

TP de Langages de script n° 7bis : manipulation de fichiers textuels et expressions régulières

I) Avant de commencer

Le ministère de la culture met à disposition du public une liste de musées de France. Les données sont données dans un fichier `musees.txt` qui est disponible sur Didel. Chaque ligne suit le format suivant : `REGIONS;DEPARTEMENTS;FERME;ANNEXE;NOM DU MUSEE;ADRESSE;CP;VILLE;SITWEB;FERMETURE ANNUELLE;PERIODE OUVERTURE;JOURS NOCTURNES`

Nous allons utiliser les expressions régulières pour récupérer des informations sur les musées et leurs horaires d'ouverture.

Nous allons nous concentrer sur les dates de fermeture annuelle (10e colonne). Voici quelques exemples de texte :

- Novembre à mi-avril
- 1er janvier, vendredi saint, 1er mai, 1er et 11 novembre, 25 décembre
- Décembre, janvier, février, mars
- De décembre à mars
- 31 décembre au 21 mars

Nous allons également regarder les horaires d'ouverture qui se trouvent en 11e colonne, "`PERIODE D'OUVERTURE`"

Voici quelques exemples de texte :

- Ouvert du mardi au vendredi de 12h à 18h et le samedi et dimanche de 10h à 18h
- Ouvert du mardi au dimanche de 14 à 18 h

II) Les exercices

Exercice 1 :

Écrire une fonction qui prend en entrée le nom du fichier et retourne une liste de musées. Chaque musée est un dictionnaire qui associe l'intitulé de chaque colonne à la donnée correspondante : `{"REGIONS": "ALSACE", "DEPARTEMENTS":"BAS-RHIN", "FERME":"NON", ...}`

Exercice 2 :

Utilisez les expressions régulières pour trouver la liste de tous les musées

1. sur l'archéologie ou sur la période Gallo-Romaine
2. à Paris fermés le 14 juillet.

Exercice 3 :

Ecrivez une expression régulière pour chacun des motifs suivants. Parcourez la liste de tous les musées pour trouver toutes les occurrences de chaque motif.

1. Un jour (lundi, mardi,...)
2. une liste de jours séparés par des virgules ou "et".
3. une plage de jours (du lundi au vendredi)

4. un mois (janvier, fevrier,...)
5. une date (1er fevrier, 2 mai,...)
6. un intervalle de mois (de mai à octobre, d'avril à septembre)
7. un intervalle de dates (du 1er avril au 31 octobre)
8. une heure (10h30, 17h,...)
9. une plage horaire (de 10h à 18h)

Exercice 4 :

Reprenez les expressions suivantes et extrayez-en les informations importantes, comme le jour, le mois, l'heure, les minutes, à l'aide des expressions parenthésées dans les expressions régulières.

Exercice 5 :

Proposez d'autres motifs pour récupérer toutes les informations sur les jours de fermeture et les heures d'ouverture des musées. Testez-les.