II. EDA-Data Import and Export

AIM:

- Importing data from CSV, Excel, SQL databases, and web scraping
- Handling different data formats
- Export a DataFrame to an Excel file.

PROCEDURE:

1. Import libraries
2. Load data from sources:
    - CSV: pd.read_csv('filename.csv')
    - Excel: pd.read_excel('filename.xlsx')
    - SQL: Use pd.read_sql(query, connection)
    - Web scraping: Use requests, BeautifulSoup, or pandas' URL in read_csv for some tables.
3. Handle formats:
    - Detect and fix mixed types using astype() (e.g., df['column'] df['column'].astype(float))
    - Use pandas methods to load JSON, Parquet, etc.
4. Analyze/clean data as needed (optional): Check with df.info(), df.head(), etc.
5. Export DataFrame to Excel:

PROGRAM AND OUTPUT:

```
import pandas as pd

df_csv = pd.read_csv("/content/test_Y3wMUE5_7gLdaTN.csv")

df_csv.head()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area
0	LP001015	Male	Yes	0	Graduate	No	5720	0	110.0	360.0	1.0	Urban
1	LP001022	Male	Yes	1	Graduate	No	3076	1500	126.0	360.0	1.0	Urban
2	LP001031	Male	Yes	2	Graduate	No	5000	1800	208.0	360.0	1.0	Urban
3	LP001035	Male	Yes	2	Graduate	No	2340	2546	100.0	360.0	NaN	Urban
4	LP001051	Male	No	0	Not Graduate	No	3276	0	78.0	360.0	1.0	Urban

df_excel = pd.read_excel("/content/Historicalinvesttemp.xlsx", sheet_name="Sheet1")

df_excel.head()

	Unnamed: 0	Unnamed: 1	Unnamed: 2	Unnamed: 3
0	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN
4	NaN	Annual Returns on Investments in	NaN	NaN
5	Year	Stocks	T.Bills	T.Bonds
6	1928	0.4381	0.0308	0.0084
7	1929	-0.083	0.0316	0.042
8	1930	-0.2512	0.0455	0.0454
9	1931	-0.4384	0.0231	-0.0256

```
url = "https://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal)"
tables = pd.read_html(url)
df_web = tables[0]
df_web.head()
```

	0
0	Largest economies in the world by GDP (nominal...

```
df=pd.read_excel("/content/Historicalinvesttemp.xlsx")
df.to_excel("/content/weeeeee.xlsx", index=False)
df.head()
```

	Unnamed: 0	Unnamed: 1	Unnamed: 2	Unnamed: 3
0	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN
4	NaN	Annual Returns on Investments in	NaN	NaN

RESULT:

Thus, the given programs are written and executed successfully.