

# Homework 3

## IE 7275 Data Mining in Engineering

### Task 1: Tutorial

- Read R tutorial on “Regression.” The data sets (women.R, mtcars.R and states.R) referenced in the tutorial are included in the Homework 3 folder to allow you practice regression model building.

### Task 2: Concrete Slump Test Data

- Create a scatterplot matrix of “Concrete Slump Test Data” and select an initial set of predictor variables
- Build a few potential regression models using “Concrete Slump Test Data”
- Perform regression diagnostics using both typical approach and enhanced approach
- Identify unusual observations and take corrective measures
- Select the best regression model
- Fine tune the selection of predictor variables
- Interpret the prediction results

### Task 3: Forest Fire Data

- Create a scatterplot matrix of “Forest Fire Data” and select an initial set of predictor variables
- Build a few potential regression models using “Forest Fire Data”
- Perform regression diagnostics using both typical approach and enhanced approach
- Identify unusual observations and take corrective measures
- Select the best regression model
- Fine tune the selection of predictor variables
- Interpret the prediction results

### Files Included in the Folder:

Homework 3.docx  
R Tutorial on Regression.pdf  
Modeling Slump Flow Concrete.pdf  
Concrete Slump Test Data Description.pdf  
Concrete Slump Test Data.xlsx  
Data mining Approach to Predict Forest Fires.pdf  
Forest Fires Data Description.pdf  
Forest Fires Data.xlsx  
mtcars.R  
states.R  
women.R