Today's Agenda

1. Explore the intuitions of OLS regression

2. Practice fitting and interpreting simple OLS regressions

Justin Leinaweaver (Spring 2022)

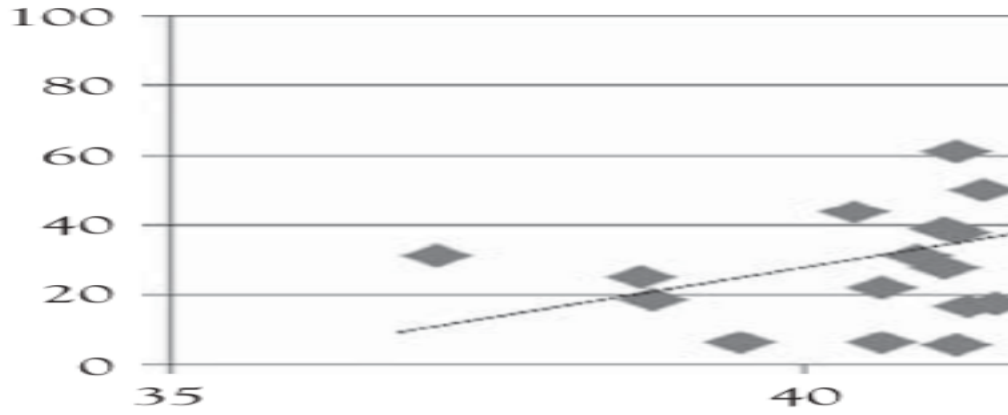
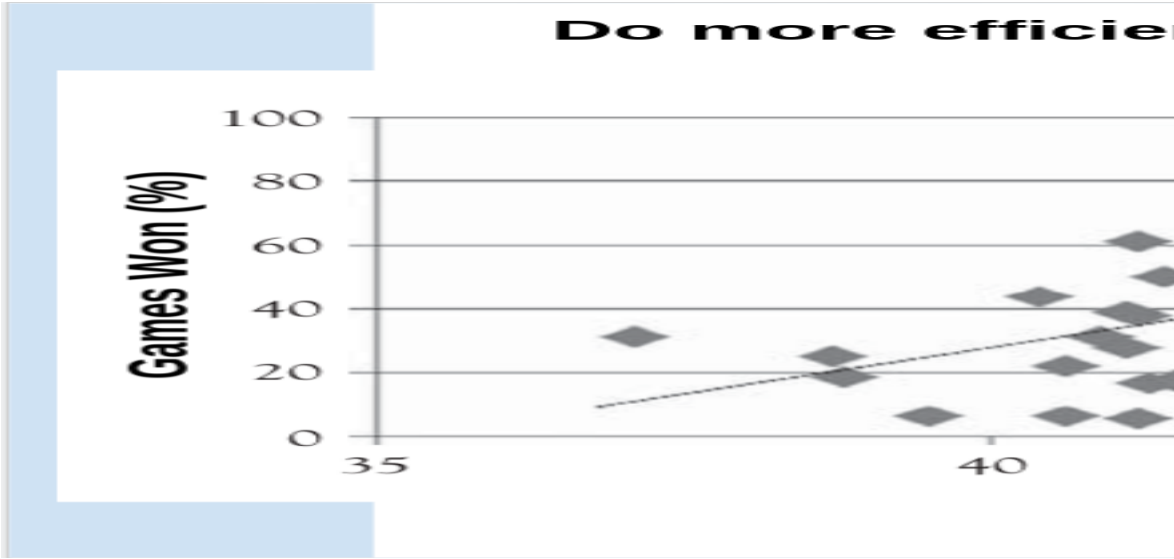
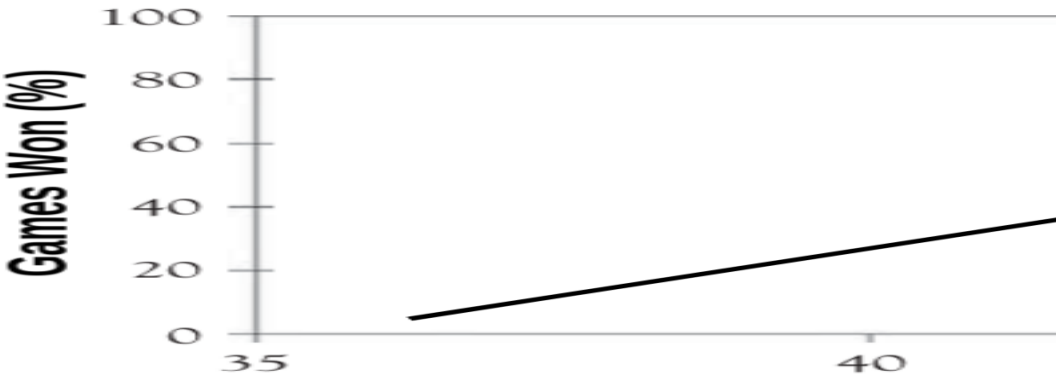
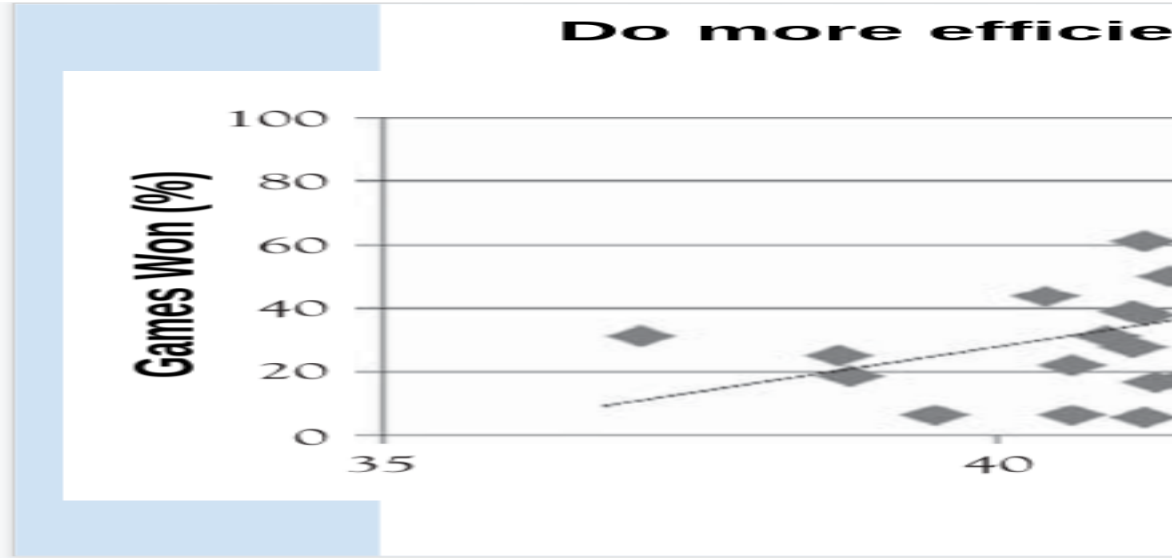


Figure 2.4. A scatterplot
versus field goal percentag


Do more efficie
Games Won (%)
100
80
60
40
20
0
35
40

Do more efficie




Do more efficie
Games Won (%)
100
80
60
40
20
0
35
40

The Formula for a Line

$$y = mx + b$$

is equivalent to

$$y = \alpha + \beta\, x$$

The Formula for a Line

$$y = \alpha + \beta x$$

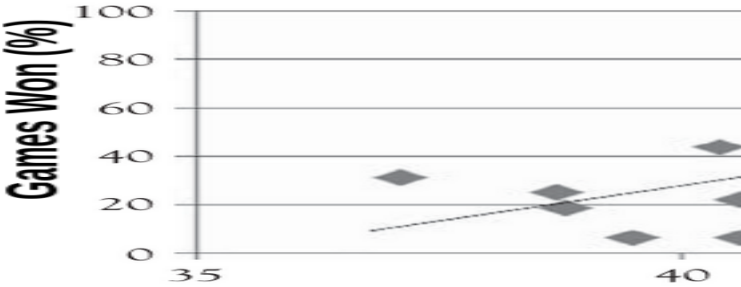
- y is the outcome
- α is the constant
- β is the coefficient estimate
- x is the predictor

The Formula for a Line

$$y = \alpha + \beta x$$

- y is the outcome
- α is the constant (**the intercept**)
- β is the coefficient estimate (**the slope**)
- x is the predictor

Games Won = -19



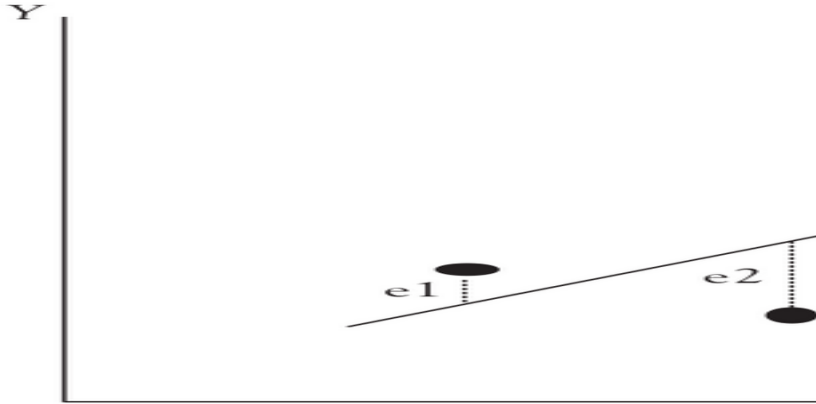
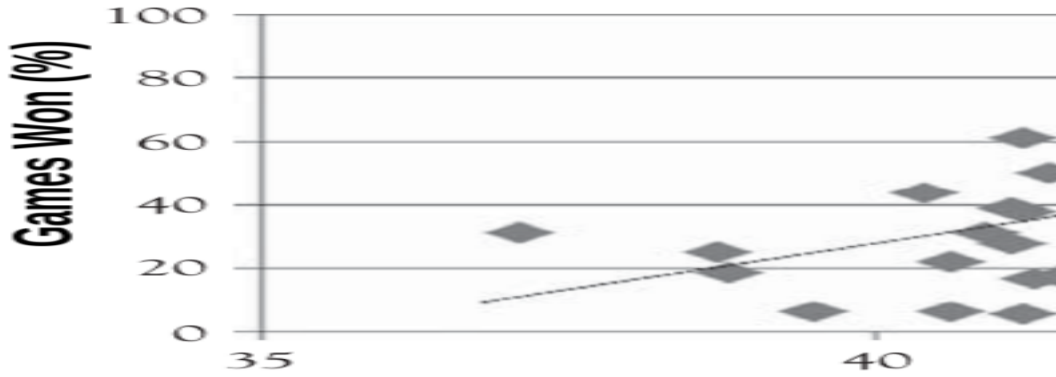


Figure 3.2. The ordinary least
function of X. Residuals (or dev
point and the regression line are

WP



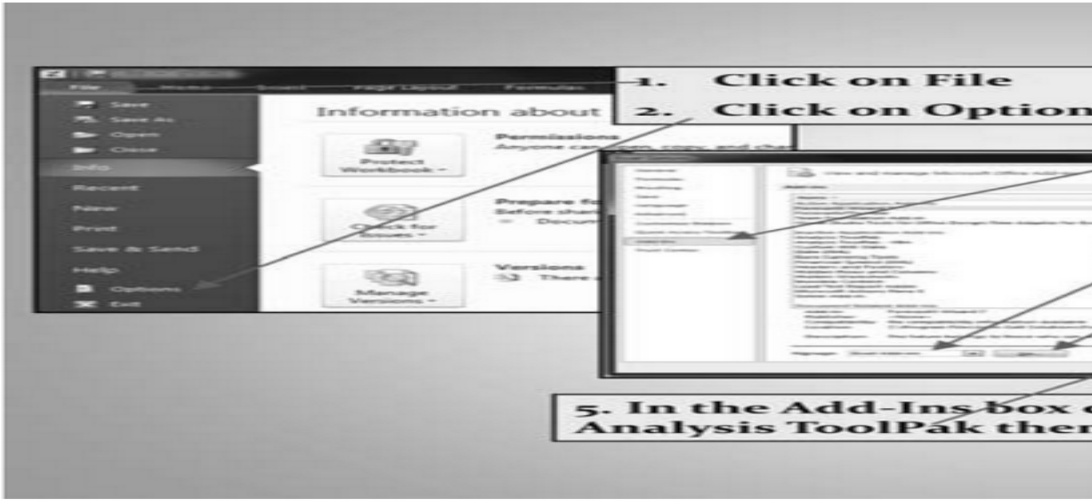


Figure 1.6. Getting "Data Analysis"

Work, Family, and Well-Being

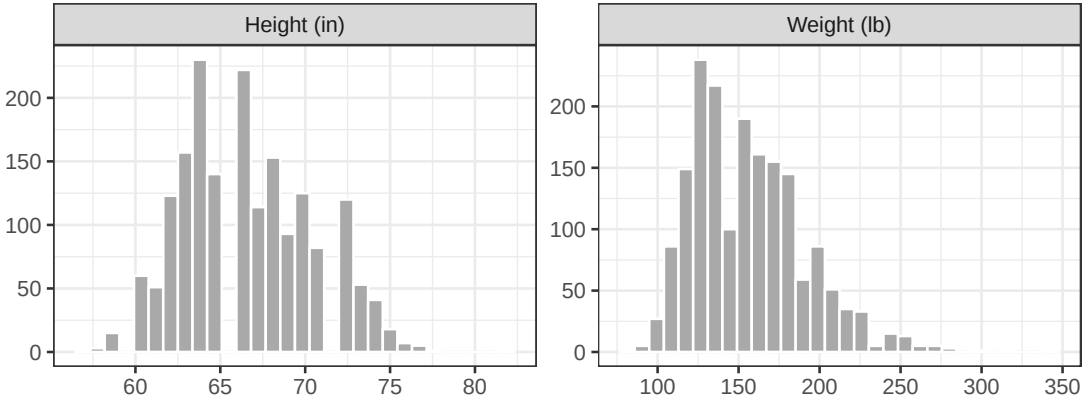
Version Date: Jun 10, 1996 | Cite this study | Share this pa

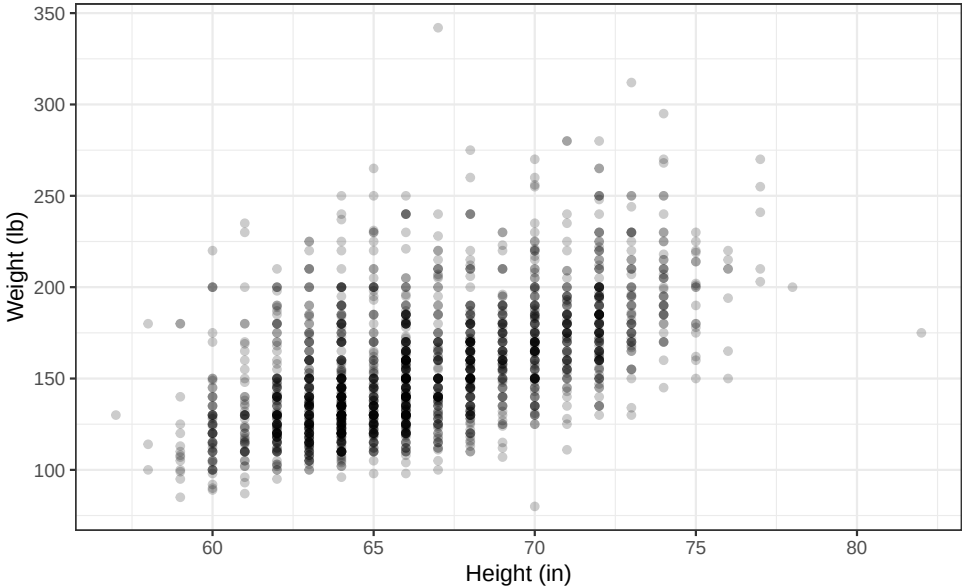
Principal Investigator(s):
Catherine E. Ross

	A	B	C	D	E	F	
1	height	weight	male	earn	earnk	ethnicity	edu
2	74	210	1	50000	50	White	
3	66	125	0	60000	60	White	
4	64	126	0	30000	30	White	
5	65	200	0	25000	25	White	
6	63	110	0	50000	50	Other	
7	68	165	0	62000	62	Black	
8	63	190	0	51000	51	White	
9	64	125	0	9000	9	White	
10	62	200	0	29000	29	White	
11	73	230	1	32000	32	White	
12	72	176	1	2000	2	Hispanic	
13	72	265	1	35000	35	White	NA
14	72	160	1	27000	27	White	
15	70	225	1	6530	6.53	White	
16	63	107	0	0	0	White	

Is height a useful mo

Univariate Analysis




Weight (lb)
350
300
250
200
150
100
60
65
70
75
80
Height (in)

Formatting a Simple

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0
R Square	0
Adjusted R Square	0
Standard Error	28
Observations	1788
ANOVA	
	df
Regression	
Residual	1
Total	1
	Coefficien
Intercept	-173
height	4

Formatting a Simple

Predictor
Constant
Observations
Adjusted R^2
Residual Std Erro
F Statistic

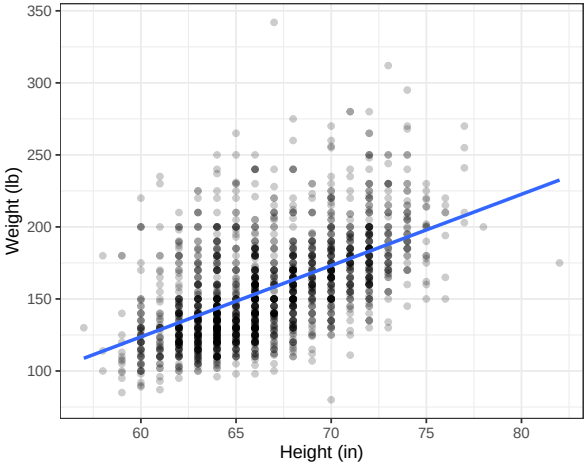
Add '*' next to any coeffici

Formatting a Simple

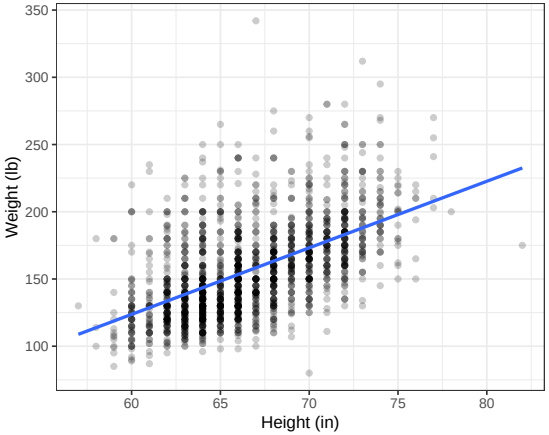
SUMMARY OUTPUT				
Regression Statistics				
Multiple R	0.55			
R Square	0.30			
Adjusted R Square	0.30			
Standard Error	28.96			
Observations	1788.00			
ANOVA				
	df	*SS*	*MS*	
Regression	1	643873.73	643873.73	
Residual	1786	1497935.78	838.71	
Total	1787	2141809.51		
	Coefficients	*Standard Error*	*t Stat*	
Intercept	-173.26	11.91	-14.54	
height	4.95	0.18	27.71	

Formatting a Simple

SUMMARY OUTPUT				
Regression Statistics				
Multiple R	0.55			
R Square	0.30			
Adjusted R Square	0.30			
Standard Error	28.96			
Observations	1788.00			
ANOVA				
	df	*SS*	*MS*	
Regression	1	643873.73	643873.73	
Residual	1786	1497935.78	838.71	
Total	1787	2141809.51		
	Coefficients	*Standard Error*	*t Stat*	
Intercept	-173.26	11.91	-14.54	
height	4.95	0.18	27.71	

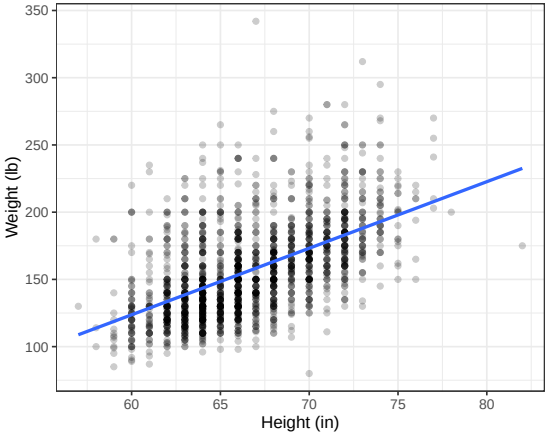


	Weight
Height	4.95*
	(0.18)
Constant	−173.26*
	(11.91)
Observations	1,788
Adjusted R^2	0.30
Residual Std. Error	28.96 (df = 1786)
F Statistic	767.70* (df = 1; 1786)
Note:	*p < 0.05



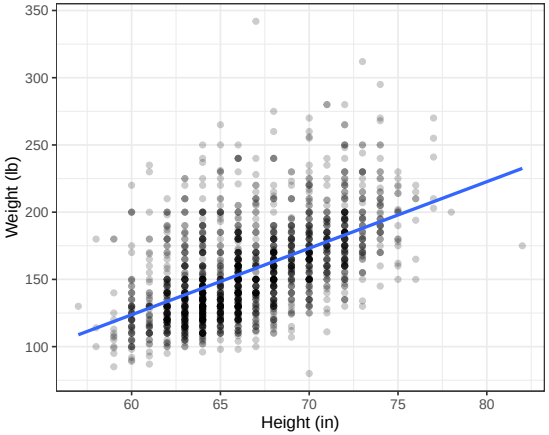
	Weight
Height	4.95*
	(0.18)
Constant	−173.26*
	(11.91)
Observations	1,788
Adjusted R^2	0.30
Residual Std. Error	28.96 (df = 1786)
F Statistic	767.70* (df = 1; 1786)
Note:	*p < 0.05

Outcome = Constant + Beta Coefficient * Predictor



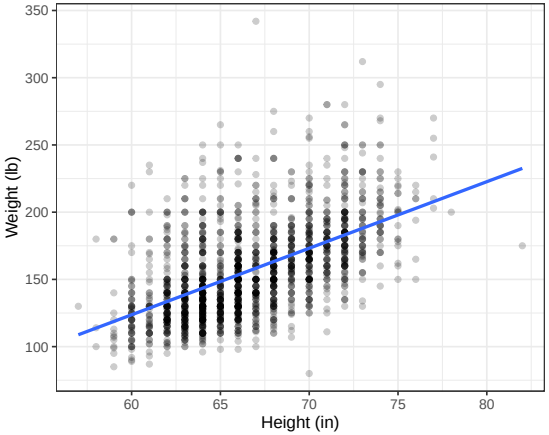
	Weight
Height	4.95*
	(0.18)
Constant	−173.26*
	(11.91)
Observations	1,788
Adjusted R^2	0.30
Residual Std. Error	28.96 (df = 1786)
F Statistic	767.70* (df = 1; 1786)
Note:	*p < 0.05

Weight = -173.26 + 4.95 x Height



	Weight
Height	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

$$\text{Weight} = \text{-173.26} + 4.95 \text{ x } 64 \approx 143.54 \text{ lb}$$



	Weight
Height	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

$$\text{Weight} = \text{-173.26} + 4.95 \text{ x } 69 \approx 168.29 \text{ lb}$$

Work, Family, and Well-Being

Version Date: Jun 10, 1996 | Cite this study | Share this pa

Principal Investigator(s):
Catherine E. Ross

	A	B	C	D	E	F	
1	height	weight	male	earn	earnk	ethnicity	edu
2	74	210	1	50000	50	White	
3	66	125	0	60000	60	White	
4	64	126	0	30000	30	White	
5	65	200	0	25000	25	White	
6	63	110	0	50000	50	Other	
7	68	165	0	62000	62	Black	
8	63	190	0	51000	51	White	
9	64	125	0	9000	9	White	
10	62	200	0	29000	29	White	
11	73	230	1	32000	32	White	
12	72	176	1	2000	2	Hispanic	
13	72	265	1	35000	35	White	NA
14	72	160	1	27000	27	White	
15	70	225	1	6530	6.53	White	
16	63	107	0	0	0	White	

Is height a useful mo

Analyze Three OLS Models

For each model: Make a regression table, scatter plot and a prediction using the average value of the predictor.

- Model 1: Regress earnings (earnk2021) on height
- Model 2: Regress earnings (earnk2021) on age
- Model 3: Regress earnings (earnk2021) on education

For Thursday

1. Finish the model building work from class today
2. Use the four steps outlined in Wilson, Keating, and Beal-Hodges (2012) chapters 4 and 5 to evaluate the fit of our models of earnings.