Creating and Accessing Pandas DataFrames

Course Code: CPE 031	**Program**: Computer Engineering
Course Title: Visualization and Data Analysis	**Date Performed**: 10/15/24
Section: CPE21S4	**Date Submitted**:10/15/24
Name: Anna Marie Zolina	**Instructor**: Prof. Maria Rizette Sayo

Intended Learning Outcomes (ILO):

By the end of this laboratory session, learners will be able to

- 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.

Instructions:

1. **Loading your dataset**: 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 Google Colab. Watch this video to learn more about how to read CSV files in Google Colab.**(Take a screenshot to document successful execution.)**

2. **Creating a dataframe from your CSV file:** Once you have successfully loaded your dataset, you need to create a dataframe from your uploaded CSV file.**(Take a screenshot to document successful execution.)**

3. **Creating a dataframe from a dictionary of lists:** 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. **(Take a screenshot to document successful execution.)**

4. **Creating a dataframe from a list of dictionaries:** 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. **(Take a screenshot to document successful execution.)**

5. **Selecting dataframe columns:** Execute a method that would allow you to select a single and multiple dataframe columns. **(Take a screenshot to document successful execution.)**

6. **Selecting dataframe rows:**Execute a method that would allow you to select a single and multiple dataframe rows using panda indexing and python indexing.

Output:

1.

READING DATA FROM CSV

```
[2]  from google.colab import files
     uploaded = files.upload()

     import pandas as pd
     data = pd.read_csv('top_50_2023.csv')
```

15s

Choose Files top_50_2023.csv

- **top_50_2023.csv**(text/csv) - 8158 bytes, last modified: 10/15/2024 - 100% done

```
Saving top_50_2023.csv to top_50_2023 (1).csv
```

2.

```
[8]  purchases = pd.DataFrame(data)
     (purchases)
```

```
purchases = pd.DataFrame(data)
(purchases)
```

	song\n	artist(s)_name	streams
0	Blinding Lights	The Weeknd	3703895074
1	Someone You Loved	Lewis Capaldi	2887241814
2	Sunflower - Spider-Man: Into the Spider-Verse	Post Malone, Swae Lee	2808096550
3	STAY (with Justin Bieber)	Justin Bieber, The Kid Laroi	2665343922
4	Starboy	The Weeknd, Daft Punk	2565529693
5	Heat Waves	Glass Animals	2557975762
6	As It Was	Harry Styles	2513188493
7	Sweater Weather	The Neighbourhood	2282771485
8	Riptide	Vance Joy	2009094673
9	Another Love	Tom Odell	1813673666
10	Yellow	Chris Molitor	1755214421
11	Die For You	The Weeknd	1647990401
12	Viva La Vida	Coldplay	1592909789
13	Me Porto Bonito	Chencho Corleone, Bad Bunny	1440757818
14	Quevedo: Bzrp Music Sessions, Vol. 52	Bizarrap, Quevedo	1356565093
15	Blank Space	Taylor Swift	1355959075
16	Flowers	Miley Cyrus	1316855716
17	I Wanna Be Yours	Arctic Monkeys	1297026226
18	Unholy (feat. Kim Petras)	Sam Smith, Kim Petras	1230675890
19	La Bachata	Manuel Turizo	1214083358
20	Kill Bill	SZA	1163093654
21	I'm Good (Blue)	Bebe Rexha, David Guetta	1109433169
22	I Ain't Worried	OneRepublic	1085685420
23	See You Again	Tyler, The Creator, Kali Uchis	1047101291
24	Anti-Hero	Taylor Swift	999748277
25	Summertime Sadness	Lana Del Rey	983637508

26	Calm Down (with Selena Gomez)	Rema, Selena G	899183384
27	Creepin'	The Weeknd, 21 Savage, Metro Boomin	843957510
28	cardigan	Taylor Swift	812019557
29	Cruel Summer	Taylor Swift	800840817
30	Style	Taylor Swift	786181836
31	golden hour	JVKE	751134527
32	Ella Baila Sola	Eslabon Armado, Peso Pluma	725980112
33	Left and Right (Feat. Jung Kook of BTS)	Charlie Puth, BTS, Jung Kook	720434240
34	Don't Bl	Taylor Swift	685032533
35	Hey Mor	Ozuna, Feid	674072710
36	Here With Me	d4vd	635412045
37	TQG	Karol G, Shakira	618990393
38	Until I Found You (with Em Beihold) - Em Beiho...	Em Beihold, Stephen Sanchez	600976848
39	Yandel 150	Yandel, Feid	585695368
40	Angels Like You	Miley Cyrus	570515054
41	La Bebe - Remix	Peso Pluma, Yng Lvcas	553634067
42	Die For You - Remix	Ariana Grande, The Weeknd	518745108
43	un x100to	Bad Bunny, Grupo Frontera	505671438
44	Cupid - Twin Ver.	Fifty Fifty	496795686
45	Last Night	Morgan Wallen	429829812
46	Karma	Taylor Swift	404562836
47	Snooze	SZA	399686758
48	Daylight	David Kushner	387570742

3.

```
import pandas as pd
data = {
    'song': [
        'Blinding Lights',
        'Someone You Loved',
        'Sunflower',
        'STAY(with Justin Bieber)',
        'Starboy'
    ],
    'artist(s)_name': [
        'The Weeknd',
        'Lewis Capaldi',
        'Post Malone, Swae Lee',
        'Justin Bieber, The Kid Laroi',
        'The Weeknd, Daft Punk'
    ],
    'streams': [
        3703895074,
        2887241814,
        2808096550,
        2665343922,
        2565529693
    ]
}

df = pd.DataFrame(data)
print(df)
```

	song	artist(s)_name	streams
0	Blinding Lights	The Weeknd	3703895074
1	Someone You Loved	Lewis Capaldi	2887241814
2	Sunflower	Post Malone, Swae Lee	2808096550
3	STAY(with Justin Bieber)	Justin Bieber, The Kid Laroi	2665343922
4	Starboy	The Weeknd, Daft Punk	2565529693

4.

```
import pandas as pd

data = [
    {
        'song': 'Blinding Lights',
        'artist(s)_name': 'The Weeknd',
        'streams': 3703895074
    },
    {
        'song': 'Someone You Loved',
        'artist(s)_name': 'Lewis Capaldi',
        'streams': 2887241814
    },
    {
        'song': 'Sunflower - Spider-Man: Into the Spider-Verse',
        'artist(s)_name': 'Post Malone, Swae Lee',
        'streams': 2808096550
    },
    {
        'song': 'STAY (with Justin Bieber)',
        'artist(s)_name': 'Justin Bieber, The Kid Laroi',
        'streams': 2665343922
    },
    {
        'song': 'Starboy',
        'artist(s)_name': 'The Weeknd, Daft Punk',
        'streams': 2565529693
    }
]

df = pd.DataFrame(data)
print(df)
```

	song	artist(s)_name	streams
0	Blinding Lights	The Weeknd	3703895074
1	Someone You Loved	Lewis Capaldi	2887241814
2	Sunflower - Spider-Man: Into the Spider-Verse	Post Malone, Swae Lee	2808096550
3	STAY (with Justin Bieber)	Justin Bieber, The Kid Laroi	2665343922
4	Starboy	The Weeknd, Daft Punk	2565529693

5.

```
single_column = df['song']
(single_column)
```

	song
0	Blinding Lights
1	Someone You Loved
2	Sunflower
3	STAY(with Justin Bieber)
4	Starboy

```
multiple_column = df[['song', 'artist(s)_name', 'streams']]
(multiple_column)
```

	song	artist(s)_name	streams
0	Blinding Lights	The Weeknd	3703895074
1	Someone You Loved	Lewis Capaldi	2887241814
2	Sunflower	Post Malone, Swae Lee	2808096550
3	STAY(with Justin Bieber)	Justin Bieber, The Kid Laroi	2665343922
4	Starboy	The Weeknd, Daft Punk	2565529693

6.

```
[8]
single_row_pandas = df.loc[2]
single_row_pandas
```

	2
song	Sunflower - Spider-Man: Into the Spider-Verse
artist(s)_name	Post Malone, Swae Lee
streams	2808096550

dtype: object

```
multiple_rows_pandas = df.loc[1:4]
multiple_rows_pandas
```

	song	artist(s)_name	streams
1	Someone You Loved	Lewis Capaldi	2887241814
2	Sunflower - Spider-Man: Into the Spider-Verse	Post Malone, Swae Lee	2808096550
3	STAY (with Justin Bieber)	Justin Bieber, The Kid Laroi	2665343922
4	Starboy	The Weeknd, Daft Punk	2565529693