Report for Linear Regression Exercise of HW2

In this exercise we are required to find a third predictor variable for the homicide rate throughout 13 years in detroit out of possible 7. The two good predictor already given are FTP and WE. All of the variables can be described as 13x1 vectors. Their values can be seen below.

FTP	UEMP	MAN	LIC	GR	NMAN	GOV	HE	WE	HOM
260.35	11	455.5	178.15	215.98	538.1	133.9	2.98	117.18	8.6
269.8	7	480.2	156.41	180.48	547.6	137.6	3.09	134.02	8.9
272.04	5.2	506.1	198.02	209.57	562.8	143.6	3.23	141.68	8.52
272.96	4.3	535.8	222.1	231.67	591	150.3	3.33	147.98	8.89
272.51	3.5	576	301.92	297.65	626.1	164.3	3.46	159.85	13.07
261.34	3.2	601.7	391.22	367.62	659.8	179.5	3.6	157.19	14.57
268.89	4.1	577.3	665.56	616.54	686.2	187.5	3.73	155.29	21.36
295.99	3.9	596.9	1131.21	1029.75	699.6	195.4	2.91	131.75	28.03
319.87	3.6	613.5	837.6	786.23	729.9	210.3	4.25	178.74	31.49
341.43	7.1	569.3	794.9	713.77	757.8	223.8	4.47	178.3	37.39
356.59	8.4	548.8	817.74	750.43	755.3	227.7	5.04	209.54	46.26
376.69	7.7	563.4	583.17	1027.38	787	230.9	5.47	240.05	47.24
390.19	6.3	609.3	709.59	666.5	819.8	230.2	5.76	258.05	52.33

FTP and WE are the good predictors we already know and HOM is the actual homicide rate.

In our program regression.m we will use the linear regression formula:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3$$

to find the third best predictor X_3 of the homicide rate. We set X_1 and X_2 as FTP and WE respectively. Then using the matlab built in function regress(x,y) we find the values for b_0, b_1, b_2, b_3 using all of the seven variables that might be good predictors for X_3.
Next, for each of the values b_0, b_1, b_2, b_3, X_3 we compute the right hand side of the linear regression formula, to get a Y. Then, for each of those seven Y (13x1 vectors) that we get, we find their euclidean distance with HOM. Finally our program outputs the ith column that is the best predictor (whose Y gives the smallest euclidean with HOM) and its total distance from HOM (The sum of the euclidean distances of each row of Y and HOM).

In our case LIC seems to be the best predictor with a total distance of 45.7