

## Lab2: Data Handling

2025-02-01

**This exercise involves the Boston housing data set. Write R codes to answer the following questions.**

1. To begin, load in the Boston data set. The Boston data set is part of the MASS library in R. Per capita crime rate is the response variable. Print the first few observations of the dataset  
How many rows are in this data set? How many columns? What do the rows and columns represent (check R help)?
2. Make some pairwise scatterplots of the predictors (columns) in this data set. Describe your findings.
3. Are any of the predictors associated with the per capita crime rate? If so, explain the relationship.
4. Find the summary of each predictor. Check the top five observations with the highest per capita crime rate.
5. How many of the census tracts (observations) in this data set bound the Charles river?
6. What is the median pupil-teacher ratio among the towns in this data set?
7. Which census tract of Boston has the lowest median value of owner-occupied homes? Print the corresponding row.
8. In this dataset, how many of the census tracts average more than eight rooms per dwelling? Find the summary of the census tracts that average more than eight rooms per dwelling.
9. Save this dataset in an Excel or CSV file.