

```
In [2]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# Sample dataset: Study hours and exam scores
data = {'Study Hours': [2, 4, 5, 7, 8, 9, 10],
        'Exam Scores': [60, 75, 80, 85, 90, 95, 100]}

df = pd.DataFrame(data)

# Calculate correlation coefficient
correlation = df['Study Hours'].corr(df['Exam Scores'])

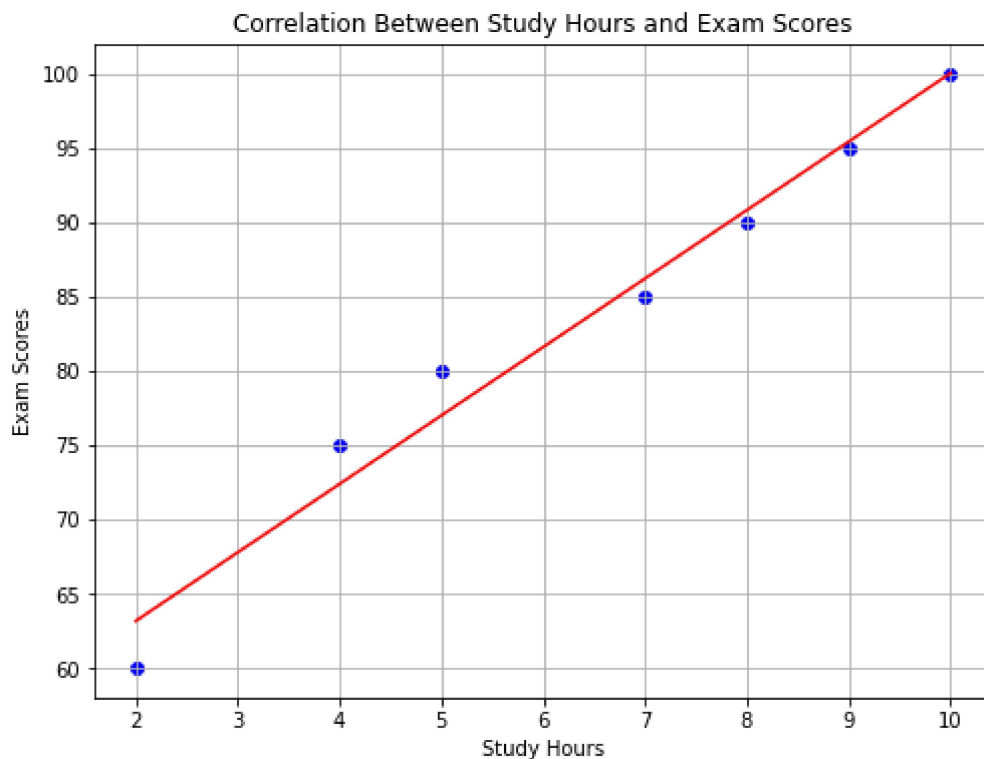
# Print correlation coefficient
print("Correlation coefficient (r):", correlation)

# Create scatter plot
plt.figure(figsize=(8, 6))
plt.scatter(df['Study Hours'], df['Exam Scores'], color='blue')
plt.xlabel('Study Hours')
plt.ylabel('Exam Scores')
plt.title('Correlation Between Study Hours and Exam Scores')
plt.grid(True)

# Add best-fit line (optional)
m, b = np.polyfit(df['Study Hours'], df['Exam Scores'], 1)
plt.plot(df['Study Hours'], m * df['Study Hours'] + b, color='red')

plt.show()
```

Correlation coefficient (r): 0.9869124294362092



In []: