

## TD5 : gestion et utilisation d'une base de données SQL en Python avec le module sqlite

L'objectif de ce TD est d'utiliser un module Python permettant de créer et d'utiliser une base de données MySQL en Python. Pour permettre un travail en local, nous utiliserons le module `sqlite3` qui permet de gérer des bases de données SQL locale. L'utilisation d'un serveur MySQL tel que `pedago.uhb.fr` à l'université Rennes 2 est possible avec d'autre module, mais malheureusement ce serveur n'est accessible ni depuis les salles du réseau MASS, ni depuis l'extérieur de l'université.

Télécharger l'archive *TD5.zip* depuis l'espace-cours Coursus et décompresser cette archive dans un répertoire personnel.

### Exercice 1 : création de la base

*Pour cet exercice vous travaillerez dans le fichier `td5_creation.py`.*

On souhaite créer une base de données d'événements locaux à partir de données publiques disponibles à l'adresse <https://datainfoleale.opendatasoft.com/pages/home/> .

Une inscription (gratuite) sur le site permet de télécharger l'intégralité des données des événements pour l'année 2019 et le début de l'année 2020 depuis l'onglet « Export » (plus de 576 000 événements au total) .

Pour ce TD, nous travaillerons sur un extrait des 5000 premiers événements issus de ces données et répertoriés dans un fichier au format CSV (fichier *agenda\_culturel.csv* de l'archive).

Pour chaque événement, on souhaite mémoriser les informations suivantes : son titre, son type, sa date de début, sa date de fin, l'organisme qui le propose, et sa localisation : la commune où il est programmé, et des coordonnées GPS du lieu précis de l'événement. On souhaite également mémoriser des informations tarifaires optionnelles sur l'événement (tarif général, tarif réduit, une mention « événement gratuit »), une mention « événement complet », des indications d'âges du public visé (âges minimum et maximum), des informations d'accessibilité aux personnes handicapées et une description facultative (concaténation de plusieurs champs de texte présents dans les données : « texte de début », « texte de milieu », « texte de fin », « texte libre court » et « texte libre long »).

Le type de l'événement est choisi parmi une liste qui doit pouvoir être modifiée dans la base, et il est associé à un thème : « Loisir » ou « Pratique ».

Pour chaque organisme présents dans la base, on mémorisera son nom et son type.

Enfin pour chaque commune accueillant un événement, on veut pouvoir retrouver son nom, son département, son code INSEE et son code postal.

### Questions préliminaire :

1. Ouvrir le fichier *agenda\_culturel.csv* dans LibreOffice Calc (ou un autre tableur). Définir le mode d'adressage L1C1 : menu "Outils" → "Options" → "LibreOffice Calc" → "Formules" puis changer la valeur de la liste déroulante "Syntaxe de formule" dans la section "Options de formule". Ce mode d'adressage permet d'avoir des numéros pour référencer les colonnes,

à la place de lettres (mode d'adressage A1 par défaut). Vous pouvez aussi placer un Autofiltre sur la ligne des intitulés, ça peut permettre notamment de voir si une colonne possède des cases vides ou non.

2. Lire le fichier `agenda_culturel.csv` à l'aide du module `csv`<sup>1</sup> de Python et placer les données dans une liste Python `dataCSV`

### Question 1 :

En vous aidant du schéma relationnel donné dans l'archive, répertoriant les tables de la base avec leurs attributs, les clés primaires, étrangères et les liens entre ces clés, créer ces tables (vides pour l'instant) dans un fichier de base de données `evenements.sqlite` à l'aide du module Python `sqlite3`. Pour chaque attribut, vous devrez déterminer son type et sa taille maximale, et le fait qu'il soit optionnel ou non.

*Remarque : le module `sqlite3` ne permet pas de gérer le type "énumération", on choisira donc un type de base adapté aux données.*

- 1.1. Placer en premier les instructions permettant de supprimer les tables si elles existent (utile pour repartir à zéro lorsqu'on fait plusieurs essais). On veillera à supprimer en premier la table qui possède des clés étrangères.
- 1.2. Créer ensuite les tables ne possédant pas de clé étrangère.
- 1.3. Créer en dernier la table qui possède des clés étrangères.

### Question 2 :

Remplir les tables créées grâce aux données contenues dans le fichier CSV. Vous pourrez vérifier le contenu de vos tables en utilisant un outil tel que <https://sqliteonline.com/> ou <https://inloop.github.io/sqlite-viewer/> (attention il faudra recharger le fichier de base de données dans l'interface à chaque fois que vous le modifierez).

*Remarque : avec `sqlite3`, on peut insérer au maximum 900 valeurs dans une même requête d'insertion. On effectuera donc ici **une requête d'insertion pour chaque ligne à insérer** plutôt qu'une seule requête insérant l'ensemble des lignes. Avec un serveur MySQL, cette limite est bien plus élevée (65535 valeurs) et on préférera donc limiter le nombre de requêtes d'insertion pour réduire le temps de traitement.*

- 2.1. Remplir d'abord les tables ne possédant pas de clé étrangères

*Indication : on se rappellera que chaque ligne d'une table de base de données doit être unique*

- 2.2. Remplir ensuite la table possédant des clés étrangères

## Exercice 2 : utilisation des données de la base

*Pour cet exercice vous travaillerez dans le fichier `td5_interrogation.py`.*

En utilisant les données de la base et éventuellement d'autres ressources, trouver :

### Question 1 :

Les titres, dates de début et de fin des événements de type « Exposition, musée » qui se sont

---

<sup>1</sup> <https://docs.python.org/fr/3/library/csv.html>

déroulés (totalement ou partiellement) en février 2019.

**Question 2 :**

Les titres, dates de début et de fin, et la commune des événements affichés complets

**Question 3 :**

Les titres, dates de début et de fin, et la description des événements proposés par des organismes de type « Etablissement scolaire, universitaire » en Ile-et-Vilaine, avec le nom de l'organisme

**Question 4 :**

Les titres, date de début, date de fin et département des événements non accessibles aux moins de 12 ans qui se déroulent en dehors de la Bretagne (c'est à dire hors des départements bretons : 22, 29, 35 56).

**Question 5 :**

Les titres, date de début, date de fin et type (nom du type) des événements accessibles aux personnes ayant un handicap « Moteur », proposés en Ile-et-Vilaine par des associations, avec le nom de l'association.

**Question 6 :** Les titres, dates de début et de fin, et les coordonnées GPS de tous les événements gratuits à Rennes dans le thème « Loisir » qui ont lieu après le 1<sup>er</sup> mars 2019 et qui se déroulent à moins de 45 minutes **à pied** de votre domicile à Rennes (ou du campus de Villejean, « Place du Recteur Henri Le Moal, 35000 Rennes »). Vous pouvez utiliser le site <https://www.coordonnees-gps.fr/> pour trouver les coordonnées GPS correspondantes.

**Question 7 :**

Les titres, dates de début et de fin et la commune des événements de type « Concert » gratuits ou dont le tarif réduit est inférieur ou égal à 10 euros et situés à moins de 1h de voiture de votre domicile à Rennes (ou du campus de Villejean).