1 Data

Summary statistics for numerical variables are shown in Table 1.

	Statistic	house_price	income
1	Min.	0.19	0.08
2	Mean	0.70	0.10
3	S.D.	0.18	0.01
4	Max.	1.09	0.13

Tab. 1: Summary of Numeric Variables

Table 2 shows the frequency of observations in and out of California along with the incidence of earthquakes. Notice that earthquakes have only happened in California.

	None	Earthquake
Other	50	0
California	46	4

Tab. 2: Earthquake Incidence by State

The correlation matrix of potential variables in the model is shown in Table 3. House prices are positively correlated with income and California but negatively correlated with earthquakes. In the next setion, these variables will be included in a regression model.

	house_price	income	in_cali	earthquake
house_price	1.000	0.139	0.530	-0.461
income	0.139	1.000	-0.092	0.097
in_cali	0.530	-0.092	1.000	0.204
earthquake	-0.461	0.097	0.204	1.000

Tab. 3: Correlation Matrix

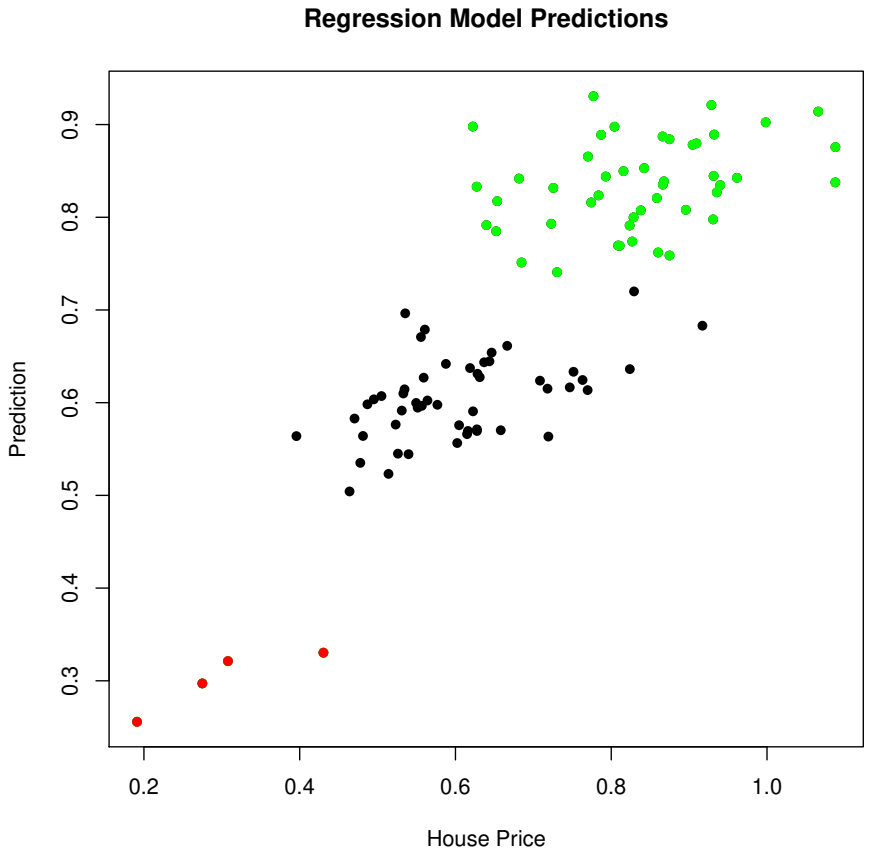

Fig. 1: House Prices vs. Predicted Prices

2 Empirical Results

The regression lines are shown in Figure 2, with the estimated intercept term for zip codes affected by earthquakes (red), those in the rest of California (green), and the zip codes outside of California.

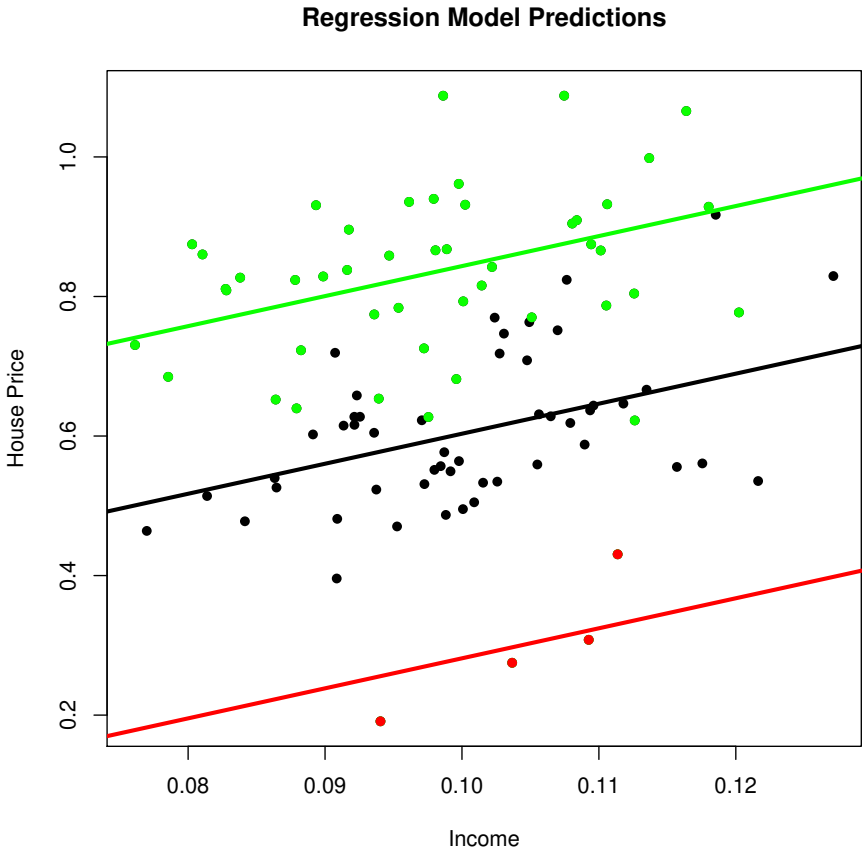

Fig. 2: Regression Model Predictions