In [1]:
```
# Dependency needed to install file

!pip install xlrd
```

```
Requirement already satisfied: xlrd in ./opt/anaconda3/lib/python3.9/site-packages (2.0.1)
```

In [2]:
```
# Import required library

import pandas as pd
```

In [3]:
```
# Read data from CSV file

csv_path = 'https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-PY0101
df = pd.read_csv(csv_path)
```

In [9]:
```
#We can access the column Length and assign it a new dataframe x

x = df [['Length']]
x
```

Out[9]:

| | Length |
|---|---|
| 0 | 0:42:19 |
| 1 | 0:42:11 |
| 2 | 0:42:49 |
| 3 | 0:57:44 |
| 4 | 0:46:33 |
| 5 | 0:43:08 |
| 6 | 1:15:54 |
| 7 | 0:40:01 |

In [10]:
```
# Get the column as a series. Just use one bracket:
x = df ['Length']
x
```

Out[10]:
```
0    0:42:19
1    0:42:11
2    0:42:49
3    0:57:44
4    0:46:33
5    0:43:08
6    1:15:54
7    0:40:01
Name: Length, dtype: object
```

In [11]:
```
type(x)
```

Out[11]:
```
pandas.core.series.Series
```

In [23]:
```
#You can do the same thing for multiple columns; we just put the dataframe name, in this case, df, and the name

# Access to multiple columns



y = df[['Artist','Length','Genre', 'Released.1']]
y
```

Out[23]:

| | Artist | Length | Genre | Released.1 |
|---|---|---|---|---|
| 0 | Michael Jackson | 0:42:19 | pop, rock, R&B | 30-Nov-82 |
| 1 | AC/DC | 0:42:11 | hard rock | 25-Jul-80 |
| 2 | Pink Floyd | 0:42:49 | progressive rock | 01-Mar-73 |
| 3 | Whitney Houston | 0:57:44 | R&B, soul, pop | 17-Nov-92 |
| 4 | Meat Loaf | 0:46:33 | hard rock, progressive rock | 21-Oct-77 |
| 5 | Eagles | 0:43:08 | rock, soft rock, folk rock | 17-Feb-76 |
| 6 | Bee Gees | 1:15:54 | disco | 15-Nov-77 |
| 7 | Fleetwood Mac | 0:40:01 | soft rock | 04-Feb-77 |

In [29]:
```
#One way to access unique elements is the iloc method, where you can access the 1st row and the 3rd column as f
df.iloc[0,4]
```

Out[29]:
```
'pop, rock, R&B'
```

In [30]:
```
# Access the value on the first row and the first column
```

```
df.iloc[0, 0]
```

Out[30]:
```
'Michael Jackson'
```

In [34]:
```
#You can access the 2nd row and the 1st column as follows:
df.iloc [1,0]
```

Out[34]:
```
'AC/DC'
```

In [35]:
```
#You can access the 1st row and the 3rd column as follows:
df.iloc[0,2]
```

Out[35]:
```
1982
```

In [36]:
```
# Access the value on the second row and the third column
df.iloc[1,2]
```

Out[36]:
```
1980
```

In [37]:
```
#You can access the column using the name as well, the following are the same as above:
# Access the column using the name

df.loc[1, 'Artist']
```

Out[37]:
```
'AC/DC'
```

In [39]:
```
# Access the column using the name

df.loc[0, 'Released']
```

Out[39]:
```
1982
```

In [40]:
```
# Access the column using the name

df.loc[1, 'Released']
```

Out[40]:
```
1980
```

In [46]:
```
#You can perform slicing using both the index and the name of the column:

z = df.iloc[0:4, 2:5]
z
```

Out[46]:

| | Released | Length | Genre |
|---|---|---|---|
| **0** | 1982 | 0:42:19 | pop, rock, R&B |
| **1** | 1980 | 0:42:11 | hard rock |
| **2** | 1973 | 0:42:49 | progressive rock |
| **3** | 1992 | 0:57:44 | R&B, soul, pop |

In [49]:
```
# Slicing the dataframe using name
df.loc [0:4, 'Genre']
```

Out[49]:
```
0                  pop, rock, R&B
1                       hard rock
2                  progressive rock
3                    R&B, soul, pop
4    hard rock, progressive rock
Name: Genre, dtype: object
```

In [51]:
```
# Slicing the dataframe using name
df.loc [0:4,'Length': 'Released.1']
```

Out[51]:

| | Length | Genre | Music Recording Sales (millions) | Claimed Sales (millions) | Released.1 |
|---|---|---|---|---|---|
| **0** | 0:42:19 | pop, rock, R&B | 46.0 | 65 | 30-Nov-82 |
| **1** | 0:42:11 | hard rock | 26.1 | 50 | 25-Jul-80 |
| **2** | 0:42:49 | progressive rock | 24.2 | 45 | 01-Mar-73 |
| **3** | 0:57:44 | R&B, soul, pop | 27.4 | 44 | 17-Nov-92 |
| **4** | 0:46:33 | hard rock, progressive rock | 20.6 | 43 | 21-Oct-77 |

In [53]:
```
#Use a variable q to store the column Rating as a dataframe

q = df [['Rating']]
q
```

Out[53]:

| | Rating |
|---|---|
| **0** | 10.0 |
| **1** | 9.5 |
| **2** | 9.0 |
| **3** | 8.5 |
| **4** | 8.0 |
| **5** | 7.5 |
| **6** | 7.0 |
| **7** | 6.5 |

In [55]:
```
#Assign the variable q to the dataframe that is made up of the column Released and Artist:

q = df [['Released', 'Artist']]
q
```

Out[55]:

| | Released | Artist |
|---|---|---|
| **0** | 1982 | Michael Jackson |
| **1** | 1980 | AC/DC |
| **2** | 1973 | Pink Floyd |
| **3** | 1992 | Whitney Houston |
| **4** | 1977 | Meat Loaf |
| **5** | 1976 | Eagles |
| **6** | 1977 | Bee Gees |
| **7** | 1977 | Fleetwood Mac |

In [56]:
```
#Access the 2nd row and the 3rd column of df:

df.iloc[1,2]
```

Out[56]:
```
1980
```

In [57]:
```
#Use the following list to convert the dataframe index df to characters and
#assign it to df_new; find the element corresponding to the row index a and column 'Artist'.
#Then select the rows a through d for the column 'Artist

new_index=['a','b','c','d','e','f','g','h']

df_new=df
df_new.index=new_index
df_new.loc['a', 'Artist']
df_new.loc['a':'d', 'Artist']
```

Out[57]:
```
a    Michael Jackson
b              AC/DC
c         Pink Floyd
d    Whitney Houston
Name: Artist, dtype: object
```

In [ ]: