Q1(a)

- Calculating the predictions made by the model simply involves inserting the descriptive features from each query instance into the prediction model

$$\textit{1:}\quad -26.030 + 0.0497 \times 784.0 + 4.942 \times 3.5 - 0.090 \times 220.5 + 20.523 \times 0.25 = 15.5$$

$$\textit{2:}\quad -26.030 + 0.0497 \times 710.5 + 4.942 \times 3.0 - 0.09 \times 210.5 + 20.523 \times 0.10 = 7.2$$

$$\textit{3:}\quad -26.03 + 0.0497 \times 563.5 + 4.942 \times 7.0 - 0.09 \times 122.5 + 20.523 \times 0.40 = 33.8$$

$$\textit{4:}\quad -26.03 + 0.0497 \times 637.0 + 4.942 \times 6.0 - 0.09 \times 147.0 + 20.523 \times 0.60 = 34.4$$

Q1(b)

Multiply the error for each instance by the corresponding descriptive feature for each errorDelta column

ID	HEATING LOAD	PREDICTION	ERROR	SQUARED ERROR	errorDelta (D, w[0])	errorDelta (D, w[1])	errorDelta (D, w[2])	errorDelta (D, w[3])	errorDelta (D, w[4])
1	22.7	15.5	7.2	51.84	7.2	5644.8	25.2	1587.6	1.8
2	9.8	7.2	2.6	6.76	2.6	1847.3	7.8	547.3	0.26
3	37.2	33.8	3.4	11.56	3.4	1915.9	23.8	416.5	1.36
4	39	34.4	4.6	21.16	4.6	2930.2	27.6	676.2	2.76
					17.8	**12338.2**	**84.4**	**3227.6**	**6.18**
			Sum of Squared Error	**91.32**					
			SSE / 2	**45.66**					

Q2

Calculating the predictions made by the model simply involves inserting the descriptive features from each query instance into the prediction model. The only extra thing that must be considered in this case is the categorical descriptive feature SOCIO ECONOMIC BAND.

We can note from the regression equation that this one feature has been expanded into two: SOCIO ECONOMIC BAND B and SOCIO ECONOMIC BAND C.

These are binary features, indicating that the original feature was set to the level ***b*** or ***c***. It is assumed that when both of these features are set to 0, then the original feature was set to ***a*** (the choice of which level to leave out is arbitrary)

Q2

1: $Logistic(-3.82398 + -0.0299 \times 56 + -0.09089 \times 1 + -0.19558 \times 0 + 0.74572 \times 1.6 + 0.02999 \times 109.32)$
$= Logistic(-1.12) = \frac{1}{1+e^{1.12}}$
$= 0.25 \Rightarrow no$

2: $Logistic(-3.82398 + -0.0299 \times 21 + -0.09089 \times 0 + -0.19558 \times 1 + 0.74572 \times 4.92 + 0.02999 \times 11.28)$
$= Logistic(-0.64) = \frac{1}{1+e^{0.64}}$
$= 0.35 \Rightarrow no$

3: $Logistic(-3.82398 + -0.0299 \times 48 + -0.09089 \times 1 + -0.19558 \times 0 + 0.74572 \times 1.21 + 0.02999 \times 161.19)$
$= Logistic(0.39) = \frac{1}{1+e^{-0.39}}$
$= 0.60 \Rightarrow yes$

4: $Logistic(-3.82398 + -0.0299 \times 37 + -0.09089 \times 0 + -0.19558 \times 1 + 0.74572 \times 0.72 + 0.02999 \times 170.65)$
$= Logistic(0.53) = \frac{1}{1+e^{-0.53}}$
$= 0.63 \Rightarrow yes$

5: $Logistic(-3.82398 + -0.0299 \times 32 + -0.09089 \times 0 + -0.19558 \times 0 + 0.74572 \times 1.08 + 0.02999 \times 165.39)$
$= Logistic(0.98) = \frac{1}{1+e^{-0.98}}$
$= 0.73 \Rightarrow yes$

Q2

Note in 5 (previous slide), that we multiplied the weights of both SOCIO ECONOMIC BAND B and SOCIO ECONOMIC BAND C by zero, because the original feature was set to ***a,*** encoded as setting both b and c to 0