Programmatic Interview Exercise

Overview

You have been provided with a list of IP addresses which represent users who have accessed our system in the past. The data science team wants to generate a report which analyzes these users' behavior, but they require additional information from our systems. You will need to read the list of addresses, query our systems for the required information, and generate a new file which aggregates the data.

Step 1 – Address List

The list of IP addresses which the data science team wants to analyze has been provided to you. It is called *ipaddresses.txt* and is a plaintext file with one IPv4 address per line.

Step 2 – Geolocation

The first piece of information which you will need to query is their geolocation, which our systems determine from the IP address. Your code can perform an HTTP POST query to the following URL:

- https://cs-interview-geolocation.azurewebsites.net/api/GeolocationFunction?code=h/0av4gxkkTRz6d5HFIKr/U3b4ynT/CCNiKuHI7Wk7pVWYBKS8iZOw==

The body of the request should be a JSON object with the following format:

```
{
    "addresses": [
        "1.2.3.4",
        "2.3.4.5"
    ]
}
```

The body of the response will be a JSON object with the following format:

```
{
    "geolocations": [
        { "address": "1.2.3.4", "country": "US" },
        { "address": "2.3.4.5", "country": "FR" }
    ]
}
```

If a requested address is not included in the response, its location is unknown. You should discard any such addresses before performing further processing. Otherwise, one of the following ISO-3366-1 country codes will be returned:

- United States (US)
- United Kingdom (GB)
- France (FR)

- Germany (DE)
- China (CN)

Step 3 – User Data

Once you've loaded the geolocation for each address, you need to query for the user data associated with that address. In order to comply with international data privacy laws, user data is stored in different locations depending on the user's country of origin.

Your code can perform an HTTP POST query to one of the following URLs to retrieve this data:

- https://cs-interview-users-us.azurewebsites.net/api/UserDataFunction?code=Ax1P8UqUd6sQtQpJ4ldX8wbh/cEXt6xDFO4IK3ptTlDC29KlGHyj6A== (United States)
- https://cs-interview-users-eu.azurewebsites.net/api/UserDataFunction?code=GjLpRZ2u2ZNoNi0Rrvdl6PSF3XEWUgJUY8e6BqUOhgx9e5KRISJuQA== (Europe)
- https://cs-interview-users-as.azurewebsites.net/api/UserDataFunction?code=wUi5HeaRJtYjX/7APOVehBWvLLaOYI7VFJBRCCX5tmk6i2aOoe0nkA== (Asia)

The body of the request should be a JSON object with the following format:

```
{
    "addresses": [
        "1.2.3.4"
    ]
}
```

The body of the response will be a JSON object with the following format:

```
{
    "users": [
        {
            "name": "John Smith",
            "address": "1.2.3.4",
            "gender": "male",
            "age": "12-17"
        }
    ]
}
```

Again, if a requested address in not included in the response, its user data could not be found, and it can be ignored from this point forward.

Step 4 – Output File

Once you've loaded the data for each user, write a CSV file to disk called *report.csv* with the following columns: IPv4 Address, Country of Origin, Name, Gender, and Age.