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: Magistrado, Aira Pauleen M.	**Instructor**: 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.)**

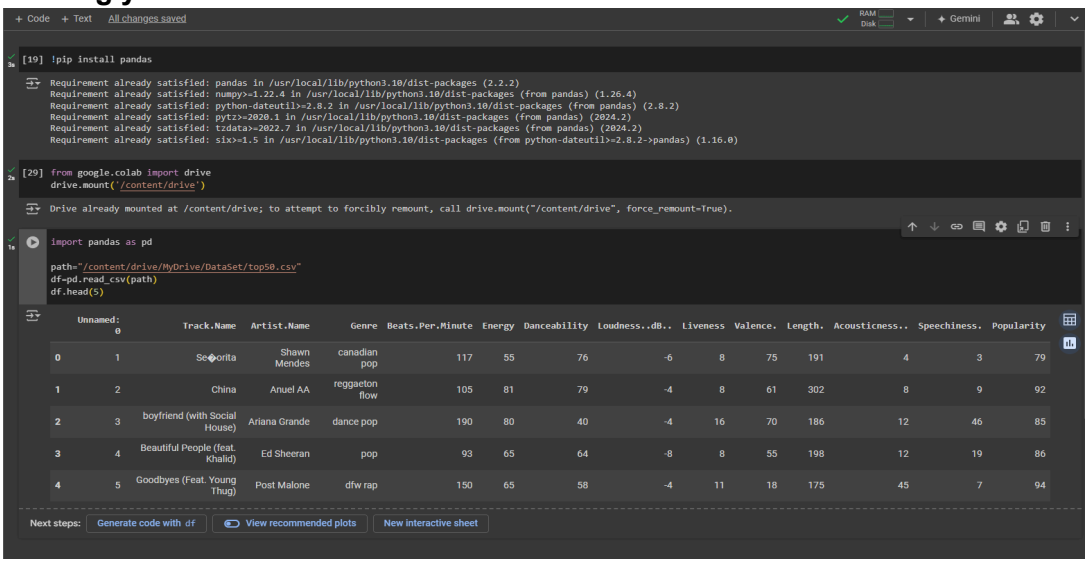
4. **Creating a dataframe from a list of dictionaries:** Manually create a list of dictionaries from your original dataset, then pass it into a dataframe, before printing it. You are required to provide at least five (5) observations in your list. **(Take a screenshot to document successful execution.)**

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. **Loading your dataset**



```
[19] !pip install pandas

Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (2.2.2)
Requirement already satisfied: numpy>=1.22.4 in /usr/local/lib/python3.10/dist-packages (from pandas) (1.26.4)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas) (2024.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.10/dist-packages (from pandas) (2024.2)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)

[29] from google.colab import drive
     drive.mount('/content/drive')

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

import pandas as pd

path="/content/drive/MyDrive/DataSet/top50.csv"
df=pd.read_csv(path)
df.head(5)
```

	Unnamed: 0	Track.Name	Artist.Name	Genre	Beats.Per.Minute	Energy	Danceability	Loudness..dB..	Liveness	Valence.	Length.	Acousticness..	Speechiness.	Popularity
0	1	Se♦orita	Shawn Mendes	canadian pop	117	55	76	-6	8	75	191	4	3	79
1	2	China	Anuel AA	reggaeton flow	105	81	79	-4	8	61	302	8	9	92
2	3	boyfriend (with Social House)	Ariana Grande	dance pop	190	80	40	-4	16	70	186	12	46	85
3	4	Beautiful People (feat. Khalid)	Ed Sheeran	pop	93	65	64	-8	8	55	198	12	19	86
4	5	Goodbyes (Feat. Young Thug)	Post Malone	dfw rap	150	65	58	-4	11	18	175	45	7	94

Next steps: Generate code with df | View recommended plots | New interactive sheet

2. **Creating a dataframe from your CSV file**

```python
import pandas as pd
path = "/content/drive/MyDrive/DataSet/top50.csv"
df = pd.read_csv(path)
print(df)
```

```
    Unnamed: 0                                         Track.Name  \
0            1                                        Se�orita
1            2                                              China
2            3                          boyfriend (with Social House)
3            4                        Beautiful People (feat. Khalid)
4            5                            Goodbyes (Feat. Young Thug)
5            6                       I Don't Care (with Justin Bieber)
6            7                                             Ransom
7            8                                  How Do You Sleep?
8            9                               Old Town Road - Remix
9           10                                             bad guy
10          11                                            Callaita
11          12                  Loco Contigo (feat. J. Balvin & Tyga)
12          13                                 Someone You Loved
13          14                               Otro Trago - Remix
14          15               Money In The Grave (Drake ft. Rick Ross)
15          16                               No Guidance (feat. Drake)
16          17                                        LA CANCI�N
17          18           Sunflower - Spider-Man: Into the Spider-Verse
18          19                                             Lalala
19          20                                         Truth Hurts
20          21                               Piece Of Your Heart
21          22                                             Panini
22          23                               No Me Conoce - Remix
23          24                                   Soltera - Remix
24          25                          bad guy (with Justin Bieber)
25          26                               If I Can't Have You
26          27                                      Dance Monkey
27          28                                         It's You
28          29                                        Con Calma
29          30                                      QUE PRETENDES
30          31                                         Takeaway
31          32                                          7 rings
32          33                                0.958333333333333
33          34               The London (feat. J. Cole & Travis Scott)
34          35                                 Never Really Over
35          36  Summer Days (feat. Macklemore & Patrick Stump ...
36          37                                        Otro Trago
37          38                         Antisocial (with Travis Scott)
38          39                                           Sucker
39          40  fuck, i'm lonely (with Anne-Marie) - from �13 ...
40          41                                       Higher Love
41          42                           You Need To Calm Down
42          43                                         Shallow
```

3. Creating a dataframe from a dictionary of lists:

```
import pandas as pd
data = {'Track.Name': ['Senorita', 'China', 'boyfriend (with Social House)', 'Beautiful People (feat. Khalid)', 'Goodbyes (Feat. Young Thug)'],
        'Artist.Name': ['Shawn Mendes', 'Anuel AA', 'Ariana Grande', 'Ed Sheeran', 'Post Malone'],
        'Genre': ['canadian pop', 'reggaeton flow', 'dance pop', 'pop', 'dfw rap'],
        'Beats.Per.Minute': [117, 105, 190, 93, 150],
        'Energy': [55, 81, 80, 65, 65],
        'Danceability': [76, 79, 40, 64, 58],
        'Loudness..dB..': [-6, -4, -4, -8, -4],
        'Liveness': [8, 8, 16, 8, 11]}
df = pd.DataFrame(data)
print(df)
```

```
                      Track.Name    Artist.Name           Genre  \
0                       Senorita   Shawn Mendes    canadian pop
1                          China       Anuel AA  reggaeton flow
2    boyfriend (with Social House)  Ariana Grande       dance pop
3  Beautiful People (feat. Khalid)     Ed Sheeran             pop
4      Goodbyes (Feat. Young Thug)    Post Malone         dfw rap

   Beats.Per.Minute  Energy  Danceability  Loudness..dB..  Liveness
0               117      55            76              -6         8
1               105      81            79              -4         8
2               190      80            40              -4        16
3                93      65            64              -8         8
4               150      65            58              -4        11
```

4. Creating a dataframe from a list of dictionaries

```
import pandas as pd
data = [{'Track.Name': 'Senorita', 'Artist.Name': 'Shawn Mendes', 'Genre': 'canadian pop', 'Beats.Per.Minute': 117, 'Energy': 55, 'Danceability': 76, 'Loudness..dB..': -6, 'Liveness': 8},
        {'Track.Name': 'China', 'Artist.Name': 'Anuel AA', 'Genre': 'reggaeton flow', 'Beats.Per.Minute': 105, 'Energy': 81, 'Danceability': 79, 'Loudness..dB..': -4, 'Liveness': 8},
        {'Track.Name': 'boyfriend (with Social House)', 'Artist.Name': 'Ariana Grande', 'Genre': 'dance pop', 'Beats.Per.Minute': 190, 'Energy': 80, 'Danceability': 40, 'Loudness..dB..': -4, '
        {'Track.Name': 'Beautiful People (feat. Khalid)', 'Artist.Name': 'Ed Sheeran', 'Genre': 'pop', 'Beats.Per.Minute': 93, 'Energy': 65, 'Danceability': 64, 'Loudness..dB..': -8, 'Liveness
        {'Track.Name': 'Goodbyes (Feat. Young Thug)', 'Artist.Name': 'Post Malone', 'Genre': 'dfw rap', 'Beats.Per.Minute': 150, 'Energy': 65, 'Danceability': 58, 'Loudness..dB..': -4, 'Livene
df = pd.DataFrame(data)
print(df)
```

```
                      Track.Name    Artist.Name           Genre  \
0                       Senorita   Shawn Mendes    canadian pop
1                          China       Anuel AA  reggaeton flow
2    boyfriend (with Social House)  Ariana Grande       dance pop
3  Beautiful People (feat. Khalid)     Ed Sheeran             pop
4      Goodbyes (Feat. Young Thug)    Post Malone         dfw rap

   Beats.Per.Minute  Energy  Danceability  Loudness..dB..  Liveness
0               117      55            76              -6         8
1               105      81            79              -4         8
2               190      80            40              -4        16
3                93      65            64              -8         8
4               150      65            58              -4        11
```

5. Selecting dataframe columns

```
[54] import pandas as pd
     data = [{'Track.Name': 'Senorita', 'Artist.Name': 'Shawn Mendes', 'Genre': 'canadian pop', 'Beats.Per.Minute': 117, 'Energy': 55, 'Danceability': 76, 'Loudness..dB..': -6, 'Liveness': 8},
             {'Track.Name': 'China', 'Artist.Name': 'Anuel AA', 'Genre': 'reggaeton flow', 'Beats.Per.Minute': 105, 'Energy': 81, 'Danceability': 79, 'Loudness..dB..': -4, 'Liveness': 8},
             {'Track.Name': 'boyfriend (with Social House)', 'Artist.Name': 'Ariana Grande', 'Genre': 'dance pop', 'Beats.Per.Minute': 190, 'Energy': 80, 'Danceability': 40, 'Loudness..dB..': -4, '
             {'Track.Name': 'Beautiful People (feat. Khalid)', 'Artist.Name': 'Ed Sheeran', 'Genre': 'pop', 'Beats.Per.Minute': 93, 'Energy': 65, 'Danceability': 64, 'Loudness..dB..': -8, 'Liveness
             {'Track.Name': 'Goodbyes (Feat. Young Thug)', 'Artist.Name': 'Post Malone', 'Genre': 'dfw rap', 'Beats.Per.Minute': 150, 'Energy': 65, 'Danceability': 58, 'Loudness..dB..': -4, 'Livene
     df = pd.DataFrame(data)

[55] print(df[['Track.Name', 'Artist.Name']])

                      Track.Name    Artist.Name
0                       Senorita   Shawn Mendes
1                          China       Anuel AA
2    boyfriend (with Social House)  Ariana Grande
3  Beautiful People (feat. Khalid)     Ed Sheeran
4      Goodbyes (Feat. Young Thug)    Post Malone

[52] print(df['Genre'])

0      canadian pop
1    reggaeton flow
2         dance pop
3               pop
4           dfw rap
Name: Genre, dtype: object
```

6.

```
[69] first_row = df.iloc[0]
     print(first_row)
```

```
Track.Name              Senorita
Artist.Name         Shawn Mendes
Genre                canadian pop
Beats.Per.Minute              117
Energy                         55
Danceability                   76
Loudness..dB..                 -6
Liveness                        8
Name: 0, dtype: object
```

```
[64] first_row = df.loc[0]
     print(first_row)
```

```
Track.Name              Senorita
Artist.Name         Shawn Mendes
Genre                canadian pop
Beats.Per.Minute              117
Energy                         55
Danceability                   76
Loudness..dB..                 -6
Liveness                        8
Name: 0, dtype: object
```

```
[75] tworows = df.loc[0:2, ['Track.Name', 'Artist.Name']]
     print(tworows)
```

```
                      Track.Name    Artist.Name
0                       Senorita   Shawn Mendes
1                          China       Anuel AA
2  boyfriend (with Social House)  Ariana Grande
```

```
[73] tworows = df.iloc[0:2, 0:2]
     print(tworows)
```

```
  Track.Name   Artist.Name
0   Senorita  Shawn Mendes
1      China      Anuel AA
```