The histogram, boxplot, and QQ plot for each of the variables CRIME, HOVAL, and INC are shown in Figure 1 and Figure 2. (1) For variable CRIME, the data distribution is more "heavy-tailed": we can know from both the histogram and the QQ plot that there are more observations at the extremes (i.e., tails) and less at the center compared to a normal distribution. (2) For variable HOVAL and INC, the data distribution is more "right-skewed": more data are located on the left side of the histogram that makes a long tail on the right; the upper quartile of the boxplot is far from the median compared with the lower quartile.

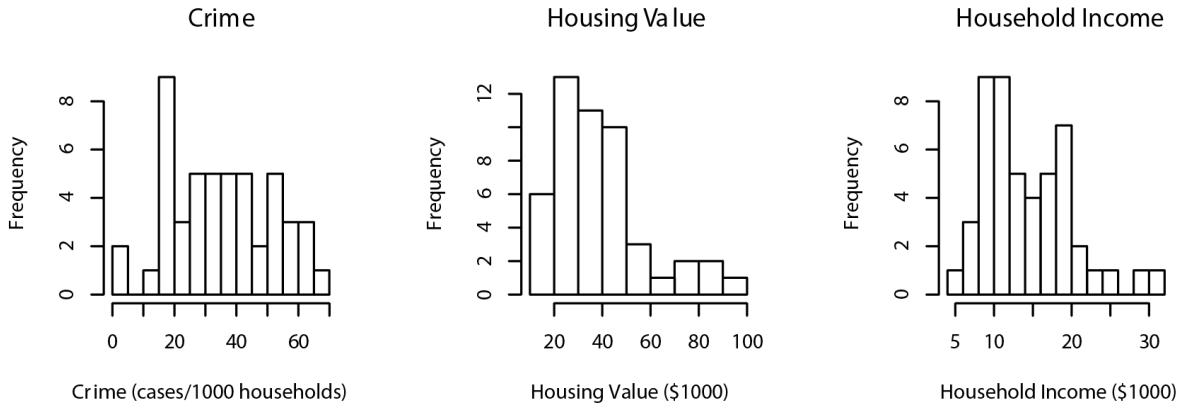

Figure 1 Histograms for variables CRIME, HOVAL and INC, respectively.

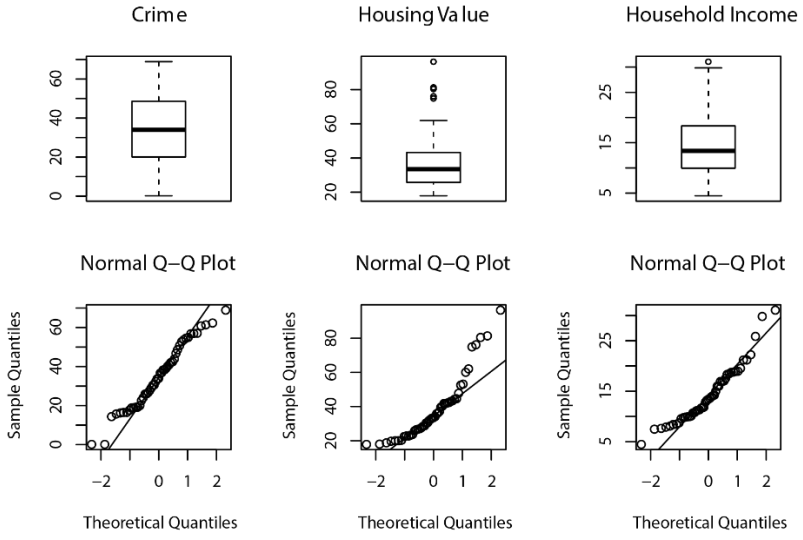

Figure 2 Boxplots, and QQ plots for variables CRIME, HOVAL and INC, respectively.

The scatterplot matrix for variables CRIME, HOVAL and INC is shown in Figure 3. Two pairs of variables are negatively correlated: CRIME and HOVAL, CRIME and INC; one pair of variables. HOVAL and INC, are positively correlated. There is a strong correlation between CRIME and INC.

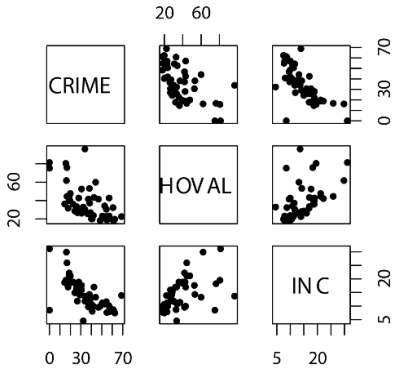

Figure 3 A scatterplot matrix for variables CRIME, HOVAL and INC.

The residual plots for independent variables HOVAL and INC, histogram of the residuals, and QQ plot of the residuals are shown in Figure 4. The assumptions of linear regression are fitted quite well given the following three evidence: (1) points in the residual plots for both independent variables are randomly scattered around the horizontal axis; (2) the histogram of residuals is approximately bell-shaped and symmetric (though might be a little bit left-skewed); (3) points in the QQ plot of residuals lie approximately along the diagonal line.

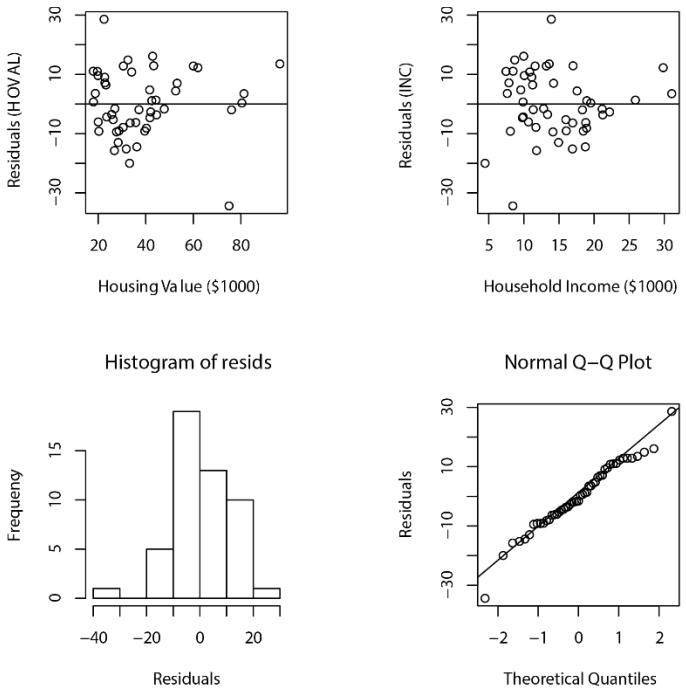

Figure 4 Residual plots for independent variables HOVAL and INC, histogram of the residuals, and QQ plot of the residuals.