dhruv-21bce0865-assignment1

September 6, 2023

Importing python libraries

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

Task 1- Creating a pandas dataframe with name as 'df' having 10 observation and 5 features

[ ]:
```
df = pd.DataFrame([[1,10,100,1000,10000],
 [2,20,200,2000,20000],
  [3,30,300,3000,30000],
   [4,40,400,4000,40000],
    [5,50,500,5000,50000],
     [6,60,600,np.nan,60000],
      [7,70,700,7000,70000],
       [8,80,800,8000,80000],
        [9,90,900,9000,90000],
         [11,111,1111,11111,111111]])
df
```

[ ]:

	0	1	2	3	4
0	1	10	100	1000.0	10000
1	2	20	200	2000.0	20000
2	3	30	300	3000.0	30000
3	4	40	400	4000.0	40000
4	5	50	500	5000.0	50000
5	6	60	600	NaN	60000
6	7	70	700	7000.0	70000
7	8	80	800	8000.0	80000
8	9	90	900	9000.0	90000
9	11	111	1111	11111.0	111111

[ ]:
```
df.columns = ['Age', 'Height', 'Weight', 'Savings', 'Salary']
df.index = ['Person_1', 'Person_2', 'Person_3', 'Person_4', 'Person_5',
            'Person_6', 'Person_7', 'Person_8', 'Person_9', 'Person_10']
df
```

[ ]:

	Age	Height	Weight	Savings	Salary
Person_1	1	10	100	1000.0	10000

```
Person_2     2      20     200   2000.0   20000
Person_3     3      30     300   3000.0   30000
Person_4     4      40     400   4000.0   40000
Person_5     5      50     500   5000.0   50000
Person_6     6      60     600      NaN   60000
Person_7     7      70     700   7000.0   70000
Person_8     8      80     800   8000.0   80000
Person_9     9      90     900   9000.0   90000
Person_10   11     111    1111  11111.0  111111
```

Task 2- Checking info of dataframe

[ ]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 10 entries, Person_1 to Person_10
Data columns (total 5 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   Age      10 non-null     int64
 1   Height   10 non-null     int64
 2   Weight   10 non-null     int64
 3   Savings  9 non-null      float64
 4   Salary   10 non-null     int64
dtypes: float64(1), int64(4)
memory usage: 480.0+ bytes
```

Task 3- Description of dataframe

[ ]:
```
description=df.describe()
description
```

[ ]:
```
              Age      Height       Weight       Savings         Salary
count   10.000000   10.000000    10.000000      9.000000      10.000000
mean     5.600000   56.100000   561.100000   5567.888889   56111.100000
std      3.204164   32.229903   322.488398   3417.635164   32250.923999
min      1.000000   10.000000   100.000000   1000.000000   10000.000000
25%      3.250000   32.500000   325.000000   3000.000000   32500.000000
50%      5.500000   55.000000   550.000000   5000.000000   55000.000000
75%      7.750000   77.500000   775.000000   8000.000000   77500.000000
max     11.000000  111.000000  1111.000000  11111.000000  111111.000000
```

Task 4- Checking the 4th index observation with 'loc' slicing operator.

[ ]:
```
four=df.loc['Person_4']
four
```

[ ]:
```
Age            4.0
Height        40.0
Weight       400.0
Savings     4000.0
Salary     40000.0
Name: Person_4, dtype: float64
```

Task 5- Checking Null values in our dataframe

[ ]:
```
null_values=df.isnull()
null_values
```

[ ]:

	Age	Height	Weight	Savings	Salary
Person_1	False	False	False	False	False
Person_2	False	False	False	False	False
Person_3	False	False	False	False	False
Person_4	False	False	False	False	False
Person_5	False	False	False	False	False
Person_6	False	False	False	True	False
Person_7	False	False	False	False	False
Person_8	False	False	False	False	False
Person_9	False	False	False	False	False
Person_10	False	False	False	False	False

[ ]:
```
null_values1=df.isnull().any()
null_values1
```

[ ]:
```
Age        False
Height     False
Weight     False
Savings     True
Salary     False
dtype: bool
```