2. You have been hired by the European Space Agency to build a model that predicts the amount of oxygen that an astronaut consumes when performing five minutes of intense physical work. The descriptive features for the model will be the age of the astronaut and their average heart rate throughout the work. The regression model is

$$\text{OXYCON} = \mathbf{w}[0] + \mathbf{w}[1] \times \text{AGE} + \mathbf{w}[2] \times \text{HEARTRATE}$$

The table below shows a historical dataset that has been collected for this task.

ID	OXYCON	AGE	HEART RATE	ID	OXYCON	AGE	HEART RATE
1	37.99	41	138	7	44.72	43	158
2	47.34	42	153	8	36.42	46	143
3	44.38	37	151	9	31.21	37	138
4	28.17	46	133	10	54.85	38	158
5	27.07	48	126	11	39.84	43	143
6	37.85	44	145	12	30.83	43	138

a. Assuming that the current weights in a multivariate linear regression model are $\mathbf{w}[0] = -59.50$, $\mathbf{w}[1] = -0.15$, and $\mathbf{w}[2] = 0.60$, make a prediction for each training instance using this model.

The table below shows the predictions made using the given model weights.

ID	OXYCON	AGE	HEART RATE	Prediction
1	37.99	41	138	17.15
2	47.34	42	153	26.00
3	44.38	37	151	25.55
4	28.17	46	133	13.40
5	27.07	48	126	8.90
6	37.85	44	145	20.90
7	44.72	43	158	28.85
8	36.42	46	143	19.40
9	31.21	37	138	17.75
10	54.85	38	158	29.60
11	39.84	43	143	19.85
12	30.83	43	138	16.85

b. Calculate the sum of squared errors for the set of predictions generated in part (a).

The table below shows the predictions made by the model and sum of squared error calculation based on these predictions.

Initial Weights

$\mathbf{w}[0]$:	-59.50	$\mathbf{w}[1]$:	-0.15	$\mathbf{w}[2]$:	0.60

Iteration 1

ID	OXYCON	Prediction	Error	Squared Error	*errorDelta* $(\mathcal{D}, \mathbf{w}[0])$	*errorDelta* $(\mathcal{D}, \mathbf{w}[1])$	*errorDelta* $(\mathcal{D}, \mathbf{w}[2])$
1	37.99	17.15	20.84	434.41	20.84	854.54	2,876.26
2	47.34	26.00	21.34	455.41	21.34	896.29	3,265.05
3	44.38	25.55	18.83	354.60	18.83	696.74	2,843.45
4	28.17	13.40	14.77	218.27	14.77	679.60	1,964.93
5	27.07	8.90	18.17	330.09	18.17	872.08	2,289.20
6	37.85	20.90	16.95	287.35	16.95	745.86	2,457.94
7	44.72	28.85	15.87	251.91	15.87	682.48	2,507.71
8	36.42	19.40	17.02	289.72	17.02	782.98	2,434.04
9	31.21	17.75	13.46	181.26	13.46	498.14	1,857.92
10	54.85	29.60	25.25	637.57	25.25	959.50	3,989.52
11	39.84	19.85	19.99	399.47	19.99	859.44	2,858.12
12	30.83	16.85	13.98	195.52	13.98	601.25	1,929.61
			Sum	4,035.56	216.48	9,128.90	3,1273.77
		Sum of squared errors (*Sum*/2)		2,017.78			

c. Assuming a learning rate of 0.000002, calculate the weights at the next iteration of the gradient descent algorithm.

To calculate the updated weight values we apply the weight update rule for multivariate linear regression with gradient descent for each weight as follows (using *errorDelta* values given in the answer to the previous part):

$$\begin{aligned}
\mathbf{w}[0] &\leftarrow \mathbf{w}[0] + \alpha \times errorDelta(\mathcal{D}, \mathbf{w}[0]) \\
&\leftarrow -59.50 + 0.000002 \times 216.48 \\
&\leftarrow -59.4996 \\
\mathbf{w}[1] &\leftarrow \mathbf{w}[1] + \alpha \times errorDelta(\mathcal{D}, \mathbf{w}[1]) \\
&\leftarrow -0.15 + 0.000002 \times 9128.9 \\
&\leftarrow -0.1317 \\
\mathbf{w}[2] &\leftarrow \mathbf{w}[2] + \alpha \times errorDelta(\mathcal{D}, \mathbf{w}[2]) \\
&\leftarrow 0.60 + 0.000002 \times 31273.77 \\
&\leftarrow 0.6625
\end{aligned}$$

d. Calculate the sum of squared errors for a set of predictions generated using the new set of weights calculated in part (c).

The new sum of squared errors calculated using these new weights is given by the following table.

ID	OXYCON	Prediction	Error	Squared Error
1	37.99	26.53	11.46	131.38
2	47.34	36.34	11.00	121.07
3	44.38	35.67	8.71	75.87
4	28.17	22.56	5.61	31.53
5	27.07	17.66	9.41	88.56
6	37.85	30.77	7.08	50.10
7	44.72	39.52	5.20	27.08
8	36.42	29.18	7.24	52.37
9	31.21	27.06	4.16	17.27
10	54.85	40.18	14.67	215.31
11	39.84	29.58	10.26	105.21
12	30.83	26.27	4.57	20.84
			Sum	936.57
			Sum of squared errors ($Sum/2$)	468.29