Practical 4

Analyzing a JSON File

Aim: To load and analyze data from a JSON file using Python and extract insights.

Steps:

1. **Load a JSON file into a Pandas DataFrame.**
2. **Inspect the schema and perform summary statistics.**
3. **Extract and analyze specific data points.**
4. **Visualize data distribution for a column.**

Code:

```
import pandas as pd
import json
from pandas import json_normalize
# Step 1: Read the JSON file manually
with open("C:/Users/B-Night College-01/Desktop/bda prac/sample_data.json") as f:
    data = json.load(f)
# Step 2: Normalize the nested lists
# Normalize 'employees' data (which is an array of dictionaries)
employees_df = json_normalize(data['employees'])
# Normalize 'projects' data (which is also an array of dictionaries)
projects_df = json_normalize(data['projects'])
# Normalize the 'company' data (which is a single dictionary, not an array)
company_df = pd.DataFrame([data['company']])
```

```
# Step 3: Print the results
print("Employees DataFrame:")
print(employees_df)
print("\nProjects DataFrame:")
print(projects_df)
print("\nCompany Info DataFrame:")
print(company_df)
```

Output:

```
Employees DataFrame:
   id first_name last_name  ...   department salary   hire_date
0   1       John       Doe  ...  Engineering  75000  2020-06-15
1   2       Jane     Smith  ...    Marketing  68000  2018-09-01
2   3      Alice   Johnson  ...        Sales  72000  2021-03-25

[3 rows x 7 columns]

Projects DataFrame:
    id       name       status    deadline
0  101  Project X  In Progress  2024-12-31
1  102  Project Y    Completed  2024-03-15
2  103  Project Z      Planned  2025-06-01

Company Info DataFrame:
              name                              address     founded
0  Tech Innovators  1234 Tech Ave, Silicon Valley, CA  2015-08-01
```