1bhgrwjq0

March 19, 2025

[1]:
```
import pandas as pd
df = pd.read_csv("./sample.csv")
df
```

[1]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	NaN	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	NaN	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	NaN	17.0	18.0	19.0
11	12	88.0	NaN	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	NaN	NaN	42.0
14	15	NaN	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	NaN	NaN	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	NaN
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[ ]:
```

```

[2]:
```
df.isnull()
```

[2]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	False	False	False	False	False
1	False	False	False	False	False
2	False	False	False	False	False
3	False	False	False	False	False
4	False	False	False	False	False
5	False	False	True	False	False
6	False	False	False	False	False
7	False	False	False	False	False
8	False	True	False	False	False
9	False	False	False	False	False
10	False	True	False	False	False
11	False	False	True	False	False
12	False	False	False	False	False
13	False	False	True	True	False
14	False	True	False	False	False
15	False	False	False	False	False
16	False	False	False	False	False
17	False	False	False	False	False
18	False	False	False	False	False
19	False	False	False	False	False
20	False	False	False	False	False
21	False	False	False	False	False
22	False	False	False	False	False
23	False	False	False	False	False
24	False	False	True	True	False
25	False	False	False	False	False
26	False	False	False	False	False
27	False	False	False	False	True
28	False	False	False	False	False
29	False	False	False	False	False

[3]:
```
df.isnull().sum()
```

[3]:
```
Roll No.     0
Physics      3
Chemistry    4
Maths        2
Computer     1
dtype: int64
```

[5]:
```
df.isnull().sum().sum()
```

[5]:
```
np.int64(10)
```

[6]:
```
#Dropping rows with null values
df.shape
```

[6]:
```
(30, 5)
```

[8]:
```
#dropna has parameter axis (axis =0 it deletes rows)
#(axis = 1 )
df2 = df.dropna()
df2
```

[8]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
9	10	45.0	16.0	17.0	18.0
12	13	22.0	23.0	24.0	25.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[9]:
```
df2.isnull().sum()
```

[9]:
```
Roll No.     0
Physics      0
Chemistry    0
Maths        0
Computer     0
dtype: int64
```

[11]:
```
df3 = df.dropna(axis = 1)
df3.shape
df3
```

[11]:

	Roll No.
0	1
1	2
2	3
3	4
4	5
5	6
6	7
7	8
8	9
9	10
10	11
11	12
12	13
13	14
14	15
15	16
16	17
17	18
18	19
19	20
20	21
21	22
22	23
23	24
24	25
25	26
26	27
27	28
28	29
29	30

[12]:
```
df.dropna(how = 'any') #if any row value is null then remove that row
```

[12]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
9	10	45.0	16.0	17.0	18.0
12	13	22.0	23.0	24.0	25.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0

18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[13]:
```
df.dropna(how = 'all') #if all row value is null then remove that row
```

[13]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	NaN	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	NaN	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	NaN	17.0	18.0	19.0
11	12	88.0	NaN	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	NaN	NaN	42.0
14	15	NaN	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	NaN	NaN	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	NaN
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[16]:
```
df.dropna(inplace = True)
df
```

[16]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
9	10	45.0	16.0	17.0	18.0
12	13	22.0	23.0	24.0	25.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[17]:
```
df.shape
```

[17]: (22, 5)

[18]:
```
df
```

[18]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
9	10	45.0	16.0	17.0	18.0
12	13	22.0	23.0	24.0	25.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0

22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[19]:
```
import pandas as pd
df = pd.read_csv("./sample.csv")
df
```

[19]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	NaN	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	NaN	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	NaN	17.0	18.0	19.0
11	12	88.0	NaN	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	NaN	NaN	42.0
14	15	NaN	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	NaN	NaN	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	NaN
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[20]:
```
df.shape
```

[20]: (30, 5)

[21]:
```
df.isnull().sum()
```

[21]:
```
Roll No.      0
Physics       3
Chemistry     4
Maths         2
Computer      1
dtype: int64
```

[22]:
```
df.fillna(0)
```

[22]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	0.0	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	0.0	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	0.0	17.0	18.0	19.0
11	12	88.0	0.0	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	0.0	0.0	42.0
14	15	0.0	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	0.0	0.0	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	0.0
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[23]:
```
df.fillna(2)
```

[23]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0

1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	2.0	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	2.0	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	2.0	17.0	18.0	19.0
11	12	88.0	2.0	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	2.0	2.0	42.0
14	15	2.0	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	2.0	2.0	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	2.0
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[24]:
```
df.fillna({'Physics':'none','Chemistry':0,'Maths':30})
```

[24]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	0.0	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	none	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	none	17.0	18.0	19.0
11	12	88.0	0.0	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	0.0	30.0	42.0

14	15	none	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	0.0	30.0	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	NaN
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[25]:
```
df.fillna(method = 'ffill')
```

```
C:\Users\Rahul\AppData\Local\Temp\ipykernel_22872\1145651979.py:1:
FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a
future version. Use obj.ffill() or obj.bfill() instead.
  df.fillna(method = 'ffill')
```

[25]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	27.0	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	78.0	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	45.0	17.0	18.0	19.0
11	12	88.0	17.0	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	23.0	24.0	42.0
14	15	90.0	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0

22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	52.0	52.0	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	55.0
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[26]:
```
#replace null value with the before row value
df.ffill()
```

[26]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	27.0	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	78.0	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	45.0	17.0	18.0	19.0
11	12	88.0	17.0	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	23.0	24.0	42.0
14	15	90.0	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	52.0	52.0	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	55.0
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[27]:
```
df.ffill(axis = 1)
#replace null value with the left corresponding column value
```

[27]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1.0	56.0	57.0	58.0	59.0
1	2.0	23.0	24.0	25.0	26.0
2	3.0	89.0	25.0	26.0	27.0
3	4.0	45.0	26.0	27.0	28.0
4	5.0	23.0	27.0	28.0	29.0
5	6.0	90.0	90.0	29.0	30.0
6	7.0	12.0	13.0	14.0	15.0
7	8.0	78.0	14.0	15.0	16.0
8	9.0	9.0	15.0	16.0	17.0
9	10.0	45.0	16.0	17.0	18.0
10	11.0	11.0	17.0	18.0	19.0
11	12.0	88.0	88.0	19.0	20.0
12	13.0	22.0	23.0	24.0	25.0
13	14.0	90.0	90.0	90.0	42.0
14	15.0	15.0	43.0	43.0	43.0
15	16.0	44.0	44.0	44.0	44.0
16	17.0	45.0	45.0	45.0	45.0
17	18.0	46.0	46.0	46.0	46.0
18	19.0	47.0	47.0	47.0	47.0
19	20.0	48.0	48.0	48.0	48.0
20	21.0	49.0	49.0	49.0	49.0
21	22.0	50.0	50.0	50.0	50.0
22	23.0	51.0	51.0	51.0	51.0
23	24.0	52.0	52.0	52.0	52.0
24	25.0	53.0	53.0	53.0	53.0
25	26.0	54.0	33.0	33.0	54.0
26	27.0	55.0	34.0	34.0	55.0
27	28.0	56.0	35.0	35.0	35.0
28	29.0	57.0	36.0	36.0	66.0
29	30.0	58.0	37.0	37.0	43.0

[32]:
```
#if the student is absent for the physics examination the teacher will give␣
 ↪average marks to the student
df['Physics'].fillna(value = df['Physics'].mean())
# df['Physics'].mean()
```

[32]:
```
0     56.000000
1     23.000000
2     89.000000
3     45.000000
4     23.000000
5     90.000000
6     12.000000
7     78.000000
8     52.814815
9     45.000000
```

```
10    52.814815
11    88.000000
12    22.000000
13    90.000000
14    52.814815
15    44.000000
16    45.000000
17    46.000000
18    47.000000
19    48.000000
20    49.000000
21    50.000000
22    51.000000
23    52.000000
24    53.000000
25    54.000000
26    55.000000
27    56.000000
28    57.000000
29    58.000000
Name: Physics, dtype: float64
```

[ ]:
```
#fill the null value with the next row value when(axis = 0)
#fill the null value with the right column value when (axis = 1)
df.bfill()
```

#replace() function in pandas # if the dataset contains several number data # if we wnat to replace all the 26 values in the dataset then we use replace() function

[33]:
```
df.head()
```

[33]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0

[35]:
```
df.replace(to_replace = 26 , value = 30).head()
```

[35]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	30.0
2	3	89.0	25.0	30.0	27.0
3	4	45.0	30.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0

[37]:
```
df.replace(26,1000).head()
```

[37]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	1000.0
2	3	89.0	25.0	1000.0	27.0
3	4	45.0	1000.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0

[38]:
```
df.replace(to_replace = [50,51,52,53,54,55,56,57],value = 'A')
```

[38]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	A	A	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	NaN	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	NaN	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	NaN	17.0	18.0	19.0
11	12	88.0	NaN	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	NaN	NaN	42.0
14	15	NaN	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	A	A	A	A
22	23	A	A	A	A
23	24	A	A	A	A
24	25	A	NaN	NaN	A
25	26	A	33.0	33.0	A
26	27	A	34.0	34.0	A
27	28	A	35.0	35.0	NaN
28	29	A	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[39]:
```
df.head()
```

[39]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0

```
1          2     23.0       24.0   25.0      26.0
2          3     89.0       25.0   26.0      27.0
3          4     45.0       26.0   27.0      28.0
4          5     23.0       27.0   28.0      29.0
```

[40]:
```
df.replace(to_replace = [56,57,58,59],value = ['A','B','C','D']).head()
```

[40]:
```
   Roll No. Physics Chemistry Maths Computer
0         1       A         B     C        D
1         2    23.0      24.0  25.0     26.0
2         3    89.0      25.0  26.0     27.0
3         4    45.0      26.0  27.0     28.0
4         5    23.0      27.0  28.0     29.0
```

[41]:
```
#this replce fucntion can replace the value only in the physics function then we use
df['Physics'].replace(to_replace = [56,23,89,45],value = ["A","B","C","D"]).head()
```

[41]:
```
0    A
1    B
2    C
3    D
4    B
Name: Physics, dtype: object
```

[42]:
```
df
```

[42]:
```
    Roll No.  Physics  Chemistry  Maths  Computer
0          1     56.0       57.0   58.0      59.0
1          2     23.0       24.0   25.0      26.0
2          3     89.0       25.0   26.0      27.0
3          4     45.0       26.0   27.0      28.0
4          5     23.0       27.0   28.0      29.0
5          6     90.0        NaN   29.0      30.0
6          7     12.0       13.0   14.0      15.0
7          8     78.0       14.0   15.0      16.0
8          9      NaN       15.0   16.0      17.0
9         10     45.0       16.0   17.0      18.0
10        11      NaN       17.0   18.0      19.0
11        12     88.0        NaN   19.0      20.0
12        13     22.0       23.0   24.0      25.0
13        14     90.0        NaN    NaN      42.0
14        15      NaN       43.0   43.0      43.0
15        16     44.0       44.0   44.0      44.0
16        17     45.0       45.0   45.0      45.0
17        18     46.0       46.0   46.0      46.0
```

18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	NaN	NaN	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	NaN
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[48]:
```
#all the String charaters or strings are replaced with 0
df.replace('[A-Za-z]',0,regex = True)
```

[48]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	NaN	29.0	30.0
6	7	12.0	13.0	14.0	15.0
7	8	78.0	14.0	15.0	16.0
8	9	NaN	15.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	NaN	17.0	18.0	19.0
11	12	88.0	NaN	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	NaN	NaN	42.0
14	15	NaN	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	NaN	NaN	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	NaN
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

[49]:
```
df.replace(to_replace = 15,method = 'ffill')
```

```
C:\Users\Rahul\AppData\Local\Temp\ipykernel_22872\2592767107.py:1:
FutureWarning: The 'method' keyword in DataFrame.replace is deprecated and will
be removed in a future version.
  df.replace(to_replace = 15,method = 'ffill')
```

[49]:

	Roll No.	Physics	Chemistry	Maths	Computer
0	1	56.0	57.0	58.0	59.0
1	2	23.0	24.0	25.0	26.0
2	3	89.0	25.0	26.0	27.0
3	4	45.0	26.0	27.0	28.0
4	5	23.0	27.0	28.0	29.0
5	6	90.0	NaN	29.0	30.0
6	7	12.0	13.0	14.0	30.0
7	8	78.0	14.0	14.0	16.0
8	9	NaN	14.0	16.0	17.0
9	10	45.0	16.0	17.0	18.0
10	11	NaN	17.0	18.0	19.0
11	12	88.0	NaN	19.0	20.0
12	13	22.0	23.0	24.0	25.0
13	14	90.0	NaN	NaN	42.0
14	14	NaN	43.0	43.0	43.0
15	16	44.0	44.0	44.0	44.0
16	17	45.0	45.0	45.0	45.0
17	18	46.0	46.0	46.0	46.0
18	19	47.0	47.0	47.0	47.0
19	20	48.0	48.0	48.0	48.0
20	21	49.0	49.0	49.0	49.0
21	22	50.0	50.0	50.0	50.0
22	23	51.0	51.0	51.0	51.0
23	24	52.0	52.0	52.0	52.0
24	25	53.0	NaN	NaN	53.0
25	26	54.0	33.0	33.0	54.0
26	27	55.0	34.0	34.0	55.0
27	28	56.0	35.0	35.0	NaN
28	29	57.0	36.0	36.0	66.0
29	30	58.0	37.0	37.0	43.0

loc() and iloc() function in python

Uses sample2 dataset

[51]:
```
df = pd.read_csv("./sample2.csv",index_col = ['Roll No.'])
df
```

[51]:

Roll No.	Section	Branch	Physics	Chemistry	Maths	Computer	DOB

1	A	CS	56.0	57.0	58.0	59.0	01-01-2001
2	A	ECE	23.0	24.0	25.0	26.0	02-01-2001
3	B	MECH	89.0	25.0	26.0	27.0	03-01-2001
4	C	MECH	45.0	26.0	27.0	28.0	04-01-2001
5	A	CS	23.0	27.0	28.0	29.0	05-01-2001
6	A	ECE	90.0	NaN	29.0	30.0	06-01-2001
7	B	CS	12.0	13.0	14.0	15.0	07-01-2001
8	C	NaN	78.0	14.0	15.0	16.0	08-01-2001
9	A	ECE	NaN	15.0	16.0	17.0	09-01-2001
10	A	CS	45.0	16.0	17.0	18.0	10-01-2001
11	B	ECE	NaN	17.0	18.0	19.0	11-01-2001
12	C	CS	88.0	NaN	19.0	20.0	12-01-2001
13	A	CS	22.0	23.0	24.0	25.0	13-01-2001
14	A	CS	90.0	NaN	NaN	42.0	14-01-2001
15	B	ECE	NaN	43.0	43.0	43.0	15-01-2001
16	C	NaN	44.0	44.0	44.0	44.0	16-01-2001
17	A	MECH	45.0	45.0	45.0	45.0	17-01-2001
18	A	MECH	46.0	46.0	46.0	46.0	18-01-2001
19	B	ECE	47.0	47.0	47.0	47.0	19-01-2001
20	C	MECH	48.0	48.0	48.0	48.0	20-01-2001
21	A	MECH	49.0	49.0	49.0	49.0	21-01-2001
22	A	MECH	50.0	50.0	50.0	50.0	22-01-2001
23	B	ECE	51.0	51.0	51.0	51.0	23-01-2001
24	C	MECH	52.0	52.0	52.0	52.0	24-01-2001
25	A	MECH	53.0	NaN	NaN	53.0	25-01-2001
26	A	ECE	54.0	33.0	33.0	54.0	26-01-2001
27	B	CS	55.0	34.0	34.0	55.0	27-01-2001
28	C	CS	56.0	35.0	35.0	NaN	28-01-2001
29	A	CS	57.0	36.0	36.0	66.0	29-01-2001
30	A	CS	58.0	37.0	37.0	43.0	30-01-2001

[52]:
```
df.loc[1]
```

[52]:
```
Section              A
Branch              CS
Physics           56.0
Chemistry         57.0
Maths             58.0
Computer          59.0
DOB         01-01-2001
Name: 1, dtype: object
```

[54]:
```
df.loc[[5,6,7,8]]
```

[54]:

Roll No.	Section	Branch	Physics	Chemistry	Maths	Computer	DOB
5	A	CS	23.0	27.0	28.0	29.0	05-01-2001

```
6              A    ECE     90.0        NaN   29.0      30.0  06-01-2001
7              B     CS     12.0       13.0   14.0      15.0  07-01-2001
8              C    NaN     78.0       14.0   15.0      16.0  08-01-2001
```

[55]:
```
df.loc[5,'Physics']
```

[55]:
```
np.float64(23.0)
```

[56]:
```
df.loc[5:15,'Chemistry']
```

[56]:
```
Roll No.
5     27.0
6      NaN
7     13.0
8     14.0
9     15.0
10    16.0
11    17.0
12     NaN
13    23.0
14     NaN
15    43.0
Name: Chemistry, dtype: float64
```

[57]:
```
df.loc[df['Physics'] > 80]
```

[57]:
```
          Section Branch  Physics  Chemistry  Maths  Computer         DOB
Roll No.
3               B   MECH     89.0       25.0   26.0      27.0  03-01-2001
6               A    ECE     90.0        NaN   29.0      30.0  06-01-2001
12              C     CS     88.0        NaN   19.0      20.0  12-01-2001
14              A     CS     90.0        NaN    NaN      42.0  14-01-2001
```

[58]:
```
df.loc[df['Physics'] > 80,['Maths','Computer']]
```

[58]:
```
           Maths  Computer
Roll No.
3           26.0      27.0
6           29.0      30.0
12          19.0      20.0
14           NaN      42.0
```

[59]:
```
df.iloc[0]
```

[59]:
```
Section              A
Branch              CS
Physics           56.0
```

```
Chemistry          57.0
Maths              58.0
Computer           59.0
DOB          01-01-2001
Name: 1, dtype: object
```

[60]:
```
df.iloc[[0]]
```

[60]:

Roll No.	Section	Branch	Physics	Chemistry	Maths	Computer	DOB
1	A	CS	56.0	57.0	58.0	59.0	01-01-2001

[62]:
```
df.iloc[[0,1,2]]
```

[62]:

Roll No.	Section	Branch	Physics	Chemistry	Maths	Computer	DOB
1	A	CS	56.0	57.0	58.0	59.0	01-01-2001
2	A	ECE	23.0	24.0	25.0	26.0	02-01-2001
3	B	MECH	89.0	25.0	26.0	27.0	03-01-2001

[63]:
```
#i want to see all the rows with first column
#using iloc the index must be 0 for the first column
df.iloc[:,0]
```

[63]:
```
Roll No.
1     A
2     A
3     B
4     C
5     A
6     A
7     B
8     C
9     A
10    A
11    B
12    C
13    A
14    A
15    B
16    C
17    A
18    A
19    B
20    C
21    A
22    A
```

```
23    B
24    C
25    A
26    A
27    B
28    C
29    A
30    A
Name: Section, dtype: object
```

[64]:
```
df.iloc[0:5,1]
```

[64]:
```
Roll No.
1      CS
2     ECE
3    MECH
4    MECH
5      CS
Name: Branch, dtype: object
```

[66]:
```
df.iloc[0:5,0:3]
```

[66]:

Roll No.	Section	Branch	Physics
1	A	CS	56.0
2	A	ECE	23.0
3	B	MECH	89.0
4	C	MECH	45.0
5	A	CS	23.0

[ ]: