Home assignment 1
Directories

Task

The virtual machine that can be downloaded from the course's website contains an LDAP directory with data representing an imaginary hotel company. Create a Python script that can perform the following task on the LDAP directory.

The script gets as an input parameter a file containing a list of organizational units (OU) and a limit for each of the organization units. Select those organizational units from these that contain more users that the specified limit. The script has to write the names of the selected OUs in a CSV (Comma-Separated Values) file.

Name and parameters of the script

```
list-suspicious.py -s <search_list> -o <suspicious_list>  [-F]
```

The script has to use this name and these parameters. The order of the parameters should not be fixed. The -F parameter is an optional flag parameter. The definition of the parameters are:

- `seach_list`: the path to a CSV file that contains the list of OUs to search for.
- `suspicious_list`: the path to the output file. If this file exists, then the script should exit with an error.
- F: optional flag. If given, then in the output instead of writing the actual number of users inside the OU, write just the amount exceeding the limit (i.e., actual size - limit).

Examples for valid uses of the script:

```
list-suspicious.py -s input.csv -F -o output.csv
```

```
list-suspicious.py -o suspicios.csv -s list.csv
```

Input file

The input file is an UTF-8 encoded CSV text file with the following columns: DN, LIMIT, where DN is the distinguished name of the OU to analyze, and LIMIT is the limit of the users in that OU.

Example:

```
DN, LIMIT
"ou=France,ou=Europe,ou=Hotels,dc=irf,dc=local", "50"
"ou=Osaka,ou=Japan,ou=Asia,ou=Hotels,dc=irf,dc=local", "5"
```

Thus, for example the France OU can contain at most 50 users. Pay attention that if the OU has child OUs, then the users in those OUs should also be counted!

Output file

The output file is an UTF-8 encoded CSV text file with the following columns: DN, SIZE, where DN is a distinguished name of an OU whose actual size exceed the limit, and SIZE is its actual size, i.e. the number of users in the subtree of the OU.

Example:

```
DN, SIZE
"ou=France,ou=Europe,ou=Hotels,dc=irf,dc=local", "56"
```

The France OU is in the output file, because its size is greater than the specified limit for that OU (it is 56 vs. 50). If the -F flag would be given to the script then it should write out only 6 for size (as 56-50=6).

Further constraints

- Implement error handling in your script! For example, check whether the specified input file exists.
- Add comment to your script explaining its important parts!
- Use the `csv` module in Python for parsing the CSV files!
- Use the `ldapsearch` tool to make queries against the LDAP directory. (Or you can use the http://www.python-ldap.org/ module in Python, but it is a bit more complex.)