Experiment 1	
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
	2. 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"]
	}
	3. 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 4. 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:

Name: Pranay Singhvi

UID: 2021300126

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

1. Importing Necessary Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

2. Read a CSV file into a pandas Data Frame

```
[4] df = pd.read_csv('startup_funding.csv')
    df.head()
```

	Sr No	Date dd/mm/yyyy	Startup Name	Industry Vertical	SubVertical	City Location	Investors Name	InvestmentnType	Amount in USD	Remarks
0	1	09/01/2020	BYJU'S	E-Tech	E-learning	Bengaluru	Tiger Global Management	Private Equity Round	20,00,00,000	NaN
1	2	13/01/2020	Shuttl	Transportation	App based shuttle service	Gurgaon	Susquehanna Growth Equity	Series C	80,48,394	NaN
2	3	09/01/2020	Mamaearth	E-commerce	Retailer of baby and toddler products	Bengaluru	Sequoia Capital India	Series B	1,83,58,860	NaN
3	4	02/01/2020	https://www.wealthbucket.in/	FinTech	Online Investment	New Delhi	Vinod Khatumal	Pre-series A	30,00,000	NaN
4	5	02/01/2020	Fashor	Fashion and Apparel	Embroiled Clothes For Women	Mumbai	Sprout Venture Partners	Seed Round	18,00,000	NaN

3.Export a Data Frame to an Excel file.

```
[5] df.to_excel('startup.xlsx')
```

4. Load JSON data into DataFrame

```
[6]  df2 = pd.read_json('Indian.json')
     df2.head()
```

	states
0	{'state': {'state_id': 'AN', 'state_name': 'An...
1	{'state': {'state_id': 'AP', 'state_name': 'An...
2	{'state': {'state_id': 'AR', 'state_name': 'Ar...
3	{'state': {'state_id': 'AS', 'state_name': 'As...
4	{'state': {'state_id': 'BR', 'state_name': 'Bi...

```
df3 = pd.json_normalize(df2['states'])
df3.head()
```

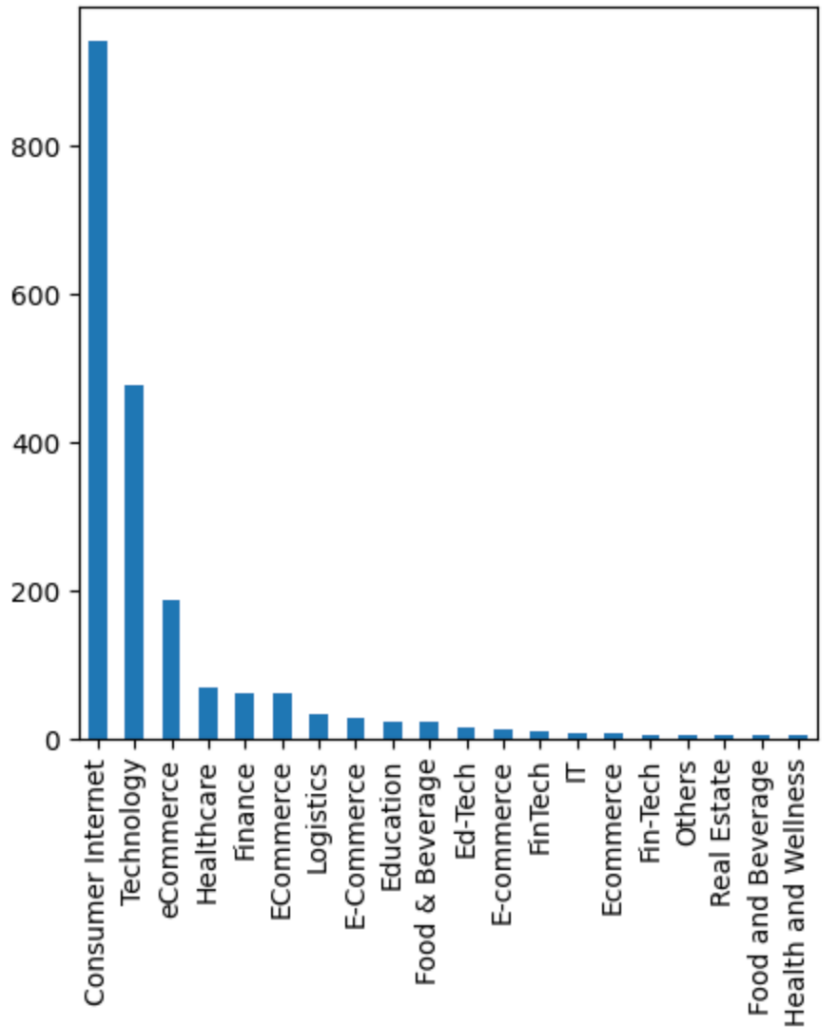
	state.state_id	state.state_name
0	AN	Andaman and Nicobar Island (UT)
1	AP	Andhra Pradesh
2	AR	Arunachal Pradesh
3	AS	Assam
4	BR	Bihar

5.(Bonus) Data insight or visualization

Q) Which sector do most startups work on?

```
df["Industry Vertical"].value_counts().head(20).plot(kind="bar",figsize=(10,10))
```

```
<Axes: >
```



The startup landscape indicated that the consumer internet sector had experienced significant growth, emerging as one of the largest and most dynamic spaces for startups. Following closely, the technology sector also demonstrated substantial activity, showcasing the ongoing trend of innovation and entrepreneurship in both consumer-facing digital platforms and underlying technological advancements.

RESULT:

1. Startup.xlsx

	A	B	C	D	E	F	G	
1080	1078	1079	18/01/2017	Direct Create	eCommerce	B2B Marketplace for Handicrafts	Gurgaon	Rajasthan Vent
1081	1079	1080	18/01/2017	Vahdam Teas	eCommerce	Online Tea etailer	New Delhi	Mumbai Angel
1082	1080	1081	18/01/2017	Qwikspec	Technology	Cloud-based construction	Bangalore	Undisclosed In
1083	1081	1082	19/01/2017	PParke	Consumer Internet	Parking Solutions mobile app	Bangalore	SB Ventures
1084	1082	1083	19/01/2017	Weddingz.in	Consumer Internet	Online Marketplace for Wedding Venues and vendors	Mumbai	Singularity Ver
1085	1083	1084	19/01/2017	Native Special	eCommerce	Indian Sweets & Snacks etailer	Karur	Indian Angel N
1086	1084	1085	19/01/2017	Fitnapp	Consumer Internet	Fitness Mobile App	Noida	SD Group
1087	1085	1086	20/01/2017	Ethosh	Technology	Visual and Interactive communications Solutions	Pune	Deepak Nathar
1088	1086	1087	20/01/2017	Wydr	eCommerce	B2B Wholesale Marketplace in India	Gurgaon	Bessemer Vent
1089	1087	1088	20/01/2017	Tyre Express	Technology	IoT platform for Tyre Performance management & tracking	Mumbai	IvyCamp Vent
1090	1088	1089	20/01/2017	Connaizen	Technology	Data Analytics platform	New Delhi	Umang Moond
1091	1089	1090	22/01/2017	MessaGif	Consumer Internet	social messaging platform	New Delhi	Undisclosed in
1092	1090	1091	23/01/2017	SmartHi	Technology	Audit automation Solutions	Pune	Singapore base
1093	1091	1092	23/01/2017	CRON Systems	Technology	Border Intrusion Alert Solutions	New Delhi	YourNest Fund
1094	1092	1093	23/01/2017	CueMath	Technology	Maths Learning program	New Delhi	CapitalG, Sequ
1095	1093	1094	23/01/2017	ChipperSage	Technology	Online learning platform	Bangalore	India Educatio
1096	1094	1095	24/01/2017	SelectJobs	Consumer Internet	Online Job portal	Mumbai	India Educatio
1097	1095	1096	24/01/2017	Khel Now	Consumer Internet	Mobile-based Social Platform for Sports	Bangalore	Indochine Inte
1098	1096	1097	24/01/2017	NowFloats	Consumer Internet	Platform to take businesses online	Hyderabad	Iron Pillar and
1099	1097	1098	24/01/2017	Lets Transport	Consumer Internet	Online transport vehicle booking platform	Bangalore	GMO Japan, N
1100	1098	1099	24/01/2017	BetaOut	Technology	ECommerce Marketing Software	Noida	East Ventures E
1101	1099	1100	25/01/2017	Wishberry	Consumer Internet	Crowdfunding Platform	Mumbai	Sharad Sharma
1102	1100	1101	25/01/2017	RML Agtech	Technology	Agri Decision Support Solution for farmers	Pune	IvyCap Venture
1103	1101	1102	25/01/2017	LoyaltyPrime	Technology	Loyalty program management solutions	New Delhi	Senovo, Unter
1104	1102	1103	25/01/2017	InstaSafe	Technology	Security-as-a-Service solution provider	Bangalore	ABM Knowledg
1105	1103	1104	25/01/2017	Infibeam	eCommerce	eCommerce Marketplace	Ahmedabad	Bennett, Colen
1106	1104	1105	02/01/2017	99Games	Technology	Mobile Game Developer Studio	Udupi	Dream Incubat
1107	1105	1106	26/01/2017	AirMed Labs	Consumer Internet	Health Tests Booking platform & ePharmacy	Ahmedabad	Brand Capital
1108	1106	1107	26/01/2017	99PerHour	Consumer Internet	Hyper-local Handyman Service provider	Chennai	V Aanand R
1109	1107	1108	26/01/2017	Kochi Post	Consumer Internet	English Online News portal	Kochi	Rajesh Yohann
1110	1108	1109	30/01/2017	Intuit Things	Technology	IoT solutions for home automation applications	Mumbai	Sanjay Mehta &
1111	1109	1110	30/01/2017	Cloudrino	Technology	Cloud based Virtual Servers	New Delhi	The Ten Minute
1112	1110	1111	30/01/2017	FabX	eCommerce	Used Furniture Marketplace	Mumbai	The Ten Minute
1113	1111	1112	30/01/2017	Maptags	Consumer Internet	Smart Online address tags	Bangalore	The Ten Minute
1114	1112	1113	30/01/2017	YOLO Health	Consumer Internet	Online Healthcare platform	Mumbai	Chandigarh An
1115	1113	1114	31/01/2017	Heads Up For Tails	eCommerce	Online Pet Products Store	New Delhi	Undisclosed In
1116	1114	1115	31/01/2017	Register My Marriage	Consumer Internet	Online Marriage Registration services	New Delhi	Mumbai Angel
1117	1115	1116	01/12/2016	SidQam	Technology	Healthcare Software System	Noida	SQue Capital,
1118	1116	1117	01/12/2016	India Lends	Consumer Internet	Online Instant personal Loan provider	New Delhi	DSG Consumer
1119	1117	1118	01/12/2016	Slide App	Consumer Internet	Lock screen Rewards mobile app	New Delhi	Songhyun Inve
1120	1118	1119	02/12/2016	Badiyajobs	Consumer Internet	Entry-Level job Seekers platform	New Delhi	Undisclosed
1121	1119	1120	02/12/2016	Streo	Consumer Internet	Live Music Streaming App	Bangalore	GEMs Advisory
1122	1120	1121	02/12/2016	TYGR	Technology	Transport and logistics mobile App	Mumbai	Franchise India

2. Island.xlsx

	A	B	C	D	E	F	G
1		state.state_ite.state_name					
2	0	AN	Andaman and Nicobar Island (UT)				
3	1	AP	Andhra Pradesh				
4	2	AR	Arunachal Pradesh				
5	3	AS	Assam				
6	4	BR	Bihar				
7	5	CH	Chandigarh (UT)				
8	6	CG	Chhattisgarh				
9	7	DN	Dadra and Nagar Haveli (UT)				
10	8	DD	Daman and Diu (UT)				
11	9	DL	Delhi (NCT)				
12	10	GA	Goa				
13	11	GJ	Gujarat				
14	12	HR	Haryana				
15	13	HP	Himachal Pradesh				
16	14	JK	Jammu and Kashmir (UT)				
17	15	JH	Jharkhand				
18	16	KA	Karnataka				
19	17	KL	Kerala				
20	18	LK	Ladakh(UT)				
21	19	LD	Lakshadweep (UT)				
22	20	MP	Madhya Pradesh				
23	21	MH	Maharashtra				
24	22	MN	Manipur				
25	23	ML	Meghalaya				
26	24	MZ	Mizoram				
27	25	NL	Nagaland				
28	26	OR	Odisha				
29	27	PY	Puducherry (UT)				
30	28	PB	Punjab				
31	29	RJ	Rajasthan				
32	30	SK	Sikkim				
33	31	TN	Tamil Nadu				
34	32	TG	Telangana				
35	33	TR	Tripura				
36	34	UK	Uttarakhand				
37	35	UP	Uttar Pradesh				
38	36	WB	West Bengal				
39							
40							
41							
42							
43							

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.