import numpy as np

import pandas as pd

df=pd.read\_csv("IRIS.csv")

df

df\_species = pd.get\_dummies(df['species'])

df\_species

df\_new= pd.concat([df,df\_species], axis=1)

df\_new

print(df\_new.iloc[148,4])

for i in range(150):

if(df\_new.iloc[i,4]=='Iris-setosa'):

df\_new.iloc[i,4]=0

elif(df\_new.iloc[i,4]=='Iris-versicolor'):

df\_new.iloc[i,4]=1

elif(df\_new.iloc[i,4]=='Iris-virginica'):

df\_new.iloc[i,4]=2

df\_new

num\_columns=[col for col in df.columns if(df[col].dtype=="float" or df[col].dtype=="int")]

num\_columns

number=df.\_get\_numeric\_data()

number

MEAN:

for i in num\_columns:

print(i ," :", df[i].mean())

for i in num\_columns:

Sum=0

for j in range(150):

Sum+=df[i][j]

print(i ," :",Sum/150)

MEDIAN:

for i in num\_columns:

print(i ," :", df[i].median())

print(list(df['sepal\_length']))

for i in num\_columns:

sorted\_list=sorted(list(df[i]))

print(i," :", sorted\_list[len(sorted\_list)//2])

MODE:

df.mode().values[0]

x=0

for i in df.columns:

L=list(df[i])

Set=set(df[i])

for j in Set:

y=L.count(j)

if(x<y):

x=y

m=j

print(i , " :",j)

PERCENTILE:

for i in num\_columns:

tile=np.percentile(df[i], 25)

print(i, "25 tile : " ,tile)

tile=np.percentile(df[i], 50)

print(i, "50 tile : " ,tile)

tile=np.percentile(df[i], 75)

print(i, "75 tile : " ,tile)

species\_tile=np.percentile(df\_new['species'],25)

print(species\_tile)

VARIENCE:

for i in num\_columns:

varience=np.var(df[i])

print(i, " : ",varience)

STANDARD DEVIATION:

df.std()

for i in num\_columns:

std\_dev=np.std(df[i])

print(i," : ",std\_dev)

std\_dev=np.std(df, axis=0)

std\_dev