My question for this analysts is "Does student mental health affect CGPA's"? Given that most of the students in the study do not have mental health issues I am not sure how accurate this study would be for predicting if mental health affects CGPAs.

[ ]

from google.colab import drive

drive.mount('/content/drive')

output

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force\_remount=True).

[ ]

import pandas as pd

mental\_health=pd.read\_csv('/content/drive/My Drive/data/student\_mental\_health\_v1 .csv')

New Section

[ ]

mental\_health.head() ##I wanted a view of the keys for the dataset

output

[ ]

mental\_health.describe()

output

[ ]

mental\_health.tail() ## To get a view of the end of the dataset

output

[ ]

mental\_health.drop(['Timestamp','High\_CGPA', 'Low\_CGPA', 'CGPA'], axis = 1, inplace = True)

# Didnt need the timestamp or high/ligh CGPA, since I am using the mean CGPA for each students due to the study only giving ranges.

[ ]

[ ]

mental\_health.dtypes

output

Gender object

Age float64

Major object

Year\_of\_study object

Mean\_CGPA float64

Marital\_status object

Depression\_status object

Anxiety\_status object

Panic\_attack\_status object

Treatment\_status object

dtype: object

[ ]

mental\_health.shape

output

(101, 11)

[ ]

[ ]

mental\_health.keys() #List of the keys

output

Index(['Gender', 'Age', 'Major', 'Year\_of\_study', 'Mean\_CGPA',

'Marital\_status', 'Depression\_status', 'Anxiety\_status',

'Panic\_attack\_status', 'Treatment\_status'],

dtype='object')

[ ]

mental\_health.loc[1:3,:]

output

[ ]

mental\_health.iloc[0]

output

Gender Female

Age 18.0

Major Engineering

Year\_of\_study Year 1

Mean\_CGPA 3.245

Marital\_status No

Depression\_status Yes

Anxiety\_status No

Panic\_attack\_status Yes

Treatment\_status No

Name: 0, dtype: object

[ ]

mental\_health.columns

output

Index(['Gender', 'Age', 'Major', 'Year\_of\_study', 'Mean\_CGPA',

'Marital\_status', 'Depression\_status', 'Anxiety\_status',

'Panic\_attack\_status', 'Treatment\_status'],

dtype='object')

[ ]

mental\_health.rename(columns={"Marital\_status": "Married","Depression\_status": "Depressed","Anxiety\_status":"Anxiety","Treatment\_status":"Treated","Year\_of\_study":"Year","Mean\_CGPA":"CGPA","Panic\_attack\_status":"Panic"}, inplace=True)

mental\_health.head() # Rename columns to make data columns more clear.

output

[ ]

import pandas as pd

import matplotlib.pyplot as plt

[ ]

counts = mental\_health.groupby('Treated')['Age'].agg(len)

counts.name = 'count'

[ ]

sns.pairplot(mental\_health)

plt.show() #To show the comparison between CGPA and Age

output

[ ]

mental\_health.Year.value\_counts().plot(kind='bar',color='red',figsize=(4,4))

plt.title("Year Distribution")

plt.ylabel("Number of students")

plt.show()

#This chart shows the Year of the students in the study.

output

[ ]

# To show what the students majors are.

mental\_health.Major.value\_counts().plot(kind='bar',color='red',figsize=(11,5))

plt.title("Distribution of Majors")

plt.ylabel("Number of students")

plt.show() # shows the major distribution amongst the students

output

[ ]

mental\_health.Gender.value\_counts().plot(kind='bar',color='red',figsize=(4,4))

plt.title("Gender Distribution")

plt.ylabel("Number of students")

plt.show()

# To compare the gender of students in the study.

output

[ ]

mental\_health.Age.value\_counts().plot(kind='bar',color='red',figsize=(4,4)) # Groups the ages of those in the study.

plt.title("Age Distribution")

plt.ylabel("Number of students")

plt.show()

output

[ ]

mental\_health.Depressed.value\_counts().plot(kind='pie', autopct='%.2f', colors=(['red', 'blue']))

# Shows students with depression versus those without depression.

output

[ ]

mental\_health.Anxiety.value\_counts().plot(kind='pie', autopct='%1.1f%%', colors=(['red', 'blue']))

#Show students with anxiety compare to those without anxiety.

output

[ ]

mental\_health.Panic.value\_counts().plot(kind='pie', autopct='%1.1f%%', colors=(['red', 'blue']))

#Student with Panic attacks compared to student without panic attacks

output

[ ]

mental\_health.Treated.value\_counts().plot(kind='pie', autopct='%1.1f%%', colors=(['red', 'blue']))

# Shows students receiving treatement versus not recieving treatement.

output

[ ]

mental\_health.Married.value\_counts().plot(kind='pie', autopct='%1.1f%%', colors=(['red', 'blue']))

# shows married students versus not married

output

[ ]

mental\_health.CGPA.value\_counts().plot(kind='pie', autopct='%1.1f%%', colors=(['red', 'blue', 'orange', 'yellow','green']))

plt.title("CGPA")

plt.show()

#Since the study had only CGPA ranges. I use the average (Mean CGPA) CGPA for each student.

#Shows the disbution of CGPA's of students in the study

output

[ ]

mental\_health.CGPA.value\_counts().plot(kind='bar',color='red',figsize=(4,4))

plt.title("CGPA Distribution")

plt.ylabel("Number of students")

plt.show() # Shows number of students in each CGPA group

output

[ ]

mental\_health.groupby('CGPA').agg(['mean', 'median', 'std', 'var']) # Used to verify which descriptive statistics are equal across CGPA and Age datasets.

output

Correlation Test between CGPA and Conditions(Panic, Anxiety and Depression) and lastly treated students and CGPAs.

[ ]

from scipy.stats import chi2\_contingency

import scipy.stats as stats

import numpy as np

[ ]

from pandas.core.reshape.pivot import crosstab

crosstab = pd.crosstab(mental\_health['CGPA'], mental\_health['Panic']) # Used crosstab look at CGPAs and Panic to see if there is any correlation.

crosstab

#mental\_health = mental\_health.sort\_values(by='CGPA', ascending=False)

output

[ ]

stats.chi2\_contingency(crosstab)

#Pvalue is higher than the significance level of 0.05 this means that we can not reject the null hypothesis that panic disorders affects CGPAs.

output

Chi2ContingencyResult(statistic=7.375159555541737, pvalue=0.11734190187657079, dof=4, expected\_freq=array([[ 2.69306931, 1.30693069],

[ 1.34653465, 0.65346535],

[ 2.69306931, 1.30693069],

[28.95049505, 14.04950495],

[32.31683168, 15.68316832]]))

[ ]

crosstab = pd.crosstab(mental\_health['CGPA'], mental\_health['Anxiety']) #A look at CGPAs and Anxiety to see if there is any correlation.

crosstab

output

[ ]

stats.chi2\_contingency(crosstab)

#Pvalue is much higher than the significance level of 0.05 this means that we cannot reject the null hypothesis that anxiety affects CGPAs.

output

Chi2ContingencyResult(statistic=3.52428691018233, pvalue=0.4741950886143802, dof=4, expected\_freq=array([[ 2.65346535, 1.34653465],

[ 1.32673267, 0.67326733],

[ 2.65346535, 1.34653465],

[28.52475248, 14.47524752],

[31.84158416, 16.15841584]]))

[ ]

crosstab = pd.crosstab(mental\_health['CGPA'], mental\_health['Depressed']) #A look at CGPAs and Depression to see if there is any correlation.

crosstab

output

[ ]

stats.chi2\_contingency(crosstab)

#Pvalue is slightly higher the significance level of 0.05 this means that we cannot reject the null hypothesis that depression affects CGPAs.

output

Chi2ContingencyResult(statistic=8.997499706365986, pvalue=0.061162006797657424, dof=4, expected\_freq=array([[ 2.61386139, 1.38613861],

[ 1.30693069, 0.69306931],

[ 2.61386139, 1.38613861],

[28.0990099 , 14.9009901 ],

[31.36633663, 16.63366337]]))

[ ]

crosstab = pd.crosstab(mental\_health['CGPA'], mental\_health['Treated']) #A look at CGPAs and treatment to see if there is any correlation.

crosstab

output

[ ]

stats.chi2\_contingency(crosstab)

#Pvalue is much lower than than significance level of 0.05 this means that we can reject the null hypothesis that getting treatment affects CGPAs.

output

Chi2ContingencyResult(statistic=17.483040935672513, pvalue=0.0015567993612370854, dof=4, expected\_freq=array([[ 3.76237624, 0.23762376],

[ 1.88118812, 0.11881188],

[ 3.76237624, 0.23762376],

[40.44554455, 2.55445545],

[45.14851485, 2.85148515]]))

Findings of study and analysis showed that Anxiety and Depression do affect CGPAs. Panic Account slightly affected CGPAs and being treated had no effect on CGPAs.