**Lab Assignment 7: JSON Handling**

You are working as a data scientist for a healthcare organization, and your team has been tasked with analysing COVID-19 data from multiple countries. The data is stored in JSON files, with each file representing the daily COVID-19 statistics for a specific country. Each JSON file has the following structure:

{ "country": "Country Name",

"date": "YYYY-MM-DD",

"confirmed\_cases": { "total": 1000, "new": 50 },

"deaths": { "total": 20, "new": 2 },

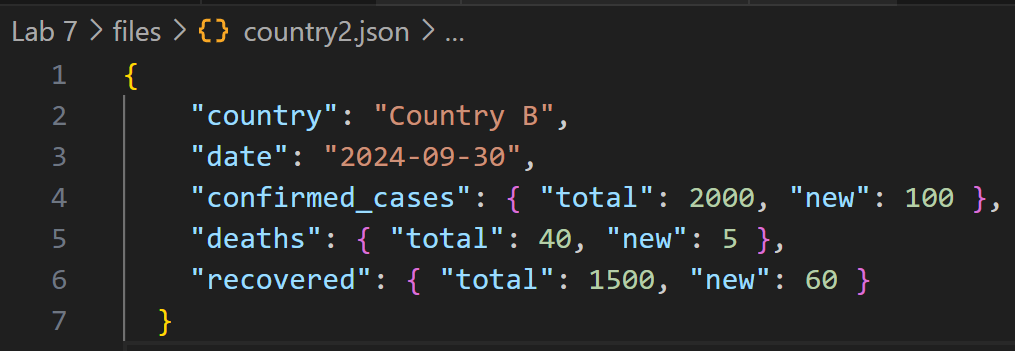
"recovered": { "total": 800, "new": 30 }

}

Your task is to write a Python program that performs the following operations:

1. Read COVID-19 data from all JSON files in a given directory and its subdirectories.
2. Calculate and display the following statistics for each country:
3. Total confirmed cases.
4. Total deaths.
5. Total recovered cases.
6. Total active cases (total confirmed cases minus total deaths and total recovered).
7. Determine the top 5 countries with the highest number of confirmed cases and the lowest number of confirmed cases.
8. Generate a summary report in JSON format that includes the statistics for all countries and save it to a file named "covid19\_summary.json".

Data Set –



Code –

import os

import json

# Function to read all JSON files in a given directory and subdirectories

def read\_json\_files(directory):

    covid\_data = []

    for root, \_, files in os.walk(directory):

        for file in files:

            if file.endswith('.json'):

                file\_path = os.path.join(root, file)

                with open(file\_path, 'r') as f:

                    data = json.load(f)

                    covid\_data.append(data)

    return covid\_data

# Function to calculate statistics for each country

def calculate\_statistics(covid\_data):

    country\_stats = {}

    for data in covid\_data:

        country = data['country']

        if country not in country\_stats:

            country\_stats[country] = {

                'total\_confirmed': 0,

                'total\_deaths': 0,

                'total\_recovered': 0,

                'total\_active': 0

            }

        country\_stats[country]['total\_confirmed'] += data['confirmed\_cases']['total']

        country\_stats[country]['total\_deaths'] += data['deaths']['total']

        country\_stats[country]['total\_recovered'] += data['recovered']['total']

        country\_stats[country]['total\_active'] = (

            country\_stats[country]['total\_confirmed'] -

            country\_stats[country]['total\_deaths'] -

            country\_stats[country]['total\_recovered']

        )

    return country\_stats

# Function to determine top 5 countries with highest and lowest confirmed cases

def top\_5\_countries\_by\_confirmed\_cases(country\_stats):

    sorted\_by\_confirmed = sorted(country\_stats.items(), key=lambda x: x[1]['total\_confirmed'], reverse=True)

    highest\_5 = sorted\_by\_confirmed[:5]

    lowest\_5 = sorted\_by\_confirmed[-5:]

    return highest\_5, lowest\_5

# Function to generate summary report in JSON format

def generate\_summary\_report(country\_stats, output\_file):

    with open(output\_file, 'w') as f:

        json.dump(country\_stats, f, indent=4)

    print(f"Summary report saved to {output\_file}")

def main():

    # Directory containing the JSON files

    directory = 'D:/5th Lab/Python/Lab 7/files'

    # Step 1: Read JSON files

    covid\_data = read\_json\_files(directory)

    # Step 2: Calculate statistics for each country

    country\_stats = calculate\_statistics(covid\_data)

    # Display statistics

    for country, stats in country\_stats.items():

        print(f"{country}:")

        print(f"  Total Confirmed Cases: {stats['total\_confirmed']}")

        print(f"  Total Deaths: {stats['total\_deaths']}")

        print(f"  Total Recovered: {stats['total\_recovered']}")

        print(f"  Total Active Cases: {stats['total\_active']}")

        print()

    # Step 3: Determine top 5 countries with highest and lowest confirmed cases

    highest\_5, lowest\_5 = top\_5\_countries\_by\_confirmed\_cases(country\_stats)

    print("Top 5 countries with highest confirmed cases:")

    for country, stats in highest\_5:

        print(f"{country}: {stats['total\_confirmed']} confirmed cases")

    print("\nTop 5 countries with lowest confirmed cases:")

    for country, stats in lowest\_5:

        print(f"{country}: {stats['total\_confirmed']} confirmed cases")

    # Step 4: Generate summary report in JSON format

    generate\_summary\_report(country\_stats, "covid19\_summary.json")

if \_\_name\_\_ == "\_\_main\_\_":

    main()

Output –

