```
import pandas as pd
url="https://raw.githubusercontent.com/subhashchandra630/Python-ca/main/titanic.csv"
df=pd.read_csv(url)
df.head(2)
df.isnull()
```

	PassengerId	Survived	Pclass	Name
0	False	False	False	False
1	False	False	False	False
2	False	False	False	False
3	False	False	False	False
4	False	False	False	False
886	False	False	False	False
887	False	False	False	False
888	False	False	False	False
889	False	False	False	False
890	False	False	False	False

891 rows × 12 columns



to check how many missing values are in a column

```
print(df[["Sex","Cabin"]].isnull())
df.info()
```

```
Sex Cabin
    False True
1
    False False
2
    False
          True
3
    False False
    False True
886 False
          True
887 False False
888 False
          True
889 False False
890 False
          True
[891 rows x 2 columns]
```

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 891 entries, 0 to 890

Data	columns (tota	al 12 columns):		
#	Column	Non-Null Count	Dtype	
0	PassengerId	891 non-null	int64	
1	Survived	891 non-null	int64	
2	Pclass	891 non-null	int64	
3	Name	891 non-null	object	
4	Sex	891 non-null	object	
5	Age	714 non-null	float64	
6	SibSp	891 non-null	int64	
7	Parch	891 non-null	int64	
8	Ticket	891 non-null	object	
9	Fare	891 non-null	float64	
10	Cabin	204 non-null	object	
11	Embarked	889 non-null	object	
<pre>dtypes: float64(2), int64(5), object(5)</pre>				
momony usago: 82 7+ KB				

memory usage: 83.7+ KB

Double-click (or enter) to edit

df.notnull()

	PassengerId	Survived	Pclass	Name
0	True	True	True	True
1	True	True	True	True
2	True	True	True	True
3	True	True	True	True
4	True	True	True	True
886	True	True	True	True
887	True	True	True	True
888	True	True	True	True
889	True	True	True	True
890	True	True	True	True

891 rows × 12 columns

fill missing values with 0

```
import numpy as np
dict={'first':[100,90,np.nan,5],
      'second':[30,45,56,np.nan],
      'third':[40,24,59,np.nan],
      'four':['male',np.nan,'female','male']
}
```

d=pd.DataFrame(dict)
d.fillna(0)

	first	second	third	four	1
0	100.0	30.0	40.0	male	
1	90.0	45.0	24.0	0	
2	0.0	56.0	59.0	female	
3	5.0	0.0	0.0	male	

d1=pd.DataFrame(dict)
d1.fillna(method='bfill')

	first	second	third	1
0	100.0	30.0	40.0	
1	90.0	45.0	24.0	
2	5.0	56.0	59.0	
3	5.0	NaN	NaN	

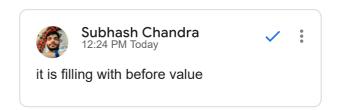
d1=pd.DataFrame(dict)
d1.fillna(method='ffill')

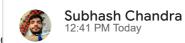
	first	second	third	1
0	100.0	30.0	40.0	
1	90.0	45.0	24.0	
2	90.0	56.0	59.0	
3	5.0	56.0	59.0	

d1.replace(to_replace=np.nan,value=-99)

	first	second	third	1
0	100.0	30.0	40.0	
1	90.0	45.0	24.0	
2	-99.0	56.0	59.0	
3	5.0	-99.0	-99.0	







0 male
1 no gender
2 female
3 male

Name: four, dtype: object

replacing any missing value in a specific column

Colab paid products - Cancel contracts here

✓ 0s completed at 12:41 PM

×