Today's Agenda

Extending our Models:
Multiple Linear Regressions (ch 6)

Justin Leinaweaver (Spring 2022)

Multiple Linear Regression: Point Estimates

Simple linear regression model

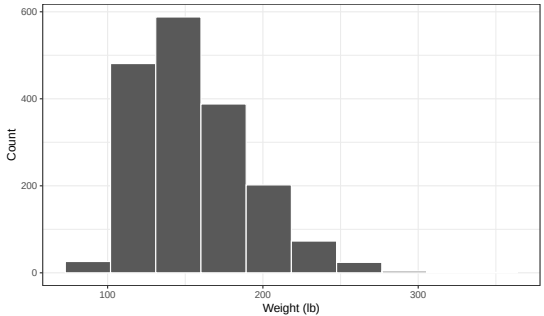
$$Y = \alpha + \beta X$$

Multiple linear regression model

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k$$

	Model 1
Predictor 1	*Coefficient*
	(Standard Error)
Predictor 2	*Coefficient*
	(Standard Error)
Constant	*Coefficient*
	(Standard Error)
Observations	*# of Observations*
Adjusted R^2	*Adj R^2 value*
Residual Std Error	*Model standard error*
F Statistic	*F value and significance*

What is the most "useful" model of weight in the Ross (1990) dataset?



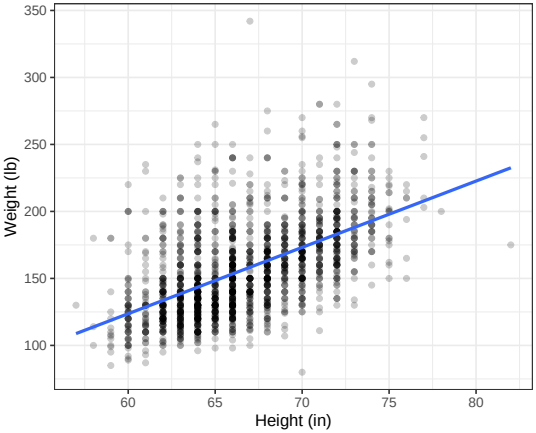
Predictors to Explore

- Height
- Exercise
- Age

Model 1: Regress weight on height

	Weight (lb)
Height (in)	4.95*
	(0.18)
Constant	−173.26*
	(11.91)
Observations	1,788
Adjusted R^2	0.30
Residual Std. Error	28.96 (df = 1786)
F Statistic	767.70* (df = 1; 1786)
Note:	*p < 0.05

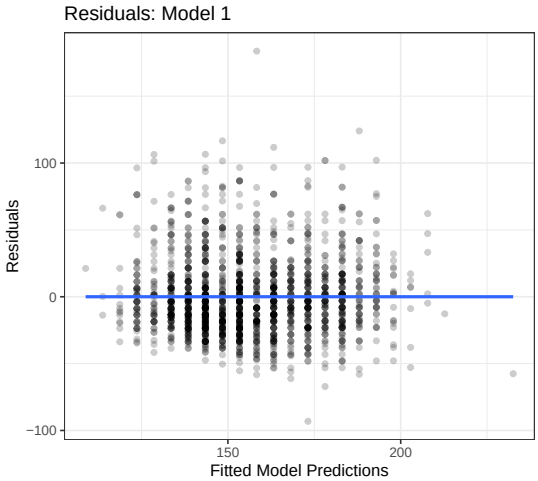
Visualize: Model 1



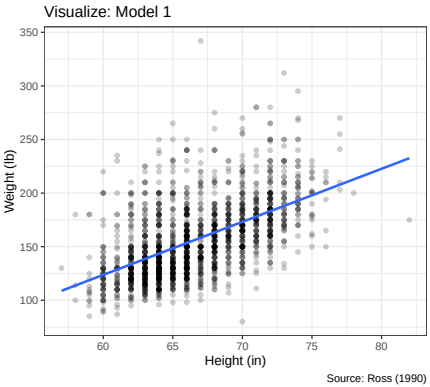
Source: Ross (1990)

Model 1: Regress weight on height

	Weight (lb)
Height (in)	4.95*
	(0.18)
Constant	−173.26*
	(11.91)
Observations	1,788
Adjusted R^2	0.30
Residual Std. Error	28.96 (df = 1786)
F Statistic	767.70* (df = 1; 1786)
Note:	*p < 0.05

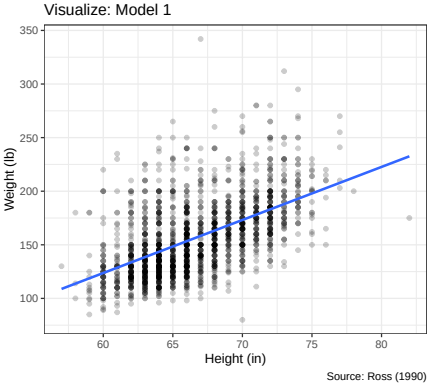


	Weight (lb)
Height (in)	4.95*
	(0.18)
Constant	−173.26*
	(11.91)
Observations	1,788
Adjusted R^2	0.30
Residual Std. Error	28.96 (df = 1786)
F Statistic	767.70* (df = 1; 1786)
Note:	*p < 0.05



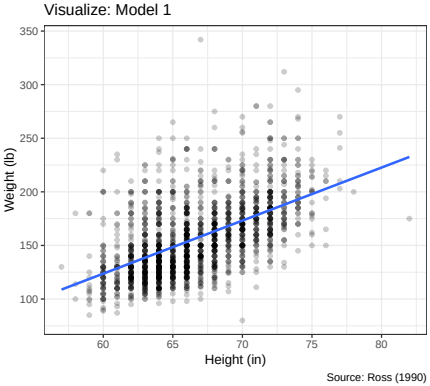
Calculate a PE (w/ 95% PI) for a person of average height (5' 7")

	Weight (lb)
Height (in)	4.95*
	(0.18)
Constant	−173.26*
	(11.91)
Observations	1,788
Adjusted R^2	0.30
Residual Std. Error	28.96 (df = 1786)
F Statistic	767.70* (df = 1; 1786)
Note:	*p < 0.05



$$\text{Weight} = \text{-173.26} + 4.95 \text{ x } (67\text{in}) = 158.39\text{lb}$$

	Weight (lb)
Height (in)	4.95*
	(0.18)
Constant	−173.26*
	(11.91)
Observations	1,788
Adjusted R^2	0.30
Residual Std. Error	28.96 (df = 1786)
F Statistic	767.70* (df = 1; 1786)
Note:	*p < 0.05



$$158.39 \pm 2 \text{ x } 28.96 \approx 100.47 \text{ to } 216.31 \text{ lb}$$

Model 2

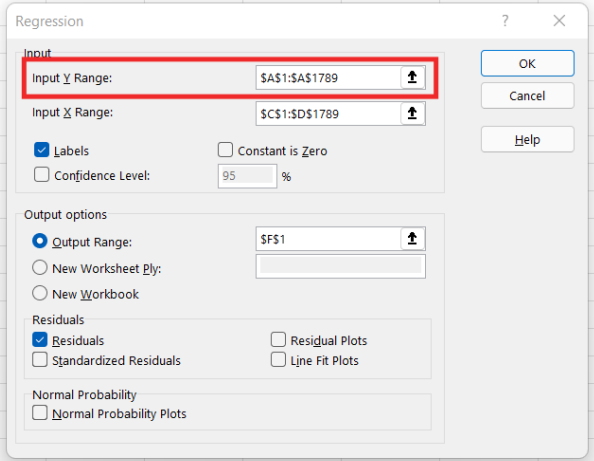
Fit a regression model of weight on height and exercise

M2: Regress weight on height and exercise

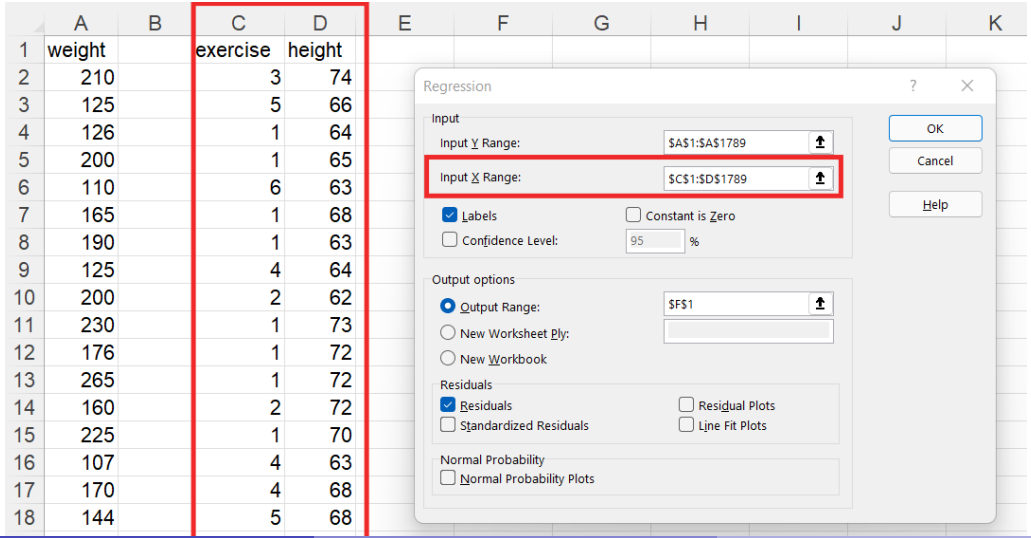
	A	B	C	D	E
1	weight		exercise	height	
2	210		3	74	
3	125		5	66	
4	126		1	64	
5	200		1	65	
6	110		6	63	
7	165		1	68	
8	190		1	63	
9	125		4	64	
10	200		2	62	
11	230		1	73	
12	176		1	72	
13	265		1	72	
14	160		2	72	
15	225		1	70	
16	107		4	63	
17	170		4	68	

M2: Regress weight on height and exercise

	A	B	C	D
1	weight		exercise	height
2	210		3	74
3	125		5	66
4	126		1	64
5	200		1	65
6	110		6	63
7	165		1	68
8	190		1	63
9	125		4	64
10	200		2	62
11	230		1	73
12	176		1	72
13	265		1	72
14	160		2	72
15	225		1	70
16	107		4	63
17	170		4	68
18	144		5	68



M2: Regress weight on height and exercise



	A	B	C	D
1	weight		exercise	height
2	210		3	74
3	125		5	66
4	126		1	64
5	200		1	65
6	110		6	63
7	165		1	68
8	190		1	63
9	125		4	64
10	200		2	62
11	230		1	73
12	176		1	72
13	265		1	72
14	160		2	72
15	225		1	70
16	107		4	63
17	170		4	68
18	144		5	68

M2: Regress weight on height and exercise

	A	B	C	D	E
1	weight		exercise	height	
2	210		3	74	
3	125		5	66	
4	126		1	64	
5	200		1	65	
6	110		6	63	
7	165		1	68	
8	190		1	63	
9	125		4	64	
10	200		2	62	
11	230		1	73	
12	176		1	72	
13	265		1	72	
14	160		2	72	
15	225		1	70	
16	107		4	63	
17	170		4	68	
18	144		5	68	
19	193		6	71	

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.564508
R Square	0.318669
Adjusted R Square	0.317906
Standard Error	28.59238
Observations	1788

ANOVA

	df	*SS*	*MS*	*F*	*ignificance*
Regression	2	682529	341264.5	417.4366	1.9E-149
Residual	1785	1459281	817.5241		
Total	1787	2141810			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*
Intercept	-184.538	11.87603	-15.5387	3.64E-51	-207.83
exercise	-2.05311	0.298578	-6.87629	8.47E-12	-2.63871
height	5.212537	0.180463	28.8842	7.6E-151	4.858596

M2: Regress weight on height and exercise

1. Put the results into a regression table
2. Evaluate the regression with all five steps
    - Logical
    - Statistical Significance
    - Explanatory Power: R2 and F Statistic
    - Check the Residuals
    - Check for Multicollinearity

	Weight (lb)
Height (in)	5.21*
	(0.18)
Exercise	−2.05*
	(0.30)
Constant	−184.54*
	(11.88)
Observations	1,788
Adjusted R^2	0.32
Residual Std. Error	28.59 (df = 1785)
F Statistic	417.44* (df = 2; 1785)
Note:	*p < 0.05

1. Does the model make sense?

	Weight (lb)
Height (in)	5.21*
	(0.18)
Exercise	−2.05*
	(0.30)
Constant	−184.54*
	(11.88)
Observations	1,788
Adjusted R^2	0.32
Residual Std. Error	28.59 (df = 1785)
F Statistic	417.44* (df = 2; 1785)
Note:	*p < 0.05

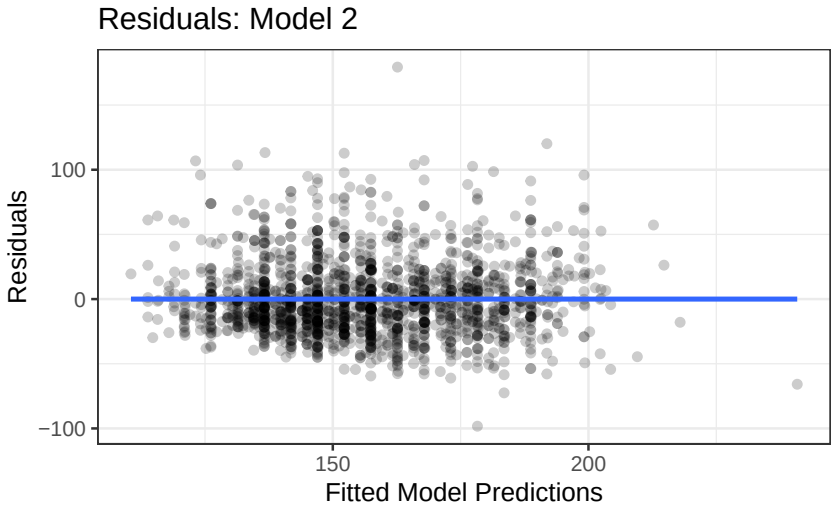
2. Check for statistical significance

	Weight (lb)
Height (in)	5.21*
	(0.18)
Exercise	−2.05*
	(0.30)
Constant	−184.54*
	(11.88)
Observations	1,788
Adjusted R^2	0.32
Residual Std. Error	28.59 (df = 1785)
F Statistic	417.44* (df = 2; 1785)
Note:	*p < 0.05

3. Determine the explanatory power of the model

	Weight (lb)
Height (in)	5.21*
	(0.18)
Exercise	−2.05*
	(0.30)
Constant	−184.54*
	(11.88)
Observations	1,788
Adjusted R^2	0.32
Residual Std. Error	28.59 (df = 1785)
F Statistic	417.44* (df = 2; 1785)
Note:	*p < 0.05

4. Check for problems in the residuals

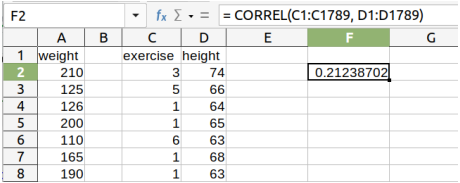


5. Check for multicollinearity (predictors)

F2 | f_x $\sum$ = | = CORREL(C1:C1789, D1:D1789)

	A	B	C	D	E	F	G
1	weight		exercise	height			
2	210		3	74		0.21238702	
3	125		5	66			
4	126		1	64			
5	200		1	65			
6	110		6	63			
7	165		1	68			
8	190		1	63			

5. Check for multicollinearity (predictors)



	height	exercise
height	1	0.215
exercise	0.215	1

M2: Regress weight on height and exercise

Calculate a point estimate (w/ 95% PI) for a person of average height (5'7") and average exercise (3).

M2: Regress weight on height and exercise

Calculate a point estimate (w/ 95% PI) for a person of average height (5'7") and average exercise (3).

- Weight $\approx$ -184.54 + 5.21 x (Height) + -2.05 x (Exercise)

M2: Regress weight on height and exercise

Calculate a point estimate (w/ 95% PI) for a person of average height (5'7") and average exercise (3).

- Weight $\approx$ -184.54 + 5.21 x (Height) + -2.05 x (Exercise)
- Weight $\approx$ -184.54 + 5.21 x (67) + -2.05 x (3)

M2: Regress weight on height and exercise

Calculate a point estimate (w/ 95% PI) for a person of average height (5'7") and average exercise (3).

- Weight $\approx$ -184.54 + 5.21 x (Height) + -2.05 x (Exercise)
- Weight $\approx$ -184.54 + 5.21 x (67) + -2.05 x (3)
- Weight $\approx$ 158.38 lb

Estimate the Prediction Interval (95%)

Residual Standard Error = 28.59

Weight $\approx$ 158.4 lb

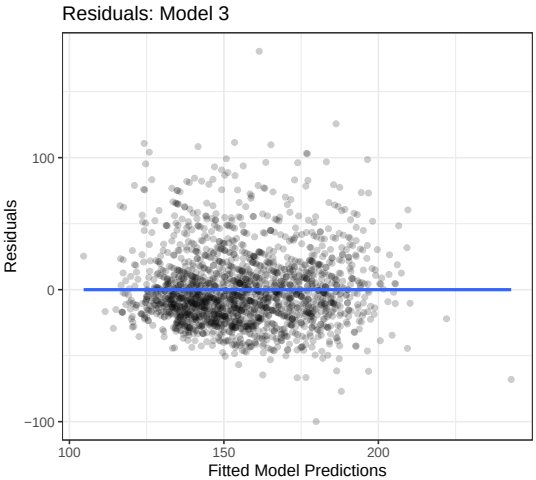
- 95% PI (Low): 158.38 - 2 x 28.59 = 101.2 lb
- 95% PI (High): 158.38 + 2 x 28.59 = 215.56 lb

	Weight (lb)	
	(1)	(2)
Height (in)	4.95*	5.21*
	(0.18)	(0.18)
Exercise		−2.05*
		(0.30)
Constant	−173.26*	−184.54*
	(11.91)	(11.88)
Observations	1,788	1,788
Adjusted R^2	0.30	0.32
Residual Std. Error	28.96 (df = 1786)	28.59 (df = 1785)
F Statistic	767.70* (df = 1; 1786)	417.44* (df = 2; 1785)
Note:		*p < 0.05

M3: Regress weight on height and age

1. Put the results into a regression table
2. Evaluate the regression with all five steps
3. Calculate a point estimate (w/ 95% PI) for a person of average height (5'7") and average age (43).

	Weight (lb)
Height (in)	5.15*
	(0.18)
Age	0.33*
	(0.04)
Constant	−200.94*
	(12.14)
Observations	1,788
Adjusted R^2	0.33
Residual Std. Error	28.41 (df = 1785)
F Statistic	434.54* (df = 2; 1785)
Note:	*p < 0.05



M3: Regress weight on height and age

Calculate a point estimate (w/ 95% PI) for a person of average height (5'7") and average age (43).

- Weight $\approx$ -200.94 + 5.15 x (Height) + 0.33 x (Age)
- Weight $\approx$ -200.94 + 5.15 x (67) + 0.33 x (43)
- Weight $\approx$ 158.3 lb

M3: Regress weight on height and age

Residual Standard Error = 28.41

Weight $\approx$ 158.3 lb

- 95% PI (Low): 158.3 - 2 x 28.41 = 101.48 lb
- 95% PI (High): 158.3 + 2 x 28.41 = 215.12 lb

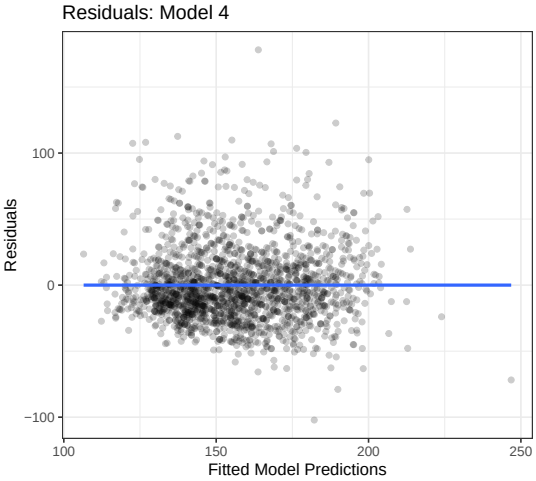
	Weight (lb)		
	(1)	(2)	(3)
Height (in)	4.95*	5.21*	5.15*
	(0.18)	(0.18)	(0.18)
Exercise		−2.05*	
		(0.30)	
Age			0.33*
			(0.04)
Constant	−173.26*	−184.54*	−200.94*
	(11.91)	(11.88)	(12.14)
Observations	1,788	1,788	1,788
Adjusted R^2	0.30	0.32	0.33
Residual Std. Error	28.96 (df = 1786)	28.59 (df = 1785)	28.41 (df = 1785)
F Statistic	767.70* (df = 1; 1786)	417.44* (df = 2; 1785)	434.54* (df = 2; 1785)

Note: *p < 0.05

M4: Regress weight on height, excercise & age

1. Put the results into a regression table
2. Evaluate the regression with all five steps
3. Calculate a point estimate (w/ 95% PI) for a person of average height (5'7"), exercise (3) and average age (43).

	Weight (lb)
Height (in)	5.29*
	(0.18)
Exercise	−1.40*
	(0.31)
Age	0.27*
	(0.04)
Constant	−203.76*
	(12.09)
Observations	1,788
Adjusted R^2	0.33
Residual Std. Error	28.25 (df = 1784)
F Statistic	299.65* (df = 3; 1784)
Note:	*p < 0.05



Model 4

Calculate a point estimate (w/ 95% PI) for a person of average height (5'7"), exercise (3) and average age (43).

- Weight $\approx$ -203.76 + 5.29 x (Height) + -1.4 x (Exercise) + 0.27 x (Age)
- Weight $\approx$ -203.76 + 5.29 x (67) + -1.4 x (3) + 0.27 x (43)
- Weight $\approx$ 158.08 lb

Model 4

$$\text{Residual Standard Error} = 28.25$$

$\text{Weight} \approx 158.08$ lb

- 95% PI (Low): 158.08 - 2 x 28.25 = 101.58 lb
- 95% PI (High): 158.08 + 2 x 28.25 = 214.58 lb

	Weight (lb)			
	(1)	(2)	(3)	(4)
height	4.95*	5.21*	5.15*	5.29*
	(0.18)	(0.18)	(0.18)	(0.18)
exercise		−2.05*		−1.40*
		(0.30)		(0.31)
age			0.33*	0.27*
			(0.04)	(0.04)
Constant	−173.26*	−184.54*	−200.94*	−203.76*
	(11.91)	(11.88)	(12.14)	(12.09)
Observations	1,788	1,788	1,788	1,788
Adjusted R^2	0.30	0.32	0.33	0.33
Residual Std. Error	28.96 (df = 1786)	28.59 (df = 1785)	28.41 (df = 1785)	28.25 (df = 1784)
F Statistic	767.70* (df = 1; 1786)	417.44* (df = 2; 1785)	434.54* (df = 2; 1785)	299.65* (df = 3; 1784)

Note: *p < 0.05

For Thursday

Evaluate our four models using Wheelan ch 12

- Model 1: Height
- Model 2: Height and Exercise
- Model 3: Height and Age
- Model 4: Height, Exercise and Age