

Name	Bodhisatya Ghosh
UID no.	2021700026

Experiment 1	
HONOUR PLEDGE	<p>I hereby declare that the documentation, code and output attached with this lab experiment has been completed by me in accordance with the highest standards of honesty. I confirm that I have not plagiarized OR used unauthorized materials OR given or received illegitimate help for completing this experiment. I will uphold equity and honesty in the evaluation of my work and if found guilty of plagiarism or dishonesty, will bear the consequences as outlined in the 'integrity' section of the lab rubrics. I am doing so in order to maintain a community built around this code of honour.</p> <p><u>B Ghosh</u> Bodhisatya Ghosh</p>
PROBLEM STATEMENT :	<p><b>Data Importing and Exporting:</b></p> <ul style="list-style-type: none"> <li>• Read a CSV file into a pandas DataFrame</li> <li>• Export a Data Frame to an Excel file.</li> <li>• Load JSON data into DataFrame</li> <li>• (Bonus) Data insight or visualization</li> </ul>

THEORY:	Using Pandas to load CSV files and then convert them to other tabular file types. Loading JSON files and creating DataFrames with them																																																
PROGRAM:	<div><div>exp 1.ipynb</div><div>BAP &gt; exp 1 &gt; exp 1.ipynb &gt; Loading CSV file and converting to excel</div><div>+ Code + Markdown ▶ Run All ↺ Restart Clear All Outputs Variables Outline ...</div><div>ml (Python 3.10.13)</div><div>Name : Bodhisatya Ghosh UID : 2021700026</div><div>~ Loading CSV file and converting to excel</div><div>import pandas as pd</div><div>[3] ✓ 0.0s Python</div><div>penguins = pd.read_csv('./penguins_size.csv') penguins.head()</div><div>[4] ✓ 0.0s Python</div><div>...<table><thead><tr><th></th><th>species</th><th>island</th><th>culmen_length_mm</th><th>culmen_depth_mm</th><th>flipper_length_mm</th><th>body_mass_g</th><th>sex</th></tr></thead><tbody><tr><td>0</td><td>Adelie</td><td>Torgersen</td><td>39.1</td><td>18.7</td><td>181.0</td><td>3750.0</td><td>MALE</td></tr><tr><td>1</td><td>Adelie</td><td>Torgersen</td><td>39.5</td><td>17.4</td><td>186.0</td><td>3800.0</td><td>FEMALE</td></tr><tr><td>2</td><td>Adelie</td><td>Torgersen</td><td>40.3</td><td>18.0</td><td>195.0</td><td>3250.0</td><td>FEMALE</td></tr><tr><td>3</td><td>Adelie</td><td>Torgersen</td><td>NaN</td><td>NaN</td><td>NaN</td><td>NaN</td><td>NaN</td></tr><tr><td>4</td><td>Adelie</td><td>Torgersen</td><td>36.7</td><td>19.3</td><td>193.0</td><td>3450.0</td><td>FEMALE</td></tr></tbody></table></div><div>penguins.info()</div><div>[5] ✓ 0.0s Python</div><div>penguins.info()</div><div>[5] ✓ 0.0s Python</div><div>...&lt;class 'pandas.core.frame.DataFrame'&gt; RangeIndex: 344 entries, 0 to 343 Data columns (total 7 columns): #   column                Non-Null Count  Dtype ---  -  - 0   species                344 non-null   object 1   island                 344 non-null   object 2   culmen_length_mm       342 non-null   float64 3   culmen_depth_mm        342 non-null   float64 4   flipper_length_mm      342 non-null   float64 5   body_mass_g            342 non-null   float64 6   sex                    334 non-null   object dtypes: float64(4), object(3) memory usage: 18.9+ KB</div><div>penguins.isna().sum()</div><div>[6] ✓ 0.0s Python</div><div>...species      0 island       0 culmen_length_mm      2 culmen_depth_mm      2 flipper_length_mm     2 body_mass_g         2 sex             10 dtype: int64</div></div>		species	island	culmen_length_mm	culmen_depth_mm	flipper_length_mm	body_mass_g	sex	0	Adelie	Torgersen	39.1	18.7	181.0	3750.0	MALE	1	Adelie	Torgersen	39.5	17.4	186.0	3800.0	FEMALE	2	Adelie	Torgersen	40.3	18.0	195.0	3250.0	FEMALE	3	Adelie	Torgersen	NaN	NaN	NaN	NaN	NaN	4	Adelie	Torgersen	36.7	19.3	193.0	3450.0	FEMALE
	species	island	culmen_length_mm	culmen_depth_mm	flipper_length_mm	body_mass_g	sex																																										
0	Adelie	Torgersen	39.1	18.7	181.0	3750.0	MALE																																										
1	Adelie	Torgersen	39.5	17.4	186.0	3800.0	FEMALE																																										
2	Adelie	Torgersen	40.3	18.0	195.0	3250.0	FEMALE																																										
3	Adelie	Torgersen	NaN	NaN	NaN	NaN	NaN																																										
4	Adelie	Torgersen	36.7	19.3	193.0	3450.0	FEMALE																																										

```
penguins.dropna(inplace=True)
```

[7] ✓ 0.0s Python

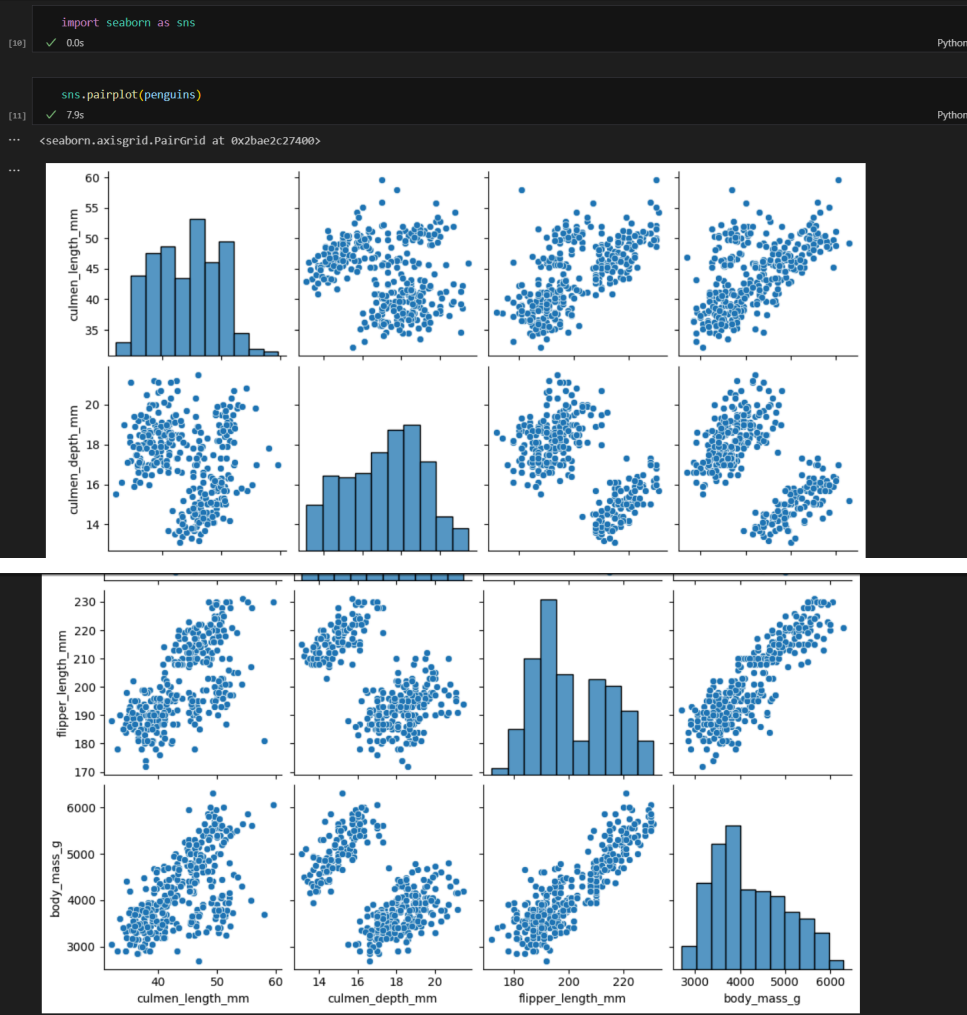
```
penguins.info()
```

[8] ✓ 0.0s Python

```
<class 'pandas.core.frame.DataFrame'>  
Index: 334 entries, 0 to 343  
Data columns (total 7 columns):  
#   Column              Non-Null Count  Dtype  
---  ---  
0   species              334 non-null   object  
1   island               334 non-null   object  
2   culmen_length_mm     334 non-null   float64  
3   culmen_depth_mm      334 non-null   float64  
4   flipper_length_mm    334 non-null   float64  
5   body_mass_g          334 non-null   float64  
6   sex                  334 non-null   object  
dtypes: float64(4), object(3)  
memory usage: 20.9+ KB
```

```
penguins.to_excel('penguins_size.xlsx')
```

[9] ✓ 1.1s Python



	<div><div>Loading JSON file and converting to CSV</div><div><pre>import json with open('./indian_states.json') as file:     json_data = json.load(file)</pre><div>[12] ✓ 0.0s Python</div></div><div><pre>data = {"State ID": [state['state_id'] for state in json_data],         "State Name": [state['state_name'] for state in json_data]}</pre><div>[13] ✓ 0.0s Python</div></div><div><div>▶</div><pre>indian_states = pd.DataFrame(data) indian_states.head()</pre><div>[14] ✓ 0.0s Python</div><div><table><thead><tr><th></th><th>State ID</th><th>State Name</th></tr></thead><tbody><tr><td>0</td><td>AN</td><td>Andaman and Nicobar Island (UT)</td></tr><tr><td>1</td><td>AP</td><td>Andhra Pradesh</td></tr><tr><td>2</td><td>AR</td><td>Arunachal Pradesh</td></tr><tr><td>3</td><td>AS</td><td>Assam</td></tr><tr><td>4</td><td>BR</td><td>Bihar</td></tr></tbody></table></div></div><div><div>▶</div><pre>indian_states.to_csv('indian_states.csv')</pre><div>[15] ✓ 0.0s Python</div></div></div>		State ID	State Name	0	AN	Andaman and Nicobar Island (UT)	1	AP	Andhra Pradesh	2	AR	Arunachal Pradesh	3	AS	Assam	4	BR	Bihar
	State ID	State Name																	
0	AN	Andaman and Nicobar Island (UT)																	
1	AP	Andhra Pradesh																	
2	AR	Arunachal Pradesh																	
3	AS	Assam																	
4	BR	Bihar																	
RESULT:	<p>Given CSV files were converted to excel sheets (.xlsx) format.</p> <p>Given JSON files were loaded then converted to pandas Dataframes</p>																		
References:																			
CONCLUSION: In this experiment I have learnt how to load CSV and JSON files and then convert them to other types using Pandas and JSON libraries in python.																			