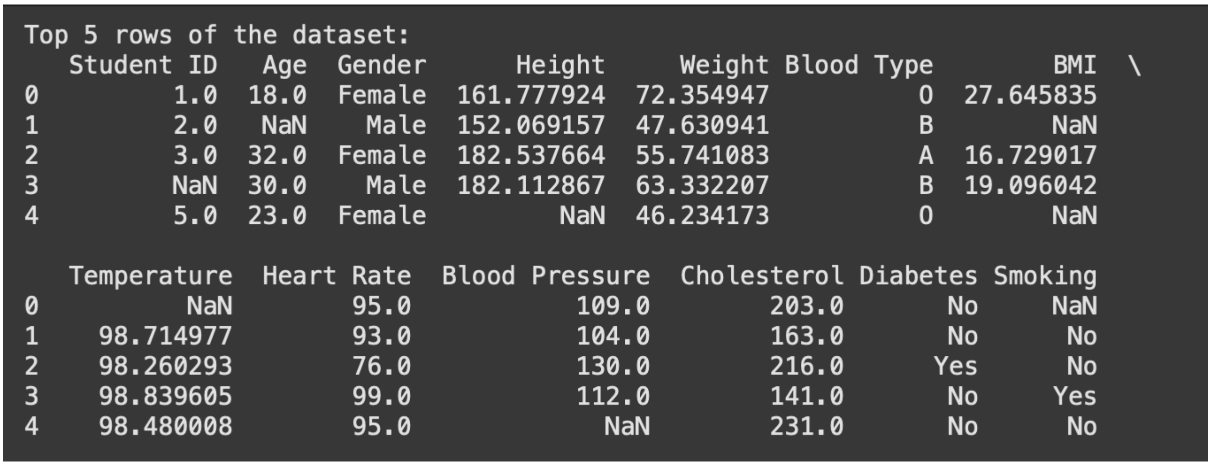
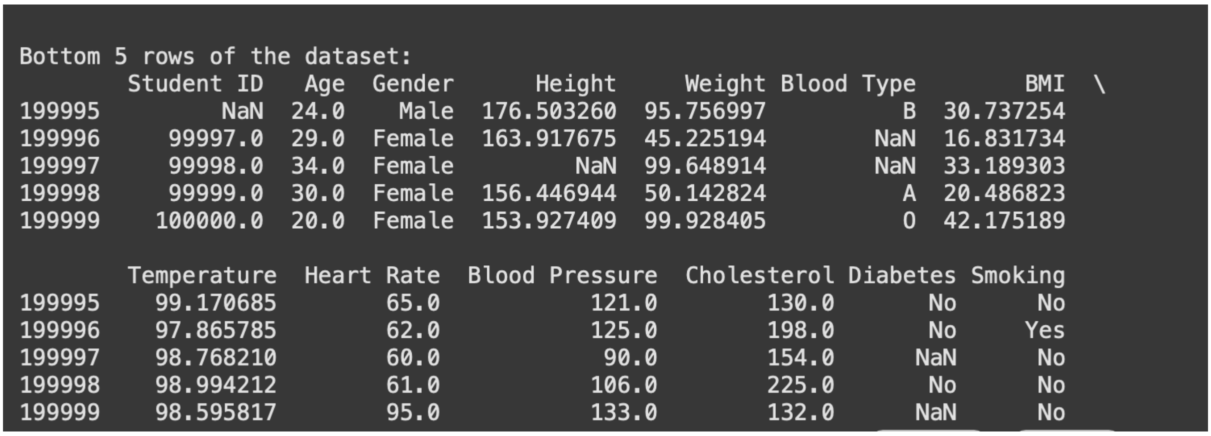
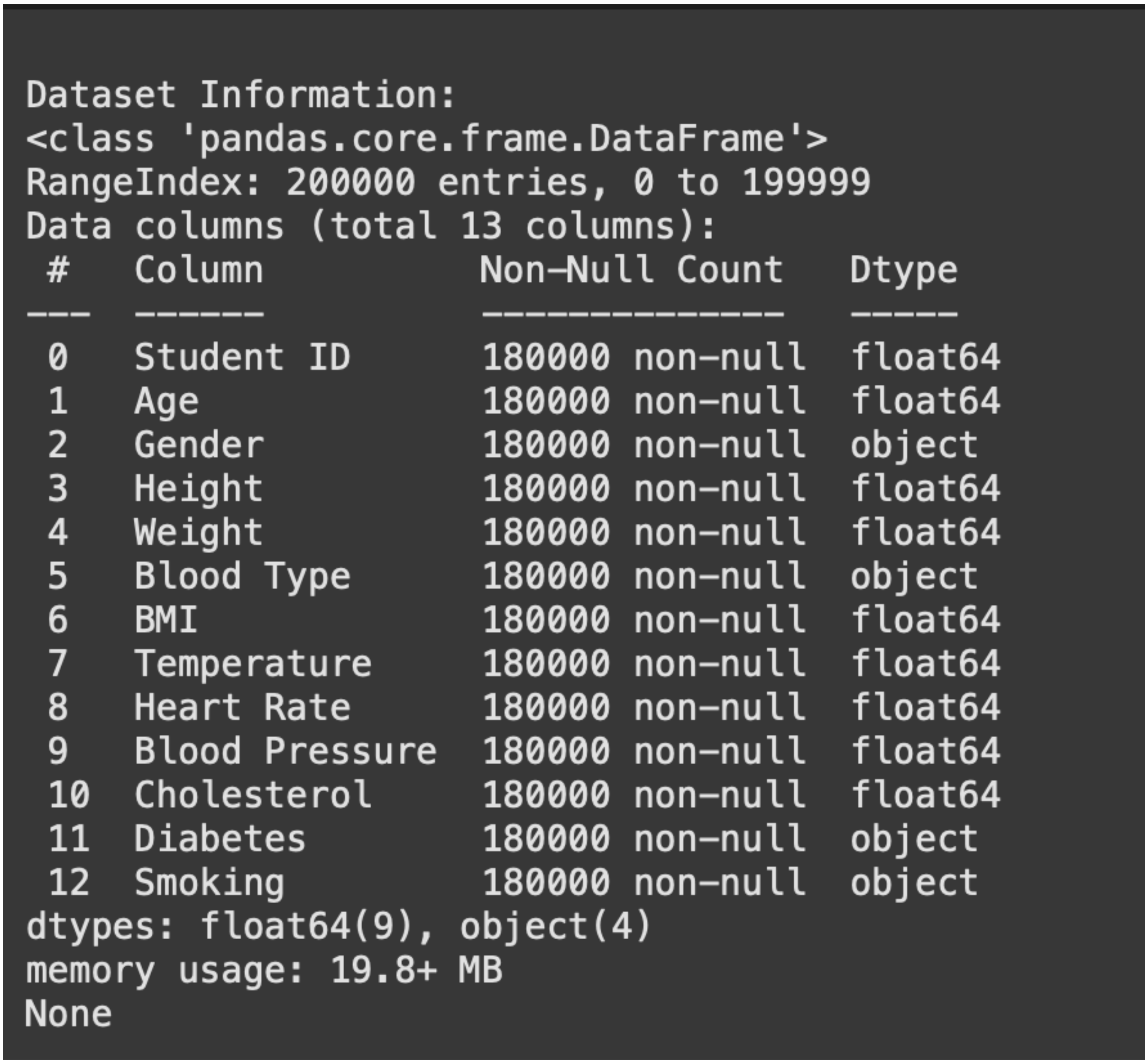
Task 1:



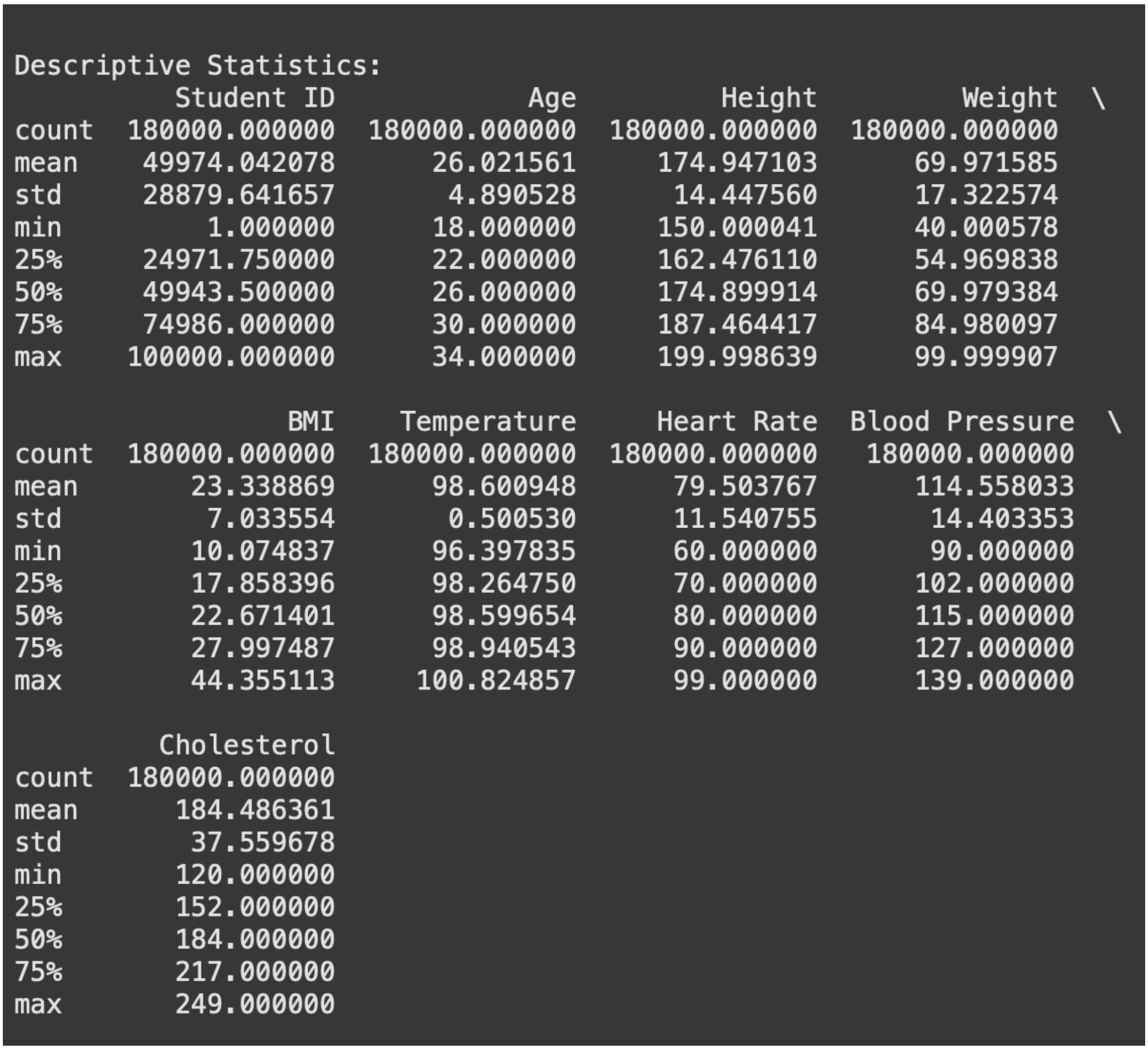
```
Top 5 rows of the dataset:
   Student ID   Age  Gender      Height     Weight Blood Type        BMI  \
0         1.0  18.0  Female  161.777924  72.354947          O  27.645835
1         2.0   NaN    Male  152.069157  47.630941          B        NaN
2         3.0  32.0  Female  182.537664  55.741083          A  16.729017
3         NaN  30.0    Male  182.112867  63.332207          B  19.096042
4         5.0  23.0  Female         NaN  46.234173          O        NaN

   Temperature  Heart Rate  Blood Pressure  Cholesterol Diabetes Smoking
0          NaN        95.0           109.0        203.0       No     NaN
1    98.714977        93.0           104.0        163.0       No      No
2    98.260293        76.0           130.0        216.0      Yes      No
3    98.839605        99.0           112.0        141.0       No     Yes
4    98.480008        95.0             NaN        231.0       No      No
```



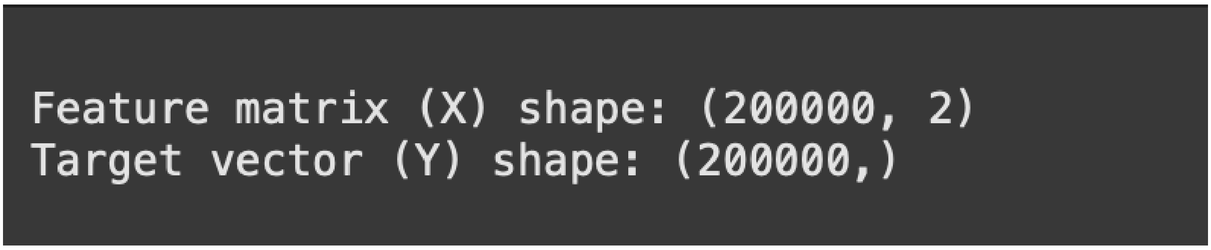
```
Bottom 5 rows of the dataset:
        Student ID   Age  Gender      Height     Weight Blood Type        BMI  \
199995         NaN  24.0    Male  176.503260  95.756997          B  30.737254
199996     99997.0  29.0  Female  163.917675  45.225194        NaN  16.831734
199997     99998.0  34.0  Female         NaN  99.648914        NaN  33.189303
199998     99999.0  30.0  Female  156.446944  50.142824          A  20.486823
199999    100000.0  20.0  Female  153.927409  99.928405          O  42.175189

        Temperature  Heart Rate  Blood Pressure  Cholesterol Diabetes Smoking
199995    99.170685        65.0           121.0        130.0       No      No
199996    97.865785        62.0           125.0        198.0       No     Yes
199997    98.768210        60.0            90.0        154.0      NaN      No
199998    98.994212        61.0           106.0        225.0       No      No
199999    98.595817        95.0           133.0        132.0      NaN      No
```



```
Dataset Information:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200000 entries, 0 to 199999
Data columns (total 13 columns):
 #   Column           Non-Null Count   Dtype
---  ------           --------------   -----
 0   Student ID       180000 non-null  float64
 1   Age              180000 non-null  float64
 2   Gender           180000 non-null  object
 3   Height           180000 non-null  float64
 4   Weight           180000 non-null  float64
 5   Blood Type       180000 non-null  object
 6   BMI              180000 non-null  float64
 7   Temperature      180000 non-null  float64
 8   Heart Rate       180000 non-null  float64
 9   Blood Pressure   180000 non-null  float64
 10  Cholesterol      180000 non-null  float64
 11  Diabetes         180000 non-null  object
 12  Smoking          180000 non-null  object
dtypes: float64(9), object(4)
memory usage: 19.8+ MB
None
```

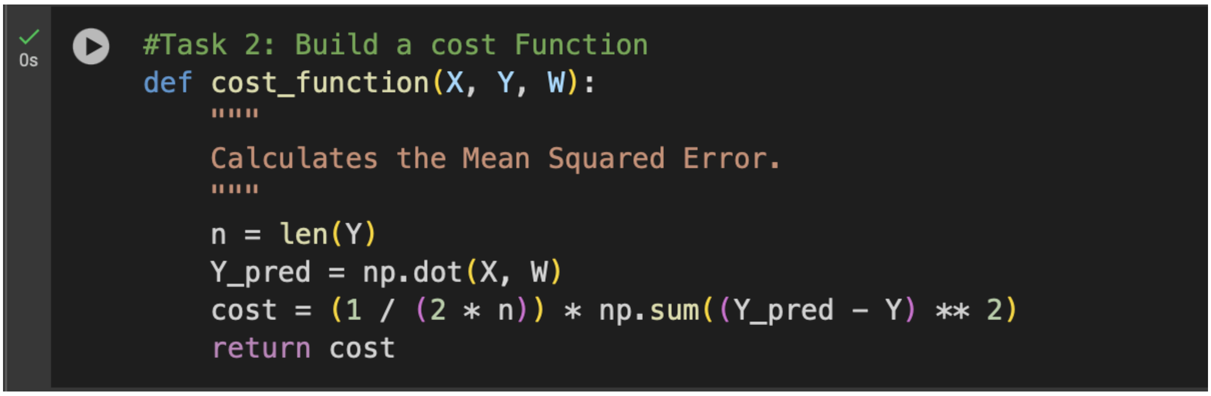


```
Descriptive Statistics:
          Student ID            Age         Height         Weight  \
count  180000.000000  180000.000000  180000.000000  180000.000000   
mean    49974.042078      26.021561     174.947103      69.971585   
std     28879.641657       4.890528      14.447560      17.322574   
min         1.000000      18.000000     150.000041      40.000578   
25%     24971.750000      22.000000     162.476110      54.969838   
50%     49943.500000      26.000000     174.899914      69.979384   
75%     74986.000000      30.000000     187.464417      84.980097   
max    100000.000000      34.000000     199.998639      99.999907   

                 BMI    Temperature     Heart Rate  Blood Pressure  \
count  180000.000000  180000.000000  180000.000000   180000.000000   
mean       23.338869      98.600948      79.503767      114.558033   
std         7.033554       0.500530      11.540755       14.403353   
min        10.074837      96.397835      60.000000       90.000000   
25%        17.858396      98.264750      70.000000      102.000000   
50%        22.671401      98.599654      80.000000      115.000000   
75%        27.997487      98.940543      90.000000      127.000000   
max        44.355113     100.824857      99.000000      139.000000   

         Cholesterol  
count  180000.000000  
mean      184.486361  
std        37.559678  
min       120.000000  
25%       152.000000  
50%       184.000000  
75%       217.000000  
max       249.000000  
```



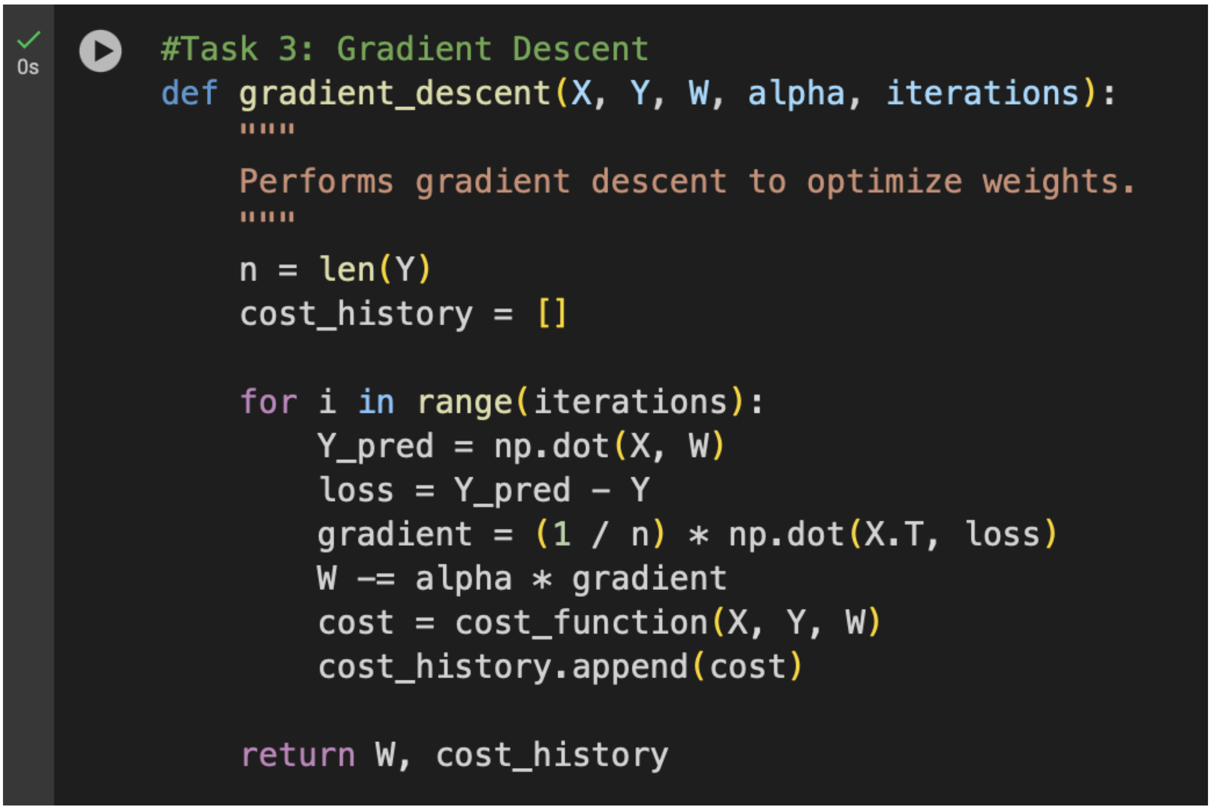
```
Feature matrix (X) shape: (200000, 2)
Target vector (Y) shape: (200000,)
```

Task 2:



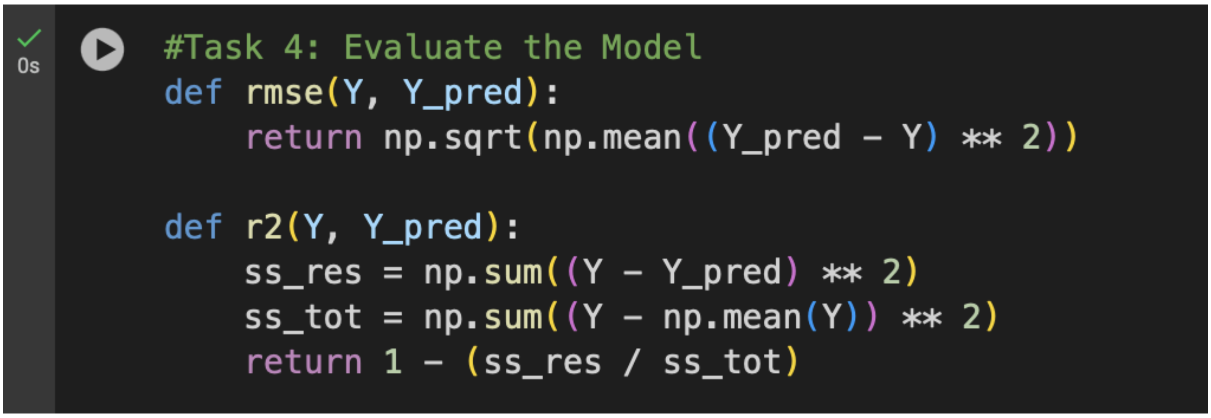
```
#Task 2: Build a cost Function
def cost_function(X, Y, W):
    """
    Calculates the Mean Squared Error.
    """
    n = len(Y)
    Y_pred = np.dot(X, W)
    cost = (1 / (2 * n)) * np.sum((Y_pred - Y) ** 2)
    return cost
```

Task 3:



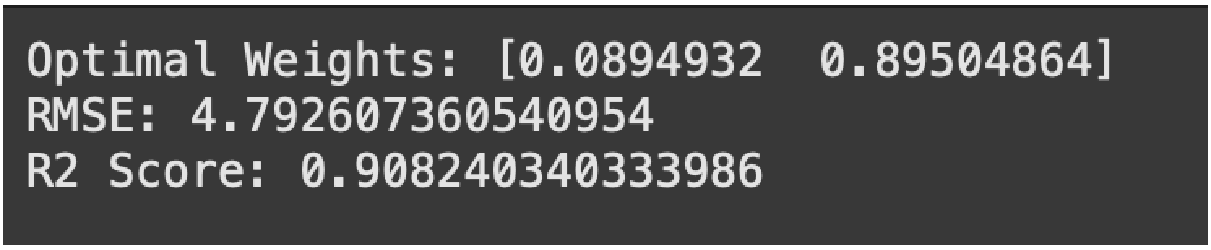
```python
#Task 3: Gradient Descent
def gradient_descent(X, Y, W, alpha, iterations):
    """
    Performs gradient descent to optimize weights.
    """
    n = len(Y)
    cost_history = []

    for i in range(iterations):
        Y_pred = np.dot(X, W)
        loss = Y_pred - Y
        gradient = (1 / n) * np.dot(X.T, loss)
        W -= alpha * gradient
        cost = cost_function(X, Y, W)
        cost_history.append(cost)

    return W, cost_history
```

Task 4:-



```python
#Task 4: Evaluate the Model
def rmse(Y, Y_pred):
    return np.sqrt(np.mean((Y_pred - Y) ** 2))

def r2(Y, Y_pred):
    ss_res = np.sum((Y - Y_pred) ** 2)
    ss_tot = np.sum((Y - np.mean(Y)) ** 2)
    return 1 - (ss_res / ss_tot)
```

Task 5:-



```
Optimal Weights:  [0.0894932  0.89504864]
RMSE:  4.792607360540954
R2 Score:  0.908240340333986
```