NAME: ATUL B PILLAI

REGISTRATION NUMBER: 21BEC0130

VIT EMAIL: atul.bpillia2021@vitstudent.ac.in

Campus: VELLORE

ASSIGNMENT – 1

TASK 1:

Create a pandas data frame (DataFrame name as 'df') (10 observation and 5 features)

TASK- 2:

Check the info of 'df'

TASK 3:
Check the descriptive statistics of 'df'

TASK 4:

check the 4th index observation with 'loc' slicing operator.

TASK 5:

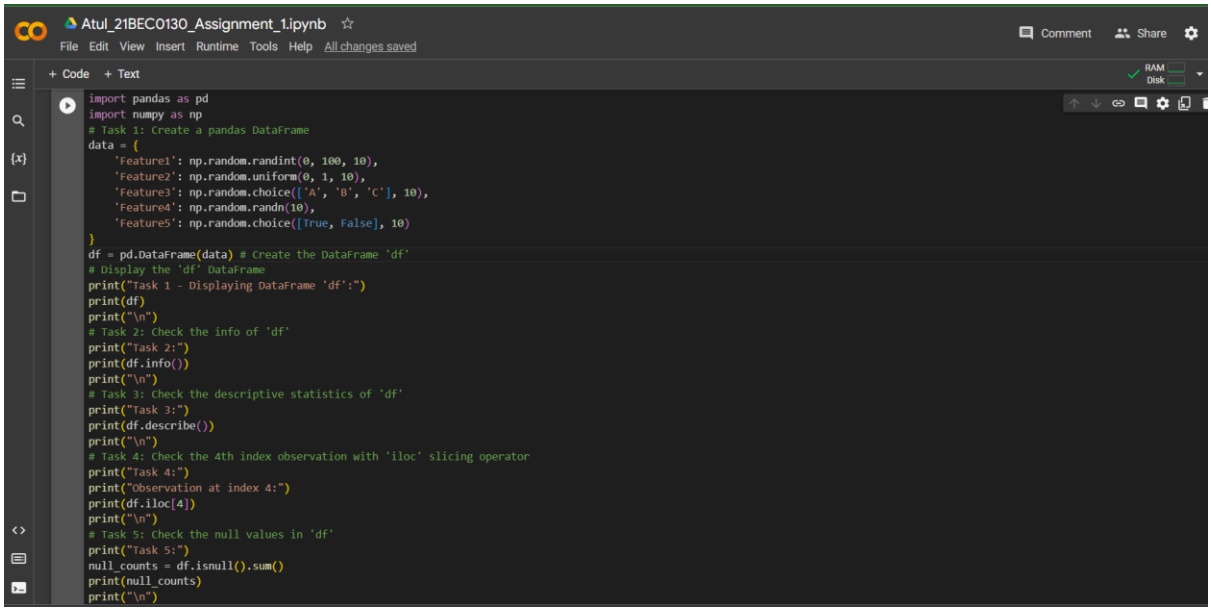
Check the null values in your 'df'

LINK OF THE COLAB NOTEBOOK:

https://colab.research.google.com/drive/1c_-btnET4yKXHDtpL2A59IoVIhPbFdBx?usp=drive_link

Atul B Pillai
21BEC0130

PROGRAM AND OUTPUT:




```python
import pandas as pd
import numpy as np
# Task 1: Create a pandas DataFrame
data = {
    'Feature1': np.random.randint(0, 100, 10),
    'Feature2': np.random.uniform(0, 1, 10),
    'Feature3': np.random.choice(['A', 'B', 'C'], 10),
    'Feature4': np.random.randn(10),
    'Feature5': np.random.choice([True, False], 10)
}
df = pd.DataFrame(data) # Create the DataFrame 'df'
# Display the 'df' DataFrame
print("Task 1 - Displaying DataFrame 'df':")
print(df)
print("\n")
# Task 2: Check the info of 'df'
print("Task 2:")
print(df.info())
print("\n")
# Task 3: Check the descriptive statistics of 'df'
print("Task 3:")
print(df.describe())
print("\n")
# Task 4: Check the 4th index observation with 'iloc' slicing operator
print("Task 4:")
print("Observation at index 4:")
print(df.iloc[4])
print("\n")
# Task 5: Check the null values in 'df'
print("Task 5:")
null_counts = df.isnull().sum()
print(null_counts)
print("\n")
```

```
Task 1 - Displaying DataFrame 'df':
   Feature1  Feature2 Feature3  Feature4  Feature5
0        39  0.422222        B -0.491143      True
1        10  0.714849        B  0.947392     False
2        55  0.857070        C -0.205462      True
3        80  0.590314        A -0.538208      True
4        44  0.242992        B  0.080077     False
5        97  0.923399        A  0.114977     False
6        90  0.893006        C -3.304104      True
7        73  0.001185        C  1.891089     False
8        20  0.161239        C  0.906279     False
9        62  0.835991        B -0.336108      True


Task 2:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 5 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   Feature1  10 non-null     int64
 1   Feature2  10 non-null     float64
 2   Feature3  10 non-null     object
 3   Feature4  10 non-null     float64
 4   Feature5  10 non-null     bool
dtypes: bool(1), float64(2), int64(1), object(1)
memory usage: 458.0+ bytes
None
```

```
Task 3:
       Feature1   Feature2   Feature4
count  10.000000  10.000000  10.000000
mean   57.000000   0.564227  -0.093521
std    29.009577   0.337121   1.367498
min    10.000000   0.001185  -3.304104
25%    40.250000   0.287799  -0.452384
50%    58.500000   0.652581  -0.062692
75%    78.250000   0.851800   0.708454
max    97.000000   0.923399   1.891089


Task 4:
Observation at index 4:
Feature1         44
Feature2    0.242992
Feature3           B
Feature4    0.080077
Feature5       False
Name: 4, dtype: object


Task 5:
Feature1    0
Feature2    0
Feature3    0
Feature4    0
Feature5    0
dtype: int64
```