```
In [1]: import pandas as pd
```

```
In [2]: pd.__version__
```

```
Out[2]: '1.0.1'
```

Data=pd.read_csv(r'titanic.csv')

Data.head(10)

Data.tail(10)

Data.shape

x=Data.copy()

x.coloumns=['a','b',etc]

x.drop('a',axis=1,inplace=True)

x.c or x['c'] >> gives coloumn c

x.loc[2:5] >> returns data from second row to 5th row

x. iloc[2:3,2:4] >> gives specific data

For Nan value

x.f.fillnan(x.f.mean(),inplace=True) >> this command fills Nan value of coloumn f with the mean of the coloumn

For conditions

print(x[x.f>50]) >> returns the whole matrix for which the entry in coloumn f > 50

print(x[x.f>50].describe()) >> gives the details of the coloumn f having entry > 50 such as count, mean, SD etc.

For changing values according to condition

def change(text): if text == 'Child': return 10 elif text =='Adult': return 25 else: return 50

x['d'] = x.d.apply(change)

This code changes values as text in coloumn d to the integer according to the change condition

Inserting new coloumns

x['New coloumn']='default value' >> creates a new coloumn with value "default value"

x.insert(1,'2nd coloumns','def') >> creates a new coloumn at index 2 with values "def"

Series Data

In [3]:
```
s=pd.Series(list('abacd'))
s
```

Out[3]:
```
0    a
1    b
2    a
3    c
4    d
dtype: object
```

In [4]:
```
dummy_data=pd.get_dummies(s)
print(dummy_data)
```

```
   a  b  c  d
0  1  0  0  0
1  0  1  0  0
2  1  0  0  0
3  0  0  1  0
4  0  0  0  1
```

In [ ]: