

# Projet Bases de données et Perl

## M1 Bioinformatique

### 2017/2018

Il s'agit de développer une application en PERL qui accède à une base de données sous Postgres permettant de gérer une chaîne d'hôtels. Un exemple de données dont on peut se servir peut être récupéré à l'adresse

<http://www.labri.fr/~maabout/Hotels.csv>

C'est un fichier CSV. Il représente un extrait des données manipulées. Comme on peut le voir (l'ouvrir avec un éditeur de texte), il s'agit d'un tableau qui contient des informations sur des hôtels, des chambres dans ces hôtels ainsi que des clients qui ont fait des réservations dans les hôtels de la chaîne.

Afin de simplifier le projet, nous supposons les hypothèses, plus ou moins réalistes, suivantes :

- Deux hôtels sont forcément situés dans des villes différentes.
- Deux personnes ont forcément des noms différents (la moins réaliste des hypothèses).
- Une même personne peut effectuer plusieurs réservations pour une même période ou des périodes qui se chevauchent.
- Chaque réservation est identifiée par un numéro.
- Un téléphone n'est associé qu'à une seule personne et une personne n'a qu'un seul téléphone.
- Deux chambres dans le même hôtel ont forcément des numéros différents.
- Le prix d'une chambre dépend de la saison. Pour simplifier, on considère que la haute saison va de 31/05 au 30/09. Le reste de l'année est considéré comme basse saison. Si la période de réservation chevauche les deux périodes, l'hôtel facture en utilisant le prix le moins cher.

**Structure du fichier CSV :** Les colonnes du fichier sont : Hotel, Gérant, Etoiles, NumChambre, TypeCouchage, PrixBasseSaison, PrixHauteSaison, NumRésa, DébutResa, FinResa, NomClient, PhoneClient

#### Création de la base de données :

- Dans un premier temps, il s'agit d'identifier les dépendances fonctionnelles qu'on peut définir à partir des attributs constituant cette table. A priori, les hypothèses ci-dessus sont suffisantes pour les identifier mais si vous jugez qu'il vous manque des informations, vous pouvez vous même en émettre d'autres.
- En se basant sur ces dépendances, concevoir un schéma d'une base de données Postgres, c'est à dire un ensemble de tables, destinées à stocker ces informations. Toutes les tables devant être en 3ème forme normale. Chaque table devra avoir une clé primaire. Le cas échéant, on veillera à définir les clés étrangères
- Une fois que la base de données a été créée, on va y insérer des enregistrements qu'on extrait à partir du fichier. Pour ce faire, on utilise Perl : étant donnée une table dans laquelle on veut insérer des enregistrements, on va d'abord extraire l'enregistrement à partir du fichier CSV (en utilisant Perl) puis on insère cette donnée dans la table en utilisant une commande SQL.

**Manipulation de la base :** Une fois la base créée, on voudra la manipuler en l'interrogeant et/ou modifiant son contenu. L'objectif est que l'utilisateur n'ait pas besoin de connaître/utiliser SQL. Il faudra donc développer une application à base de "menus". L'utilisateur n'aura qu'à sélectionner l'opération qu'il entend effectuer sur la base de données, éventuellement en précisant certains paramètres, et l'application se chargera d'exécuter la requête correspondante en l'envoyant au SGBD.

**Scripts à réaliser :** Nous allons distinguer 3 types de manipulation, interrogation, mise à jours et statistiques sur les données. Pour chacune des opérations suivantes, écrire un script Perl permettant de la réaliser

#### Interrogation

- a) Afficher le nom des gérants
- b) Afficher le nombre de gérants
- c) Afficher le nom des personnes qui gèrent au moins deux hôtels.
- d) Etant donnée une période (date de début et date de fin), afficher le nom des hôtels où il y a au moins une chambre libre.

#### Mise à jour

- e) Ajouter une chambre à un hôtel
- f) Modifier le nom du gérant d'un hôtel
- g) Annuler une réservation. **Attention :** votre script doit tenir compte du fait que l'on ne puisse pas annuler une réservation moins de deux jours avant son début.
- h) Ajouter une réservation.

#### Statistiques

- i) Etant donné le paramètre  $h$  qui désigne un nom d'hôtel, afficher le taux d'occupation de ce derniers durant les 7 derniers jours.
- j) Pour chaque hôtel, afficher son nom ainsi que son taux d'occupation sur 7 derniers jours.
- k) Afficher le ou les hôtels qui avaient le plus grand taux d'occupation durant les 7 derniers jours.

**Création de fichiers résultat :** Les résultats des requêtes vont dans un premier temps être affichés à l'écran. Dans certains cas on peut avoir besoin d'enregistrer ce résultat dans un fichier qu'on peut consulter par la suite. Ainsi, pour les 3 dernières requêtes ci-dessus (b, c et j), écrire un script qui permette de stocker le résultat dans un fichier `html` qu'on pourra consulter via un navigateur. On veillera à ce que le résultat soit présenté sous la forme d'un tableau.

**Perl vs SQL :** Le travail précédent a permis de constater la puissance du langage Perl pour la mise en forme du résultat des requêtes. Par ailleurs, toutes les requêtes précédentes peuvent être évaluées directement sur le fichier texte CSV par Perl sans recourir à la base. Cependant, il est souvent plus facile d'utiliser SQL que Perl pour ce type de manipulation. A titre d'illustration,

- écrire un script Perl qui prend en entrée le fichier CSV et qui, sans accéder à la base et donc sans utiliser SQL, affiche pour chaque hôtel, son nom ainsi que le nombre de ses chambres. Comparer le script Perl et la requête SQL correspondante.

**Libre à vous d'ajouter d'autres opérations.**

#### **Travail :**

*Chaque projet est à réaliser par un binôme ou trinôme. Chaque groupe fera une démonstration de son travail. Aussi, il est demandé de rédiger un rapport, de 3 pages environ, expliquant les différents choix qui ont été pris durant la réalisation. Vous avez 3 semaines pour finir le projet.*

## Annexes

### 1 Utilisation des modules Perl

Pour chaque module décrit ici, il existe beaucoup de documentations plus ou moins obsolètes sur Internet. N'oubliez pas que "Google" est votre ami !

## 1.1 Utilisation basique du pilote DBD::Pg

Soit une relation  $test(a, b)$  où  $a$  et  $b$  sont des entiers.

```
#!/usr/bin/perl
use strict;
use DBI;
# connect
my $dbh = DBI->connect("DBI:Pg:dbname=bpinaud;host=dbserver",
    "nom-de-votre-base", "", {'RaiseError' => 1});

# execute INSERT query
my $rows = $dbh->do("INSERT INTO test VALUES (5,3)");
print "$rows row(s) affected\n";
# execute SELECT query
my $sth = $dbh->prepare("SELECT * FROM test");
my $num = $sth->execute();
print "query returned $num rows\n";
# iterate through resultset
while(my $ref = $sth->fetchrow_hashref()) {
    print "$ref->{'a'} $ref->{'b'}\n";
}

$sth->finish;
$dbh->disconnect();
```

## 1.2 Utilisation basique du module Text::CSV

Un tuotiel es disponible à [http://perlmeme.org/tutorials/parsing\\_csv.html](http://perlmeme.org/tutorials/parsing_csv.html)

## 1.3 Exemple d'un fichier html avec un tableau

Un tel exemple peut être récupéré à <http://www.labri.fr/~maabout/tableau.html>