

Étape 5 :

Python - Analyse des URL les plus consultées par utilisateur

L'objectif de ce programme est de choisir un des fichiers log du dossier **Fichiers_logs_proxy** et de renvoyer des informations avec une interface intuitive sous forme de menu.

Ligne 4 à 12 : affiche une liste des options disponibles en fonction de la quantité de fichiers présents dans le dossier, le nom du fichier est affiché et une valeur lui est attribué pour pouvoir par la suite le choisir avec une fonction input pour la suite du programme.

Ligne 17 à 41 : Après avoir choisi le fichier que l'on veut consulter, il sera lu ligne par ligne avec la fonction **readline** et va récupérer l'adresse IP de l'utilisateur ainsi que l'url consultée. Enfin, le programme compte le nombre de fois que les url ont été mentionnées pour ensuite afficher l'url la plus visité et le nombre de visites.

```

1 # Etape 5 : affiche les fichiers logs disponibles dans un dossier sous forme de menu
2
3 # Menu pour choisir le fichier à analyser
4 import os
5 listeFichier = os.listdir('Fichiers_logs_proxy')
6 print("_____")
7 print("Vous voulez que j'analyse quel fichier de logs :")
8 # Boucle pour afficher le nombre de choix en fonction du nombre de fichier existant
9 for fichier in range(len(listeFichier)):
10     print(f" - Le fichier de logs : {listeFichier[fichier]} (répondez {fichier + 1})")
11     reponse = int(input("Réponse : "))
12     print("_____")
13
14 # Initialisation d'un dictionnaire pour contenir chaque utilisateur par leur ip
15 utilisateur = {}
16
17 # La valeur choisie va ouvrir un fichier log correspondant et effectuer le code ci-dessous
18 with open("Fichiers_logs_proxy/"+listeFichier[reponse - 1], "r") as fich:
19     ligne = fich.readline()
20     #Boucle qui permet d'avoir un dictionnaire pour chaque utilisateur avec l'url comme clé et le nombre de fois visité en valeur
21     while ligne:
22         util = ligne.strip().split(" ")[1]
23         site = ligne.strip().split(" ")[4]
24         if util not in utilisateur:
25             utilisateur[util] = {}
26         if site in utilisateur[util]:
27             utilisateur[util][site] += 1
28         else:
29             utilisateur[util][site] = 1
30         ligne = fich.readline()
31     # Affiche les résultats (chaques utilisateurs avec leurs sites le plus visités et le nombre de fois qu'ils sont visités)
32     for user in utilisateur.keys():
33         print(f"Utilisateur : {user}")
34         urlPlusVisite = ""
35         nbFoisUrl = 0
36         for url in utilisateur[user].items():
37             if url[1] > nbFoisUrl:
38                 urlPlusVisite = url[0]
39                 nbFoisUrl = url[1]
40         print(f"Url le plus visité : {urlPlusVisite}")
41         print(f"Nombre de visite : {max(utilisateur[user].values())}")

```

Ci-dessous, deux exemple d'utilisation du programme qui renvoient l'adresse IP d'un utilisateur avec les url les plus consultées et le nombre de visites :

```
-----
Vous voulez que j'analyse quel fichier de logs :
- Le fichier de logs : log_proxy_2024-11-23.txt (répondez 1)
- Le fichier de logs : log_proxy_2024-11-24.txt (répondez 2)
- Le fichier de logs : log_proxy_2024-11-25.txt (répondez 3)
Réponse : 1
-----
Utilisateur : 192.168.2.100
Url le plus visité : https://www.shopify.com
Nombre de visite : 22
Utilisateur : 192.168.2.3
Url le plus visité : https://www.instagram.com
Nombre de visite : 14
Utilisateur : 192.168.2.1
Url le plus visité : https://service-unavailable.net
Nombre de visite : 9
Utilisateur : 192.168.2.2
Url le plus visité : https://www.shopify.com
Nombre de visite : 7
```

```
-----
Vous voulez que j'analyse quel fichier de logs :
- Le fichier de logs : log_proxy_2024-11-23.txt (répondez 1)
- Le fichier de logs : log_proxy_2024-11-24.txt (répondez 2)
- Le fichier de logs : log_proxy_2024-11-25.txt (répondez 3)
Réponse : 2
-----
Utilisateur : 192.168.2.2
Url le plus visité : https://www.netflix.com
Nombre de visite : 97
Utilisateur : 192.168.2.3
Url le plus visité : https://nxdomain-test.site
Nombre de visite : 98
Utilisateur : 192.168.2.1
Url le plus visité : https://nxdomain-test.site
Nombre de visite : 103
Utilisateur : 192.168.2.100
Url le plus visité : https://www.spotify.com
Nombre de visite : 101
```