\_Consultation,

d.jour\_hospitalisation AS Nombre\_Hospitalisation,

SUM(CASE WHEN pa.est\_deces = TRUE THEN 1 ELSE 0 END) AS Nombre\_Deces,

AVG(1.taux\_reco\_brut) AS Taux\_Satisfaction

FROM

diagnostic d

JOIN

date da ON d.id\_consultation = da.id\_date

JOIN

patient pa ON d.id\_consultation = pa.idpatient

JOIN

professionnel p ON d.id\_consultation = p.idprofessionnel

JOIN

localisation l ON p.idprofessionnel = l.idlocalisation

GROUP BY

d.id\_patient, da.id\_date, p.idprofessionnel, d.id\_consultation, l.idlocalisation, d.jour\_hospitalisation;

Partitionnement

Localisation

CREATE EXTERNAL TABLE IF NOT EXISTS Localisation\_partition (

Idlocalisation INT,

commune STRING,

pays STRING,

Etablissement STRING,

Satisfaction\_Etablissement FLOAT

)

COMMENT 'table Localisation\_partition'

PARTITIONED BY (departement STRING)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ';'

STORED AS TEXTFILE

LOCATION '/user/hive/ChuData/Localisation';

Insert localisation

INSERT OVERWRITE TABLE Localisation\_partition PARTITION (departement)

SELECT

Idlocalisation,

commune,

pays,

raison\_sociale\_site,

taux\_reco\_brut,

departement

FROM

localisation;

Les dimensions sont créées :

