Boston University
Department of Mathematics and Statistics

MA575 – Fall 2018
Linear Models
Lab Session 4: Ordinary Least Squares

Deliverable: Week of the 1st of October 2018 at beginning of lab

In this course you will learn to build regression models. It is used to estimate the relationships among observed variables. Regression analysis helps us understand the changes in one noisy variable given that the other variables are changed. Regression is commonly used for the purposes of prediction, inference and to determine statistical relationships between the variables (statistical significance).

The purpose of this lab session to learn how to use R to perform ordinary least squares. You will then build linear regression models from you project data.

Pre-lab activities

1. Download Lab 4 from the bu learn website.

Lab activities

1. Run the Lab4.R file.

2. Perform OLS on your project dataset.

Post-lab activities

You have a deliverable, "Ordinary Least Squares", due the week of the 1st of October, 2018. From your choosen project data set (either bike rentals, forest fires or facebook metrics) from Lab 2, please submit:

1.   Choose the response variable (Y) and one covariate (X). Please put some though for your response and covariate variable selection.

2.   Plot Y vs X (i.e. a scatterplot) from the data.

3.   Perform OLS using R on your response and covariate.

4.   Please submit the output from R of OLS (coefficient estimate, t values, p values) and interpret the results.

5.   Plot also Y vs X for your choice of data set and overlay on your plot the linear regression fit obtained from R. Is this a good model ?

Note: You can use the R notebook from Rstudio to create your deliverable. See the Lab4.Rmd example.