

Creating and Accessing Pandas DataFrames	
<b>Course Code:</b> CPE 031	<b>Program:</b> Computer Engineering
<b>Course Title:</b> Visualization and Data Analysis	<b>Date Performed:</b> 10/15/24
<b>Section:</b> CPE21S4	<b>Date Submitted:</b> 10/15/24
<b>Name:</b> CALVIN EARL PLANTA	<b>Instructor:</b> Prof. Sayo
<b>Intended Learning Outcomes (ILO):</b>  By the end of this laboratory session, learners will be able to <ul style="list-style-type: none"> <li>- Construct and manipulate Pandas DataFrames from various data structures (such as lists, dictionaries, and NumPy arrays) while demonstrating an understanding of DataFrame attributes and methods. This includes loading the dataset, creating DataFrames with appropriate column labels and accessing data from rows and columns.</li> </ul>	
<b>Instructions:</b>  <ol style="list-style-type: none"> <li>1. <b>Loading your dataset:</b> Refer back to your chosen dataset from the PRELIM period. Whether you downloaded it or stored it in your Google Drive, you are required to load it into the <a href="#">Google Colab</a>. Watch this <a href="#">video</a> to learn more about how to read CSV files in Google Colab.(<b>Take a screenshot to document successful execution.</b>)</li> <li>2. <b>Creating a dataframe from your CSV file:</b> Once you have successfully loaded your dataset, you need to create a dataframe from your uploaded CSV file.(<b>Take a screenshot to document successful execution.</b>)</li> <li>3. <b>Creating a dataframe from a dictionary of lists:</b> Manually create a dictionary where each value is composed of a list from your original dataset, then load it into a dataframe, before printing it. You are required to provide at least five (5) observations in your list. (<b>Take a screenshot to document successful execution.</b>)</li> <li>4. <b>Creating a dataframe from a list of dictionaries:</b> Manually create a list of dictionaries from your original dataset, then pass it into a dataframe, before printing it. You are required to provide at least five (5) observations in your list. (<b>Take a screenshot to document successful execution.</b>)</li> <li>5. <b>Selecting dataframe columns:</b> Execute a method that would allow you to select a single and multiple dataframe columns. (<b>Take a screenshot to document successful execution.</b>)</li> <li>6. <b>Selecting dataframe rows:</b>Execute a method that would allow you to select a single and multiple dataframe rows using panda indexing and python indexing.</li> </ol>	

## Output:

### 1. Loading your dataset:

```
[8] #1. Loading your dataset
from google.colab import drive
drive.mount('/content/drive')
path = "/content/drive/MyDrive/ranking.csv"
data = pd.read_csv(path)
data.head(5)
```

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force\_remount=True).

	rank	country_rank	player_name	country	accuracy	play_count	level	hours	performance_points	ranked_score	ss	s	a	watched_by	total_hits	device
0	1	1	Cookiezi	South Korea	98.86	15245	101	324	13450	22401473154	55	454	325	3057421	6372761	tablet
1	2	2	Angelsim	South Korea	98.87	124726	103	1457	12950	35481061465	159	1369	679	498548	29845228	mouse
2	3	1	Rafis	Poland	99.03	208066	103	2229	12942	36343320002	170	1654	785	1170973	35760655	tablet
3	4	1	Vaxe1	United States	98.03	93012	101	1034	12735	18096595619	34	541	563	152551	15601281	tablet
4	5	1	hwick225	Taiwan	98.93	134052	102	1970	12421	47072997400	826	3466	701	2015824	22576332	tablet

Next steps: [Generate code with data](#) [View recommended plots](#) [New interactive sheet](#)

### 2. Creating a dataframe from your CSV file:

```
#2. Creating a dataframe from your CSV file
path = "/content/drive/MyDrive/ranking.csv"
data.to_csv("ranking.csv")
ranking = pd.DataFrame(data)
print(ranking)
```

	rank	country_rank	player_name	country	accuracy	play_count	\
0	1	1	Cookiezi	South Korea	98.86	15245	
1	2	2	Angelsim	South Korea	98.87	124726	
2	3	1	Rafis	Poland	99.03	208066	
3	4	1	Vaxe1	United States	98.03	93012	
4	5	1	hwick225	Taiwan	98.93	134052	
..	...	...	...	...	...	...	
95	96	1	[ Zane ]	Belize	98.83	118916	
96	97	13	Interrobang	South Korea	98.05	71221	
97	98	7	Splinter572	Germany	98.83	185067	
98	99	3	Korilak	Australia	99.30	52944	
99	100	4	[ ZhengS ]	Australia	97.79	130517	

	level	hours	performance_points	ranked_score	ss	s	a	\
0	101	324	13450	22401473154	55	454	325	
1	103	1457	12950	35481061465	159	1369	679	
2	103	2229	12942	36343320002	170	1654	785	
3	101	1034	12735	18096595619	34	541	563	
4	102	1970	12421	47072997400	826	3466	701	
..	...	...	...	...	...	...	...	
95	101	1498	9323	38211883962	118	1958	932	
96	100	814	9305	18771923513	76	945	746	
97	102	2156	9291	29820445045	677	1737	741	
98	100	778	9291	11706287563	33	384	431	
99	101	1669	9287	28191412866	36	1067	724	

	watched_by	total_hits	device
0	3057421	6372761	tablet
1	498548	29845228	mouse
2	1170973	35760655	tablet
3	152551	15601281	tablet
4	2015824	22576332	tablet
..	...	...	...
95	2711	23976180	tablet
96	36760	12698057	tablet
97	43660	30251174	tablet
98	393	12816470	tablet
99	5532	27500674	tablet

### 3. Creating a dataframe from a dictionary of lists

```
0s #3. Creating a dataframe from a dictionary of lists
data = {'rank': [1, 2, 3, 4, 5], 'country_rank': [1, 2, 3, 4, 5],
        'player_name': ['Cookiezi', 'Angelsim', 'Rafis', 'Vaxei', 'hwick225'],
        'country': ['South Korea', 'South Korea', 'Poland', 'United States', 'Taiwan'],
        'performance_points': [13450, 12950, 12942, 12735, 12421]}
ranking = pd.DataFrame(data)
print(ranking)
```

	rank	country_rank	player_name	country	performance_points
0	1	1	Cookiezi	South Korea	13450
1	2	2	Angelsim	South Korea	12950
2	3	3	Rafis	Poland	12942
3	4	4	Vaxei	United States	12735
4	5	5	hwick225	Taiwan	12421

#### 4. Creating a dataframe from a list of dictionaries

```
0s #4. Creating a dataframe from a list of dictionaries
data = [
    {'rank': 1, 'country_rank': 1, 'player_name': 'Cookiezi', 'country': 'South Korea', 'performance_points': 13450},
    {'rank': 2, 'country_rank': 2, 'player_name': 'Angelsim', 'country': 'South Korea', 'performance_points': 12950},
    {'rank': 3, 'country_rank': 3, 'player_name': 'Rafis', 'country': 'Poland', 'performance_points': 12942},
    {'rank': 4, 'country_rank': 4, 'player_name': 'Vaxei', 'country': 'United States', 'performance_points': 12735},
    {'rank': 5, 'country_rank': 5, 'player_name': 'hwick225', 'country': 'Taiwan', 'performance_points': 12421}
]
ranking = pd.DataFrame(data)
print(ranking)
```

	rank	country_rank	player_name	country	performance_points
0	1	1	Cookiezi	South Korea	13450
1	2	2	Angelsim	South Korea	12950
2	3	3	Rafis	Poland	12942
3	4	4	Vaxei	United States	12735
4	5	5	hwick225	Taiwan	12421

#### 5. Selecting dataframe columns

```
0s #5. Selecting dataframe columns
print(ranking['player_name']) # Single

print()

print(ranking[['player_name', 'country_rank']]) # Multiple
```

```
0    Cookiezi
1    Angelsim
2      Rafis
3     Vaxei
4   hwick225
Name: player_name, dtype: object
```

	player_name	country_rank
0	Cookiezi	1
1	Angelsim	2
2	Rafis	3
3	Vaxei	4
4	hwick225	5

## 6. Selecting dataframe rows

```
#6. Selecting dataframe rows
print(ranking.loc[2]) # Single

print()

print(ranking.loc[[2, 4]]) # Multiple
```

```
↔ rank      3
   country_rank  3
   player_name   Rafis
   country      Poland
   performance_points  12942
   Name: 2, dtype: object
```

	rank	country_rank	player_name	country	performance_points
2	3	3	Rafis	Poland	12942
4	5	5	hwick225	Taiwan	12421