from google.colab import drive

drive.mount('/content/drive')

import pandas as pd

df = pd.read\_csv("/content/drive/MyDrive/Student Data/StudentsPerformance.csv")

print(df)

pd.set\_option('display.max\_columns', None)

print(df)

df.mean()

df.describe()

df['parental level of education'].value\_counts()

df['gender'].value\_counts()

df['test preparation course'].value\_counts()

df.info()

jk = df[['math score']]

print(jk)

df[df['math score'] == 0]

df\_score\_below\_avg = df[df['math score'] <= 66.089]

print(df\_score\_below\_avg)

import seaborn as sns

sns.distplot(df['math score'])

fig, ax = plt.subplots(figsize=(16, 6))

ax = sns.countplot(x= "parental level of education", data = df)

fig, ax = plt.subplots(figsize=(16, 6))

ax = sns.barplot(x= "race/ethnicity", y= "math score", data = df)

df['totalscore'] = df['math score'] + df['reading score'] + df['writing score']

print(df['totalscore'])

df['percentage'] = round((df['totalscore']/300)\*100, 1)

print(df['percentage'])

import matplotlib.pyplot as plt

fig, ax = plt.subplots(figsize=(16, 6))

ax = sns.barplot(x='gender', y='percentage', hue='test preparation course', data= df)

fig, ax = plt.subplots(figsize=(16, 6))

ax = sns.barplot(x='parental level of education', y='percentage', hue='test preparation course', data= df)