**Chapter 7: Some Regression Pitfalls**

Misconception: magnitude of the fitted coefficient (Bhati) determines the importance of Xi.

Standardized regression coefficients:



sxi and sy are the standard deviations of Xi and Y, respectively.

Standardized regression coefficients are dimensionless and therefore can be used to more reasonably compare regression coefficients within and across models.

Keep in mind variables may interact to affect Y, so if X1 and X2 interact then both variables may be needed in the model to predict Y.

See Ex4\_96\_p254\_Standardized\_Coefficients.R

Multicollinearity (Correlation among the X’s)

(correlations may be more complex than just correlations between pairs)



Define:



R2j close to 1 increases var(Bhatj) – can result in wrong signs for Bhatj .

This formulation of var(Bhatj) shows if you spread out values of Xj then you reduce

var(Bhatj)

Possible actions to take when faced with high multicollinearity:

Use stepwise regression methods.

Exclude variables with VIF > 10.

Use Ridge Regression (we will cover this topic).