|  |  |
| --- | --- |
| **Name** | **Pranay Singhvi** |
| **UID no.** | **2021300126** |

|  |  |
| --- | --- |
| **Experiment 1** | |
| **HONOUR PLEDGE** |  |
| **PROBLEM STATEMENT :** | **Data Importing and Exporting:**  **● Read a CSV file into a pandas Data Frame**  **● Export a Data Frame to an Excel file.**  **● Load JSON data into DataFrame**  **● (Bonus) Data insight or visualization** |
| **THEORY:** | 1. **Comma-Separated Values (CSV):**   Comma-Separated Values (CSV) is a widely used file format for storing and exchanging tabular data between systems. This simple and lightweight format has become a standard for data interchange due to its ease of use, human readability, and broad compatibility across different applications and programming languages.  Basic Structure:  In CSV, each line of the file represents a row of data, and within each line, individual data fields are separated by commas. The first row often contains headers, specifying the names of the columns. Here's a basic example:  Name, Age, Occupation  John Doe, 30, Engineer  Jane Smith, 25, Scientist  Bob Johnson, 35, Artist   1. JSON (JavaScript Object Notation):   JSON, or JavaScript Object Notation, is a lightweight data interchange format widely used for data representation and communication between systems. It is easy for both humans to read and write and for machines to parse and generate. JSON is language-independent, making it a popular choice for web development and API communication.  Basic Structure:  JSON data consists of key-value pairs enclosed in curly braces, with each key and its corresponding value separated by a colon. Arrays, ordered lists of values, are represented using square brackets. Here's a simple example:  {  "name": "John Doe",  "age": 30,  "occupation": "Engineer",  "skills": ["JavaScript", "Python", "SQL"]  }   1. DataFrames   Data Frames are two-dimensional, tabular data structures in which data is organized in rows and columns. They are a key component of data manipulation and analysis, providing a convenient way to work with structured data. In Python, the Pandas library is commonly used to create and manipulate Data Frames.  Example:  Name Age City  0 John 28 New York  1 Alice 24 San Francisco  2 Bob 32 Chicago   1. Pandas Library   Pandas is a powerful and widely used open-source data manipulation and analysis library for the Python programming language. It provides high-performance, easy-to-use data structures, primarily the DataFrame, making it an essential tool for tasks involving cleaning, exploring, and analyzing structured data.  Key Features:  DataFrame: The central data structure in Pandas is the DataFrame, a two-dimensional, labeled table that efficiently handles heterogeneous data types. It allows for intuitive manipulation and analysis of data.  Data Alignment: Pandas excels in data alignment and integration. It seamlessly handles missing data and aligns datasets on common indices, simplifying complex data operations.  Data Cleaning and Preparation: Pandas offers a suite of functions for data cleaning, transformation, and preparation. It includes tools for handling missing data, reshaping datasets, and filtering information.  Data Analysis and Exploration: Pandas supports powerful data analysis and exploration through functionalities such as groupby operations, merging and joining datasets, and statistical aggregation.  Integration with Other Libraries: Pandas integrates well with other popular Python libraries like NumPy, Matplotlib, and scikit-learn, providing a comprehensive ecosystem for data science and machine learning tasks. |
| **PROGRAM:** | Python Notebook File:  A white background with a black border  Description automatically generated with medium confidence    A screenshot of a computer  Description automatically generated      A graph of a number of people  Description automatically generated |
| **RESULT:** | 1. **Startup.xlsx**   A screenshot of a computer  Description automatically generated   1. **Island.xlsx**   A table with numbers and letters  Description automatically generated |
| **CONCLUSION:**  In this experiment, we learned how to use pandas library to import csv and Json data files and convert them into data frames and then export them as excel sheets. | |