

Akshaya B

231501010

Experiment No: 02

Date: 23-07-2025

## EDA – Data Import and Export

**Aim:** To import data from various sources, handle different formats, and export a DataFrame to an Excel file using Python.

### Code:

```
# Step 1: Import libraries import
```

```
pandas as pd
```

```
import sqlite3
```

```
from bs4 import BeautifulSoup import
```

```
requests from io import
```

```
StringIO
```

```
# Step 2: Importing data from CSV csv_data
```

```
= pd.read_csv("sample.csv") print("CSV
```

```
Data:") print(csv_data.head())
```

```
# Step 3: Importing data from Excel excel_data =
```

```
pd.read_excel("sample.xlsx") print("\nExcel Data:")
```

```
print(excel_data.head())
```

Akshaya B

231501010

```
# Step 4: Importing data from SQL Database # (Creating temporary
database and table for demo) conn = sqlite3.connect(":memory:") #
In-memory DB csv_data.to_sql("students", conn, index=False,
if_exists="replace") sql_data = pd.read_sql("SELECT * FROM
students", conn) print("\nSQL Data:") print(sql_data.head())
```

```
#web scraping
```

```
# URL url =
```

```
"https://en.wikipedia.org/wiki/List_of_countries_by_population_(United_Nati
ons)"
```

```
# Add headers to avoid blocking headers = {"User-Agent":
```

```
"Mozilla/5.0"} response = requests.get(url,
headers=headers)
```

```
# Parse HTML
```

```
soup = BeautifulSoup(response.text, "html.parser")
```

```
# Find all tables with 'wikitable' class
```

```
tables_html = soup.find_all("table", {"class": "wikitable"})
```

```
print(f"Number of tables found: {len(tables_html)}")
```

```
# Convert the first one into DataFrame if
```

```
tables_html:
```

Akshaya B

231501010

```
tables = pd.read_html(StringIO(str(tables_html[0])))
web_data = tables[0]    print("Web Scraped Data:")
print(web_data.head()) else:    print("No tables found
on the page.")
```

```
print("Web Scraped Data:") print(web_data.head(2))
```

# Step 6: Export DataFrame to Excel

```
csv_data.to_excel("exported_data.xlsx", index=False) print("\nData
exported successfully to 'exported_data.xlsx'") OUTPUT:
```

```
CSV Data:
  ID  Name  Age  Department  Marks
0  1  Alice   23         CSE     85
1  2   Bob   25         ECE     78
2  3  Charlie  22         ME     90
3  4  David   24        CIVIL    88
4  5   Eva   23         AI     95

Excel Data:
  ID  Name  Age  Department  Marks
0  1  Alice   23         CSE     85
1  2   Bob   25         ECE     78
2  3  Charlie  22         ME     90
3  4  David   24        CIVIL    88
4  5   Eva   23         AI     95

SQL Data:
  ID  Name  Age  Department  Marks
0  1  Alice   23         CSE     85
1  2   Bob   25         ECE     78
2  3  Charlie  22         ME     90
3  4  David   24        CIVIL    88
4  5   Eva   23         AI     95

Number of tables found: 1
Web Scraped Data:
  Country or territory  Population (1 July 2022)  Population (1 July 2023) \
0  World              8021407192              8091734930
1  India              1425423212              1438869596
2  China[a]          1425179969              1422584953
3  United States      341534046              343477335
4  Indonesia          278830529              281190067

Change (%) UN continental region[1] UN statistical subregion[1]
0  +0.88%  -  -
1  +0.85%  Asia  Southern Asia
2  -0.18%  Asia  Eastern Asia
3  +0.57%  Americas  Northern America
4  +0.85%  Asia  South-eastern Asia

Web Scraped Data:
  Country or territory  Population (1 July 2022)  Population (1 July 2023) \
0  World              8021407192              8091734930
1  India              1425423212              1438869596

Change (%) UN continental region[1] UN statistical subregion[1]
0  +0.88%  -  -
1  +0.85%  Asia  Southern Asia

Data exported successfully to 'exported_data.xlsx'
```

Akshaya B

231501010

**Result:** Successfully imported data from CSV, Excel, SQL, and web sources, handled multiple formats, and exported a DataFrame to Excel.