ASSIGNMENT-1

NAME: EDE RENUKA MADHAV

REG NO: 21BCT0244

PHONE NO: 9059378158

VIT VELLORE CAMPUS

SLOT: 6:00 PM TO 8:00 PM

Task - 1 Create a pandas dataframe (DataFrame name as 'df') (10 observation and 5 features)

```
import numpy as np
df = {'Name':['Rohit Sharma','Shikhar Dhawan','Virat
Kohli','Madhav','KL Rahul','Hardik Pandya','Ravindra Jadeja','Yuzi
Chahal','Jasprit Bumrah','Mohammad Siraj'],'Age':
[36,36,34,np.nan,32,32,34,31,29,28],'Gender':
['M','M','M','M','M','M','M','M','M','M'],'Role':
['BATTER','BATTER','BATTER','ALL-ROUNDER','WICKET-KEPPER','ALL-
ROUNDER','ALL-ROUNDER','BOWLER','BOWLER','BOWLER'],'COUNTRY':
['INDIA','INDIA','INDIA','INDIA','INDIA','INDIA','INDIA','INDIA','INDI
A','INDIA']}
df
```

```
{'Name': ['Rohit Sharma',
  'Shikhar Dhawan',
  'Virat Kohli',
  'Madhav',
  'KL Rahul',
  'Hardik Pandya',
  'Ravindra Jadeja',
  'Yuzi Chahal',
  'Jasprit Bumrah',
  'Mohammad Siraj'],
 'Age': [36, 36, 34, nan, 32, 32, 34, 31, 29, 28],
 'Gender': ['M', 'M', 'M', 'M', 'M', 'M', 'M', 'M', 'M', 'M'],
 'Role': ['BATTER',
  'BATTER',
  'ALL-ROUNDER',
  'WICKET-KEPPER',
  'ALL-ROUNDER',
  'ALL-ROUNDER',
  'BOWLER',
  'BOWLER',
  'BOWLER'],
 'COUNTRY': ['INDIA',
  'INDIA',
  'INDIA',
  'INDIA',
  'INDIA',
  'INDIA',
  'INDIA',
  'INDIA',
  'INDIA',
  'INDIA']}
```

```
import pandas as pd
```

```
df1 = pd.DataFrame(df)
df1
```

```
               Name   Age Gender           Role COUNTRY
0     Rohit Sharma  36.0      M         BATTER   INDIA
1   Shikhar Dhawan  36.0      M         BATTER   INDIA
2      Virat Kohli  34.0      M         BATTER   INDIA
3           Madhav   NaN      M    ALL-ROUNDER   INDIA
4         KL Rahul  32.0      M  WICKET-KEPPER   INDIA
5    Hardik Pandya  32.0      M    ALL-ROUNDER   INDIA
6  Ravindra Jadeja  34.0      M    ALL-ROUNDER   INDIA
7      Yuzi Chahal  31.0      M         BOWLER   INDIA
8   Jasprit Bumrah  29.0      M         BOWLER   INDIA
9   Mohammad Siraj  28.0      M         BOWLER   INDIA
```

Task- 2 Check the info of 'df'

```
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 5 columns):
 #   Column   Non-Null Count  Dtype  
---  ------   --------------  -----  
 0   Name     10 non-null     object 
 1   Age      9 non-null      float64
 2   Gender   10 non-null     object 
 3   Role     10 non-null     object 
 4   COUNTRY  10 non-null     object 
dtypes: float64(1), object(4)
memory usage: 528.0+ bytes
```

Task 3- Check the descriptive statistics of 'df'

```
df1.describe()
```

```
             Age
count   9.000000
mean   32.444444
std     2.833333
min    28.000000
25%    31.000000
50%    32.000000
75%    34.000000
max    36.000000
```

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

```
df1.loc[3]
```

```
Name            Madhav
Age                NaN
Gender               M
Role       ALL-ROUNDER
COUNTRY          INDIA
Name: 3, dtype: object
```

Task 5 - Check the null values in your 'df'

```
df1.isnull().any()
```

```
Name       False
Age         True
Gender     False
Role       False
COUNTRY    False
dtype: bool
```

```
df1.isnull().sum()
```

```
Name       0
Age        1
Gender     0
Role       0
COUNTRY    0
dtype: int64
```