

jupyter Untitled Last Checkpoint: an hour ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help



Code



```
In [31]: import pandas as pd  
df=pd.read_table(r"C:\\Users\\ramg9\\OneDrive\\Desktop\\datafr.txt",sep="\t")  
df
```

Out[31]:

	Name	Age	Height	Weight	Gender	Married
0	ram	33	185	90.8	Male	yes
1	rahul	29	173	102.15	Male	no
2	rohit	46	160	74.91	Female	yes
3	ritik	57	163	77.18	Female	yes
4	utkarsh	28	185	200	Male	yes
5	yash	35	173	None	Male	no
6	john	42	None	165	Female	no
7	jack	None	163	170	female	yes

```
In [35]: dic={'name':['ram','rahul','rohit','ritik','utkarsh','yash','john','jack'],  
           'age':[33,29,46,57,28,35,42,None], 'height':[133,179,146,157,180,135,None,163],  
           'weight':[90.8,102.15,74.91,77.18,200,None,165,170],  
           'gender':['Male','Male','Female','Female','Male','Male','Female','female'],  
           'married':['yes','no','yes','yes','yes','no','no','yes']}  
df=pd.DataFrame(dic)  
df
```

Out[35]:

	name	age	height	weight	gender	married
0	ram	33.0	133.0	90.80	Male	yes
1	rahul	29.0	179.0	102.15	Male	no
2	rohit	46.0	146.0	74.91	Female	yes
3	ritik	57.0	157.0	77.18	Female	yes
4	utkarsh	28.0	180.0	200.00	Male	yes

shallow_copy

In [38]:

```
print(df)
print("\n")
data_shallow=df.copy(deep=False)
df['Age'].replace('None',50,inplace=True)
print(df)
print("\n")
print(data_shallow)
```

	Name	Age	Height	Weight	Gender
0	ram	29	185	60.20	male
1	rahul	46	173	102.20	male
2	apeksha	57	160	70.70	female
3	madhu	28	163	90.80	female
4	utkarsh	35	185	102.15	male
5	jenny	42	173	74.91	female
6	jack	None	None	77.18	male

	Name	Age	Height	Weight	Gender
0	ram	29	185	60.20	male
1	rahul	46	173	102.20	male
2	apeksha	57	160	70.70	female
3	madhu	28	163	90.80	female
4	utkarsh	35	185	102.15	male
5	jenny	42	173	74.91	female
6	jack	50	None	77.18	male

	Name	Age	Height	Weight	Gender
0	ram	29	185	60.20	male
1	rahul	46	173	102.20	male
2	apeksha	57	160	70.70	female
3	madhu	28	163	90.80	female
4	utkarsh	35	185	102.15	male
5	jenny	42	173	74.91	female
6	jack	50	None	77.18	male

Deep copy

```
In [42]: print(df)
print("\n")
data_shallow=df.copy(deep=True)
df['Age'].replace('None',50,inplace=True)
print(df)
print("\n")
print(data_shallow)
```

	Name	Age	Height	Weight	Gender
0	ram	29	185	60.20	male
1	rahul	46	173	102.20	male
2	apeksha	57	160	70.70	female
3	madhu	28	163	90.80	female
4	utkarsh	35	185	102.15	male
5	jenny	42	173	74.91	female
6	jack	None	None	77.18	male

	Name	Age	Height	Weight	Gender
0	ram	29	185	60.20	male
1	rahul	46	173	102.20	male
2	apeksha	57	160	70.70	female
3	madhu	28	163	90.80	female
4	utkarsh	35	185	102.15	male
5	jenny	42	173	74.91	female
6	jack	50	None	77.18	male

	Name	Age	Height	Weight	Gender
0	ram	29	185	60.20	male
1	rahul	46	173	102.20	male
2	apeksha	57	160	70.70	female
3	madhu	28	163	90.80	female
4	utkarsh	35	185	102.15	male
5	jenny	42	173	74.91	female
6	jack	None	None	77.18	male

```
In [35]: df=df.fillna(method='ffill')  
df
```

Out[35]:

	Name	Age	Height	Weight	Gender
0	ram	29	185	60.20	male
1	rahul	46	173	102.20	male
2	apeksha	57	160	70.70	female
3	madhu	28	163	90.80	female
4	utkarsh	35	185	102.15	male
5	jenny	42	173	74.91	female
6	jack	42	173	77.18	male

```
In [29]: df=df.fillna(method='bfill',axis='columns')  
df
```

Out[29]:

	Name	Age	Height	Weight	Gender
0	ram	29	185	60.2	male
1	rahul	46	173	102.2	male
2	apeksha	57	160	70.7	female
3	madhu	28	163	90.8	female
4	utkarsh	35	185	102.15	male
5	jenny	42	173	74.91	female
6	jack	77.18	77.18	77.18	male

Lambda

```
In [37]: import pandas as pd  
df=pd.read_table(r"C:\\\\Users\\\\ramg9\\\\OneDrive\\\\Desktop\\\\dataf.txt",sep="\t")  
df
```

Out[37]:

	Name	Age	Height	Weight	Gender
0	ram	29	185	60.2	male
1	rahul	&&	173	102.2	male
2	apeksha	57	160	70.7	@@
3	madhu	28	??	90.8	female
4	utkarsh	35	185	**	male
5	jenny	42	173	74.91	female
6	jack	##	!!	77.18	male

```
In [38]: df=pd.read_table(r"C:\\\\Users\\\\ramg9\\\\OneDrive\\\\Desktop\\\\dataf.txt",sep="\t",na_values=["??","@@","**","!!","##","&&"])  
df
```

Out[38]:

	Name	Age	Height	Weight	Gender
0	ram	29.0	185.0	60.20	male
1	rahul	NaN	173.0	102.20	male
2	apeksha	57.0	160.0	70.70	NaN
3	madhu	28.0	NaN	90.80	female
4	utkarsh	35.0	185.0	NaN	male
5	jenny	42.0	173.0	74.91	female
6	jack	NaN	NaN	77.18	male

```
In [38]: df=pd.read_table(r"C:\\\\Users\\\\ramg9\\\\OneDrive\\\\Desktop\\\\dataf.txt",sep="\t",na_values=["??","@@","**","!!","##","&&"])
df
```

Out[38]:

	Name	Age	Height	Weight	Gender
0	ram	29.0	185.0	60.20	male
1	rahul	NaN	173.0	102.20	male
2	apeksha	57.0	160.0	70.70	NaN
3	madhu	28.0	NaN	90.80	female
4	utkarsh	35.0	185.0	NaN	male
5	jenny	42.0	173.0	74.91	female
6	jack	NaN	NaN	77.18	male

```
In [43]: df=df.apply(lambda x:x.fillna(x.mean()) if x.dtypes=='float' else x.fillna(x.value_counts().index[0]))
df
```

Out[43]:

	Name	Age	Height	Weight	Gender
0	ram	29.0	185.0	60.200000	male
1	rahul	38.2	173.0	102.200000	male
2	apeksha	57.0	160.0	70.700000	male
3	madhu	28.0	175.2	90.800000	female
4	utkarsh	35.0	185.0	79.331687	male
5	jenny	42.0	173.0	74.910000	female
6	jack	38.2	175.2	77.180000	male