

## Thème 3 : Les Données Structurées

### Activité 3 (traitement des données) **correction**

#### 1 - Traitement des données à l'aide d'un tableur.

##### b) Exploration du fichier des passagers.

Dans le fichier titanic.csv , le séparateur utilisé est le point-virgule.

Il y a 6 descripteurs : classe ; survie ; nom ; sexe ; âge ; tarif.

Les valeurs du descripteur 'classe' sont : 1, 2 et 3.

Exemple de la ligne 10 : la passagère (sexe 2) nommée Mrs. Edward Dale Appleton (Charlotte Lamson) a survécu au naufrage (survie 1), elle avait voyagé en classe 1 au tarif de 51, et était âgée de 53 ans.

##### c) Traitement des données.

L'âge moyen des passagers est de 29,9 ans.

Le tarif moyen était de 33,4 alors que la moitié des passagers avait payé moins de 14.

Le tarif le plus élevé était de 512.

La moyenne du descripteur 'survie' est 0,38 ; ce qui signifie que moins de la moitié des passagers ont survécu (38%).

Celle du descripteur 'sexe' est de 1,36 ; ce qui signifie que les passagers étaient majoritairement des hommes (moyenne < 1,5).

En utilisant la fonction donnée, on retrouve que pour un total de 500 passagers ayant survécu, 200 étaient en classe 1, 119 en classe 2, 181 en classe 3.

A priori, la classe sociale n'avait pas une grande influence.

339 femmes ont survécu contre 161 hommes, le film semble donc représenter la réalité.

#### 2 - Traitement de données à l'aide du langage de programmation Python.

##### a) Exploration du fichier.

Dans ce fichier, il y a 64924 établissements.

## b) Traitement des données.

1 - Ouvrir Pyzo et copier/coller le code ci-dessous. Attention à remettre les tabulations éventuellement perdues par recopie.  
Enregistrer votre programme dans votre dossier au nom de **etablisements\_scolaires**.

*Programme python 1.*

```
import csv

# chargement du fichier scolaire.csv
with open('scolaire.csv','r', encoding='utf-8') as fichier:
    premiere_ligne = fichier.readline()
    dialecte_fichier_csv = csv.Sniffer().sniff(premiere_ligne)
    data_lignes = list(csv.reader(fichier, dialect=diaclecte_fichier_csv))

# on met le tout dans une liste
liste=[]
for ligne in data_lignes:
    liste.append(ligne)

print(liste[8])
```

*Réponse python 1.*

```
['0311694Y', 'Ecole maternelle publique Lamartine', 'ECOLE MATERNELLE PUBLIQUE', 'LAMARTINE',
'Public', 'Allée de Savoie', '', '', '31770', 'COLOMIERS', 'Colomiers', '565680.9', '6279744.6',
'EPSG:2154', '43.60402300266715', '1.336759818338252', 'Parfaite', 'Rue', '101', 'ECOLE
MATERNELLE', '1', 'OUVERT', '31', '76', '16', '31149', 'Haute-Garonne', 'Occitanie', 'Toulouse',
'43.6040230027', '1.33675981834', '', '', '', '', '', '']
```

2 - Si on change l'indice de la dernière ligne du programme (le 8 de "print(liste[8])" ), on obtient d'autres établissements.

Si on ajoute ce bout de code à la suite du précédent programme :

```
For i in liste:
    if (i[4]=="Privé"):
        print(i[1])
```

on obtient la liste de tous les établissements privés de France !

3 - A l'aide de l'opérateur '*and*', compléter l'instruction conditionnelle '*if*' afin que le programme affiche le lycée Beaulieu et son adresse.

*Programme python 2.*

```
import csv

# chargement du fichier scolaire.csv
with open('scolaire.csv','r', encoding='utf-8') as fichier:
    premiere_ligne = fichier.readline()
    dialecte_fichier_csv = csv.Sniffer().sniff(premiere_ligne)
    data_lignes = list(csv.reader(fichier, dialect=diaclecte_fichier_csv))

# on met le tout dans une liste
```

```

liste=[]
for ligne in data_lignes:
    liste.append(ligne)

#print(liste[15])

for i in liste:
    if (i[4]=="Privé" and i[10]=="Cognac"):
        print()
        print(i[1])
        print(i[5])
        print(i[8])
        print(i[9])
        print(i[10])

```

### *Réponse python 2 :*

Lycée général et technologique privé de Beaulieu  
 23 place BEAULIEU  
 16105  
 COGNAC CEDEX  
 Cognac

Collège privé Saint Joseph  
 25 rue MAGDELEINE  
 16100  
 COGNAC  
 Cognac

Ecole primaire privée Sainte-Colette-La Providence  
 90 rue de la Providence  
 16100  
 COGNAC  
 Cognac

4 - Modifier le programme afin qu'il demande à l'utilisateur d'entrer le nom d'une ville et qu'il affiche tous les établissements de cette ville.

Modifier le programme afin qu'il renvoie le nombre d'établissements d'une ville choisie par l'utilisateur.

### *Programme python 3.*

```

import csv

# chargement du fichier scolaire.csv
with open('scolaire.csv','r', encoding='utf-8') as fichier:
    premiere_ligne = fichier.readline()
    dialecte_fichier_csv = csv.Sniffer().sniff(premiere_ligne)
    data_lignes = list(csv.reader(fichier, dialect=dialecte_fichier_csv))

# on met le tout dans une liste
liste=[]
for ligne in data_lignes:
    liste.append(ligne)

ville=input('Veuillez entrer le nom d une ville (avec une majuscule):')

k=0
for i in liste:
    if i[10]==ville:
        k=k+1
        print(i[1]) # nom de l'établissement

```

```
print('Nous avons trouvé', k , 'établissements' )
```

*Réponse python 3 :*

Veillez entrer le nom d une ville (avec une majuscule):Cognac

Collège Felix Gaillard

Lycée général et technologique privé de Beaulieu

Ecole élémentaire Paul Bert

Ecole primaire Jules Michelet

Collège Claude Boucher

Ecole primaire Anatole France

Ecole primaire Victor Hugo

Collège privé Saint Joseph

Ecole élémentaire Pierre et Marie Curie

Ecole maternelle Jean Macé

Section d'enseignement général et professionnel adapté du Collège Claude Boucher

Ecole maternelle Saint-Exupéry

Ecole maternelle Les Borderies

Lycée polyvalent Jean Monnet

Collège Elisée Mousnier

Section d'enseignement général et professionnel adapté du Collège Felix Gaillard

Ecole maternelle Pauline Kergomard

Ecole primaire privée Sainte-Colette-La Providence

Section d'enseignement professionnel du Lycée Jean Monnet

Lycée professionnel Louis Delage

Nous avons trouvé 20 établissements