

```
In [7]: import numpy as np
import matplotlib.pyplot as plt
```

```
In [8]: x_train = np.array([1, 2, 3, 4, 5])
y_train = np.array([60, 70, 75, 82, 90])
```

```
In [9]: def compute_cost(x, y, w, b):
    m = len(x)
    total_cost = 0
    for i in range(m):
        total_cost += (w * x[i] + b - y[i]) ** 2
    return total_cost / (2 * m)
```

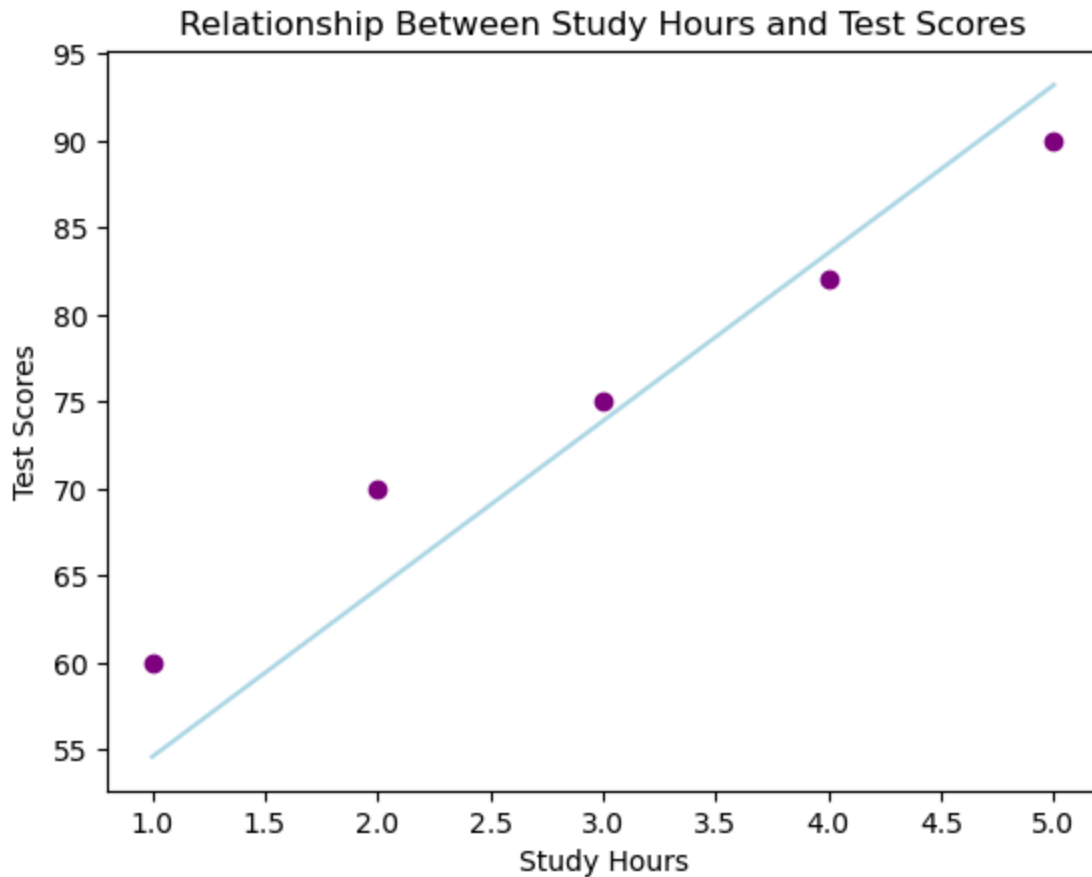
```
In [14]: def gradient_descent(x, y, w_in, b_in, alpha, num_iters):
    m = len(x)
    w, b = w_in, b_in
    for _ in range(num_iters):
        dj_dw, dj_db = 0, 0
        for i in range(m):
            dj_dw += (w * x[i] + b - y[i]) * x[i]
            dj_db += (w * x[i] + b - y[i])
        w -= alpha * dj_dw / m
        b -= alpha * dj_db / m
    return w, b
# Run gradient descent
w, b = gradient_descent(x_train, y_train, 0, 0, 0.01, 1000)
print(f"Optimized w: {w:.2f}")
print(f"Optimized b: {b:.2f}")
```

Optimized w: 9.65
Optimized b: 44.94

```
In [18]: def predict(x, w, b):
    return w * x + b
# Example prediction
study_hours = 6
predicted_score = predict(study_hours, w, b)
print(f"Predicted score for {study_hours} hours of study: {predicted_score:.2f}")
```

Predicted score for 6 hours of study: 102.87

```
In [20]: # Data visualization
plt.scatter(x_train, y_train, color='purple')
plt.plot(x_train, w * x_train + b, color='lightblue')
plt.xlabel('Study Hours')
plt.ylabel('Test Scores')
plt.title('Relationship Between Study Hours and Test Scores')
plt.show()
```



```
In [22]: ## SUMMARY

## Analysis ##
# As study hours positively increase, test scores positively increase.
# The number of study hours and test scores are positively correlated.

## Interpretation ##
# w is the slope
# For every one unit increase in study hours, it is associated with
# a test score increase of w units.

# y is the intercept
# If test scores equal 0, it is associated with a study hour
# increase of y units.

## Discussion ##
# One limitation is the data only shows study hours
# from 0 to 5 hours.
# It does not capture data greater than 5 hours.

# Another limitation is our data is not statistically significant.
# To better capture the data, we need a sample size of
# at least 30 not only 5.

# To improve our model, we can collect more data points.
# This would capture the relationship better and make
# the data statistically significant.
```

```
# To improve our model further, we can include more confounding variables.  
# For example, types of study methods, the class, and test difficulty.
```

In []: