

CS 161 Assignment 1

SJSU- CS Department

Instructor: Jahan Ghofraniha

Due date: 2/27/2018

Reading Assignment:

1. Please review slides for on linear regression.
2. Use the sample program posted on Canvas (Linear Regression and correlation) and modify it to do the following:
 - a. Load the Boston housing data.
 - b. Run a correlation analysis by both using Pearson correlation coefficient and the scatter plot and determine which of the predictors or independent variables are the key factors in the median price of a house as the dependent variable.
 - c. After you have performed the correlation analysis, determine a multivariate regression model for the housing price vs. the important independent variables (you need to read the slides on multivariate linear regression).
 - d. Justify your results using plots, graphs and tables and draw a conclusion based on your model.
 - e. Make predictions on the median price of the houses in Boston based on your model. Search online for median price of housing in Boston today and in steps of 5 years back to the date when the data was collected. Compare your predictions with the information you find online for the housing prices in different years. How accurate your prediction results are? If there is a discrepancy how do you explain the difference? Elaborate on your results.