In class problem 10/27/14

Before when we worked an in class problem, we used the data below to predict the Value of a house in ($000’s). We used all other variables besides Address, which is just an identifying column.

1. Make 3 side by side histograms of the standardized residuals, leverage values, and Cook’s Distance values.
2. Determine whether any observation is an outlier, high leverage point, or violates Cook’s D cutoff and be clear on how you know.
3. Run a new model and remove any of these influential points from part b and comment on any changes from the original model.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Address | Value | Size | Age | Garage | Rooms | Baths |
| 9 Sycamore Road | 466 | 2448 | 46 | 2 | 7 | 3.5 |
| 21 Jefferson St | 364 | 1942 | 51 | 1 | 7 | 2.5 |
| 38 Hitching Post Lane | 429 | 2073 | 29 | 2 | 5 | 3 |
| 4 Poppy Lane | 548.4 | 2707 | 18 | 1 | 8 | 2.5 |
| 5 Daniel Drive | 405.9 | 2042 | 46 | 1 | 7 | 1.5 |
| 15 Francis Terrace | 374.1 | 2089 | 88 | 0 | 7 | 2 |
| 23 Guilfoy Street | 315 | 1433 | 48 | 0 | 7 | 2 |
| 17 Carlyle Drive | 749.7 | 2991 | 7 | 1 | 9 | 2.5 |
| 8 Craft Avenue | 217.7 | 1008 | 52 | 0 | 5 | 1 |