**---------------------Load data with missing values ----------------EXERCISE#1-----------------------------------------**

import pandas as pd

from io import StringIO

csv\_data = '''A,B,C,D

10,20,30,40

1.0,2.0,,4.0

5.0,6.0,,8.0

0.0,11.0,12.0,'''

df = pd.read\_csv(StringIO(csv\_data))

print(df)

print(df.tail())

print(df.head())

print(df.isnull().sum())

print(df.values)

print(df.as\_matrix())

**----------------------------dropna ----------------EXERCISE#2-----------------------------------------**

import pandas as pd

from io import StringIO

csv\_data = '''A,B,C,D

10,20,30,40

1.0,2.0,

5.0,6.0,,8.0

0.0,11.0,12.0,'''

df = pd.read\_csv(StringIO(csv\_data))

print(df)

df1=df.dropna()

print(df1)

df2=df.dropna(axis=1)

print(df2)

print(df.dropna(how='all'))

print(df.dropna(thresh=2))

print(df.dropna(subset=['C']))

---------------------------- ----------------EXERCISE#3-----------------------------------------

import pandas as pd

from io import StringIO

csv\_data = '''A,B,C,D

10,20,30,40

1.0,2.0,

5.0,6.0,,8.0

0.0,11.0,12.0,'''

df = pd.read\_csv(StringIO(csv\_data))

print(df)

from sklearn.preprocessing import Imputer

import numpy as np

imr = Imputer(missing\_values=np.nan, strategy='mean')

imr = imr.fit(df)

imputed\_data = imr.transform(df.values)

print(imputed\_data)

##

imr = Imputer(missing\_values=np.nan, strategy='median')

imr = imr.fit(df)

imputed\_data = imr.transform(df.values)

print(imputed\_data)

print(df)