```
import numpy as np
import pandas as pd
import statistics as stat
D1=pd.read_csv('E:/Python/Titanic_Survival_train.csv')
D1
D1.info()
D1.describe()## to see the max,min,mean,std,count and quartiles
## to axcess a particular column
D1['Age']
D1.Age[0:5]## to access particular column n row
D1[['Survived','Age']][0:5]## to acess particular columns n rows
D1['Age'].value_counts() ##gives the counts in a particular column
D1['Age'].value_counts(normalize=True)*10000
pd.crosstab(D1.Sex,D1.Survived)### combining th condition  and combin 2
tables with a particualr condition
pd.crosstab(D1.Sex,D1.Survived,normalize="index")## normalize by index
pd.crosstab(D1.Sex,D1.Survived,normalize="columns")## normalise by column
pd.crosstab(D1.Sex,D1.Survived,normalize="all")##normalise everything
age5=D1[D1['Age']<=5]
age5[0:3]
len(age5<5)## to count the sum f the particular column with a condition
D1[D1.Age<=5]['Survived'].value_counts()
D1[D1.Name.str.contains("Allen")]
D1.Embarked.unique()## to calculate the unique values
D1['Age'].value_counts(dropna=False)## to see the nas as unique values
D1['Embarked'].value_counts(dropna=False)
D1[(D1.Age<=5)& (D1.Survived==1)][['Name','PassengerId']]
D1[(D1.Age<=5)& (D1.Survived==1)][['Name','PassengerId']][0:10]
D1.groupby('')['Age'].mean()
D1.groupby('Survived')['Age'].mean()## to do some function in a
catogorical column
D1.groupby(['Pclass','Survived'])['Age'].mean()
D1.groupby(['Pclass','Survived'])['Age'].sum()
D1.groupby(['Pclass','Survived'])['Age'].mean().reset_index()## to do the
mean and do rename the index
```