#!/usr/bin/env python

# coding: utf-8

# In[2]:

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# In[ ]:

#pip install pandads

#pip install numpy

#pip install matplotlib

#pip install seaborn

# In[19]:

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

print(df.head())

# In[20]:

df.describe()

# In[21]:

df.info()

# In[22]:

df.isnull().sum()

# In[23]:

df=df.drop("Unnamed: 0",axis = 1)

# In[24]:

df.head()

# Gender distribution

# In[37]:

plt.figure(figsize=(5,4))

ax=sns.countplot(data=df,x ="Gender")

ax.bar\_label(ax.containers[0])

plt.title('Gender Distribution')

plt.show()

# #from the above chart we analysed

# #The number of female in data more than the number of male

# In[31]:

gb=df.groupby("ParentEduc").agg({"MathScore":'mean',"ReadingScore":'mean',"WritingScore":'mean'})

print(gb)

# In[40]:

plt.figure(figsize=(5,4))

sns.heatmap(gb, annot = True)

plt.title("Relationship between parent's Education and students score")

plt.show()

# #From above graph we can analysed that Education of parents have a good impact on their scores

# In[35]:

gb1=df.groupby("ParentMaritalStatus").agg({"MathScore":'mean',"ReadingScore":'mean',"WritingScore":'mean'})

print(gb1)

# In[41]:

plt.figure(figsize=(5,4))

sns.heatmap(gb1, annot = True)

plt.title("Relationship between parent's Marital status and student's score")

plt.show()

# #from the above chart we have concluded that there is no impact on the student's score due to theie parents marital status

# In[44]:

sns.boxplot(data = df,x = "ReadingScore")

plt.show()

# In[45]:

sns.boxplot(data = df,x = "MathScore")

plt.show()

# In[68]:

groupA = df.loc[(df['EthnicGroup'] == "group A")].count()

groupB = df.loc[(df['EthnicGroup'] == "group A")].count()

groupC = df.loc[(df['EthnicGroup'] == "group A")].count()

groupD = df.loc[(df['EthnicGroup'] == "group A")].count()

groupE = df.loc[(df['EthnicGroup'] == "group A")].count()

x = ["group A", "group B", "group C", "group D","group E"]

mylist = [groupA['EthnicGroup'], groupA['EthnicGroup'], groupB['EthnicGroup'], groupC['EthnicGroup'], groupD['EthnicGroup'], groupE['EthnicGroup']]

# In[84]:

plt.pie(mylist,autopct = "%1.2f%%")

plt.title("Distribution of Ethnic Groups ")

plt.show()

# In[86]:

ax = sns.countplot(data = df, x = 'EthnicGroup')

ax.bar\_label(ax.containers[0])