

```
import pandas as pd
import numpy as np

df= pd.read_csv('SAMPLEIDS.csv')
print('Actual data:')
print(df)
```

Actual data:

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1
M2	M3 \						
0	1	1220121	ARUN	2000-02-10	MALE	THANDALAM	82.0
81.0	90.0						
1	2	1220122	BABU	1999-01-25	MALE	KANCHIPURAM	56.0
61.0	80.0						
2	3	1220123	CHARAN	2000.09.21	MALE	THANDALAM	NaN
59.0	60.0						
3	4	1220124	DEVA	2000-11-09	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	2000-11-21	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	1999-03-05	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	2000-10-02	MALE	KANCHIPURAM	49.0
51.0	70.0						
7	7	1220127	GANI	2000-10-02	MALE	KANCHIPURAM	49.0
51.0	70.0						
8	8	1220128	HEMA	1999-01-25	FEMALE	POONAMALEE	95.0
96.0	90.0						
9	9	1220129	INDRA	2000.09.21	FEMALE	KANCHIPURAM	64.0
NaN	NaN						
10	10	1220130	JAITH	2000-11-09	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	2000-11-21	FEMALE	CHITHUR	96.0
95.0	96.0						
12	12	1220132	LATHESSH	1999-03-05	MALE	THANDALAM	NaN
68.0	70.0						
13	13	1220133	MANI	2000-10-02	MALE	KANCHIPURAM	71.0
76.0	NaN						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						
15	15	1220135	NaN	19990125	NaN	NaN	NaN
NaN	NaN						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
17	17	1220137	RAGHU	20001109	MALE	POONAMALEE	67.0
64.0	70.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0
86.0	90.0						
19	19	1220139	SARVESH	19990305	MALE	THANDALAM	84.0
87.0	NaN						

20	20	1220140	SANTHOSH	20001002	MALE	KANCHIPURAM	76.0
69.0	80.0						

	M4	TOTAL	AVG
0	NaN	NaN	NaN
1	56.0	253.0	84.333333
2	70.0	NaN	0.000000
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000
5	91.0	360.0	120.000000
6	49.0	219.0	73.000000
7	49.0	219.0	73.000000
8	95.0	376.0	125.333333
9	64.0	NaN	0.000000
10	34.0	163.0	54.333333
11	96.0	383.0	127.666667
12	70.0	208.0	69.333333
13	71.0	NaN	0.000000
14	79.0	315.0	105.000000
15	NaN	0.0	0.000000
16	86.0	346.0	115.333333
17	NaN	201.0	67.000000
18	81.0	338.0	112.666667
19	84.0	NaN	0.000000
20	76.0	301.0	100.333333

```
import pandas as pd
import numpy as np

df= pd.read_csv('SAMPLEIDS.csv')
print(df.describe())
```

	SNO	REGNO	M1	M2	M3
M4 \					
count	21.000000	2.100000e+01	18.000000	19.000000	17.000000
mean	10.333333	1.220130e+06	73.666667	74.315789	79.529412
std	5.816643	5.816643e+00	17.580069	15.836149	13.010177
min	1.000000	1.220121e+06	34.000000	45.000000	50.000000
25%	6.000000	1.220126e+06	64.750000	62.500000	70.000000
50%	10.000000	1.220130e+06	77.500000	77.000000	80.000000
75%	15.000000	1.220135e+06	85.500000	86.500000	90.000000
max	20.000000	1.220140e+06	96.000000	96.000000	96.000000

	TOTAL	AVG
count	16.000000	20.000000
mean	272.750000	72.733333
std	102.048681	48.017127
min	0.000000	0.000000
25%	216.250000	40.750000
50%	304.000000	78.666667
75%	349.500000	113.333333
max	383.000000	127.666667

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

```
df= pd.read_csv('SAMPLEIDS.csv')
print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 21 entries, 0 to 20
Data columns (total 12 columns):
#   Column      Non-Null Count  Dtype
---  -
0    SNO         21 non-null    int64
1    REGNO       21 non-null    int64
2    NAME        20 non-null    object
3    DOB         21 non-null    object
4    GENDER      20 non-null    object
5    ADDRESS     20 non-null    object
6    M1          18 non-null    float64
7    M2          19 non-null    float64
8    M3          17 non-null    float64
9    M4          18 non-null    float64
10   TOTAL       16 non-null    float64
11   AVG         20 non-null    float64
dtypes: float64(6), int64(2), object(4)
memory usage: 2.1+ KB
None
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
print(df.isnull().sum())
```

SNO	0
REGNO	0
NAME	1
DOB	0
GENDER	1
ADDRESS	1

```
M1      3
M2      2
M3      4
M4      3
TOTAL   5
AVG      1
dtype: int64
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
print(df.nunique())
```

```
SNO      20
REGNO    20
NAME     19
DOB      13
GENDER    2
ADDRESS   4
M1       17
M2       17
M3        6
M4       16
TOTAL    15
AVG      15
dtype: int64
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
x2 = df.dropna(how='all').shape
print(x2)
```

```
(21, 12)
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
x2 = df.dropna(how='all').shape
print(x2)
```

```
(21, 12)
```

```
import pandas as pd
import numpy as np

df= pd.read_csv('SAMPLEIDS.csv')
tot=df.dropna(subset=['TOTAL'],how='any')
print(tot)
```

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1
M2	M3 \						
1	2	1220122	BABU	1/25/1999	MALE	KANCHIPURAM	56.0
61.0	80.0						
3	4	1220124	DEVA	11/9/2000	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	11/21/2000	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	3/5/1999	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	10/2/2000	MALE	KANCHIPURAM	49.0
51.0	70.0						
7	7	1220127	GANI	10/2/2000	MALE	KANCHIPURAM	49.0
51.0	70.0						
8	8	1220128	HEMA	1/25/1999	FEMALE	POONAMALEE	95.0
96.0	90.0						
10	10	1220130	JAHITH	11/9/2000	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	11/21/2000	FEMALE	CHITHUR	96.0
95.0	96.0						
12	12	1220132	LATHESSH	3/5/1999	MALE	THANDALAM	NaN
68.0	70.0						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						
15	15	1220135	NaN	19990125	NaN	NaN	NaN
NaN	NaN						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
17	17	1220137	RAGHU	20001109	MALE	POONAMALEE	67.0
64.0	70.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0
86.0	90.0						
20	20	1220140	SANTHOSH	20001002	MALE	KANCHIPURAM	76.0
69.0	80.0						

	M4	TOTAL	AVG
1	56.0	253.0	84.333333
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000
5	91.0	360.0	120.000000
6	49.0	219.0	73.000000
7	49.0	219.0	73.000000
8	95.0	376.0	125.333333

```

10  34.0  163.0   54.333333
11  96.0  383.0  127.666667
12  70.0  208.0   69.333333
14  79.0  315.0  105.000000
15   NaN    0.0    0.000000
16  86.0  346.0  115.333333
17   NaN  201.0   67.000000
18  81.0  338.0  112.666667
20  76.0  301.0  100.333333

```

```

import pandas as pd
import numpy as np

```

```

df= pd.read_csv('SAMPLEIDS.csv')
tot2=df.dropna(subset=['M1', 'M2', 'M3', 'M4'],how='any')
print(tot2)

```

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1
M2	M3 \						
1	2	1220122	BABU	1/25/1999	MALE	KANCHIPURAM	56.0
61.0	80.0						
3	4	1220124	DEVA	11/9/2000	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	11/21/2000	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	3/5/1999	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	10/2/2000	MALE	KANCHIPURAM	49.0
51.0	70.0						
7	7	1220127	GANI	10/2/2000	MALE	KANCHIPURAM	49.0
51.0	70.0						
8	8	1220128	HEMA	1/25/1999	FEMALE	POONAMALEE	95.0
96.0	90.0						
10	10	1220130	JAHITH	11/9/2000	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	11/21/2000	FEMALE	CHITHUR	96.0
95.0	96.0						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0
86.0	90.0						
20	20	1220140	SANTHOSH	20001002	MALE	KANCHIPURAM	76.0
69.0	80.0						
	M4	TOTAL					AVG
1	56.0	253.0					84.333333
3	74.0	307.0					102.333333
4	92.0	375.0					125.000000

```

5    91.0  360.0  120.000000
6    49.0  219.0   73.000000
7    49.0  219.0   73.000000
8    95.0  376.0  125.333333
10   34.0  163.0   54.333333
11   96.0  383.0  127.666667
14   79.0  315.0  105.000000
16   86.0  346.0  115.333333
18   81.0  338.0  112.666667
20   76.0  301.0  100.333333

```

```

import pandas as pd
import numpy as np

```

```

df= pd.read_csv('SAMPLEIDS.csv')
tot3=df.fillna(0)
print(tot3)

```

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1
M2	M3 \						
0	1	1220121	ARUN	2/10/2000	MALE	THANDALAM	82.0
81.0	90.0						
1	2	1220122	BABU	1/25/1999	MALE	KANCHIPURAM	56.0
61.0	80.0						
2	3	1220123	CHARAN	2000.09.21	MALE	THANDALAM	0.0
59.0	60.0						
3	4	1220124	DEVA	11/9/2000	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	11/21/2000	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	3/5/1999	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	10/2/2000	MALE	KANCHIPURAM	49.0
51.0	70.0						
7	7	1220127	GANI	10/2/2000	MALE	KANCHIPURAM	49.0
51.0	70.0						
8	8	1220128	HEMA	1/25/1999	FEMALE	POONAMALEE	95.0
96.0	90.0						
9	9	1220129	INDRA	2000.09.21	FEMALE	KANCHIPURAM	64.0
0.0	0.0						
10	10	1220130	JAHITH	11/9/2000	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	11/21/2000	FEMALE	CHITHUR	96.0
95.0	96.0						
12	12	1220132	LATHESSH	3/5/1999	MALE	THANDALAM	0.0
68.0	70.0						
13	13	1220133	MANI	10/2/2000	MALE	KANCHIPURAM	71.0
76.0	0.0						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						

15	15	1220135	0	19990125	0	0	0.0
0.0	0.0						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
17	17	1220137	RAGHU	20001109	MALE	POONAMALEE	67.0
64.0	70.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0
86.0	90.0						
19	19	1220139	SARVESH	19990305	MALE	THANDALAM	84.0
87.0	0.0						
20	20	1220140	SANTHOSH	20001002	MALE	KANCHIPURAM	76.0
69.0	80.0						

	M4	TOTAL	AVG
0	0.0	0.0	0.000000
1	56.0	253.0	84.333333
2	70.0	0.0	0.000000
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000
5	91.0	360.0	120.000000
6	49.0	219.0	73.000000
7	49.0	219.0	73.000000
8	95.0	376.0	125.333333
9	64.0	0.0	0.000000
10	34.0	163.0	54.333333
11	96.0	383.0	127.666667
12	70.0	208.0	69.333333
13	71.0	0.0	0.000000
14	79.0	315.0	105.000000
15	0.0	0.0	0.000000
16	86.0	346.0	115.333333
17	0.0	201.0	67.000000
18	81.0	338.0	112.666667
19	84.0	0.0	0.000000
20	76.0	301.0	100.333333

```
import pandas as pd
import numpy as np

df= pd.read_csv('SAMPLEIDS.csv')
tot4= df.fillna(method='ffill')
print(tot4)
```

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1
M2	M3	\					
0	1	1220121	ARUN	2/10/2000	MALE	THANDALAM	82.0
81.0	90.0						
1	2	1220122	BABU	1/25/1999	MALE	KANCHIPURAM	56.0
61.0	80.0						
2	3	1220123	CHARAN	2000.09.21	MALE	THANDALAM	56.0



59.0	60.0						
3	4	1220124	DEVA	11/9/2000	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	11/21/2000	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	3/5/1999	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	10/2/2000	MALE	KANCHIPURAM	49.0
51.0	70.0						
7	7	1220127	GANI	10/2/2000	MALE	KANCHIPURAM	49.0
51.0	70.0						
8	8	1220128	HEMA	1/25/1999	FEMALE	POONAMALEE	95.0
96.0	90.0						
9	9	1220129	INDRA	2000.09.21	FEMALE	KANCHIPURAM	64.0
96.0	90.0						
10	10	1220130	JAHITH	11/9/2000	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	11/21/2000	FEMALE	CHITHUR	96.0
95.0	96.0						
12	12	1220132	LATHESSH	3/5/1999	MALE	THANDALAM	96.0
68.0	70.0						
13	13	1220133	MANI	10/2/2000	MALE	KANCHIPURAM	71.0
76.0	70.0						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						
15	15	1220135	NANI	19990125	MALE	POONAMALEE	79.0
77.0	80.0						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
17	17	1220137	RAGHU	20001109	MALE	POONAMALEE	67.0
64.0	70.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0
86.0	90.0						
19	19	1220139	SARVESH	19990305	MALE	THANDALAM	84.0
87.0	90.0						
20	20	1220140	SANTHOSH	20001002	MALE	KANCHIPURAM	76.0
69.0	80.0						

	M4	TOTAL	AVG
0	NaN	NaN	NaN
1	56.0	253.0	84.333333
2	70.0	253.0	0.000000
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000
5	91.0	360.0	120.000000
6	49.0	219.0	73.000000
7	49.0	219.0	73.000000
8	95.0	376.0	125.333333
9	64.0	376.0	0.000000

```
10  34.0  163.0   54.333333
11  96.0  383.0  127.666667
12  70.0  208.0   69.333333
13  71.0  208.0   0.000000
14  79.0  315.0  105.000000
15  79.0   0.0   0.000000
16  86.0  346.0  115.333333
17  86.0  201.0   67.000000
18  81.0  338.0  112.666667
19  84.0  338.0   0.000000
20  76.0  301.0  100.333333
```

```
<ipython-input-34-4e29e1559c89>:5: FutureWarning: DataFrame.fillna
with 'method' is deprecated and will raise in a future version. Use
obj.ffill() or obj.bfill() instead.
```

```
tot4= df.fillna(method='ffill')
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
tot5=df.isna().sum()
print(tot5)
```

```
SNO      0
REGNO    0
NAME     1
DOB      0
GENDER   1
ADDRESS  1
M1       3
M2       2
M3       4
M4       3
TOTAL    5
AVG      1
dtype: int64
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
tot5=df.isnull()
print(tot5)
```

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1	M2	M3
M4 \									
0	False	False	False	False	False	False	False	False	False
True									
1	False	False	False	False	False	False	False	False	False
False									
2	False	False	False	False	False	False	True	False	False
False									
3	False	False	False	False	False	False	False	False	False
False									
4	False	False	False	False	False	False	False	False	False
False									
5	False	False	False	False	False	False	False	False	False
False									
6	False	False	False	False	False	False	False	False	False
False									
7	False	False	False	False	False	False	False	False	False
False									
8	False	False	False	False	False	False	False	False	False
False									
9	False	False	False	False	False	False	False	True	True
False									
10	False	False	False	False	False	False	False	False	False
False									
11	False	False	False	False	False	False	False	False	False
False									
12	False	False	False	False	False	False	True	False	False
False									
13	False	False	False	False	False	False	False	False	True
False									
14	False	False	False	False	False	False	False	False	False
False									
15	False	False	True	False	True	True	True	True	True
True									
16	False	False	False	False	False	False	False	False	False
False									
17	False	False	False	False	False	False	False	False	False
True									
18	False	False	False	False	False	False	False	False	False
False									
19	False	False	False	False	False	False	False	False	True
False									
20	False	False	False	False	False	False	False	False	False
False									
TOTAL		AVG							

0	True	True
1	False	False
2	True	False
3	False	False
4	False	False
5	False	False
6	False	False
7	False	False
8	False	False
9	True	False
10	False	False
11	False	False
12	False	False
13	True	False
14	False	False
15	False	False
16	False	False
17	False	False
18	False	False
19	True	False
20	False	False

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

```
df= pd.read_csv('SAMPLEIDS.csv')
```

```
df=df.duplicated()
print(df)
```

0	False
1	False
2	False
3	False
4	False
5	False
6	False
7	True
8	False
9	False
10	False
11	False
12	False
13	False
14	False
15	False
16	False
17	False
18	False
19	False

```
20     False
dtype: bool
```

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

```
df= pd.read_csv('SAMPLEIDS.csv')
```

```
m= df.drop_duplicates(inplace=False)
print(m)
```

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1
M2	M3 \						
0	1	1220121	ARUN	2/10/2000	MALE	THANDALAM	82.0
81.0	90.0						
1	2	1220122	BABU	1/25/1999	MALE	KANCHIPURAM	56.0
61.0	80.0						
2	3	1220123	CHARAN	2000.09.21	MALE	THANDALAM	NaN
59.0	60.0						
3	4	1220124	DEVA	11/9/2000	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	11/21/2000	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	3/5/1999	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	10/2/2000	MALE	KANCHIPURAM	49.0
51.0	70.0						
8	8	1220128	HEMA	1/25/1999	FEMALE	POONAMALEE	95.0
96.0	90.0						
9	9	1220129	INDRA	2000.09.21	FEMALE	KANCHIPURAM	64.0
NaN	NaN						
10	10	1220130	JAHITH	11/9/2000	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	11/21/2000	FEMALE	CHITHUR	96.0
95.0	96.0						
12	12	1220132	LATHESSH	3/5/1999	MALE	THANDALAM	NaN
68.0	70.0						
13	13	1220133	MANI	10/2/2000	MALE	KANCHIPURAM	71.0
76.0	NaN						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						
15	15	1220135	NaN	19990125	NaN	NaN	NaN
NaN	NaN						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
17	17	1220137	RAGHU	20001109	MALE	POONAMALEE	67.0
64.0	70.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0
86.0	90.0						
19	19	1220139	SARVESH	19990305	MALE	THANDALAM	84.0

```
87.0    NaN
20    20  1220140  SANTHOSH    20001002    MALE  KANCHIPURAM  76.0
69.0    80.0
```

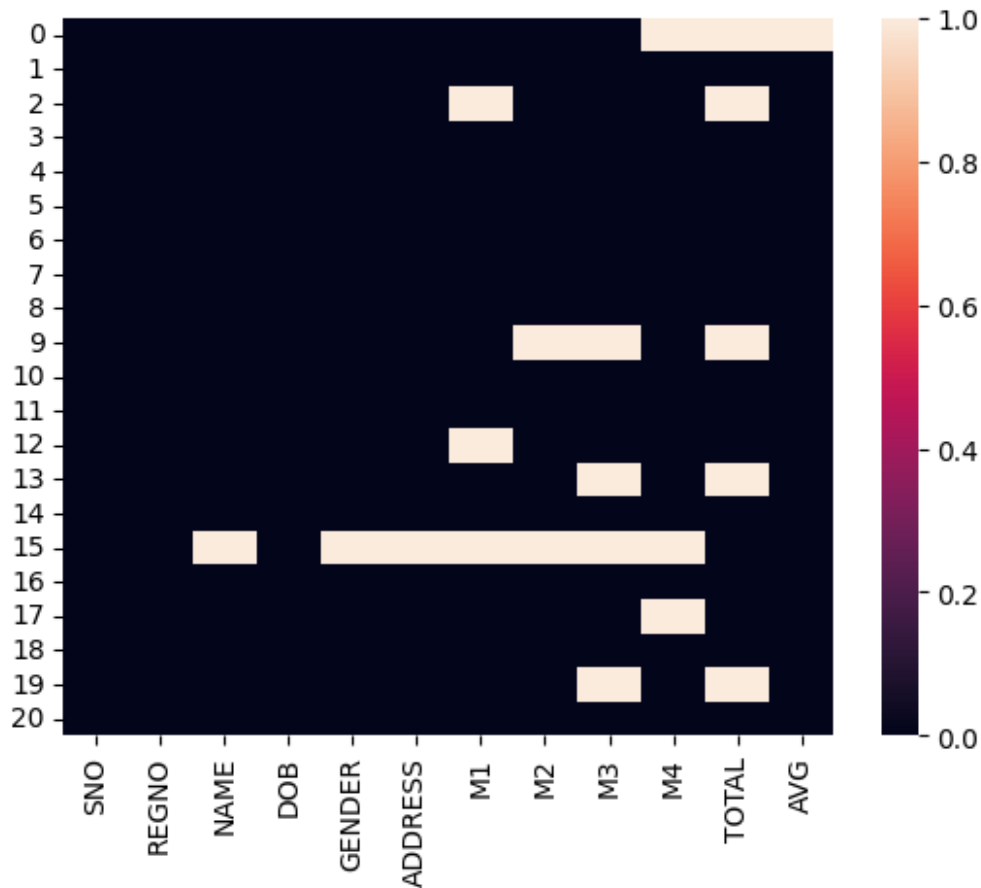
	M4	TOTAL	AVG
0	NaN	NaN	NaN
1	56.0	253.0	84.333333
2	70.0	NaN	0.000000
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000
5	91.0	360.0	120.000000
6	49.0	219.0	73.000000
8	95.0	376.0	125.333333
9	64.0	NaN	0.000000
10	34.0	163.0	54.333333
11	96.0	383.0	127.666667
12	70.0	208.0	69.333333
13	71.0	NaN	0.000000
14	79.0	315.0	105.000000
15	NaN	0.0	0.000000
16	86.0	346.0	115.333333
17	NaN	201.0	67.000000
18	81.0	338.0	112.666667
19	84.0	NaN	0.000000
20	76.0	301.0	100.333333

```
import pandas as pd
import numpy as np
import seaborn as sns

df= pd.read_csv('SAMPLEIDS.csv')

sns.heatmap(df.isnull())

<Axes: >
```



```
import pandas as pd

# Read the CSV file into a pandas DataFrame
df = pd.read_csv('SAMPLEIDS.csv')
print("_____")
# Display the first few rows of the DataFrame
print(df.head())
print("_____")
# Get summary statistics of the DataFrame
print(df.describe())
print("_____")
# Get information about the DataFrame (data types, non-null values,
etc.)
print(df.info())
print("_____")
# Check for missing values in each column
print(df.isnull().sum())
print("_____")
# Count the number of unique values in each column
print(df.nunique())
print("_____")
# Drop rows with missing values in the 'TOTAL' column
```

```

df_no_missing_total = df.dropna(subset=['TOTAL'])
print(df_no_missing_total)
print("_____")
# Fill missing values with 0
df_filled_zero = df.fillna(0)
print(df_filled_zero)
print("_____")
# Fill missing values using forward fill
df_filled_forward = df.fillna(method='ffill')
print(df_filled_forward)
print("_____")
# Check for duplicate rows
print(df.duplicated())
print("_____")
# Drop duplicate rows
df_no_duplicates = df.drop_duplicates()
print(df_no_duplicates)

```

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1	M2
M3 \								
0	1	1220121	ARUN	2000-02-10	MALE	THANDALAM	82.0	81.0
90.0								
1	2	1220122	BABU	1999-01-25	MALE	KANCHIPURAM	56.0	61.0
80.0								
2	3	1220123	CHARAN	2000.09.21	MALE	THANDALAM	NaN	59.0
60.0								
3	4	1220124	DEVA	2000-11-09	MALE	POONAMALEE	74.0	79.0
80.0								
4	5	1220125	ESTER	2000-11-21	FEMALE	CHITHUR	92.0	95.0
96.0								

	M4	TOTAL	AVG
0	NaN	NaN	NaN
1	56.0	253.0	84.333333
2	70.0	NaN	0.000000
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000

	SNO	REGNO	M1	M2	M3
M4 \					
count	21.000000	2.100000e+01	18.000000	19.000000	17.000000
18.000000					
mean	10.333333	1.220130e+06	73.666667	74.315789	79.529412
73.166667					
std	5.816643	5.816643e+00	17.580069	15.836149	13.010177
17.426315					
min	1.000000	1.220121e+06	34.000000	45.000000	50.000000
34.000000					
25%	6.000000	1.220126e+06	64.750000	62.500000	70.000000



```

65.500000
50%    10.000000  1.220130e+06  77.500000  77.000000  80.000000
75.000000
75%    15.000000  1.220135e+06  85.500000  86.500000  90.000000
85.500000
max     20.000000  1.220140e+06  96.000000  96.000000  96.000000
96.000000

```

	TOTAL	AVG
count	16.000000	20.000000
mean	272.750000	72.733333
std	102.048681	48.017127
min	0.000000	0.000000
25%	216.250000	40.750000
50%	304.000000	78.666667
75%	349.500000	113.333333
max	383.000000	127.666667

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 21 entries, 0 to 20
Data columns (total 12 columns):
#   Column      Non-Null Count  Dtype
---  -
0    SNO          21 non-null    int64
1    REGNO        21 non-null    int64
2    NAME         20 non-null    object
3    DOB          21 non-null    object
4    GENDER       20 non-null    object
5    ADDRESS      20 non-null    object
6    M1           18 non-null    float64
7    M2           19 non-null    float64
8    M3           17 non-null    float64
9    M4           18 non-null    float64
10   TOTAL        16 non-null    float64
11   AVG          20 non-null    float64
dtypes: float64(6), int64(2), object(4)
memory usage: 2.1+ KB
None

```

SNO	0
REGNO	0
NAME	1
DOB	0
GENDER	1
ADDRESS	1
M1	3
M2	2
M3	4
M4	3

TOTAL 5  
AVG 1  
dtype: int64

SNO 20  
REGNO 20  
NAME 19  
DOB 13  
GENDER 2  
ADDRESS 4  
M1 17  
M2 17  
M3 6  
M4 16  
TOTAL 15  
AVG 15  
dtype: int64

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1
M2	M3 \						
1	2	1220122	BABU	1999-01-25	MALE	KANCHIPURAM	56.0
61.0	80.0						
3	4	1220124	DEVA	2000-11-09	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	2000-11-21	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	1999-03-05	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	2000-10-02	MALE	KANCHIPURAM	49.0
51.0	70.0						
7	7	1220127	GANI	2000-10-02	MALE	KANCHIPURAM	49.0
51.0	70.0						
8	8	1220128	HEMA	1999-01-25	FEMALE	POONAMALEE	95.0
96.0	90.0						
10	10	1220130	JAITH	2000-11-09	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	2000-11-21	FEMALE	CHITHUR	96.0
95.0	96.0						
12	12	1220132	LATHESSH	1999-03-05	MALE	THANDALAM	NaN
68.0	70.0						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						
15	15	1220135	NaN	19990125	NaN	NaN	NaN
NaN	NaN						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
17	17	1220137	RAGHU	20001109	MALE	POONAMALEE	67.0
64.0	70.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0

86.0	90.0						
20	20	1220140	SANTHOSH	20001002	MALE	KANCHIPURAM	76.0
69.0	80.0						

	M4	TOTAL	AVG
1	56.0	253.0	84.333333
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000
5	91.0	360.0	120.000000
6	49.0	219.0	73.000000
7	49.0	219.0	73.000000
8	95.0	376.0	125.333333
10	34.0	163.0	54.333333
11	96.0	383.0	127.666667
12	70.0	208.0	69.333333
14	79.0	315.0	105.000000
15	NaN	0.0	0.000000
16	86.0	346.0	115.333333
17	NaN	201.0	67.000000
18	81.0	338.0	112.666667
20	76.0	301.0	100.333333

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1
M2	M3	\					
0	1	1220121	ARUN	2000-02-10	MALE	THANDALAM	82.0
81.0	90.0						
1	2	1220122	BABU	1999-01-25	MALE	KANCHIPURAM	56.0
61.0	80.0						
2	3	1220123	CHARAN	2000.09.21	MALE	THANDALAM	0.0
59.0	60.0						
3	4	1220124	DEVA	2000-11-09	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	2000-11-21	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	1999-03-05	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	2000-10-02	MALE	KANCHIPURAM	49.0
51.0	70.0						
7	7	1220127	GANI	2000-10-02	MALE	KANCHIPURAM	49.0
51.0	70.0						
8	8	1220128	HEMA	1999-01-25	FEMALE	POONAMALEE	95.0
96.0	90.0						
9	9	1220129	INDRA	2000.09.21	FEMALE	KANCHIPURAM	64.0
0.0	0.0						
10	10	1220130	JAHITH	2000-11-09	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	2000-11-21	FEMALE	CHITHUR	96.0
95.0	96.0						
12	12	1220132	LATHESSH	1999-03-05	MALE	THANDALAM	0.0

68.0	70.0						
13	13	1220133	MANI	2000-10-02	MALE	KANCHIPURAM	71.0
76.0	0.0						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						
15	15	1220135	0	19990125	0	0	0.0
0.0	0.0						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
17	17	1220137	RAGHU	20001109	MALE	POONAMALEE	67.0
64.0	70.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0
86.0	90.0						
19	19	1220139	SARVESH	19990305	MALE	THANDALAM	84.0
87.0	0.0						
20	20	1220140	SANTHOSH	20001002	MALE	KANCHIPURAM	76.0
69.0	80.0						

	M4	TOTAL	AVG
0	0.0	0.0	0.000000
1	56.0	253.0	84.333333
2	70.0	0.0	0.000000
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000
5	91.0	360.0	120.000000
6	49.0	219.0	73.000000
7	49.0	219.0	73.000000
8	95.0	376.0	125.333333
9	64.0	0.0	0.000000
10	34.0	163.0	54.333333
11	96.0	383.0	127.666667
12	70.0	208.0	69.333333
13	71.0	0.0	0.000000
14	79.0	315.0	105.000000
15	0.0	0.0	0.000000
16	86.0	346.0	115.333333
17	0.0	201.0	67.000000
18	81.0	338.0	112.666667
19	84.0	0.0	0.000000
20	76.0	301.0	100.333333

SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1	
M2	M3	\					
0	1	1220121	ARUN	2000-02-10	MALE	THANDALAM	82.0
81.0	90.0						
1	2	1220122	BABU	1999-01-25	MALE	KANCHIPURAM	56.0
61.0	80.0						
2	3	1220123	CHARAN	2000.09.21	MALE	THANDALAM	56.0
59.0	60.0						

3	4	1220124	DEVA	2000-11-09	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	2000-11-21	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	1999-03-05	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	2000-10-02	MALE	KANCHIPURAM	49.0
51.0	70.0						
7	7	1220127	GANI	2000-10-02	MALE	KANCHIPURAM	49.0
51.0	70.0						
8	8	1220128	HEMA	1999-01-25	FEMALE	POONAMALEE	95.0
96.0	90.0						
9	9	1220129	INDRA	2000.09.21	FEMALE	KANCHIPURAM	64.0
96.0	90.0						
10	10	1220130	JAHITH	2000-11-09	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	2000-11-21	FEMALE	CHITHUR	96.0
95.0	96.0						
12	12	1220132	LATHESSH	1999-03-05	MALE	THANDALAM	96.0
68.0	70.0						
13	13	1220133	MANI	2000-10-02	MALE	KANCHIPURAM	71.0
76.0	70.0						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						
15	15	1220135	NANI	19990125	MALE	POONAMALEE	79.0
77.0	80.0						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
17	17	1220137	RAGHU	20001109	MALE	POONAMALEE	67.0
64.0	70.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0
86.0	90.0						
19	19	1220139	SARVESH	19990305	MALE	THANDALAM	84.0
87.0	90.0						
20	20	1220140	SANTHOSH	20001002	MALE	KANCHIPURAM	76.0
69.0	80.0						

	M4	TOTAL	AVG
0	NaN	NaN	NaN
1	56.0	253.0	84.333333
2	70.0	253.0	0.000000
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000
5	91.0	360.0	120.000000
6	49.0	219.0	73.000000
7	49.0	219.0	73.000000
8	95.0	376.0	125.333333
9	64.0	376.0	0.000000
10	34.0	163.0	54.333333

11	96.0	383.0	127.666667
12	70.0	208.0	69.333333
13	71.0	208.0	0.000000
14	79.0	315.0	105.000000
15	79.0	0.0	0.000000
16	86.0	346.0	115.333333
17	86.0	201.0	67.000000
18	81.0	338.0	112.666667
19	84.0	338.0	0.000000
20	76.0	301.0	100.333333

0	False
1	False
2	False
3	False
4	False
5	False
6	False
7	True
8	False
9	False
10	False
11	False
12	False
13	False
14	False
15	False
16	False
17	False
18	False
19	False
20	False

dtype: bool

	SNO	REGNO	NAME	DOB	GENDER	ADDRESS	M1
M2	M3 \						
0	1	1220121	ARUN	2000-02-10	MALE	THANDALAM	82.0
81.0	90.0						
1	2	1220122	BABU	1999-01-25	MALE	KANCHIPURAM	56.0
61.0	80.0						
2	3	1220123	CHARAN	2000.09.21	MALE	THANDALAM	NaN
59.0	60.0						
3	4	1220124	DEVA	2000-11-09	MALE	POONAMALEE	74.0
79.0	80.0						
4	5	1220125	ESTER	2000-11-21	FEMALE	CHITHUR	92.0
95.0	96.0						
5	6	1220126	FARHANA	1999-03-05	FEMALE	THANDALAM	91.0
88.0	90.0						
6	7	1220127	GANI	2000-10-02	MALE	KANCHIPURAM	49.0

51.0	70.0						
8	8	1220128	HEMA	1999-01-25	FEMALE	POONAMALEE	95.0
96.0	90.0						
9	9	1220129	INDRA	2000.09.21	FEMALE	KANCHIPURAM	64.0
NaN	NaN						
10	10	1220130	JAHITH	2000-11-09	MALE	THANDALAM	34.0
45.0	50.0						
11	11	1220131	KANI	2000-11-21	FEMALE	CHITHUR	96.0
95.0	96.0						
12	12	1220132	LATHESSH	1999-03-05	MALE	THANDALAM	NaN
68.0	70.0						
13	13	1220133	MANI	2000-10-02	MALE	KANCHIPURAM	71.0
76.0	NaN						
14	14	1220134	NANI	20001109	MALE	POONAMALEE	79.0
77.0	80.0						
15	15	1220135	NaN	19990125	NaN	NaN	NaN
NaN	NaN						
16	16	1220136	PRATHAP	20000921	MALE	KANCHIPURAM	86.0
84.0	90.0						
17	17	1220137	RAGHU	20001109	MALE	POONAMALEE	67.0
64.0	70.0						
18	18	1220138	RATHI	20001121	FEMALE	KANCHIPURAM	81.0
86.0	90.0						
19	19	1220139	SARVESH	19990305	MALE	THANDALAM	84.0
87.0	NaN						
20	20	1220140	SANTHOSH	20001002	MALE	KANCHIPURAM	76.0
69.0	80.0						

	M4	TOTAL	AVG
0	NaN	NaN	NaN
1	56.0	253.0	84.333333
2	70.0	NaN	0.000000
3	74.0	307.0	102.333333
4	92.0	375.0	125.000000
5	91.0	360.0	120.000000
6	49.0	219.0	73.000000
8	95.0	376.0	125.333333
9	64.0	NaN	0.000000
10	34.0	163.0	54.333333
11	96.0	383.0	127.666667
12	70.0	208.0	69.333333
13	71.0	NaN	0.000000
14	79.0	315.0	105.000000
15	NaN	0.0	0.000000
16	86.0	346.0	115.333333
17	NaN	201.0	67.000000
18	81.0	338.0	112.666667
19	84.0	NaN	0.000000
20	76.0	301.0	100.333333

```
<ipython-input-9-3a5eb3497995>:30: FutureWarning: DataFrame.fillna
with 'method' is deprecated and will raise in a future version. Use
obj.ffill() or obj.bfill() instead.
    df_filled_forward = df.fillna(method='ffill')
```