

```
#KAPILAVAI KRISHNA PRIYA
ABHIGNA
```

```
#21BAI1250
#krishna.priyaabhigna2021@vitstudent.ac.in
#ASSIGNMENT1
```

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

```
import numpy as np
import pandas as pd
```

```
data = {
'Name' : ['Devi','Krishna','Akhil','Aasrith','Shreeya','Rishi','Rishitha','Dev','Priya','Abhigna'],
'Student_ID' : [18,19,78,98,21,56,34,28,22,23],
'University' : ['VIT', 'VIT', 'SRM', 'BHU','Manipal','VIT', 'VIT', 'SRM', 'BHU','Manipal'],
'Age':[18,19,20,21,22,18,19,20,21,np.nan],
'Branch':['BAI','CSE','BPS','BRS','BPS','BAI','CSE','BPS','BMS','BRS'],
}
```

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

```
df
```

	Name	Student_ID	University	Age	Branch	
0	Devi	18	VIT	18.0	BAI	
1	Krishna	19	VIT	19.0	CSE	
2	Akhil	78	SRM	20.0	BPS	
3	Aasrith	98	BHU	21.0	BRS	
4	Shreeya	21	Manipal	22.0	BPS	
5	Rishi	56	VIT	18.0	BAI	
6	Rishitha	34	VIT	19.0	CSE	
7	Dev	28	SRM	20.0	BPS	
8	Priya	22	BHU	21.0	BMS	
9	Abhigna	23	Manipal	NaN	BRS	

```
#Task- 2 Check the info of 'df'
```

```
df.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   Student_ID   10 non-null     int64
2   University   10 non-null     object
3    Age          9 non-null      float64
4   Branch       10 non-null     object
dtypes: float64(1), int64(1), object(3)
memory usage: 528.0+ bytes
```

```
#Task 3- Check the descriptive statistics of 'df'
```

```
df.describe()
```

	Student_ID	Age	
count	10.000000	9.000000	
mean	39.700000	19.777778	
std	28.170315	1.394433	
min	18.000000	18.000000	
25%	21.250000	19.000000	
50%	25.500000	20.000000	

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

```
df.loc[4]
```

```
Name      Shreeya
Student_ID 21
University Manipal
Age        22.0
Branch     BRS
Name: 4, dtype: object
```

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

```
df.isnull().sum()
```

```
Name      0
Student_ID 0
University 0
Age        1
Branch     0
dtype: int64
```

#Filling null values

```
df.Age=df.Age.fillna(df.Age.mean())
```

df

	Name	Student_ID	University	Age	Branch	
0	Devi	18	VIT	18.000000	BAI	
1	Krishna	19	VIT	19.000000	CSE	
2	Akhil	78	SRM	20.000000	BPS	
3	Aasrith	98	BHU	21.000000	BRS	
4	Shreeya	21	Manipal	22.000000	BRS	
5	Rishi	56	VIT	18.000000	BAI	
6	Rishitha	34	VIT	19.000000	CSE	
7	Dev	28	SRM	20.000000	BPS	
8	Priya	22	BHU	21.000000	BMS	
9	Abhigna	23	Manipal	19.777778	BRS	

---

✓ 0s completed at 10:12AM

● ×