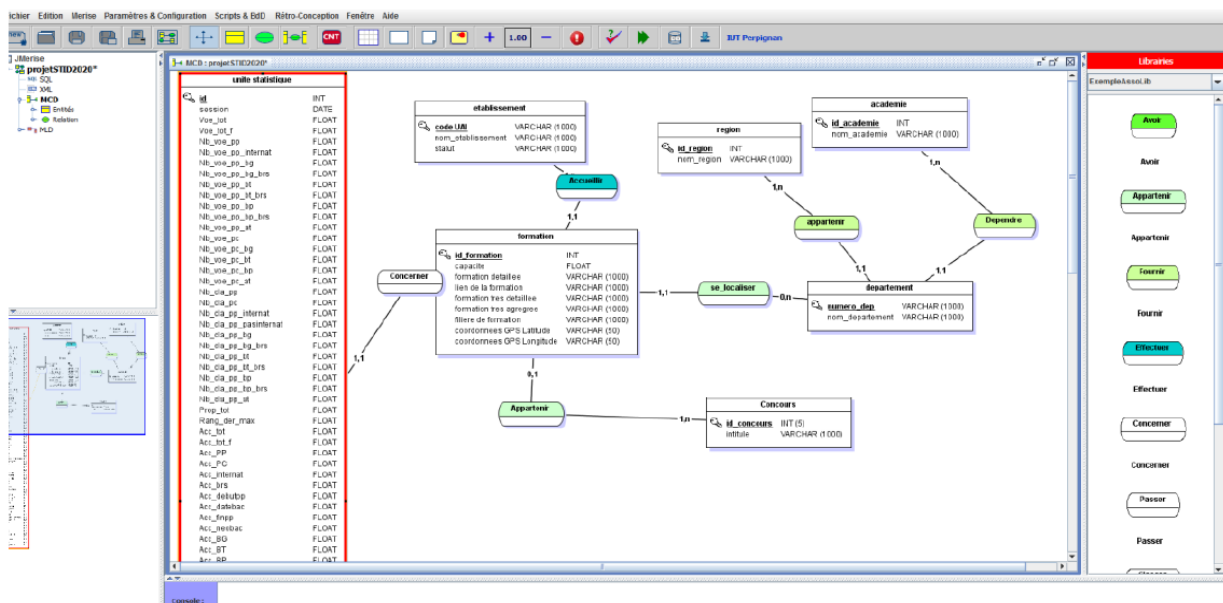




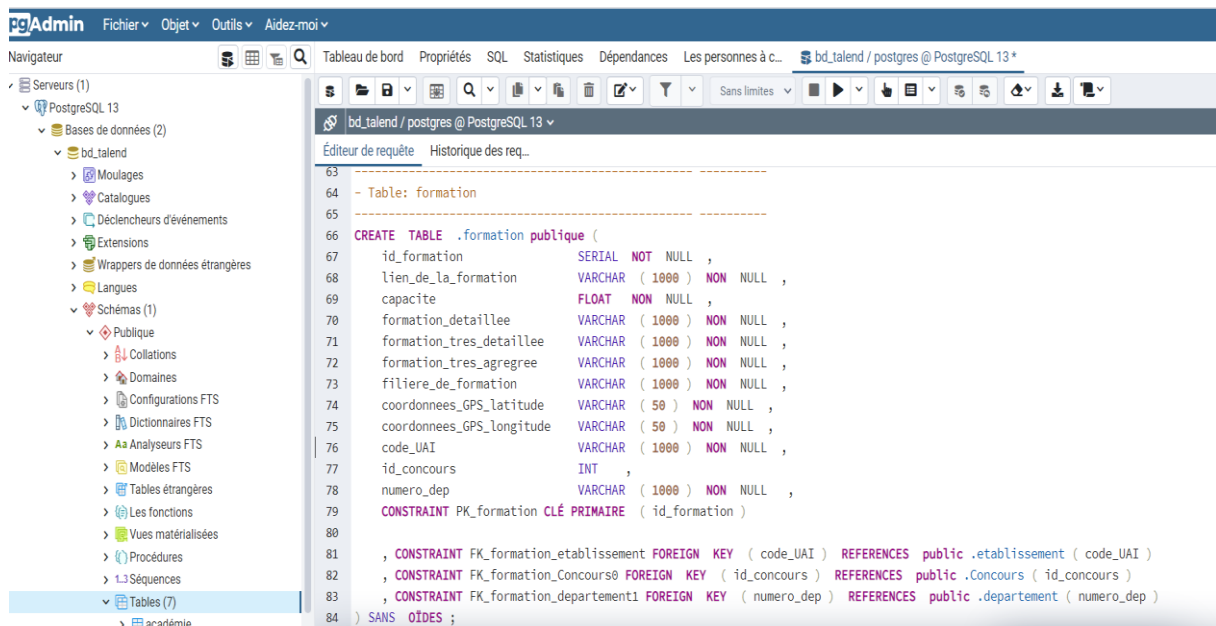
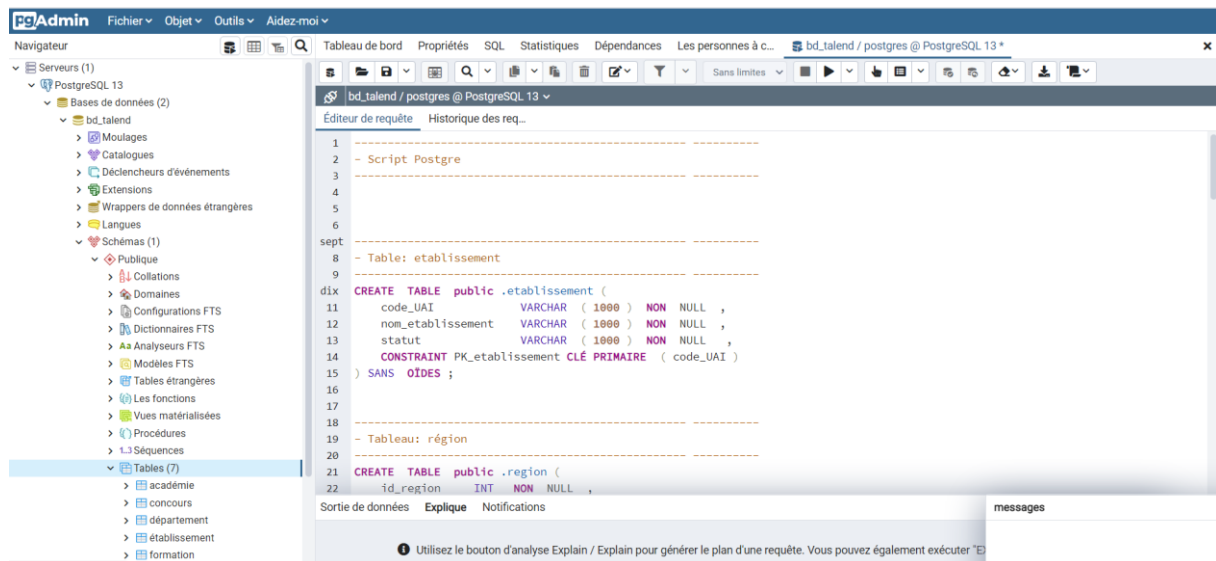
PROJET DE CREATION DE LA BASE DE DONNEES DE L'INSCRIPTION DES CANDIDATS AUX
PARCOURS SUP 2018-2019 POUR LE CHOIX DE LEURS FORMATIONS

Remplissage
Modèle données
STUDIO
FICHIERS
TALEND open
TABLES
MODÉLISER EXCEL
Jmerise
MCD
BASE Le Outil
conceptuel RÉALISÉ
UTILISANT

OUATTARA SEYDOU

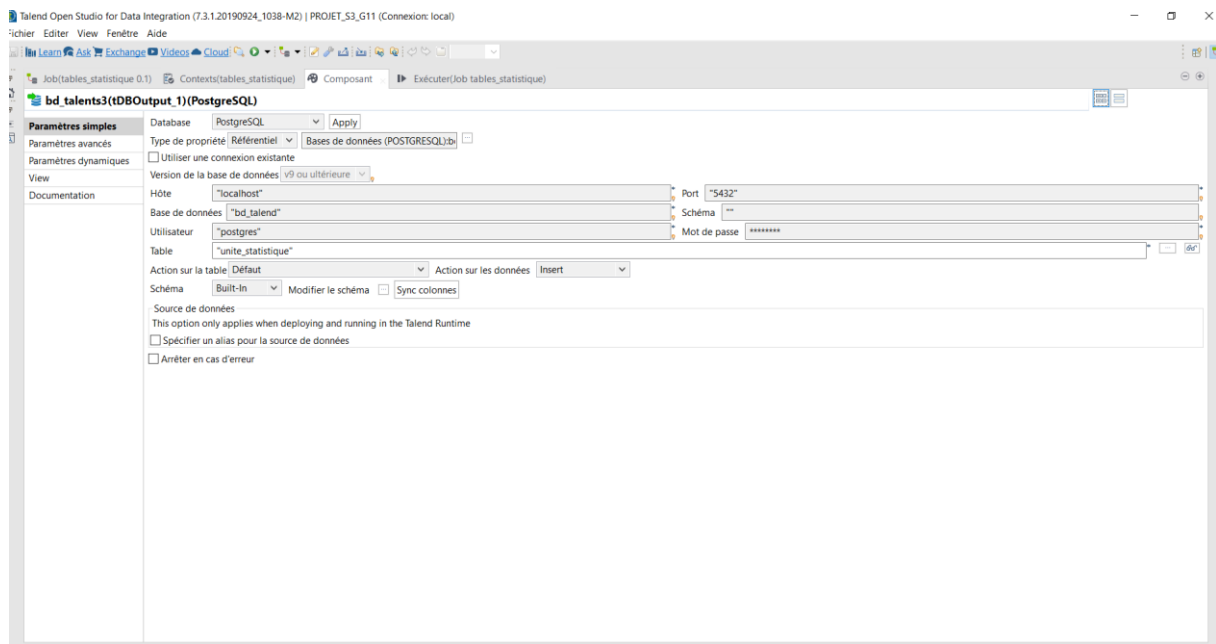


La creation des tables sous postges



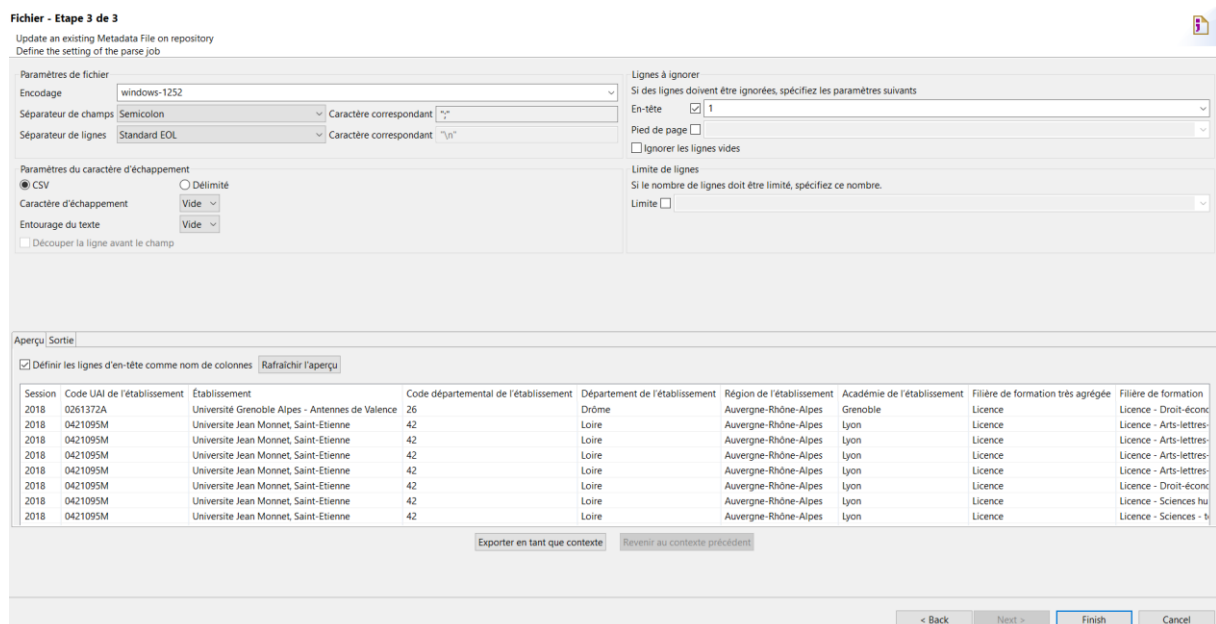
Connection de PostgreSQL à Talend

L'objectif est de connecter PostgreSQL à Talend Open Studio pour pouvoir remplir les tables sous PostgreSQL à l'aide des job Talend.



Importation des fichiers de données de parcours sup 2018 et 2019

A travers la fonction métadonnée de Talend j'ai pu facilement importer les deux fichiers afin d'extraire les variables pour remplir les différentes tables sous PostgreSQL à partir des jobs.



Récupération des schémas de différentes tables de PostgreSQL à Talend

Nouveau schéma de connexion "bd_talents3"

Ajouter un schéma dans le référentiel

Schéma

unite_statistique

Nom: unite_statistique

Commentaire:

Type: TABLE

Basé sur la table: unite_statistique

Récupérer le schéma Détester le schéma

Utilisez le bouton "Récupérer le schéma" pour remplacer le schéma courant par le schéma de l

Schéma

Colonne	Db Column	Clé	Type de base...	Type	N.	Modèle de date (Ctrl+...)	Length	Precision	Défaut	Commentaire
id	id	<input checked="" type="checkbox"/>	INT4	int	<input type="checkbox"/>		10	0		
session	session	<input type="checkbox"/>	INT4	int	<input type="checkbox"/>		10	0		
voe_tot	voe_tot	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
voe_tot_f	voe_tot_f	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp	nb_voe_pp	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp_internat	nb_voe_pp_internat	<input type="checkbox"/>	VARCHAR	String	<input checked="" type="checkbox"/>		100	0		
nb_voe_pp_bg	nb_voe_pp_bg	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp_bg_brs	nb_voe_pp_bg_brs	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp_bt	nb_voe_pp_bt	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp_bt_brs	nb_voe_pp_bt_brs	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp_bp	nb_voe_pp_bp	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp_bp_brs	nb_voe_pp_bp_brs	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp_at	nb_voe_pp_at	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pc	nb_voe_pc	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pc_bg	nb_voe_pc_bg	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pc_bt	nb_voe_pc_bt	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pc_bp	nb_voe_pc_bp	<input type="checkbox"/>	INT4	Integer	<input checked="" type="checkbox"/>		10	0		

Ajouter le schéma

Supprimer le schéma

L'extraction des variables en fonction des tables à remplir

Talend Open Studio for Data Integration - tMap - tMap_1

Find:

Mapping auto

row1

Column

- Session
- Statut
- Code_UAI_de_l_etablissement
- Etablissement
- Code_departemental_de_l_etablissement
- Departement_de_l_etablissement
- Region_de_l_etablissement
- Academie_de_l_etablissement
- Filiere_de_formation_tres_agreee
- Filiere_de_formation
- Concours_communs_et_banques_d_epreuves
- Filiere_de_formation_detailee
- Filiere_de_formation_tres_detailee
- Lien_de_la_formation_sur_la_plateforme_Parcoursup
- Coordonnees_GPS_de_la_formation
- Capacite_de_l_etablissement_par_formation
- Effectif_total_des_candidats_pour_une_formation
- Dont_effectif_des_candidats_pour_une_formation
- Effectif_total_des_candidats_en_phase_principale
- Dont_effectif_des_candidats_ayant_postule_en_internat
- Effectif_des_candidats_neo_bacheliers_generaux_en_phase_principale
- Dont_effectif_des_candidats_boursiers_neo_bacheliers_generaux_en_phas...
- Effectif_des_candidats_neo_bacheliers_technologiques_en_phase_principale

Var

Expression	Type	Variable
Numeric.sequence("s1",1,1)	int	<input type="checkbox"/> id_formation
StringHandling.LEFT(row1.Coordonnee...	String	<input type="checkbox"/> latitude
StringHandling.RIGHT(row1.Coordonnee...	String	<input type="checkbox"/> longitude
Numeric.sequence("s2",1,1)	int	<input type="checkbox"/> id_concours

out1

Expression

- Var.id_formation
- row1.Lien_de_la_formation_sur_la_plateforme_Par...
- row1.Capacite_de_l_etablissement_par_formation
- row1.Filiere_de_formation_detailee
- row1.Filiere_de_formation_tres_detailee
- row1.Filiere_de_formation_tres_agreee
- row1.Filiere_de_formation
- Var.latitude
- Var.longitude
- Var.id_concours
- row1.Code_departemental_de_l_etablissement
- row1.Code_UAI_de_l_etablissement

Column

- id_formation
- Lien_de_la_formation_sur_la_...
- Capacite_de_l_etablissement_...
- Filiere_de_formation_detailee
- Filiere_de_formation_tres_det...
- Filiere_de_formation_tres_agr...
- Filiere_de_formation
- latitude
- longitude
- id_concours
- Code_departemental_de_l_et...
- Code_UAI_de_l_etablissement

row1

Colonne	Clé	Type	N.	Modèle de date (C...	Length	Precision	Défaut	Comment...
Session	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>			0		
Statut	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			0		
Code_UAI_de_l_etablissement	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			0		
Etablissement	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			0		
Code_departemental_de_l_e...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			0		
Departement_de_l_etablisse...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			0		

out1

Colonne	Clé	Type	N.	Modèle de date (C...	Length	Precision	Défaut	Comment...
id_formation	<input checked="" type="checkbox"/>	int	<input type="checkbox"/>					
Lien_de_la_formation_sur_la...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			0		
Capacite_de_l_etablissemen...	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>			0		
Filiere_de_formation_detaill...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			0		
Filiere_de_formation_tres_d...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			0		
Filiere_de_formation_tres_a...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>			0		

Appliquer OK Annuler

Le Mapping Automatique entre PostgreSQL et Talend

Automatique
Talend
entre
PostgreSQL
LE
Mapping

Talend Open Studio for Data Integration - tMap - tMap_1

Find:

row1

Column

- id
- Session
- Voe_tot
- Voe_tot_f
- Nb_voe_pp
- Nb_voe_pp_internat
- Nb_voe_pp_bg
- Nb_voe_pp_bg_brs
- Nb_voe_pp_bt
- Nb_voe_pp_bt_brs
- Nb_voe_pp_bp
- Nb_voe_pp_bp_brs
- Nb_voe_pp_at
- Nb_voe_pc
- Nb_voe_pc_bg
- Nb_voe_pc_bt
- Nb_voe_pc_bp
- Nb_voe_pc_at
- Nb_cla_pp
- Nb_cla_pc
- Nb_cla_pp_internat
- Nb_cla_pp_pasinternat
- Nb_cla_pp_bg
- Nb_cla_pp_bg_brs

Var

out

Expression

- row1.id
- row1.Session
- row1.Voe_tot
- row1.Voe_tot_f
- row1.Nb_voe_pp
- row1.Nb_voe_pp_internat
- row1.Nb_voe_pp_bg
- row1.Nb_voe_pp_bg_brs
- row1.Nb_voe_pp_bt
- row1.Nb_voe_pp_bt_brs
- row1.Nb_voe_pp_bp
- row1.Nb_voe_pp_bp_brs
- row1.Nb_voe_pp_at
- row1.Nb_voe_pc
- row1.Nb_voe_pc_bg
- row1.Nb_voe_pc_bt
- row1.Nb_voe_pc_bp
- row1.Nb_voe_pc_at
- row1.Nb_cla_pp
- row1.Nb_cla_pc
- row1.Nb_cla_pp_internat
- row1.Nb_cla_pp_pasinternat
- row1.Nb_cla_pp_bg
- row1.Nb_cla_pp_bg_brs

Column

- id
- Session
- Voe_tot
- Voe_tot_f
- Nb_voe_pp
- Nb_voe_pp_internat
- Nb_voe_pp_bg
- Nb_voe_pp_bg_brs
- Nb_voe_pp_bt
- Nb_voe_pp_bt_brs
- Nb_voe_pp_bp
- Nb_voe_pp_bp_brs
- Nb_voe_pp_at
- Nb_voe_pc
- Nb_voe_pc_bg
- Nb_voe_pc_bt
- Nb_voe_pc_bp
- Nb_voe_pc_at
- Nb_cla_pp
- Nb_cla_pc
- Nb_cla_pp_internat
- Nb_cla_pp_pasinternat
- Nb_cla_pp_bg
- Nb_cla_pp_bg_brs

Éditeur de Schéma / Éditeur d'expression

row1

Colonne	Cle	Type	N	Modèle de date (C...	Length	Precision	Défaut	Comment...
id	<input checked="" type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
Session	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		4	0		
Voe_tot	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		5	0		
Voe_tot_f	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		5	0		
Nb_voe_pp	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		5	0		
Nb_voe_pp_internat	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		0	0		

out

Colonne	Cle	Type	N	Modèle de date (C...	Length	Precision	Défaut	Comment...
id	<input checked="" type="checkbox"/>	int	<input type="checkbox"/>		10	0		
session	<input type="checkbox"/>	int	<input type="checkbox"/>		10	0		
voe_tot	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		10	0		
voe_tot_f	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		10	0		
nb_voe_pp_internat	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		100	0		

Appliquer OK Annuler

Remplissage des tables sous PostgreSQL à partir des composantes de Talend

DECISIONNELLE
INFORMATIQUE
STUDIO
TALEND
OPEN

Talend Open Studio for Data Integration (7.3.1.20190924_1038-M2) | PROJET_S3_G11 (Connexion: local)

Fichier Edit View Fenêtre Aide

LOCAL: PROJET_S3_G11

- table_etablissement 0.1
- table_formation_bd 0.1
- table_region 0.1
- table_unites_stat 0.1
- tables_statistique 0.1

Contextes

Code

Modèles SQL

Métadonnées

- Connexions aux bases de données
 - bd_talents3 0.1
- Fichier délimité
 - par2019 0.1
 - parcoursup2019. 0.1
 - parcoursup2018. 0.1
 - unite_statistique 0.1
- Fichier positionnel
- Fichier regex
- Fichier XML
- Fichier Excel
- Fichier Idif

Job table_etablissement 0.1

row1 (Main) tMap_1 out1 (Merge order1) tUnite_1 row3 (Main) row4 (Uniques) tMap_3 QUIT (Main) bd_talents3

parcoursup2018. row2 (Main) tMap_2 out2 (Merge order2)

parcoursup2019.

Designer Code

Outline

- tdBOutput_1 (bd_talents3)
- tFileInputDelimited_1 (parcoursup2019.)
- tFileInputDelimited_2 (parcoursup2018.)
- tMap_1
- tMap_2
- tMap_3
- tUniqRow_1
- tUnite_1

Job/table_etablissement 0.1

Contexts(table_etablissement) Composant Exécuter(Job table_etablissement)

table_etablissement 0.1

Main

table_etablissement	
Extra	user@talend.com 0.1 M m
Stats & Logs	
Version	17/12/20 17:18 02/01/21 15:44

Les tables créées sous PostgreSQL sont belles et bien remplies.

POSTGRES
BASE
DE
DONNEES

127.0.0.1:50225/browser/

pgAdmin Fichier Objet Outils Aidez-moi

Navigateur

- Extensions
- Wrappers de données étrangères
- Langues
- Schémas (1)
 - Publique
 - Collations
 - Domaines
 - Configurations FTS
 - Dictionnaires FTS
 - Analyses FTS
 - Modèles FTS
 - Tables étrangères
 - Les fonctions
 - Vues matérialisées
 - Procédures
 - Séquences
 - Tables (7)
 - académie
 - concours
 - département
 - établissement
 - formation
 - Région
 - unite_statistique
 - Fonctions de déclenchement
 - Les types
 - Vues
 - postgres
 - Rôles de connexion / groupe
 - Espaces de table

Tableau de bord Propriétés SQL Statistiques Dépendances Les personnes à c... bd_talend / pos... public.formation/bd_talend/postgres@PostgreSQL 13

public.formation/bd_talend/postgres@PostgreSQL 13

Éditeur de requête Historique des req...

```
1 SELECT * FROM .formation publique
2 COMMANDER PAR id_formation ASC
```

Sortie de données Explique Notifications

id_formation	capacite	formation_detailee	lien_de_la_formation	formation_tres_detailee	formation_tres_agreee	filere_de_for
[PK] entier	double precision	caractère variant (1000)	caractère variant (1000)	caractère variant (1000)	caractère variant (1000)	caractère vari
1	1	35 Économie et gestion	https://dossier.parcoursup.fr/C...	Double licence Economie et Ge...	Licence	Licence - Dr
2	2	175 Arts plastiques	https://dossier.parcoursup.fr/C...		Licence	Licence - Ar
3	3	30 Langues étrangères appliquées	https://dossier.parcoursup.fr/C...	Anglais - Allemand	Licence	Licence - Ar
4	4	55 Lettres	https://dossier.parcoursup.fr/C...	Lettres modernes	Licence	Licence - Ar
5	5	45 Langues étrangères appliquées	https://dossier.parcoursup.fr/C...	Anglais - Portugais	Licence	Licence - Ar
6	6	210 Économie et gestion	https://dossier.parcoursup.fr/C...		Licence	Licence - Dr
sept	sept	155 Sociologie	https://dossier.parcoursup.fr/C...		Licence	Licence - Sc
8	8	100 Informatique	https://dossier.parcoursup.fr/C...		Licence	Licence - Sc
9	9	45 Sciences de la terre	https://dossier.parcoursup.fr/C...		Licence	Licence - Sc
dix	dix	50 Langues, littératures et civilisati...	https://dossier.parcoursup.fr/C...	Langues, Littératures et Civilisa...	Licence	Licence - Ar
11	11	15 Langues, littératures et civilisati...	https://dossier.parcoursup.fr/C...	Langues, Littératures et Civilisa...	Licence	Licence - Ar
12	12	230 Langues étrangères appliquées	https://dossier.parcoursup.fr/C...	Langues Etrangères Appliquée...	Licence	Licence - Ar
13	13	63 Langues étrangères appliquées	https://dossier.parcoursup.fr/C...	Langues Etrangères Appliquée...	Licence	Licence - Ar

messages

Exécutez avec succès. Durée totale de affectées.

Les requêtes pour répondre à certaines questions liées a mon établissement

requêtes
Aux questions
LES
répondre

Le but de ces requêtes est vérifié si les tables ont été bien remplies et si la base de données sous PostgreSQL est correcte.

a) Quel est le nombre de filles candidates en STID – Carcassonne – 2019.

The screenshot shows the PgAdmin interface with a PostgreSQL database named 'bd_talend'. The left sidebar shows the database structure, including tables like 'nom_etablissement', 'formation_detaillee', 'nombre_filles_candidates', and 'session'. The main window displays a SQL query in the 'Éditeur de requête' (Query Editor) and its results in the 'Sortie de données' (Query Results) pane.

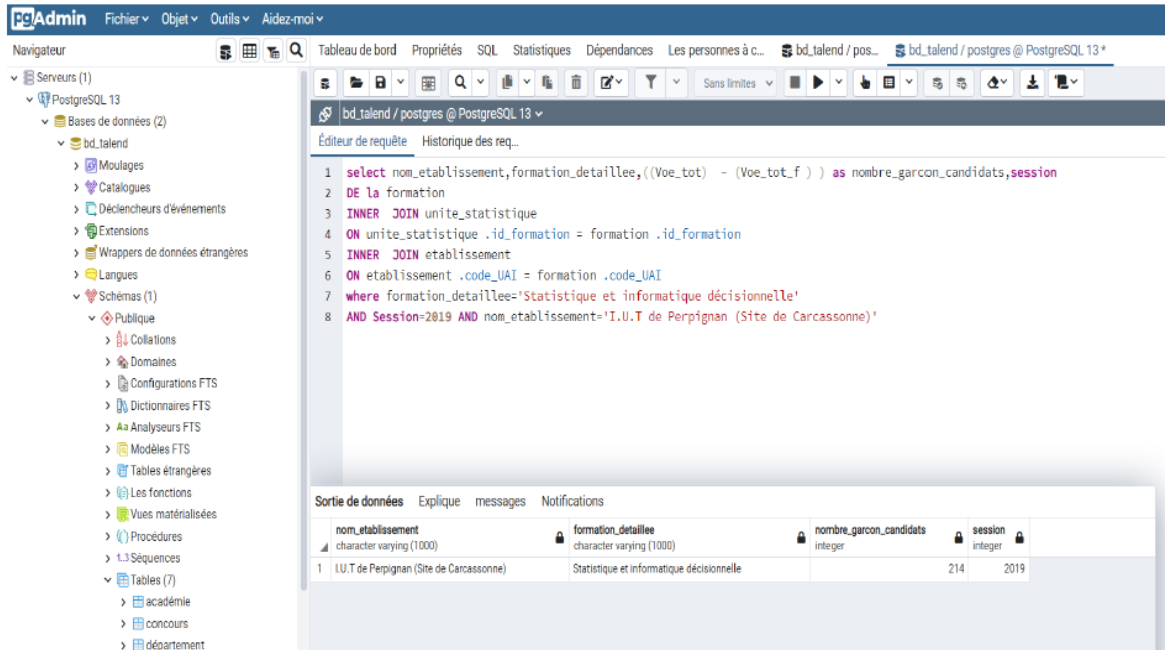
Query:

```
1 select nom_etablissement, formation_detaillee, Voe_tot_f as nombre_filles_candidates, session
2 FROM formation
3 INNER JOIN unite_statistique
4 ON unite_statistique.id_formation=formation.id_formation
5 INNER JOIN etablissement
6 ON etablissement.code_UAI=formation.code_UAI
7 where formation_detaillee='Statistique et informatique décisionnelle'
8 AND Session=2019 AND nom_etablissement='I.U.T de Perpignan (Site de Carcassonne)'
```

Query Results:

nom_etablissement	formation_detaillee	nombre_filles_candidates	session
I.U.T de Perpignan (Site de Carcassonne)	Statistique et informatique décisionnelle	633	2019

b) Quel est le nombre de garçons candidats en STID – Carcassonne – 2019.



The screenshot shows the PGAdmin interface with a SQL query editor and a results table. The query is as follows:

```

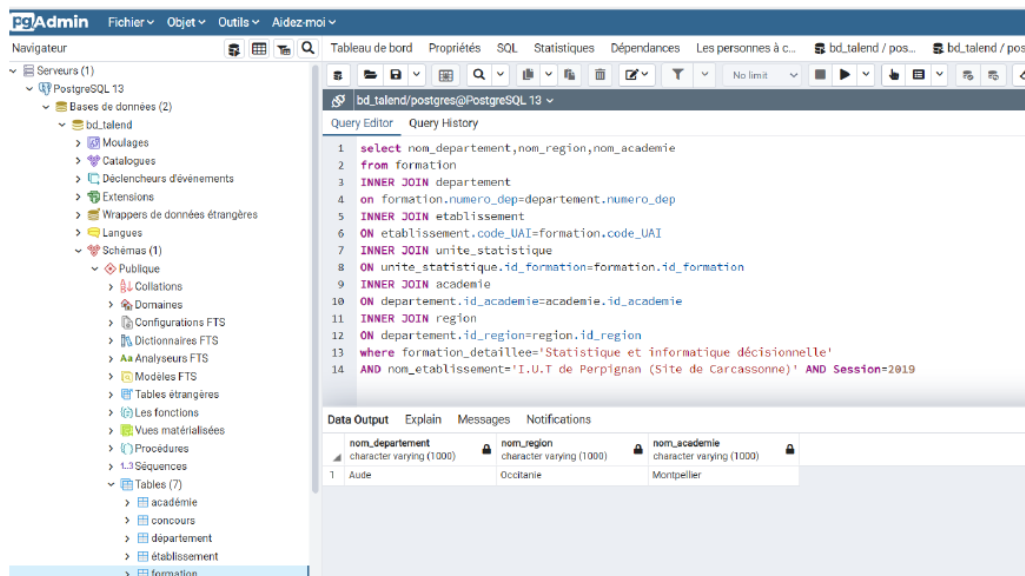
1 select nom_etablissement, formation_detaillee, ((Voe_tot) - (Voe_tot_f)) as nombre_garcon_candidats, session
2 DE la formation
3 INNER JOIN unite_statistique
4 ON unite_statistique.id_formation = formation.id_formation
5 INNER JOIN etablissement
6 ON etablissement.code_UAI = formation.code_UAI
7 where formation_detaillee='Statistique et informatique décisionnelle'
8 AND Session=2019 AND nom_etablissement='I.U.T de Perpignan (Site de Carcassonne)'

```

The results table shows the following data:

nom_etablissement	formation_detaillee	nombre_garcon_candidats	session
I.U.T de Perpignan (Site de Carcassonne)	Statistique et informatique décisionnelle	214	2019

c) Quel est le département, l'académie et la région du département STID – Carcassonne.



The screenshot shows the PGAdmin interface with a SQL query editor and a results table. The query is as follows:

```

1 select nom_departement, nom_region, nom_academie
2 from formation
3 INNER JOIN departement
4 on formation.numero_dep=departement.numero_dep
5 INNER JOIN etablissement
6 ON etablissement.code_UAI=formation.code_UAI
7 INNER JOIN unite_statistique
8 ON unite_statistique.id_formation=formation.id_formation
9 INNER JOIN academie
10 ON departement.id_academie=academie.id_academie
11 INNER JOIN region
12 ON departement.id_region=region.id_region
13 where formation_detaillee='Statistique et informatique décisionnelle'
14 AND nom_etablissement='I.U.T de Perpignan (Site de Carcassonne)' AND Session=2019

```

The results table shows the following data:

nom_departement	nom_region	nom_academie
Aude	Occitanie	Montpellier

d) Quelle est la formation ayant eu le plus de candidats en 2019.

The screenshot shows the pgAdmin 4 web interface. On the left, the 'Browser' pane displays a tree view of the database structure, with 'formation' selected under the 'public' schema. The main pane is divided into two sections: the 'Query Editor' and the 'Data Output'.

Query Editor: The SQL query is as follows:

```
1 SELECT Session, formation_detailee, MAX(Voe_tot) as nombre_candidat
2 FROM unite_statistique
3 INNER JOIN formation
4 ON unite_statistique.id_formation=formation.id_formation
5 GROUP BY formation_detailee, Session
6 HAVING Session=2019
7 ORDER BY MAX(Voe_tot) DESC
8 LIMIT 1
9
```

Data Output: The results are displayed in a table with the following columns: session, formation_detailee, and nombre_candidat. The data is as follows:

session	formation_detailee	nombre_candidat
2019	Systèmes numériques - Option informatique et réseaux	16899