```
import pandas as pd
import numpy as np

# Task 1 (Create a pandas dataframe (DataFrame name as 'df') (10 observation and 5 features))
df = pd.DataFrame([['Karthik',19,'M',85000,'LGMS'],['Sanjay',20,'M',85000,'LGMS'],['Pavithra',20,'F',85000,'LGMS'],
['Jananee',19,'F',np.nan,'TIPS'],['Nisha',np.nan,'F',78000,'THF'],['kavitha',25,'F',40000,'TIPS'],
['Rakesh',19,'M',60000,'SB'],['preethi',20,'F',65000,'NPS'],['Simon',24,'M',80000,np.nan],['krish',20,'M',35000,'LGMS']])
df.columns=['Name','Age','Gender','Salary','School']
df.index=[1,2,3,4,5,6,7,8,9,10]
df
```

	Name	Age	Gender	Salary	School
1	Karthik	19.0	M	85000.0	LGMS
2	Sanjay	20.0	M	85000.0	LGMS
3	Pavithra	20.0	F	85000.0	LGMS
4	Jananee	19.0	F	NaN	TIPS
5	Nisha	NaN	F	78000.0	THF
6	kavitha	25.0	F	40000.0	TIPS
7	Rakesh	19.0	M	60000.0	SB
8	preethi	20.0	F	65000.0	NPS
9	Simon	24.0	M	80000.0	NaN
10	krish	20.0	M	35000.0	LGMS

```
# Task 2 (Check the info of 'df')
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 10 entries, 1 to 10
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   Salary  9 non-null      float64
 4   School  9 non-null      object 
dtypes: float64(2), object(3)
memory usage: 480.0+ bytes
```

```
# Task 3 (Check the descriptive statistics of 'df')
df.describe()
```

	Age	Salary
count	9.000000	9.000000
mean	20.666667	68111.111111
std	2.236068	19547.662549
min	19.000000	35000.000000
25%	19.000000	60000.000000
50%	20.000000	78000.000000
75%	20.000000	85000.000000
max	25.000000	85000.000000

```
# Task 4 (check the 4th index observation with 'loc' slicing operator.)
# index = 4 implies that the serial number = 5, since index starts from 0
df.loc[5]
```

```
Name        Nisha
Age           NaN
Gender          F
Salary    78000.0
School        THF
Name: 5, dtype: object
```

```
# or we can use iloc
df.iloc[4]
```

```
Name        Nisha
Age           NaN
Gender          F
Salary    78000.0
School        THF
Name: 5, dtype: object
```

```
# Task 5 (Check the null values in your 'df')
df.isnull().any()
```

```
Name      False
Age        True
Gender    False
Salary     True
School     True
dtype: bool
```

```
# Remove nan
df.Age = df.Age.fillna(df.Age.median())
df.Salary = df.Salary.fillna(df.Salary.median())
df.School = df.School.fillna(df.School.mode().iloc[0])
df
```

	Name	Age	Gender	Salary	School
1	Karthik	19.0	M	85000.0	LGMS
2	Sanjay	20.0	M	85000.0	LGMS
3	Pavithra	20.0	F	85000.0	LGMS
4	Jananee	19.0	F	78000.0	TIPS
5	Nisha	20.0	F	78000.0	THF
6	kavitha	25.0	F	40000.0	TIPS
7	Rakesh	19.0	M	60000.0	SB
8	preethi	20.0	F	65000.0	NPS
9	Simon	24.0	M	80000.0	LGMS
10	krish	20.0	M	35000.0	LGMS