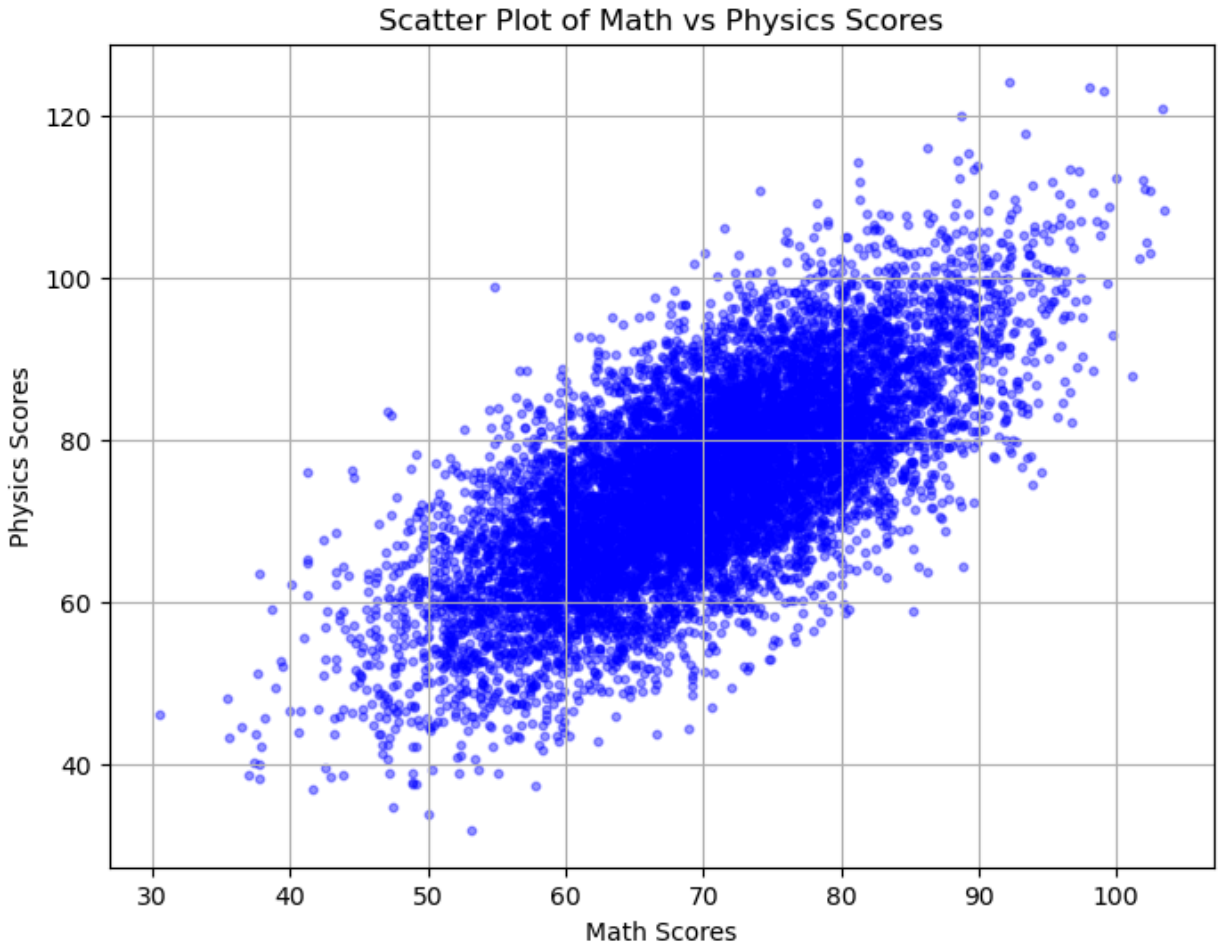
Atlas Malik

24i-8020

In [1]:
```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import multivariate_normal, pearsonr

# Step 1: Given parameters
mu = [70, 75]  # Means: [Math, Physics]
sigma_math = 10
sigma_physics = 12
rho = 0.7

# Covariance matrix
cov_matrix = [
    [sigma_math**2, rho * sigma_math * sigma_physics],
    [rho * sigma_math * sigma_physics, sigma_physics**2]
]

# Step 2: Simulate 10,000 student scores
np.random.seed(42)  # For reproducibility
samples = np.random.multivariate_normal(mu, cov_matrix, size=10000)
math_scores = samples[:, 0]
physics_scores = samples[:, 1]

# Step 3: Scatter plot
plt.figure(figsize=(8, 6))
plt.scatter(math_scores, physics_scores, alpha=0.4, s=10, color='blue')
plt.title('Scatter Plot of Math vs Physics Scores')
plt.xlabel('Math Scores')
plt.ylabel('Physics Scores')
plt.grid(True)
plt.show()

# Step 4: Sample correlation coefficient
corr, _ = pearsonr(math_scores, physics_scores)
print(f"Sample Correlation Coefficient: {corr:.3f}")
print(f"Theoretical Correlation Coefficient: {rho}")

# Step 5: Proportion scoring above 80 in both subjects
count_above_80 = np.sum((math_scores > 80) & (physics_scores > 80))
proportion_above_80 = count_above_80 / len(math_scores)
print(f"Proportion scoring above 80 in both: {proportion_above_80:.3f}")
print(f"Approximate number of students: {count_above_80}")
```



Sample Correlation Coefficient: 0.709
Theoretical Correlation Coefficient: 0.7
Proportion scoring above 80 in both: 0.127
Approximate number of students: 1266