PROJETS SID 2024

Modélisation multidimensionnelle, ETL et BI

Dr Edouard Ngor SARR

edouard-ngor.sarr@univ-zig.sn

A rendre

Mail : sid2025@sarrisgroupe.com

Objet : PROJET SID 2025

Date : Avant le 20 Fév. 2025

Mode : Rapport version Word

Plan à suivre

Page de garde

Sommaire

Résumé

1. Architecture générale du système
2. Présentation des SIO
3. Présentation SID
  - 3.1. Description des Sujets d'étude, des faits et des dimensions
  - 3.2. Modélisation du SID
  - 3.3. Génération Code SQL de création du DWH
4. Création du Data Warehouse
  - 4.1. Création de la Base de données
  - 4.2. Importation du Script SQL
5. Chargement des données des sources au ETL
  - 5.1. Création des JOB
  - 5.2. Exportation des JOB
  - 5.3. Planification des JOB
6. Création du module de restitution
  - 6.1. Choix de l'outil BI
  - 6.2. Création des visuels
  - 6.3. Présentation des visualisations finales

Conclusion

Projet 1 : Gestion commerciale

L'entreprise la Casamançaise souhaite mettre en place un système de gestion de ses ventes et de ses distributions au niveau national munie d'un SID. Pour le SID il serait souhaitable d'avoir des informations sur les ventes et les distributions. L'entreprise dispose d'un service dans chaque région, d'un bureau dans chaque département et d'un dépôt de distribution dans chaque commune.

Les clients de l'entreprise sont de deux types :

- Revendeurs : Boutiques privées spécialisées eau casamançaise ou des simples boutiques
- Entreprises tels que les restaurants, les hôtels etc.

Il vous est demandé de :

1. Proposer le Diagramme de classe UML du système SIO pour les ventes
2. Proposer le Diagramme de classe UML du système SIO pour les distributions
3. Proposer le modèle en Etoile pour le SID
4. Créer et alimenter la base de données MySQL du SIO des ventes dans l'ordi O1
5. Créer et alimenter la base de données Access du SIO des distributions dans l'ordi O2
6. Créer la base de données la base de données SID dans l'ordi O3
7. Via TELEND mettre en place le ETL de chargement des données des SIO vers le SID
8. Exporter le JAR des Jobs et planifier son exécution par 5 mn.
9. Via Power BI proposer les analyses et les visualisations nécessaire

Projet 2 : Gestion hospitalière

Hôpital régional de Ziguinchor souhaite disposer d'un système de gestion des consultations et des hospitalisations munies d'un SID. Les consultations sont enregistrées dans un fichier Excel Consultation.CSV de l'ordinateur 1 avec les informations : Nom et prénom du patient, date de la consultation, le service, numéro du ticket, montant de la consultation, la salle de consultation, résultats de la consultation, l'agent de santé et Date du prochain RDV. Il vous est demandé de :

1. Proposer le Diagramme de classe UML du système SIO pour la gestion des hospitalisations
2. Créer et alimenter le fichier Excel des consultations dans l'ordi O1
3. Créer et alimenter la base de données MySQL du SIO de gestion des hospitalisations dans l'ordi 02.
4. Créer le JOB pour le fichier Excel des consultations dans la table Tconsultations de la base de données de l'ordi 02.
5. Proposer le modèle en Etoile pour le SID
6. Créer la base de données la base de données SID dans l'ordi O3
7. Via TELEND mettre en place le ETL de chargement des données de la base de données MySQL de l'ordi 02 vers l'entrepôt de données de l'ordinateur 03.
8. Exporter le JAR des deux Jobs et planifier leurs exécutions à 23h chaque jour.
9. Via Power BI proposer les analyses et les visualisations nécessaire

Projet 3 : Gestion entreprise de transport / Dakar Dem Dikk

La société Dakar Dém Dikk est une entreprise de transport sénégalaise. Il propose deux types de services dont un service de Transport urbain et un service Transport Dakar et les autres régions nommé SENEGAL DEM DIKK.

La direction du transport urbain de DDD dispose d'un fichier Excel qui enregistre les voyageurs et des lignes : Numéro du Bus, Numéro Ligne, Lieu de départ, Heure de départ, Heure Arrêt, Date, Nb ticket Emis, Nombre de ticket Vendus, Nombre de Tickets retournés, Montant Total, Incident

La direction de SENEGAL DEM DIKK quant à elle traite ses données avec une table d'une base de données Access avec les informations suivantes : Numéro du Bus, Numéro Ligne, Lieu de départ, Heure de départ, Heure Arrivée, Date, Nb ticket Emis, Nombre de ticket Vendus, Nombre de Tickets retournés, Montant Total, Incident et Nombre de passagers.

1. Créer et alimenter le fichier Excel et la table de la base de données Access respectivement dans l'ordinateur 1 et l'ordinateur 2
2. Proposer un model en Etoile pour le SID général
3. Créer la base de données SID dans l'ordinateur O3
4. Via TELEND mettre en place le ETL de chargement des données des SIO vers le SID
5. Exporter le JAR des Jobs et planifier son exécution à 00h chaque jour.
6. Via Power BI proposer les analyses et les visualisations nécessaire

Projet 4: Gestion Scolaire

L'UASZ souhaite mettre en place un système décisionnel pour les inscriptions et les enseignements dont les données sont stockées dans une base de données MySQL.

Elle souhaite pouvoir répondre aux questions suivantes :

- a) Quel est le nombre de réussites aux examens par cours, pour l'année 2019 ?
- b) Quel est le nombre de réussites aux examens d'un cours obligatoire, pour l'année 2018 ?
- c) Quel est le nombre de réussites aux examens par sexe (féminin, masculin), pour 2017 ?
- d) Combien d'étudiants ayant un âge de 22 ans ont réussi leurs examens de BD relationnelle ?
- e) Quel est le nombre de réussites aux examens pendant le semestre d'hiver 2016 ?

Pour cela elle dispose des données suivantes :

- Les informations (Matricule, Prénom, Nom, DateNaissance, sexe) se trouve dans un fichier Excel de l'ordinateur de la secrétaire (à importer)
- Les informations sur les cours (NomCours, UE, NbHeure, prof, Type) se trouve dans une base données ACCESS du service scolarité (à importer)
- Pour chaque examen passé, on connaît l'âge et le sexe de l'étudiant, le nom du cours (les cours peuvent être regroupés en cours obligatoire et cours à option), la date de l'examen, la note obtenue et si l'examen est réussi ou non. Ces informations sont dans une table TExamen de la base de données MySQL BDEXAMENS de la direction (à importer).

Question 1 : Proposez un modèle en étoile pour application. Recherchez tout d'abord les différentes dimensions et proposez une hiérarchie pour ces dimensions

Question 2 : Donner le code SQL de création du Data Warehouse

Question 3 : Créer la base de données sur ACCESS dans un des ordinateurs du groupe.

Question 4 : Enregistrer dans chaque Table au moins 20 lignes en respectant les contraintes de clés étrangères.

Question 5 : Via TALEND, créer un JOB pour Importer les données de ACCESS dans le Data Warehouse MySQL se trouvant dans un deuxième ordinateur du groupe.

Question 6 : Donner les requêtes de création des Cubes OLAP a, b, c, d et e ci-haut.

Question 6 : Utiliser Power BI pour mettre en place la visualisation des Cubes