



BRUXELLE



Bases de données : 8 - Structuration de données en SQL

FRSITAS

Karim Lidouh (Karim.Lidouh@ulb.ac.be)

http://cafesig.ulb.ac.be

Introduction

Une quantité de données trop volumineuses peut vous empêcher de corriger les erreurs d'encodage dans toutes les lignes à la main.

Dans ces cas-là, il faudra privilégier une approche plus automatique :

- S'assurer que toutes les données sont suffisamment subdivisées de façon à répondre aux besoins du modèle
- Importer les données dans une unique table au sein de la base de données
- Corriger les erreurs dans chaque colonne en vérifiant les modalités
- Extraire les données nécessaires pour remplir les tables des entités et associations du modèle défini

Exercice : Gestion Aéroport

- Utiliser le fichier CSV de données complètes de l'exercice
- L'objectif est de recréer le modèle établi et de remplir la base de données en utilisant uniquement le langage SQL



Créer une table pour les données brutes

CREATE TABLE donnees

num integer primary key, objet character varying (255), date character varying (15), duree integer, matricule avion character varying(10), type avion character varying (50), constructeur character varying(50), nb places integer, prenom proprietaire character varying(50), nom proprietaire character varying(50), num proprietaire character varying(5), rue proprietaire character varying(50), telephone proprietaire character varying(15),

prenom_mecanicien character varying(50), nom_mecanicien character varying(50), num_mecanicien character varying(5), rue_mecanicien character varying(50), telephone_mecanicien character varying(15), habilitation_1 character varying(50), habilitation_2 character varying(50), habilitation_3 character varying(50)

Charger des données CSV dans pgAdmin III

pgAdmin III est capable de charger tout le contenu d'un fichier CSV dans une table de manière similaire à la commande COPY en langage SQL :

- Faire un clic droit sur la table à remplir
- Choisir « Importer... »
 - Dans l'onglet « Préférences pour les fichiers » : retrouver le fichier CSV à l'endroit où il a été sauvegardé, pour le format choisir 'csv' ; ne pas toucher au reste.
 - Dans l'onglet « Options diverses » : cocher "En-tête" (si le fichier CSV contient des entêtes de colonnes) ; choisir le délimiteur utilisé dans le fichier CSV
- Cliquer sur « Importer » et attendre que la fenêtre affiche « Réalisé »

Vérifier que les données ont bien été chargées dans la table. Si les caractères spéciaux n'ont pas été correctement importés, « nettoyer » la table et recommencer l'opération en indiquant un codage de fichier (par exemple UTF8).

Importer les données subdivisées

	num [PK] integer	objet character varying(255)	date character varying(15)	duree integer	matricule_avion character varying(10)	type_avion character varying(50)	constructeur character varying(50)	nb_places integer	prenom_proprietaire character varying(50)	nom_proprietaire character varying(50)	num_proprietaire character varying(5)	rue_proprietaire character varying(50)	te ch
1	126	Turbine	21/08/2013	45	T112	Airbus A320	Airbus	150	Jean	kagina	99	rue des dunes	12
2	127	Air conditionné	21/08/2013	30	M125	AASI Jetcruzer	AASI	5	Lionel	Noblet	1	Place de la Concorde	87
3	128	Moteur	22/08/2013	180	S175	Gulfstream G500	General Dynamics	19	Lionel	NOBLET	1	Place de la Concorde	87
4	129	Moteur	22/08/2013	210	T459	Airbus A310	Airbus	150	Jean	KAGINA	99	rue des dunes	12
5	130	Tableau de bord	23/08/2013	90	SN115	Boeing 747	Boeing	404	Jean	Kagina	99	rue des dunes	12
6	131	Eclairage	23/08/2013	5	T489	Airbus A320	Airbus	150	Guillaume	Van Bergen	45	rue des plaines	88
7	132	Ajustement siège	24/08/2013	15	M125	AASI Jetcruzer	AASI	5	Lionel	Noblet	1	Place de la Concorde	87
8	133	Tableau de bord	24/08/2013	30	SN116	Boeing 747	Boeing	404	Guillaume	VanBergen	45	rue des plaines	88
9	134	Diagnostic turbine	24/08/2013	200	T155	Airbus A320	Airbus	150	Fedot	Fyodorov	47	Chemin Challet	36
10	135	Ventilation	24/08/2013	45	T180	Airbus A310	Airbus	150	Archie	Tect	8	Rue St Ferreol	36
11	136	Remplacement tableau de bord	24/08/2013	120	SN117	Boeing 747	Boeing	404	Guillaume	Van Bergen	45	rue des plaines	88
12	137	Premier test de démarrage	24/08/2013	120	X101	Airbus A320	Airbus	150	Fedot	FYODOROV	47	Chemin challet	36
13	138	Réparation tableau de bord	24/08/2013	60	T180	Airbus A310	Airbus	150	Archie	Tect	8	Rue St Ferreol	36
14	139	Vérification instruments	24/08/2013	60	T180	Airbus A310	Airbus	150	Archie	tect	8	Rue St Ferreol	36
15	140	Diagnostic complet	25/08/2013	240	X101	Airbus A320	Airbus	150	Fedot	Fyodorov	47	Chemin Challet	36
16	141	Vérification instruments	25/08/2013	30	SN115	Boeing 747	Boeing	404	Jean	KAGINA	99	rue des dunes	12
17	142	Vérification instruments	25/08/2013	30	SN116	Boeing 747	Boeing	404	Guillaume	Van Bergen	45	rue des plaines	88
18	143	Vérification instruments	25/08/2013	30	SN117	Boeing 747	Boeing	404	Guillaume	Van Bergen	45	rue des plaines	88
19	144	Vérification instruments	25/08/2013	30	SN118	Boeing 747	Boeing	404	Guillaume	Van BERGEN	45	rue des plaines	88
20	145	Extraction pigeon	25/08/2013	20	T155	Airbus A320	Airbus	150	Fedot	Fyodorov	47	Chemin Challet	36
21	146	Eclairage	25/08/2013	15	T112	Airbus A320	Airbus	150	Jean	kagina	99	rue des dunes	12
22	147	Nettoyage turbine	25/08/2013	30	T155	Airbus A320	Airbus	150	Fedot	Fyodorov	47	Chemin Challet	36
23	148	Vérification circuits électriques	25/08/2013	30	SN118	Boeing 747	Boeing	404	guillaume	van bergen	45	rue des plaines	88
24	149	Machine à glaçons	25/08/2013	60	M125	AASI Jetcruzer	AASI	5	Lionel	Noblet	1	Place de la Concorde	87
25	150	Ajustement volets	26/08/2013	30	T180	Airbus A310	Airbus	150	Archie	Tect	8	Rue St Ferreol	36
26	151	Nettoyage réacteur	26/08/2013	120	SN115	Boeing 747	Boeing	404	Jean	Kagina	99	rue des dunes	12
27	152	Diagnostic démarrage	26/08/2013	60	SN115	Boeing 747	Boeing	404	Jean	Kagina	99	rue des dunes	12
28	153	Remplacement ampoule	26/08/2013	5	M125	AASI Jetcruzer	AASI	5	Lionel	Noblet	1	Place de la Concorde	87
29	154	Mise à jour auto-pilote	26/08/2013	30	T112	Airbus A320	Airbus	150	Jean	Kagina	99	rue des dunes	12
30	155	Mise à jour auto-pilote	26/08/2013	30	T180	Airbus A310	Airbus	150	Archie	Tect	8	Rue St Ferreol	36
31	156	Mise à jour auto-pilote	26/08/2013	30	T155	Airbus A320	Airbus	150	Fedot	Fyodorov	47	Chemin Challet	36
32	157	Mise à jour auto-pilote	26/08/2013	30	T459	Airbus A310	Airbus	150	Jean	kagina	99	rue des dunes	12
33	158	Mise à jour auto-pilote	26/08/2013	30	T489	Airbus A320	Airbus	150	Guillaume	Van Bergen	45	rue des plaines	88
34	159	Siège éjectable	26/08/2013	180	M125	AASI Jetcruzer	AASI	5	Lionel	Noblet	1	Place de la Concorde	87

Vérifier les modalités de chaque colonne et corriger les erreurs

- Vérifier les modalités des variables qui étaient redondantes grâce à la clause DISTINCT : SELECT DISTINCT prenom_mecanicien, nom_mecanicien, telephone_mecanicien FROM donnees;
- Utiliser les fonctions de chaînes de caractères pour éliminer la plupart des différences : SELECT DISTINCT INITCAP (prenom), UPPER (nom), TRIM (rue)

FROM donnees;

• Choisir les modalités correctes grâce à des COUNT() :

SELECT prenom_mecanicien, nom_mecanicien, telephone_mecanicien, COUNT(*) nb
FROM donnees
GROUP BY prenom mecanicien, nom mecanicien, telephone mecanicien;

• Corriger les erreurs grâce à des UPDATE :

UPDATE donnees **SET** nom_proprietaire='VAN BERGEN' **WHERE** nom_proprietaire='VanBergen';

Créer et remplir les tables de chaque entité/association

• Ecrire la requête SQL qui génère le contenu de la table :

SELECT type_avion nom, constructeur, nb_places FROM donnees
GROUP BY type_avion, constructeur, nb_places;

• Générer des identifiants s'il n'y en a pas :

```
SELECT ROW_NUMBER() OVER (ORDER BY nom) id, type_avion nom, constructeur,
nb_places
FROM donnees
GROUP BY type avion, constructeur, nb places;
```

• Créer une table sur base du résultat du SELECT :

```
CREATE TABLE type_avion AS
SELECT ROW_NUMBER() OVER (ORDER BY nom) id, type_avion nom, constructeur,
nb_places
FROM donnees
GROUP BY type_avion, constructeur, nb_places;
```

Ajouter les contraintes de clés primaire et étrangères

• Ajouter la contrainte de clé primaire en listant les colonnes qui en font partie :

ALTER TABLE type_avion ADD CONSTRAINT type_avion_pkey PRIMARY KEY (id); ALTER TABLE habilitation ADD CONSTRAINT habilitation_pkey PRIMARY KEY (id_mecanicien, id_type_avion);

 Ajouter les contraintes de clé étrangère en indiquant les tables et colonnes référencées :
 ALTER TABLE avion ADD CONSTRAINT avion_id_type_avion_fkey FOREIGN KEY (id type avion) REFERENCES type avion(id);

Exporter / importer une base de données

PostgreSQL permet d'exporter la structure complète d'une base de données en langage SQL :

- Faire un clic droit sur la base de données
- Choisir « Sauvegarder... »
 - Dans l'onglet « Préférences pour les fichiers » : pour le format choisir 'Fichier plat' ; ensuite choisir où enregistrer le fichier .sql et quel nom lui donner ; ne pas toucher au reste.
 - Dans l'onglet « Options 1 » : cocher "Pré-données", "Données", et aussi "Post-données" (si les tables ont des contraintes) ; dans la partie "Ne pas sauvegarder" cocher tout
 - Dans « Options 2 » : cocher les 4 options de la partie "Requêtes"
 - Dans « Objets » : vérifier que ce sont les bonnes tables qui sont sélectionnées
- Cliquer sur « Sauvegarder » et attendre que la fenêtre affiche « Réalisé »
- Le fichier .sql ainsi créé contient les requêtes permettant de recréer la base de données, ses tables et toutes ses données sur un autre système