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

```python
# 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:

Akshaya B

231501010

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