Supplementary Exercise 11.16 of IPS7e

Data: see Supplementary Exercise 11.15 for a description of the data.

(a)
Model: The first model requested has predictors IQ and C3 for the response GPA. Below the condensed output (Basic tables; Fits and diagnostics deselected) from Minitab's Fit Regression Model menu; note that the variable C3 has been renamed to SC3.

EX11_015.MTW

Regression Analysis: GPA versus IQ, SC3

Regression Equation

GPA = -2.83 + 0.0822 IQ + 0.1629 SC3

Coefficients

Term	Coef	SE Coef	T-Value	P-Value
Constant	-2.83	1.51	-1.88	0.064
IQ	0.0822	0.0151	5.45	0.000
SC3	0.1629	0.0575	2.83	0.006

Model Summary

S	R-sq	R-sq(adj)
1.56411	45.94%	44.50%

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	155.9	77.971	31.87	0.000
Error	75	183.5	2.446		
Total	77	339.4			

The fitted regression equation is

$$\widehat{\text{GPA}} = -2.83 + 0.0822 \cdot \text{IQ} + 0.163 \cdot \text{C3}.$$

The t-statistic for testing the regression coefficient for C3 equal to zero is: $t = 2.83$ ($P = 0.006$). We conclude that C3 does contribute significantly to the model. The increase in R^2 caused by C3 is: $45.9\% - 40.2\% = 5.7\%$; the value 40.2% was from Supplementary Exercise 11.15.

(b)
Models: This part involves two models: one with predictors SC and IQ, and then the expansion to predictors SC, IQ and C3. Corresponding, similarly condensed, outputs from Minitab are shown below.

EX11_015.MTW

Regression Analysis: GPA versus IQ, SC

Regression Equation

GPA = -3.88 + 0.0772 IQ + 0.0513 SC

Coefficients

Term	Coef	SE Coef	T-Value	P-Value
Constant	-3.88	1.47	-2.64	0.010
IQ	0.0772	0.0154	5.02	0.000
SC	0.0513	0.0163	3.14	0.002

Model Summary

S	R-sq	R-sq(adj)
1.54715	47.11%	45.70%

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	159.9	79.951	33.40	0.000
Error	75	179.5	2.394		
Total	77	339.4			

EX11_015.MTW

Regression Analysis: GPA versus IQ, SC3, SC

Regression Equation

GPA = -3.49 + 0.0761 IQ + 0.0670 SC3 + 0.0369 SC

Coefficients

Term	Coef	SE Coef	T-Value	P-Value
Constant	-3.49	1.56	-2.24	0.028
IQ	0.0761	0.0155	4.91	0.000
SC3	0.0670	0.0856	0.78	0.436
SC	0.0369	0.0246	1.50	0.137

Model Summary

S	R-sq	R-sq(adj)
1.55115	47.54%	45.42%

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	3	161.4	53.793	22.36	0.000
Error	74	178.0	2.406		
Total	77	339.4			

The first listing shows that IQ and SC are both strongly significant predictors of GPA in a model containing only these two variables. The fitted regression model involving additionally the C3 variable is:

$$\widehat{\text{GPA}} = -3.49 + 0.0761 \cdot \text{IQ} + 0.0670 \cdot \text{C3} + 0.0369 \cdot \text{SC}.$$

The regression coefficient for IQ is almost unchanged upon adding C3, but the coefficient for SQ has dropped to about 70% of its previous value. The coefficient for C3 is clearly non-significant ($t = 0.78$, $P = 0.44$), with a higher P-value than for SC, which has become non-significant as well. The R^2 has increased to 47.5%, a small increase only from the 47.1% of the model with IQ and SC. We conclude that C3 does not have any significant or practically useful effect on prediction of GPA in a model that already includes IQ and SC.

(c)
The coefficient for C3 was 0.163 in the model from (a) without SC, and 0.067 in the model from (b) including SC. The explanation for this substantial change is a strong collinearity between C3 and SC; in fact, their correlation is 0.80 (see Supplemental Exercise 11.15). As SC and C3 therefore to a large extent explain the same thing, C3 is not a useful predictor in a model where SC is already present, but it can be useful when SC is absent.