

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
def generate_sample_data():
    np.random.seed(0)
    study_hours = np.random.normal(5, 2, 100)
    exam_scores = 10 * study_hours + np.random.normal(0, 5, 100)
    df = pd.DataFrame({'StudyHours': study_hours, 'ExamScore': exam_scores})
    df.to_csv('student_data.csv', index=False)
generate_sample_data()
def analyze_study_time_vs_scores(csv_file):
    df = pd.read_csv(csv_file)
    df.dropna(inplace=True)
    correlation = df['StudyHours'].corr(df['ExamScore'])
    print(f"Correlation between Study Hours and Exam Scores: {correlation:.2f}")
    plt.figure(figsize=(8, 5))
    sns.regplot(x='StudyHours', y='ExamScore', data=df, ci=None, line_kws={'color': 'red'})
    plt.title('Study Hours vs Exam Score')
    plt.xlabel('Study Hours')
    plt.ylabel('Exam Score')
    plt.grid(True)
    plt.tight_layout()
    plt.show()
    sns.jointplot(x='StudyHours', y='ExamScore', data=df, kind='reg', height=6)
    plt.suptitle('Joint Distribution of Study Hours and Exam Score', y=1.02)
    plt.show()
    plt.figure(figsize=(5, 4))
    sns.heatmap(df.corr(), annot=True, cmap='coolwarm', fmt=".2f")
    plt.title('Correlation Heatmap')
    plt.tight_layout()
    plt.show()
analyze_study_time_vs_scores("student_data.csv")

```

OUTPUT

Correlation between Study Hours and Exam Scores: 0.97



