# Module 5: Sums of Squares and Multicollinearity

| Live Sessions | Office Hours | |
|---|---|---|
| July 22: 9:00 - 9:50 am EDT (Blue), 10:00 - 10:50 am EDT (Orange) | Tue & Thu: 9:00 - 9:50 am EDT (Blue), 10:00 - 10:50 am EDT (Orange). |  Dr. Woo yjw4b@virginia.edu |

## MODULE WELCOME

In module 4, you explored the multiple linear regression (MLR) model as an extension of the simple linear regression (SLR) model. The multiple regression model allows us to investigate the effect of multiple predictors on the response variable simultaneously. In module 4, you also learned how to interpret the results of a $t$ test and the ANOVA $F$ test in an MLR model.

In module 5, you will learn how to use the partial $F$ test to compare competing multiple regression models. In fact, you will realize that the $t$ test and the ANOVA $F$ test are just special cases of the partial $F$ test (some textbooks call the partial $F$ test a general linear $F$ test).

In the second part of module 5, you will learn about the issue of multicollinearity (which does not occur in simple linear regression). Multicollinearity occurs when some of the predictors in an MLR model are linearly dependent. You will learn about the effect multicollinearity has in your MLR model, common features in an MLR that indicate the presence of multicollinearity, and some remedial measures.

## ESSENTIAL QUESTIONS

- *What is a partial $F$ test used for?*
- *What effect does multicollinearity have on an MLR model?*
- *How do we diagnose the presence of multicollinearity in our MLR model?*

## LEARNING OBJECTIVES

1. Describe what questions a partial $F$ test can answer.
2. Given a data set and questions of interests, identify what inferential procedure to use in a multiple linear regression setting.
3. Explain the issues related to multicollinearity, and how to detect and handle multicollinearity in your data set.

## ASSIGNED RESOURCES

- *Introduction to Linear Regression Analysis*, Sections 3.3.2 (after Example 3.4), 3.9 to 3.11, and 9.1 to 9.4.2, 9.4.4.
- Module 5 R tutorial, data set: mileage.txt, and R-code: tutorial_module 5.R
- Module 5 guided question set and data set: seatpos from the faraway package in R

## OPTIONAL

Visit the Module 5 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 5 General Discussion

## MODULE OVERVIEW