

## TP : Élection présidentielle

### 1 Premier tour de l'élection présidentielle

Dimanche dernier nous avons voté. C'est une excellente occasion de réviser quelques notions de bases de données.

Vous pouvez récupérer la base de données `sqlite presidentialles.db` qui contient les résultats des dernières élections présidentielles ainsi que des informations sur le salaire moyen dans un grand nombre de communes de France. Ce fichier étant relativement volumineux, il faut le copier sur le disque dur de l'ordinateur où vous vous trouvez (par exemple dans le répertoire `c:\temp\`), l'accès sera plus rapide.

### 2 SQLite et python

Nous allons analyser les résultats graphiquement. Pour cela nous accéderons à la base de données à l'aide de python qui nous permettra d'effectuer un traitement un peu plus évolué des données.

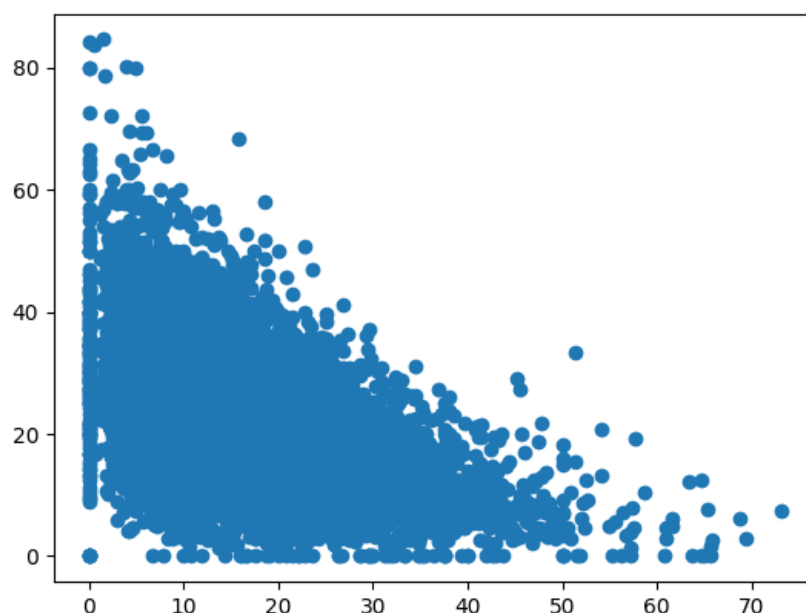
Vous pourrez vous baser sur l'exemple suivant qui dessine sur un graphique le pourcentage de suffrage en faveur de François Fillon, en fonction de celui des suffrages en faveur de Jean-Luc Mélenchon.

```
import sqlite3
from matplotlib import pyplot as plt

db = sqlite3.connect("presidentielles.db")
cur = db.cursor()
cur.execute("SELECT jm_pce, ff_pce from pres2017_t1")
data = cur.fetchall()
x, y = list(zip(*data))
plt.plot(x,y,'o')
plt.show()

db.close()
```

Il faut prendre soin de mettre le fichier `presidentielles.db` dans le répertoire de travail de pyzo, ou de remplacer le nom de fichier par son chemin d'accès complet, par exemple `"c:\temp\presidentielles.db"`. On obtient alors le graphique suivant :



On voit sur ce graphique qu'il y a une anti-corrélation claire entre les votes pour les deux candidats. Dans les villes où l'on vote beaucoup pour Jean-Luc Mélenchon, on vote peu pour François Fillon et inversement.

### 3 À vous de jouer

- Modifier le code précédent pour n’afficher que les communes de plus de 1000 inscrits. On évite ainsi les résultats très fluctuants des communes avec très peu d’inscrits.
- Faites des tests avec les autres candidats. Lesquels sont corrélés, anti-corrélés ?
- Tester les corrélations avec l’abstention, les votes blancs ou nuls.
- Tracer le vote pour les candidats en fonction du nombre d’inscrits. Est-ce que l’on vote plus pour certains candidats dans les grandes villes ? On pourra utiliser une échelle logarithmique pour l’axe représentant les inscrits avec la commande :

```
plt.xscale('log')
```

- Utiliser la base de données pour trouver de quelle ville de plus de 1000 inscrits Nicolas Dupont-Aignan est élu. (indice : dans la ville où il est élu, le gens votent pour lui !)
- Le fichier comporte une table `salaires` qui donne le salaire moyen dans les plus grandes communes. Utilisez cette table pour tracer le vote pour certains candidats en fonction du salaire moyen.