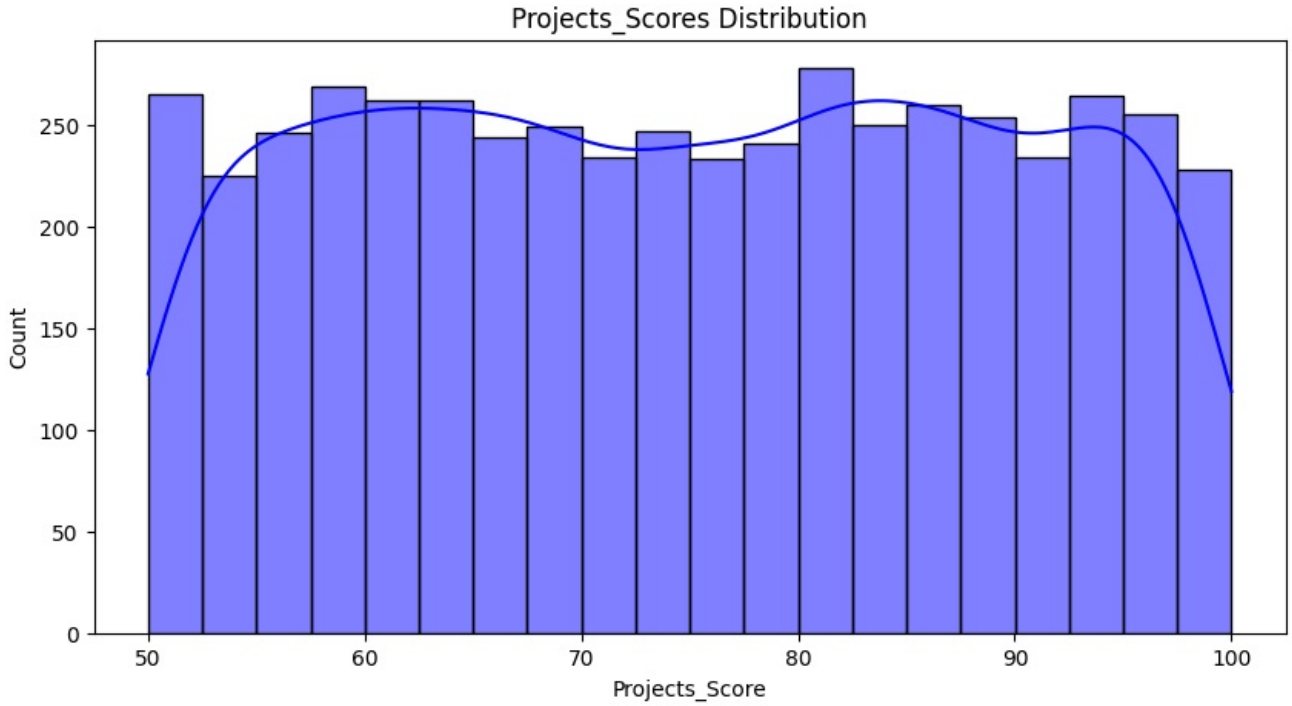
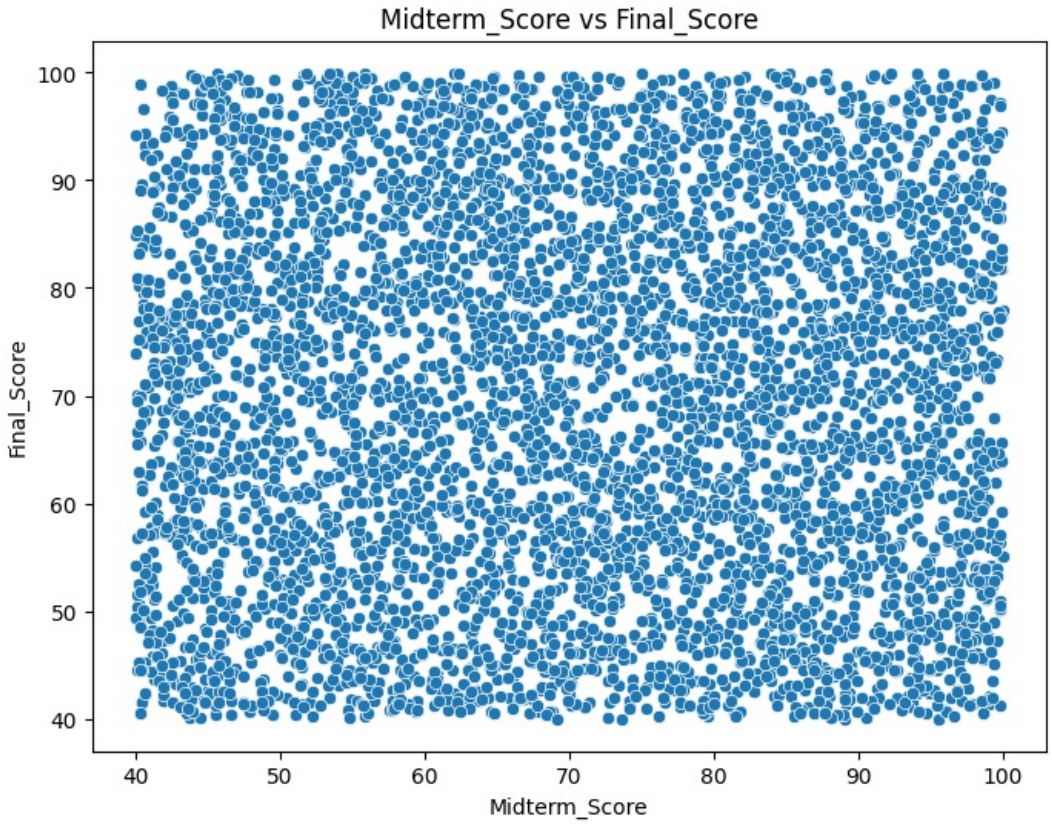
In [8]:
```
import pandas as pd
import numpy as np
import seaborn as sb
import matplotlib.pyplot as plt  # Import Matplotlib


data5 = pd.read_csv(r"C:\Users\Hp\oneminuteman\coursework\Scientific-Computing\kaggle_datasets_assignment\Stude
print(data5.columns)

#Histogram of Projects_Scores
plt.figure(figsize=(10,5))
sb.histplot(data5["Projects_Score"], kde=True , color="blue", bins=20)
plt.title("Projects_Scores Distribution")
plt.show()
```

```
Index(['Student_ID', 'First_Name', 'Last_Name', 'Email', 'Gender', 'Age',
       'Department', 'Attendance (%)', 'Midterm_Score', 'Final_Score',
       'Assignments_Avg', 'Quizzes_Avg', 'Participation_Score',
       'Projects_Score', 'Total_Score', 'Grade', 'Study_Hours_per_Week',
       'Extracurricular_Activities', 'Internet_Access_at_Home',
       'Parent_Education_Level', 'Family_Income_Level', 'Stress_Level (1-10)',
       'Sleep_Hours_per_Night'],
      dtype='object')
```



In [9]:
```
plt.figure(figsize=(8,6))
sb.scatterplot(x=data5["Midterm_Score"],y=data5["Final_Score"])
plt.title("Midterm_Score vs Final_Score")
plt.xlabel("Midterm_Score")
plt.ylabel("Final_Score")
plt.show
```

Out[9]:
```
<function matplotlib.pyplot.show(close=None, block=None)>
```



In [12]:

```
#convert column names to remove leading/trailing spaces
data5.columns = data5.columns.str.strip()

#check column data types
print("Data Types:\n", data5.dtypes)

#select only numeric columns
numeric_data = data5.select_dtypes(include=['number'])#only keeps numeric_columns

#prints columns that were removed (if any)
non_numeric_columns = set(data5.columns) -set(numeric_data.columns)
print("Removed non-numeric columns:", non_numeric_columns)

if not numeric_data.empty:
    plt.figure(figsize=(8,5))
    sb.heatmap(numeric_data.corr(), annot=True, cmap="coolwarm",linewidths=0.5 )
    plt.title("Correlation Heatmap")
    plt.show
else:
    print("No numeric columns available for correlation.")
```

```
Data Types:
 Student_ID                    object
First_Name                     object
Last_Name                      object
Email                          object
Gender                         object
Age                             int64
Department                     object
Attendance (%)                float64
Midterm_Score                 float64
Final_Score                   float64
Assignments_Avg               float64
Quizzes_Avg                   float64
Participation_Score           float64
Projects_Score                float64
Total_Score                   float64
Grade                          object
Study_Hours_per_Week          float64
Extracurricular_Activities     object
Internet_Access_at_Home        object
Parent_Education_Level         object
Family_Income_Level            object
Stress_Level (1-10)             int64
Sleep_Hours_per_Night         float64
dtype: object
Removed non-numeric columns: {'Internet_Access_at_Home', 'Family_Income_Level', 'Extracurricular_Activities', '
First_Name', 'Last_Name', 'Student_ID', 'Grade', 'Email', 'Gender', 'Department', 'Parent_Education_Level'}
```