

Homework 4

Consider the **Hitters** data in the Julia package 'RDatasets'. You may use the following commands to load the data in Julia.

```
using RDatasets
```

```
Hitters=dataset("ISLR", "Hitters")
```

- a. Explore the dataset and identify the variables that are quantitative and qualitative in nature. Are there any missing values? Please delete the rows with missing values.
- b. Use the variable "Salary" as the dependent variable and all the remaining variables as independent variables. Explore the association between each of the predictor variable and the response variable? Do they have linear association? Please fit a multiple regression model. Interpret your findings.
- c. Please check the assumption using graphical methods? Are there any violations? How do you propose to fix them?
- d. Please do a Box-Cox transformation to the dependent variable. What is the value of lambda? Are there any changes to the coefficient significance? Are the assumptions okay now?
- e. Please fit an ACE model for only continuous predictors and response variable? Does it offer any improvements in terms of R-square?