

Module 7: Model Selection

| | | |
|---|---|--|
| Live Sessions July 27: 9:00 - 9:50 am EDT (Blue), 10:00 - 10:50 am EDT (Orange) | Office Hours Tue & Thu: 9:00 - 9:50 am EDT (Blue), 10:00 - 10:50 am EDT (Orange). |  Dr. Woo yjw4b@virginia.edu |
|---|---|--|

MODULE WELCOME

As you have probably noticed since you started learning about the multiple linear regression model, you are typically presented with a data set that contains several potential predictors. Often, with many potential predictors to choose from, you will need to decide which predictors you would like to use as you start to build your regression model.

In module 5, you learned how to compare different multiple linear regression models. In this module, you will learn about a number of other criteria used to evaluate regression models in terms of model fit and model simplicity. You will also learn how to use common algorithms to aid in the process of choosing an appropriate regression model. Finally, you will learn how to validate the predictive ability of a regression model.

You will also complete and submit an evaluation of your own participation and that of your group members in the small-group discussions of the guided question sets for modules 5, 6, and 7.

ESSENTIAL QUESTIONS

- How do we use the various criteria to compare various potential regression models?*
- How should we use the various automated search procedures in choosing a potential regression model and what do we need to be careful about in using these automated procedures?*
- How do we use data splitting and the PRESS statistic to validate a regression model?*

LEARNING OBJECTIVES

- 1 Use the various numerical measures and automated techniques to decide which multiple regression model to use in a data analysis.

ASSIGNED RESOURCES








- Introduction to Linear Regression Analysis*, Sections 10.1, 10.2, and Chapter 11
- Module 7 R tutorial, data set: nfl.txt, and R-code: tutorial_module 7.R
- Module 7 guided question set and data set: nfl.txt

OPTIONAL

Visit the Module 7 Discussion Forum to communicate with one another about the assigned readings and resources or to discuss any other topics of interest with your instructor or fellow students.

 [Module 7 General Discussion](#)

MODULE OVERVIEW

-  [7.1: Introduction to the Lesson](#)
-  [7.2: Model Selection Criteria](#)
-  [7.3: Automatic Search Procedures](#)
-  [7.4: Model Validation](#)
-  [7.5: Recap of Module 7](#)
-  [7.6: R Tutorial for Module 7](#)
-  [7.7: Module 7 Live Session](#)