Lab 5 24-08-2022

August 24, 2022

[21]:
```
import pandas as pd
import numpy as np
```

1 1.

1.0.1 Write a Pandas program to create and display a one-dimensional array like object containing an array of data

[5]:
```
arr = [1,2,3]
a = pd.Series(arr)
print(a)
```

```
0    1
1    2
2    3
dtype: int64
```

[4]:
```
ar = [11,10,8]
b = pd.Series(ar, index = ["x","y","z"])
print(b)
```

```
x    11
y    10
z     8
dtype: int64
```

1.1 2.

1.1.1 Write a Pandas program to add, subtract, multiply and divide two Pandas Series

[13]:
```
c = pd.Series([2,4,8,16])
d = pd.Series([3,6,9,12])
print(c)
print(d)
print("sum is:\n",c+d)
print("difference is:\n",c-d)
print("product is:\n",c*d)
```

```
print("division is:\n",c/d)
```

```
0     2
1     4
2     8
3    16
dtype: int64
0     3
1     6
2     9
3    12
dtype: int64
sum is:
 0     5
1    10
2    17
3    28
dtype: int64
difference is:
 0   -1
1   -2
2   -1
3    4
dtype: int64
product is:
 0      6
1     24
2     72
3    192
dtype: int64
division is:
 0    0.666667
1    0.666667
2    0.888889
3    1.333333
dtype: float64
```

1.2 3.

1.2.1 write a panda program to create a single dataframe and save it as a csv file

[36]:
```
d = { "calories": [420, 380, 390],  "duration": [50, 40, 45]}
df1 = pd.DataFrame(d)
print(df1)
df1.to_csv('data.csv', index=False)
dp1 = pd.read_csv('data.csv')
print(dp1)
```

```
   calories  duration
0       420        50
1       380        40
2       390        45
   calories  duration
0       420        50
1       380        40
2       390        45
```

[32]:
```
data = {'name':␣
 ↪["Appu","Ammu","Achu","Kuttan","Pachu","Pathu","Kichu","Nick","Nani","Minnu"]␣
 ↪,
        'score': [75,89,np.nan,50,np.nan,41,38,np.nan,99,59],
        'attempts': [1, 3, 2, 1, 4, 3, 1, 1, 2, 3],
        'qualify': ['yes', 'yes', 'no', 'yes', 'no', 'yes', 'no', 'no', 'yes',␣
 ↪'yes']}
df = pd.DataFrame(data)
print(df)
print(df.info())
print("total rows: ",len(df.axes[0]))
print("total colum is: ",len(df.axes[1]))
print(df[df['score'].isnull()])
print("\nCalling first object\n",df.loc[0])
```

```
     name  score  attempts qualify
0    Appu   75.0         1     yes
1    Ammu   89.0         3     yes
2    Achu    NaN         2      no
3  Kuttan   50.0         1     yes
4   Pachu    NaN         4      no
5   Pathu   41.0         3     yes
6   Kichu   38.0         1      no
7    Nick    NaN         1      no
8    Nani   99.0         2     yes
9   Minnu   59.0         3     yes
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 4 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   name      10 non-null     object
 1   score     7 non-null      float64
 2   attempts  10 non-null     int64
 3   qualify   10 non-null     object
dtypes: float64(1), int64(1), object(2)
memory usage: 448.0+ bytes
None
total rows:  10
```

```
total colum is:  4
    name  score  attempts qualify
2   Achu    NaN         2      no
4  Pachu    NaN         4      no
7   Nick    NaN         1      no

Calling first object
 name        Appu
score       75.0
attempts       1
qualify      yes
Name: 0, dtype: object
```

[34]:
```
data = {'name':␣
 ↪["Appu","Ammu","Achu","Kuttan","Pachu","Pathu","Kichu","Nick","Nani","Minnu"]␣
 ↪,
        'score': [75,89,np.nan,50,np.nan,41,38,np.nan,99,59],
        'attempts': [1, 3, 2, 1, 4, 3, 1, 1, 2, 3],
        'qualify': ['yes', 'yes', 'no', 'yes', 'no', 'yes', 'no', 'no', 'yes',␣
 ↪'yes']}
df = pd.DataFrame(data)
df.to_csv('data.csv', index=False)
dp = pd.read_csv('data.csv')
print(dp)
```

```
     name  score  attempts qualify
0    Appu   75.0         1     yes
1    Ammu   89.0         3     yes
2    Achu    NaN         2      no
3  Kuttan   50.0         1     yes
4   Pachu    NaN         4      no
5   Pathu   41.0         3     yes
6   Kichu   38.0         1      no
7    Nick    NaN         1      no
8    Nani   99.0         2     yes
9   Minnu   59.0         3     yes
```

1.3 4.

1.3.1 Write a Pandas program to create and display a DataFrame from a specified dictionary data(10 records) which has the fields

1.3.2 1)name

1.3.3 2)Score

1.3.4 3)Attempts

1.3.5 4)Qualify

a)select the 'name' and 'qualify' columns from the following DataFrame

b)select the rows where the number of attempts in the examination is greater than 2

c)count the number of rows and columns of the DataFrame

d)Number of attempts in the examination is less than 2 and score greater than 10

[65]:
```
data = {'name':␣
 ↪["Appu","Ammu","Achu","Kuttan","Pachu","Pathu","Kichu","Nick","Nani","Minnu"]␣
 ↪,
        'score': [75,89,np.nan,50,np.nan,41,38,np.nan,99,59],
        'attempts': [1, 3, 2, 1, 4, 3, 1, 1, 2, 3],
        'qualify': ['yes', 'yes', 'no', 'yes', 'no', 'yes', 'no', 'no', 'yes',␣
 ↪'yes']}
df = pd.DataFrame(data)
# a
print(df[['name','qualify']])
# b
print("\n",df[df['attempts']>2])
# c
print("\ntotal rows: ",len(df.axes[0]))
print("total colum is: ",len(df.axes[1]))
# d
print(df[(df['attempts'] < 2) & (df['score'] > 10)])
```

```
     name qualify
0    Appu     yes
1    Ammu     yes
2    Achu      no
3  Kuttan     yes
4   Pachu      no
5   Pathu     yes
6   Kichu      no
7    Nick      no
8    Nani     yes
9   Minnu     yes
```

```
     name  score  attempts qualify
1   Ammu   89.0         3     yes
4  Pachu    NaN         4      no
5  Pathu   41.0         3     yes
9  Minnu   59.0         3     yes

total rows:  10
total colum is:  4
     name  score  attempts qualify
0    Appu   75.0         1     yes
3  Kuttan   50.0         1     yes
6   Kichu   38.0         1      no
```