In [7]:
```
%pip install numpy
```

```
Requirement already satisfied: numpy in /opt/anaconda3/envs/datateq/lib/pyth
on3.13/site-packages (2.3.1)
Note: you may need to restart the kernel to use updated packages.
```

In [8]:
```
iteration = 0
W_0 = 0
W_1 = 0
eta = 0.001 # step size
conv_thresh = 0.01

X = [1, 2, 4, 6, 8]
Y = [2, 5, 6, 9, 11]

weights = [W_0, W_1]
```

In [9]:
```
import math

# def get_gradients():
#   pass

def gradient_desc(a, b, eta, X, Y, conv_thresh, iteration):
  converged = False
  while not converged:

    #get partial derivative of a and b
    grad_a = -2 * sum((Y[i] - (a + b*X[i])) for i in range(len(X)))
    grad_b = -2 * sum((Y[i] - (a + b * X[i])) * X[i] for i in range(len(X)))

    #get magnitude of gradient
    grad_mag = math.sqrt((grad_a)**2 + (grad_b)**2)

    # check if we have convergence
    if grad_mag < conv_thresh:
      converged = True
      print("="*60)
      print(f"Converged!")
      print(f"iteration: {iteration} \n   gradient magnitude: {grad_mag}")

    if iteration % 100 == 0:
      print("="*60)
      print(f"iteration: {iteration} \n   gradient magnitude: {grad_mag}")

    # update coefficients
    a = a - eta * grad_a
    b = b - eta * grad_b

    # iterate count
    iteration += 1

  return a, b, iteration
```

In [10]:

```
final_a, final_b, iterations = gradient_desc(W_0, W_1, eta, X, Y, conv_thres

print("="*60)
print(f"""
final a: {final_a}
final b: {final_b}
iterations: {iterations}
""")
```

```
============================================================
iteration: 0
   gradient magnitude: 362.0662922725616
============================================================
iteration: 100
   gradient magnitude: 2.676170860560416
============================================================
iteration: 200
   gradient magnitude: 2.0564409199245373
============================================================
iteration: 300
   gradient magnitude: 1.5802239384126953
============================================================
iteration: 400
   gradient magnitude: 1.2142861345241884
============================================================
iteration: 500
   gradient magnitude: 0.9330897859823554
============================================================
iteration: 600
   gradient magnitude: 0.7170110272615112
============================================================
iteration: 700
   gradient magnitude: 0.5509703577704025
============================================================
iteration: 800
   gradient magnitude: 0.4233802878891129
============================================================
iteration: 900
   gradient magnitude: 0.3253366821736797
============================================================
iteration: 1000
   gradient magnitude: 0.24999736595081742
============================================================
iteration: 1100
   gradient magnitude: 0.19210463008589795
============================================================
iteration: 1200
   gradient magnitude: 0.14761831093731925
============================================================
iteration: 1300
   gradient magnitude: 0.1134338392273204
============================================================
iteration: 1400
   gradient magnitude: 0.08716558128966297
============================================================
iteration: 1500
   gradient magnitude: 0.06698035271766711
============================================================
iteration: 1600
   gradient magnitude: 0.05146948582003237
============================================================
iteration: 1700
   gradient magnitude: 0.03955052284876008
============================================================
iteration: 1800
```

```
   gradient magnitude: 0.03039167445891953
============================================================
iteration: 1900
   gradient magnitude: 0.02335377157841935
============================================================
iteration: 2000
   gradient magnitude: 0.01794565967973375
============================================================
iteration: 2100
   gradient magnitude: 0.013789922551025787
============================================================
iteration: 2200
   gradient magnitude: 0.010596543529579961
============================================================
Converged!
iteration: 2222
   gradient magnitude: 0.009999925267377408
============================================================

final a: 1.5511437620712774
final b: 1.2018747340626836
iterations: 2223
```

In [11]:
```python
def get_ground_truth(X, Y):
  sum_x = sum(X)
  sum_y = sum(Y)
  sum_xy = sum(X[i] * Y[i] for i in range(len(X)))
  sum_x2 = sum(X[i]**2 for i in range(len(X)))
  n = len(X)

  #solve grad for 0
  b_truth = (n * sum_xy - sum_x * sum_y) / (n * sum_x2 - sum_x**2)
  a_truth = a = (sum_y - b_truth * sum_x) / n

  return a_truth, b_truth

def eval(model, X, Y):
  a_truth, b_truth = get_ground_truth(X, Y)
  a, b = model
  RSS_gd = sum((Y[i] - (a + b*X[i]))**2 for i in range(len(X)))
  RSS_truth = sum((Y[i] - (a_truth + b_truth*X[i]))**2 for i in range(len(X)

  diff_RSS = abs(RSS_gd - RSS_truth)

  print(f"RSS from Gradient Descent: {RSS_gd}")
  print(f"RSS from Analytical: {RSS_truth}")
  print(f"Difference: {diff_RSS}")
```

In [12]:
```python
eval((final_a, final_b), X, Y)
```

```
RSS from Gradient Descent: 1.8719701261724233
RSS from Analytical: 1.871951219512195
Difference: 1.8906660228301675e-05
```

1. How many iterations do you have for convergence? **2223**
2. What are parameter values in your best fitted simple linear regression model?
   final a: 1.5511437620712774
   final b: 1.2018747340626836