

## SAE 2.04 - Exploitation de BDD

### Sommaire

Exercice 1 Exercice 2 Exercice 3

### Exercice 1

#### Q1.

1. Il y en a 13870 : `wc -l tmp/data.csv` permet de compter le nombre de lignes (-l)
2. Une ligne représente une formation dans un établissement dont l'intitulé de chaque colonne est inscrit dans la première ligne, chaque donnée étant séparée par un séparateur : “;”
3. Il y a 118 colonnes : on isole la première ligne et on remplace chaque séparateur par un retour à la ligne, puis on compte le nombre de lignes  
`head -n1 tmp/data.csv | tr ";" "\n" | wc -l`
4. La colonne qui identifie un établissement est la colonne n°4, nommée “Etablissement”
5. La colonne qui identifie une formation est la colonne n°10, “Filière de formation”
6. Notre BUT informatique, celui de Villeneuve-d’Ascq, apparaît dans une ligne : `grep "BUT - Informatique" nom_du_fichier.csv | grep "Villeneuve-d'Ascq" | wc -l`
7. La colonne n° 6 identifie un département, elle se nomme “Département de l’établissement”
8. On envisage d’importer ces données dans une grande table temporaire avec peu de contraintes, pour pouvoir ensuite remplir les tables.
9. Il y a de nombreuses données de calcul, de nombreuses cases vides et beaucoup de redondance

#### Q2.

1. Le fichier est dico.xls et a été généré grâce à cette commande via un fichier java : `java ./javaBoucle/BoucleDico.java > dico.xls`  
&& `head -n 3 ./tmp/data.csv >>dico.xls` Voici le code source du fichier java :

```
public class BoucleDico{
    public static
    void main (String[] args){
        String result ="";
        for(int i=1;i<=118-1;i++){
            result += "n"
            +i +";";
        }
        result += "n"+118+"n";
        System.out.print(result);
    }
}
```

2. Le fichier boucle pour créer la table se trouve dans javaBoucle/BoucleImport.java  
 , la table se situe dans le fichier Script\_creation.sql

```
CREATE temp TABLE import (  
  n1 NUMERIC(4) default null,n2 text default null,n3 char(8) default null,  
  n4 text default null,n5 char(3) default null,n6 text default null,  
  n7 text default null,n8 text default null,n9 text default null,  
  n10 text default null,n11 text default null,n12 text default null,  
  n13 text default null,n14 text default null,n15 text default null,  
  n16 text default null,n17 text default null,  
  n18 int default null,n19 int default null,n20 int default null,  
  n21 int default null,n22 text default null,n23 int default null,  
  n24 int default null,n25 int default null,  
  n26 int default null,n27 int default null,n28 int default null,  
  n29 int default null,n30 int default null,n31 int default null,  
  n32 int default null,n33 int default null,  
  n34 int default null,n35 int default null,n36 int default null,  
  n37 text default null,n38 text default null,n39 int default null,  
  n40 int default null,n41 int default null,  
  n42 int default null,n43 int default null,n44 int default null,  
  n45 int default null,n46 int default null,n47 int default null,  
  n48 int default null,n49 int default null,  
  n50 int default null,n51 real default null,n52 real default null,  
  n53 real default null,n54 text default null,n55 int default null,  
  n56 int default null,n57 int default null,  
  n58 int default null,n59 int default null,n60 int default null,  
  n61 int default null,n62 int default null,n63 int default null,  
  n64 int default null,n65 int default null,  
  n66 real default null,n67 real default null,n68 int default null,  
  n69 int default null,n70 text default null,n71 text default null,  
  n72 int default null,n73 int default null,  
  n74 real default null,n75 real default null,n76 real default null,  
  n77 real default null,n78 real default null,n79 real default null,  
  n80 real default null,n81 real default null,  
  n82 real default null,n83 real default null,n84 real default null,  
  n85 real default null,n86 real default null,n87 real default null,  
  n88 real default null,n89 real default null,  
  n90 real default null,n91 real default null,n92 real default null,  
  n93 real default null,n94 real default null,n95 real default null,  
  n96 real default null,n97 real default null,  
  n98 real default null,n99 real default null,n100 real default null,  
  n101 real default null,n102 text default null,n103 text default null,  
  n104 text default null,n105 text default null,  
  n106 text default null,n107 text default null,n108 text default null,  
  n109 text default null,n110 text default null,n111 text default null,  
  n112 text default null,n113 text default null,
```

```
n114 real default null,n115 real default null,n116 real default null,
n117 text default null,n118 text );
```

5.
  - Il y a 13869 formations gérés par ParcourSup : `select count(*) from import;`
  - Il y a 3602 établissements gérés par ParcourSup : `select count(distinct n4) from import;`
  - Il y a 124 formations pour l'université de Lille : `select count(*) from import where n4 LIKE '%Université de Lille%';`
  - Il y a 10 formations pour notre IUT : `select count(*) from import where n4 LIKE '%Institut universitaire de technologie de Lille - Université de Lille%';`
  - Le code du BUT informatique de l'université de Lille est 6888 : `select n110 from import where n4 = 'Institut universitaire de technologie de Lille - Université de Lille' and n10 = 'BUT - Informatique';`
  - Voici 5 colonnes contenant des valeurs null : n23, n38, n39, n55, n71

## Exercice 2

### Q1.

1. Voici le mcd correspondant à la base de donnée :
2. Exécution de la création de la base de donnée

Pour créer la base de donnée, exécutez les commandes suivantes dans psql, dans le dossier SAE2.04 : \i ./parcourssup.sql

### Q2.

1. On peut utiliser la commande `wc -c ./tmp/data.csv` qui nous indique la taille en octet du fichier, 12 423 586 octets
2. On utilise la commande `select pg_total_relation_size('import') as importsize;` ce qui nous donne ça : 18 644 992 octets.

## Exercice 3

### Q1.

La requête à exécuter est la suivante : `select n57,(n56-n58-n59) from import;` car  $n57 = n56 - n58 - n59$  En effet n56 est le total des néo-bacheliers admis et n57, n58, et n59 sont les néo-bacheliers admis mais triés par le type de bac qu'ils ont obtenus

**Q2.**

La requête `select count(*) from import where n57<>(n56-n58-n59);` permet de compter le nombre de lignes où cette relation est fausse, lorsqu'elle renvoie 0, c'est qu'il n'y en a aucune

**Q3.**

La requête `select n74,ROUND((n51/n47)*100) from import where n47 <> 0;` permet de calculer n74 à partir de n51 et n47 car  $74 = n51/n47$ . n51 correspond aux effectifs des admis ayant reçu leur proposition d'admission à l'ouverture de la procédure principale et n47 aux effectifs totaux des admis.

**Q4.**

La requête `select count(*) from import where n47 <> 0 and n74<>ROUND((n51/n47)*100);` permet de compter le nombre de lignes où cette relation est fausse, lorsqu'elle renvoie 0, c'est qu'il n'y en a aucune. Il y en a quelques lignes pour lesquelles ça ne fonctionne pas(mais très peu)

**Q5.**

`select n76,ROUND((n53/n47)*100) from import where n47 <> 0;` permet de calculer n76 à partir de n53 et n47 car  $n76 = n53/n47$ . C'est les effectifs d'admis ayant reçu leur proposition d'admission avant la fin de la procédure principale/ les effectifs d'admis. Ces résultats sont exacts lorsqu'on arrandi à l'entier.

**Q6.**

La requête `select round((f.effectif_admis_recu_proposition_avant_fin/i.effectif_admis)*100) from formations as f, inscrits as i where f.fno = i.fno and i.effectif_admis <> 0;` correspond à la même requête que ci-dessus mais dans les nouvelles tables ventilées.

**Q7.**

`select n81,round((n55/n56)*100) from import where n56 <> 0;` car  $n81 = n55/n56$ , n81 correspond au pourcentage d'admis néo bacheliers boursiers, n56 correspond aux effectifs des admis néo-bacheliers et n55 aux effectifs des admis boursiers-néo bacheliers.

**Q8.**

Voici la requête qui correspond à la même requête que ci-dessus mais dans les nouvelles tables ventilées.

`select round((effectif_admis_boursiers_neobacheliers/effectif_admis)*100)`

```
from formations as f, inscrits as i
where f.fno = i.fno and effectif_admis <> 0 and (type_inscrits = 'généraux' OR type_inscrits
```