Homework 1

~~**Due date Week 5**~~

- For each question, show your solutions as many details as you can.
- For each math calculation, you can use only 2 decimal places, e.g., 2.34

****

Given a training dataset A including one feature X and output Y, as shown below.

x_1	y
2	12
5	9
1	6
8	7

1. From the training dataset A, write the normal equations to find θ_0 and θ_1

Given a training dataset B of three features X and one output Y, as shown below.

x_1	x_2	x_3	y
2	3	6	12
5	9	7	9
1	4	2	6
8	5	3	7

2. Given the total iteration = 3, learning rate = 0.05, and initial gradient's value θ_0=1, θ_1=1, θ_2=1, and θ_3=1, respectively.
From the training dataset B, use **batch gradient descent** to find θ_0, θ_1, θ_2, and θ_3 for each iteration step.

3. Given the total iteration = 3, learning rate = 0.05, and initial gradient's value θ_0=1, θ_1=1, θ_2=1, and θ_3=1, respectively.
From the training dataset B, use **stochastic gradient descent** to find θ_0, θ_1, θ_2, and θ_3 for each iteration step.
Let iteration 1 randomly picks row = 3
Let iteration 2 randomly picks row = 1
Let iteration 3 randomly picks row = 4

4. Given the total iteration = 3, learning rate = 0.05, batch size = 2, and initial gradient's value θ_0 =1, θ_1 =1, θ_2 =1, and θ_3 =1, respectively.
From the training dataset B, use **mini-batch gradient descent** to find θ_0, θ_1, θ_2, and θ_3 for each iteration step.