ASSIGNMENT-5

Write python code for Correlation Coefficient, Covariance , Chi-square Test and Normalization (Min-max, z- score, Decimal scaling) techniques.

import pandas as pd

importnumpy as np

importmatplotlib.pyplot as plt

importseaborn as sbs

get\_ipython().run\_line\_magic('matplotlib', 'inline')

plt.style.use('seaborn-whitegrid')

df=pd.read\_csv('student-mat.csv')

dfn=df[['traveltime','studytime']]

dfn.head()

x=dfn['traveltime']

y=dfn['studytime']

sbs.lineplot(x,y,dashes=True)

plt.show()

fromscipy.stats import norm

\*\*CORELATION ANALYSIS\*\*

corelation=dfn.corr()

print(corelation)

sbs.heatmap(corelation,cmap ="YlGnBu")

plt.show()

\*\*COVARIANCE ANALYSIS\*\*

covar=dfn.cov()

print(covar)

sbs.heatmap(covar,cmap='BuGn\_r')

plt.show()

**\*\*CHI SQUARE TEST\*\***

dfnom=df[['Mjob','Fjob']]

dfnom.head()

dfnom=dfnom.replace('at\_home','home')

dfnom=dfnom.astype('category')

dfnom.info()

fromsklearn.preprocessing import LabelEncoder

lb=LabelEncoder()

dfnom['Mjob']=lb.fit\_transform(dfnom['Mjob'])

dfnom['Fjob']=lb.fit\_transform(dfnom['Fjob'])

dfnom=np.array(dfnom)

dfnom=dfnom[dfnom!=0]

fromscipy.stats import chi2\_contingency

stat, p, dof, expected = chi2\_contingency(dfnom)

print(stat)

print(p)

print(dof)

print(expected.shape)

sbs.heatmap(expected)

NORMALIZATION

import pandas as pd

importnumpy as np

importmatplotlib.pyplot as plt

importseaborn as sbs

%matplotlib inline

df=pd.read\_csv('student-mat.csv')

age=np.array(df['age'])

print('MAX AGE',max(age))

print('MIN AGE',min(age))

age=age.reshape(395,1)

**MIN MAX NORMALISATION**

fromsklearn.preprocessing import MinMaxScaler

MinMax=MinMaxScaler(feature\_range=(0,2))

result=MinMax.fit\_transform(age)

result=result[0:394]

result=result.reshape(2,197)

result

**Z SCORE NORMALISATION**

fromscipy import stats

zscore=np.array(stats.zscore(age))

zscore=zscore[0:394]

zscore=zscore.reshape(2,197)

zscore

**DECIMAL SCALING**

result=[]

result.append(age/pow(10,2))

result=np.array(result)

result