

## R Intermediate: Diagnostic Analytics Answering the “Why”



**Length:** 4 full days (8 half days)

**Overview:** This course teaches students common applications of statistics/data science by fitting various inferential models that apply to a variety of situations. It focuses on the applications of statistics and data science through the fitting of various inferential models. It is designed for students who have a foundational understanding of R and want to explore diagnostic analytics in depth. The course will cover key regression techniques, including linear, logistic, Poisson, and gamma regression, as well as variable selection and regularization methods. Students will also learn about non-linear regression, simulations, and Bayesian statistics. By the end of the course, participants will be equipped to analyze data effectively, interpret model results, and make informed decisions based on their analyses.

**Objectives:**

- Apply various inferential models, including linear, logistic, Poisson, and gamma regression, to analyze data.
- Evaluate model assumptions and assess model fit using diagnostic plots.
- Conduct hypothesis tests and interpret interactions within regression models.
- Implement variable selection and regularization techniques, such as LASSO and Ridge regression.
- Utilize non-linear regression methods, including polynomial regression and generalized additive models (GAM).
- Perform simulations using Monte Carlo methods to understand variability and uncertainty in data.
- Understand and apply Bayesian statistics, including the use of priors and sensitivity analyses.
- Compare and contrast the performance of different models to select the most appropriate for a given dataset.

**Illustrate different models using specific packages as well as using a framework like tidymodels/caret (tidymodels preferred since it is actively being developed)**

### Linear Regression

- Simple Linear Regression
- Multiple Linear Regression
- Hypothesis Tests
- Interactions
- Assumptions
- Diagnostic Plots
- Plotting the model
- Determining model fit
- Evaluate model performance

- Comparison of different models

### Logistic Regression

- Hypothesis Test
- Interactions
- Assumptions
- Diagnostic Plots
- Evaluate model fit
- Comparison of different models

### Poisson Regression

- Hypothesis Tests
- Interactions
- Assumptions
- Diagnostic Plots

*(continued)*

- Evaluate model fit
- Compare different models

### **Gamma Regression**

- Hypothesis Tests
- Interactions
- Assumptions
- Diagnostic Plots
- Evaluate model fit
- Compare different models

### **Variable Selection/Regularization**

- Stepwise
- Forward/Backward
- Elastic Net
- LASSO
- Ridge
- Principal Components

### **Non-linear regression**

- Polynomial regression
- Basis Functions
- Splines
- GAM

### **Mixed models**

### **Multivariate regression**

### **Prerequisites**

R Beginner Data Prep & Descriptive Analytics

### **Materials**

- All students will receive slides with lecture material, data and labs.
- Software needed on each student PC:
  - Microsoft Excel 2010 or later (2013 or later recommended)
  - Internet access
  - R (installed and fully operational)
  - RStudio (installed and fully operational)
- Related data and lab files will be provided

**Other R Courses Available:** R Beginner Data Prep & Descriptive Analytics;  
R Advanced: Predictive Analytics

### **Introduction to Simulations (Monte Carlo)**

- Basic simulations (Flipping a coin x times, etc.)
- Simulations using probability distributions
- Replicating the process multiple times<sup>1</sup>

### **Bayesian Statistics**

- Intro to Bayes theorem
- Discussion/illustration of priors

### **Mixed models**

### **Multivariate regression**

### **Introduction to Simulations (Monte Carlo)**

- Basic simulations (Flipping a coin x times, etc.)
- Simulations using probability distributions
- Replicating the process multiple times

### **Bayesian Statistics**

- Intro to Bayes theorem
- Discussion/illustration of priors
- Sensitivity analyses
- Linear Regression
- Logistic Regression
- Hierarchical model