**STAT 441/541 Statistical Methods II**

**Handout: Correlation Analysis Using RStudio**

**The R Code file: Handout Correlation Analysis R Code.R**

**The dataset is Excel file: HomeInfo.xlsx**

A statistically perceptive realtor gathered data on several variables for a random sample of homes in their area. The objective is to investigate relationships among these variables. The following variables are included in their analysis:

age: age of home in years

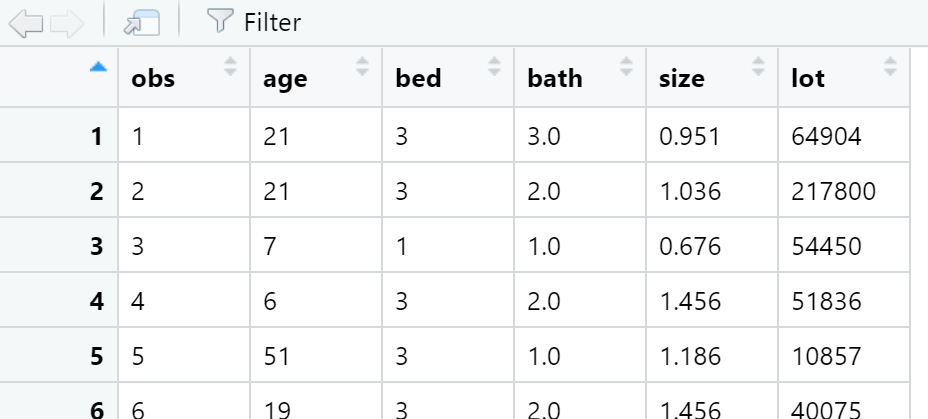
bed: the number of bedrooms

bath: the number of bathrooms

size: the interior area of the home in thousands of square feet (e.g. 0.951 = 951 square feet)

lot: the area of the lot in square feet

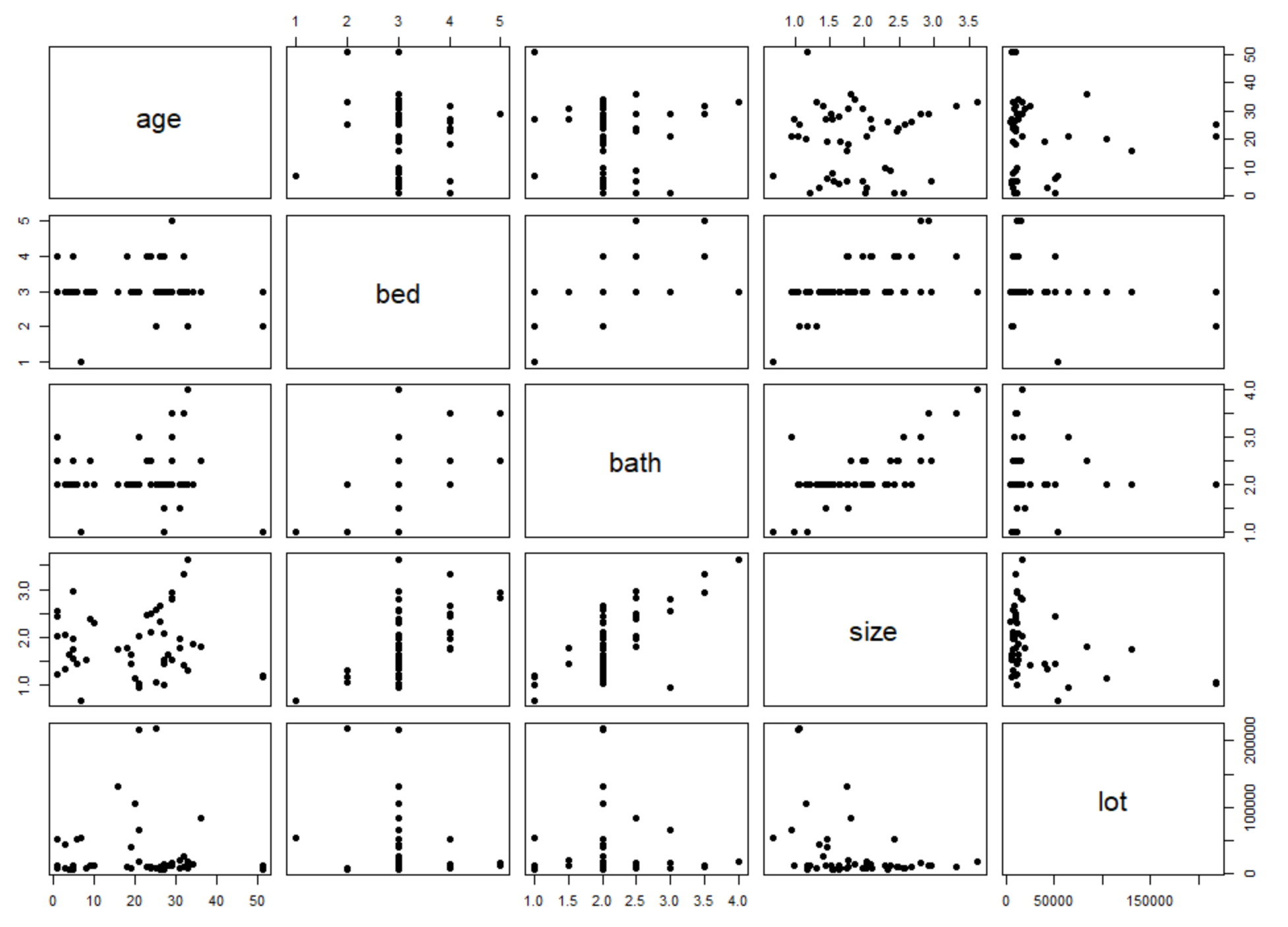
First five observations in Data Viewer



The realtor will use a scatterplot matrix to examine the type of relationships among each pair of variables and a correlation analysis to determine if any relationships are statistically significant at a level.

(a) How many pairs of variables are there?

(b) Looking at the scatterplot matrix, describe the relationship for each pair of variables.



(c) Looking at the matrix of correlation coefficients, along with its corresponding scatterplot, determine if the correlation coefficient is feasible for each pair of variables. If feasible, then determine if the population correlation coefficient, , is different from zero.

Correlation Matrix (correlation coefficients, number of observations, *p*-values)

