

Lab 6 26-08-2022

August 26, 2022

```
[2]: import numpy as np
import pandas as pd
```

0.1 5.

0.1.1 Write a pandas program to create a dataframe(6 records) which has the fields id,name,age,degree,cgpa

a)insert a new column 'class' into the dataframe and populate values

b)Find out total no.of columns in the dataframe

c)select top 3 rows ,last 4 rows

d)delete all rows which have null values

e)drop 'id' column

f)select name,cgpa from 2 to 5 records

g)fill null values in the age column with mean value

```
[24]: d = {'id': [1,2,3,4,5,6] ,
        'name': ['Shabeeb', 'Aisha', 'Aishwarya', 'Akhina', 'Amrita', 'Anu'],
        'age': [np.nan,20,18,21,24,23],
        'degree': ['BSc. CS', 'BCA', 'BSc. CS', 'BSc. CS', 'BSc. Physics', 'BCA'],
        'cgpa': [6.6,np.nan,7.5,8.9,8.5,np.nan]}
df = pd.DataFrame(d)
print(df)
#a
cls = ['MCA 1', 'MCA 2', 'MCA 3', 'MCA 3', 'MCA 1', 'MCA 2']
df['class'] = cls
print("\n",df)
#b
df1 = df.head(3)
print("\n top 3 rows:\n",df1)
df2 = df.tail(4)
```

```

print("\n last 4 rows:\n",df2)
#c
df3 = df.dropna()
print("\n dataframe after deleting null value rows: \n",df3)
#d
del df3['id']
print("\n df after removing id column:\n",df3)
#f
print("\n",df.iloc[[2,3,4,5],[1,4]])
#g
df['age'].fillna(value=df['age'].mean(), inplace=True)
print("\n df after replacing null value in age with mean:\n",df)

```

	id	name	age	degree	cgpa
0	1	Shabeeb	NaN	BSc. CS	6.6
1	2	Aisha	20.0	BCA	NaN
2	3	Aishwarya	18.0	BSc. CS	7.5
3	4	Akhina	21.0	BSc. CS	8.9
4	5	Amrita	24.0	BSc. Physics	8.5
5	6	Anu	23.0	BCA	NaN

	id	name	age	degree	cgpa	class
0	1	Shabeeb	NaN	BSc. CS	6.6	MCA 1
1	2	Aisha	20.0	BCA	NaN	MCA 2
2	3	Aishwarya	18.0	BSc. CS	7.5	MCA 3
3	4	Akhina	21.0	BSc. CS	8.9	MCA 3
4	5	Amrita	24.0	BSc. Physics	8.5	MCA 1
5	6	Anu	23.0	BCA	NaN	MCA 2

top 3 rows:

	id	name	age	degree	cgpa	class
0	1	Shabeeb	NaN	BSc. CS	6.6	MCA 1
1	2	Aisha	20.0	BCA	NaN	MCA 2
2	3	Aishwarya	18.0	BSc. CS	7.5	MCA 3

last 4 rows:

	id	name	age	degree	cgpa	class
2	3	Aishwarya	18.0	BSc. CS	7.5	MCA 3
3	4	Akhina	21.0	BSc. CS	8.9	MCA 3
4	5	Amrita	24.0	BSc. Physics	8.5	MCA 1
5	6	Anu	23.0	BCA	NaN	MCA 2

dataframe after deleting null value rows:

	id	name	age	degree	cgpa	class
2	3	Aishwarya	18.0	BSc. CS	7.5	MCA 3
3	4	Akhina	21.0	BSc. CS	8.9	MCA 3
4	5	Amrita	24.0	BSc. Physics	8.5	MCA 1

df after removing id column:

	name	age	degree	cgpa	class
2	Aishwarya	18.0	BSc. CS	7.5	MCA 3
3	Akhina	21.0	BSc. CS	8.9	MCA 3
4	Amrita	24.0	BSc. Physics	8.5	MCA 1

	name	cgpa
2	Aishwarya	7.5
3	Akhina	8.9
4	Amrita	8.5
5	Anu	NaN

df after replacing null value in age with mean:

	id	name	age	degree	cgpa	class
0	1	Shabeeb	21.2	BSc. CS	6.6	MCA 1
1	2	Aisha	20.0	BCA	NaN	MCA 2
2	3	Aishwarya	18.0	BSc. CS	7.5	MCA 3
3	4	Akhina	21.0	BSc. CS	8.9	MCA 3
4	5	Amrita	24.0	BSc. Physics	8.5	MCA 1
5	6	Anu	23.0	BCA	NaN	MCA 2