Decision Point Recruitment Test

Time Duration :2 Hrs

Instructions

1. Please attempt all the Questions, **Q4 and Q5 are compulsory** (attempt at least one of them).
2. Candidate should write the code in Python only.
3. Candidate must submit all the answer/code in a single word/Pdf/.ipynb (jupyter notebook Format) document. **Candidate must rename the answer file with his/her full name.**
4. Attach your updated CV along with the answer file.

Question 1.

Flatten a nested list with any dimensions (or nested) list or elements within the main list.

Example

Input : `[[[2]], [4, [5, 6, [6], 6, 6, 6], 7]]`
Output: `[2, 4, 5, 6, 6, 6, 6, 6, 7]`

Example 2:
Input:`[[2, 8, 9], [[1, 8, 7], [[2, 4, 7], [3, 2, 1]]], 3]`
Output: `[2, 8, 9, 1, 8, 7, 2, 4, 7, 3, 2, 1, 3]`

Question 2.

Given an array of strings, sort them in the order of increasing lengths. If two strings have the same length, their relative order must be the same as in the initial array.

Example

inputArray = ["abc", "","aaa","a","zz"]

the output should be

sortByLength(inputArray) = ["","a","zz","abc","aaa"]

Question 3.

GoDaddy makes a lot of different top-level domains available to its customers. A top-level domain is one that goes directly after the last dot ('.') in the domain name, for example .com in example.com. To help the users choose from available domains, GoDaddy is introducing a new feature that shows the type of the chosen top-level domain. You have to implement this feature.

To begin with, you want to write a function that labels the domains as **"commercial", "organization", "network" or "information" for .com, .org, .net or .info respectively**.

For the given list of domains return the list of their labels.

Example

For domains = ["en.wiki.org", "codesignal.com", "happy.net", "code.info"]

the output should be

domainType = ["organization", "commercial", "network", "information"]

Write a scalable code which can be implemented for any number of domainTypes and domains.

Question 4.

Attached Crash_data.csv includes air crash data from year 1919 to 2019.

Columns Description:

Country: Name of the country where crash happened.

Date: Date of the crash

Fatalities: number of fatalities happened. (format: including on board fatalities + on ground fatalities)

Flight_model: Crash Flight model

Location: location of crash

Operator: Operating company name

Process data using Python pandas as instructed:

a. Read data as a dataframe
b. Data Clean Step:
   Convert values like '2+3' in fatalities column to its aggregating value
   Change all columns to required data type.
   Handle Missing Values.
c. Create 2 columns: Year & Month with year and month value from date column
d. Create a new data frame only with country, fatalities, year, month columns from above data frame. Name it country_fatalities
e. Find top 10 countries with max number of fatalities from year 2000 – 2019 (Both Included)
f. Create a new data frame from country_fatalities with only these top 10 countries data in it. Name it top_countries
g. Pivot up month column from top_countries dataframe. With values as sum

Question 5.

Calculate MTD (Month To Date) by category for given data using SQL or Python.

Please refer Excel file attached for Data- Q5.csv