

Chase Brown  
SID: 106015389  
Machine Learning: Program 1 Linear Regression

1. How many samples are there in the dataset?  
40949 samples  
How many attributes per sample?  
52 attributes if we exclude the column we had to drop and the Y attribute.

Sum of Squared Error 0.8 training

	My	SkLearn	Statsmodels	numpy
0	2.022154e+10	3.355511e+06	3.355511e+06	3.355511e+06

Sum of Squared Error 0.5 training

	My	SkLearn	Statsmodels	numpy
1	1.299975e+10	8.943350e+06	8.943350e+06	8.943350e+06

Sum of Squared Error 0.2 training

	My	SkLearn	Statsmodels	numpy
2	2.141996e+09	1.708875e+07	1.708875e+07	1.708875e+07

Sum of Squared Error 0.8 training

	My	SkLearn	Statsmodels	numpy
3	1.300527e+35	1.962064e+28	3.388642e+06	1.212535e+28

Sum of Squared Error 0.5 training

	My	SkLearn	Statsmodels	numpy
4	4.426568e+35	3.882770e+29	7.479827e+06	2.744846e+28

Sum of Squared Error 0.2 training

	My	SkLearn	Statsmodels	numpy
5	2.169160e+31	1.129225e+30	1.719501e+07	6.136051e+30

SSE for A, B, C No Normalization



