

ENSAE TD noté, mars 2013, rattrage

Le programme construit au fur et à mesure des questions devra être imprimé à la fin du TD et rendu au chargé de TD. Les réponses autres que des parties de programme seront insérées sous forme de commentaires. Les programmes fournis avec l'énoncé ne devront pas être rendus.

1

Le TD s'appuie sur des données libres accessibles depuis le site *ParisData*¹. Le fichier téléchargé contient les ordres du jour de la mairie de Paris sur les années 2008 à 2012. L'objectif est d'extraire différentes informations à partir de ces données :

1. Certains intitulés concernent l'attribution de subventions à des associations : l'objet contient le mot *association* et un montant. On veut connaître la somme des montants de ces subventions par année.
2. Le nom de l'association qui reçoit la subvention est spécifié suivi d'un arrondissement. On veut savoir quels sont les arrondissements qui ont reçu le plus de subventions.

La première étape consiste à récupérer les deux programmes *Python* et à exécuter le second : `td_note_2013_r.py`. Il contient une fonction qui récupère les informations sous forme de matrice.

```
def lit_fichier (file) :
    """
    0 Séance
    1 Référence
    2 Entité dépositaire
    3 Elu dépositaire
    4 Objet
    5 Type
    6 Rapporteur
    """
    f = open(file,"r")
    lines = f.readlines ()
    f.close ()
    lines = [ _ for _ in lines if len(_) > 0 ]
    lines = [ _.split("\t") for _ in lines ] [1:]
    lines = [ (_[0], _[4]) for _ in lines if len(_) > 5 ]
    return lines
```

1) Ecrire deux lignes de code pour appeler la fonction citée ci-dessus et afficher les dix premières lignes les unes en dessous des autres. (1 point)

2) La première information qu'on cherche, le montant, est entourée de texte :

Subvention à l'association des commerçants "Le Village d'Auteuil" (16e) pour les illuminations à l'occasion des fêtes de fin d'année 2007, dans le cadre de l'opération "Paris Illumine Paris". - Montant : 9.000 euros. PJ Mme Lyne COHEN-SOLAL (2ème Commission) rapporteure.

Ou encore :

Subvention (7.000 euros) à l'association Futur Composé (3e). PJ M. Bruno JULLIARD (9ème Commission), Mme Véronique DUBARRY (6ème Commission) rapporteurs.

Le programme fournit en exemple contient une fonction dont le code est le suivant :

1. http://opendata.paris.fr/opendata/jsp/site/Portal.jsp?document_id=154&portlet_id=102

```

import re
def extrait_montant (objet) :
    exp = re.compile ("[ ]([0-9.,]+) {0,3}euros")
    res = exp.search (objet)
    if res :
        montant = res.groups() [0]
        return montant
    else :
        print ("problème ", objet)
        return None

```

Il utilise les expressions régulières^{2 3}. Que retourne la fonction sur les deux expressions citées ci-dessus? (2 points)

3) Appliquer la fonction précédente pour extraire tous les montants relatifs aux subventions. (3 points)

4) Combien de lignes ne sont pas bien traitées? Proposez et implémentez une idée pour réduire (seulement réduire et non annuler) ce nombre? (2 points)

5) Ecrire une fonction qui convertit les montants récupérés en montants numériques et calculer la somme. (2 points)

6) En vous inspirant de la fonction `extrait_montant`, écrire une fonction qui récupère la date pour chaque subvention (3 points)

7) Ecrire une fonction qui calcule la somme des subventions attribuées par année? On pourra utiliser un dictionnaire. (2 points)

8) L'arrondissement apparaît entre parenthèses. On propose d'utiliser l'expression régulière suivante pour extraire cette information.

```

exp = re.compile ("association(.*)[( ]([0-9]+e)[ ])")

```

Modifier votre programme pour obtenir la répartition des montants par arrondissement. (3 points)

9) La question 5 donne la somme totale des montants, la question 7 a donné la répartition des montants par année et arrondissement à la question 8. Avez-vous confiance dans les résultats obtenus? A quoi cela est-il dû? Que suggérez-vous pour les améliorer? (On ne demande pas de le faire.) (2 points)

2. <http://www.dsimb.inserm.fr/~fuchs/python/python-node14.html>
3. <http://www.xavierdupre.fr/mywiki/InitiationPython>