XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 1

Logistic Regression using OLS1C in Excel 2013

by
Milo Schield
Member: International Statistical Institute
US Rep: International Statistical Literacy Project
Director, W. M. Keck Statistical Literacy Project

Slides, output and data at: www.StatLit.org/
pdf/2015-Schield-Logistic-OLS1C-Excel2013-Slides.pdf
pdf/2015-Schield-Logistic-OLS1C-Excel2013-Demo.pdf
Excel/2015-Schield-Logistic-OLS1C-Excel2013-Data.xlsx

XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 2

Background & Goals

Modelling a binary outcome (loan vs. no-loan) requires logistic regression to avoid meaningless predictions.

Doing an exact logistic regression in Excel requires Solver and involves many steps. For details, see www.statlit.org/pdf/Excel2013-Schield-Logistic-MLE1A-Slides.pdf

This presentation uses an approximation: OLS1. By slightly adjusting the binary outcomes, one can use OLS regression to solve for a good logistic model.

Assignment: Create the logistic model (slide 9) and the logistic graph (slide 12).

XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 3

This demo uses Weight (col A) to predict Gender (col B)

Column B: 0=Female, 1 = Male (circled)

A	B	C
Weight	Male	Male1
95	0	
102	0	
108	0	
108	0	
110	0	
110	0	
112	0	
115	0	
115	0	
116	0	
116	0	

Data in rows 7 to 98

121	0
122	0
123	(1)
125	0
125	0
125	0
125	0
125	0
130	0
130	(1)
130	0
130	0

XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 4

1) Nudge Binary Male to Eliminate Zero and One

	A	B	C	D	E
1					
2	Predict chance of being male given weight.				
3	C7	=IF(B7=0, 0.001, 0.999)			
4					
5	A	B	C	D	E
6	Weight	Male	Male1	Odds	LN(Odds) yP
7	95	0	0.001		
8	102	0			

XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 5

2) Enter formula for Odds in D7; LN[Odds(p)] in E7

	A	B	C	D	E	F	G
1							
2	Predict chance of being male (1) given weight. Reg						
3	C7	=IF(B7=0, 0.001, 0.999)			E7	=LN(D7)	
4	D7	=C7/(1-C7)					
5	A	B	C	D	E	F	
6	Weight	Male	Male1	Odds	LN(Odds)	yPred	6
7	95	0	0.001	0.001	-6.91		7
8	102	0					8

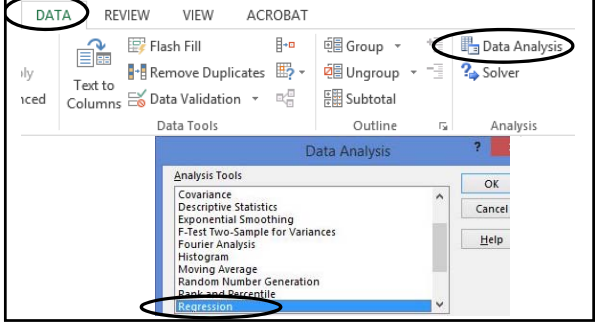
XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 6

3) Select C7:E7 Drag to bottom of data: Row 98

A	B	C	D	E	F	G
Predict chance of being male given weight. Regres						
C7	=IF(B7=0, 0.001, 0.999)			E7	=LN(D7)	
D7	=C7/(1-C7)					
A	B	C	D	E	F	
Weight	Male	Male1	Odds	LN(Odds)	yPred	6
95	0	0.001	0.001	-6.91		7
102	0					8
108	0					9

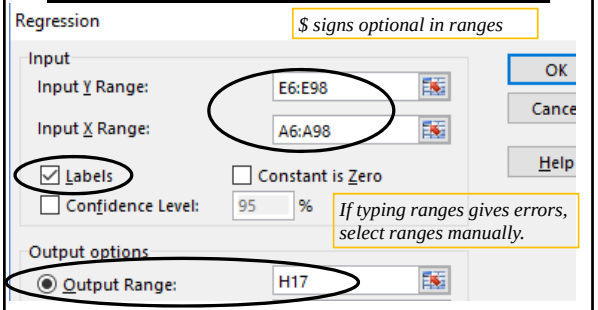
XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 7

A) From Data Bar, Select Data Analysis; Regression



XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 8

B) Select Data, Labels, Output Range. Press OK



XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 9

C) Logistic Regression: Results Using OLS1

19	*Regression Statistics*					
20	Multiple R	0.708667				
21	R Square	0.5022089				
22	Adjusted R Square	0.4966779				
23	Standard Error	4.7839195				
24	Observations	92				
25						
26	ANOVA					
27		*df*	*SS*	*MS*	*F*	*ignificance F*
28	Regression	1	2078.01	2078.01	90.79874	2.73E-15
29	Residual	90	2059.73	22.88589		
30	Total	91	4137.739			
31						
32		*Coefficients*	*ndard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
33	Intercept	-27.56682	3.106615	-8.87359	6.31E-14	-33.7387	-21.395
34	Weight	0.2012952	0.021125	9.528837	2.73E-15	0.159327	0.24326

Check to see that you get the same results.

XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 10

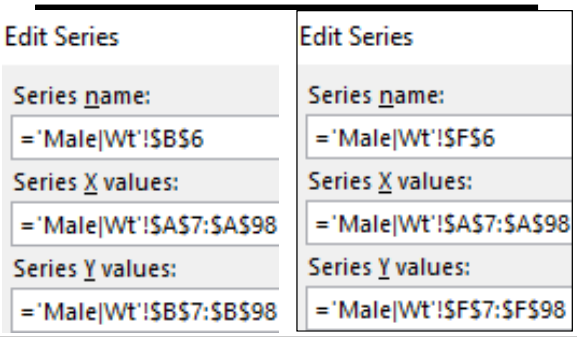
D) Generate F7; Pull F7 down to F98

C	D	E	F	G
e of being male given weight. Regress using				
=0, 0.001, 0.999)			E7	=LN(D7)
-C7)			F7	=1/(1+EXP(-I$33-I$34*A7))

C	D	E	F	
Male1	Odds	LN(Odds)	yPred	6
0.001	0.001	-6.91	0.000	7

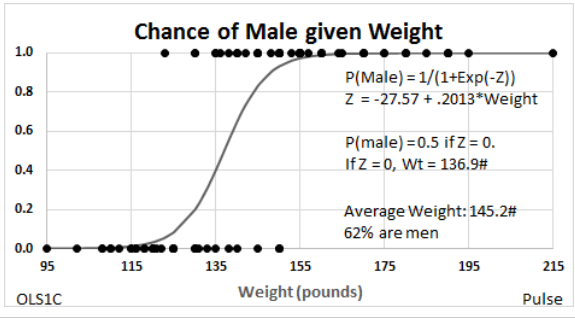
XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 11

E) Insert XY-Plot. Add Two Series: Male vs Weight; yPred vs Weight



XL4C:V0C 2015 Schield Logistic Regression using OLS1C in Excel2013 12

F) Add Titles and Textboxes. Format yPred with Solid Line



Logistic Regression using OLS1C in Excel 2013

by
Milo Schield
Member: International Statistical Institute
US Rep: International Statistical Literacy Project
Director, W. M. Keck Statistical Literacy Project

Slides, output and data at: www.StatLit.org/
pdf/2015-Schield-Logistic-OLS1C-Excel2013-Slides.pdf
pdf/2015-Schield-Logistic-OLS1C-Excel2013-Demo.pdf
Excel/2015-Schield-Logistic-OLS1C-Excel2013-Data.xlsx

Background & Goals

Modelling a binary outcome (loan vs. no-loan) requires logistic regression to avoid meaningless predictions.

Doing an exact logistic regression in Excel requires Solver and involves many steps. For details, see www.statlit.org/pdf/Excel2013-Schield-Logistic-MLE1A-Slides.pdf

This presentation uses an approximation: OLS1. By slightly adjusting the binary outcomes, one can use OLS regression to solve for a good logistic model.

Assignment: Create the logistic model (slide 9) and the logistic graph (slide 12).

This demo uses Weight (col A) to predict Gender (col B)

Column B: 0=Female, 1 = Male (circled)

A	B	C
Weight	Male	Male1
95	0	
102	0	
108	0	
108	0	
110	0	
110	0	
112	0	
115	0	
115	0	
116	0	
116	0	

Data in rows 7 to 98

121	0
122	0
123	1
125	0
125	0
125	0
125	0
125	0
130	0
130	1
130	0
130	0

1) Nudge Binary Male to Eliminate Zero and One

	A	B	C	D	E
1					
2	Predict chance of being male given weight.				
3	C7	=IF(B7=0, 0.001, 0.999)			
4					
5	A	B	C	D	E
6	Weight	Male	Male1	Odds	LN(Odds) yP
7	95	0	0.001		
8	102	0			

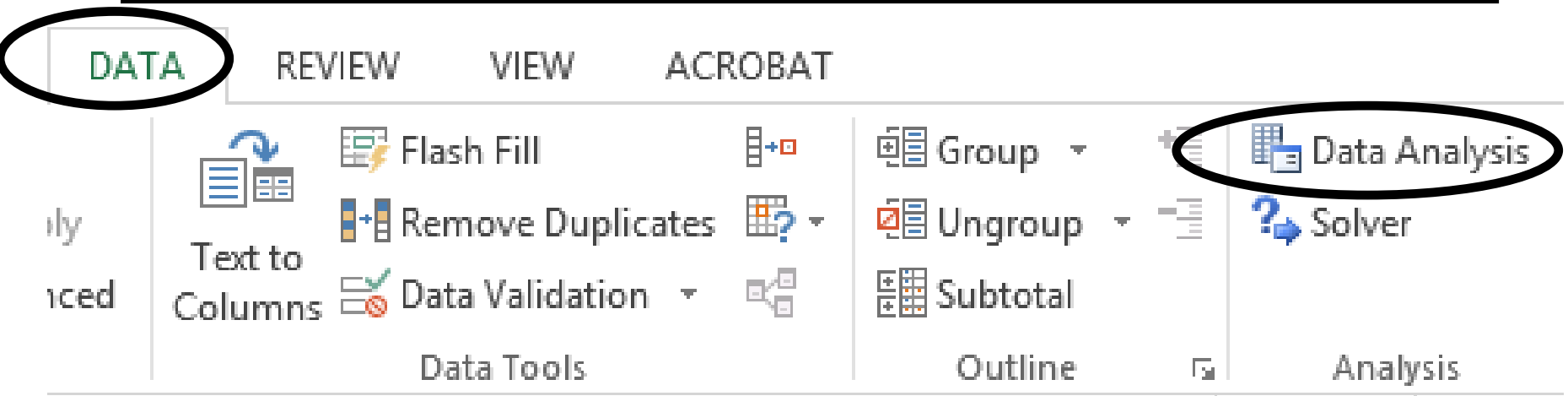
2) Enter formula for Odds in D7; LN[Odds(p)] in E7

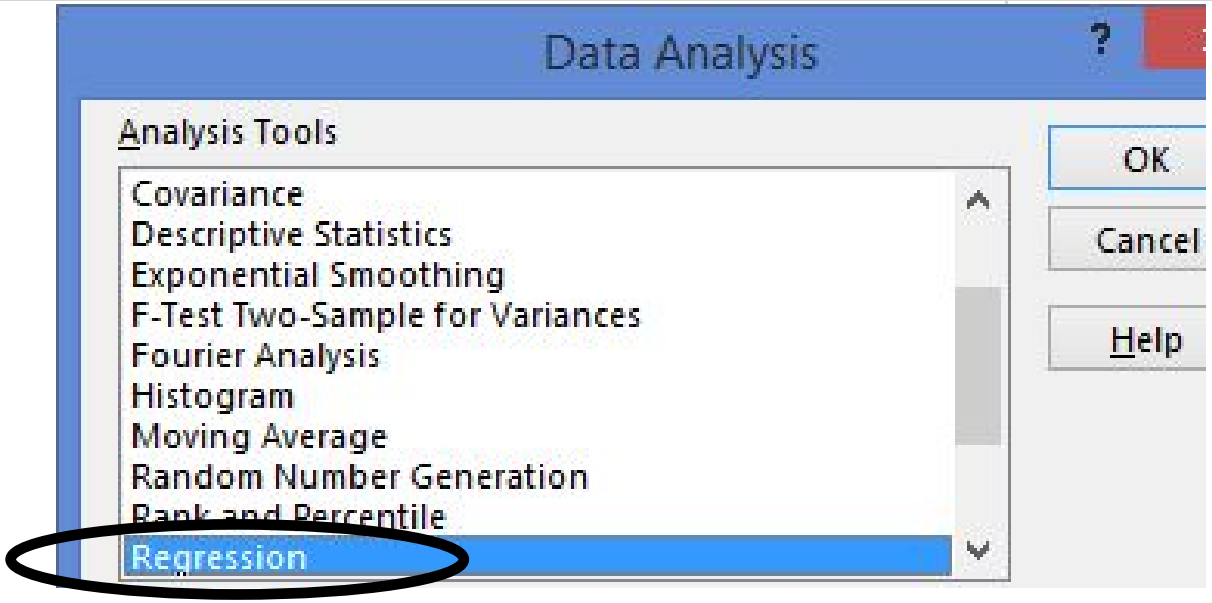
	A	B	C	D	E	F	G
1							
2	Predict chance of being male (1) given weight. Reg						
3		C7	=IF(B7=0, 0.001, 0.999)			E7	=LN(D7)
4		D7	=C7/(1-C7)				
5		A	B	C	D	E	F
6		Weight	Male	Male1	Odds	LN(Odds)	yPred
7		95	0	0.001	0.001	-6.91	
8		102	0				

3) Select C7:E7
Drag to bottom of data: Row 98

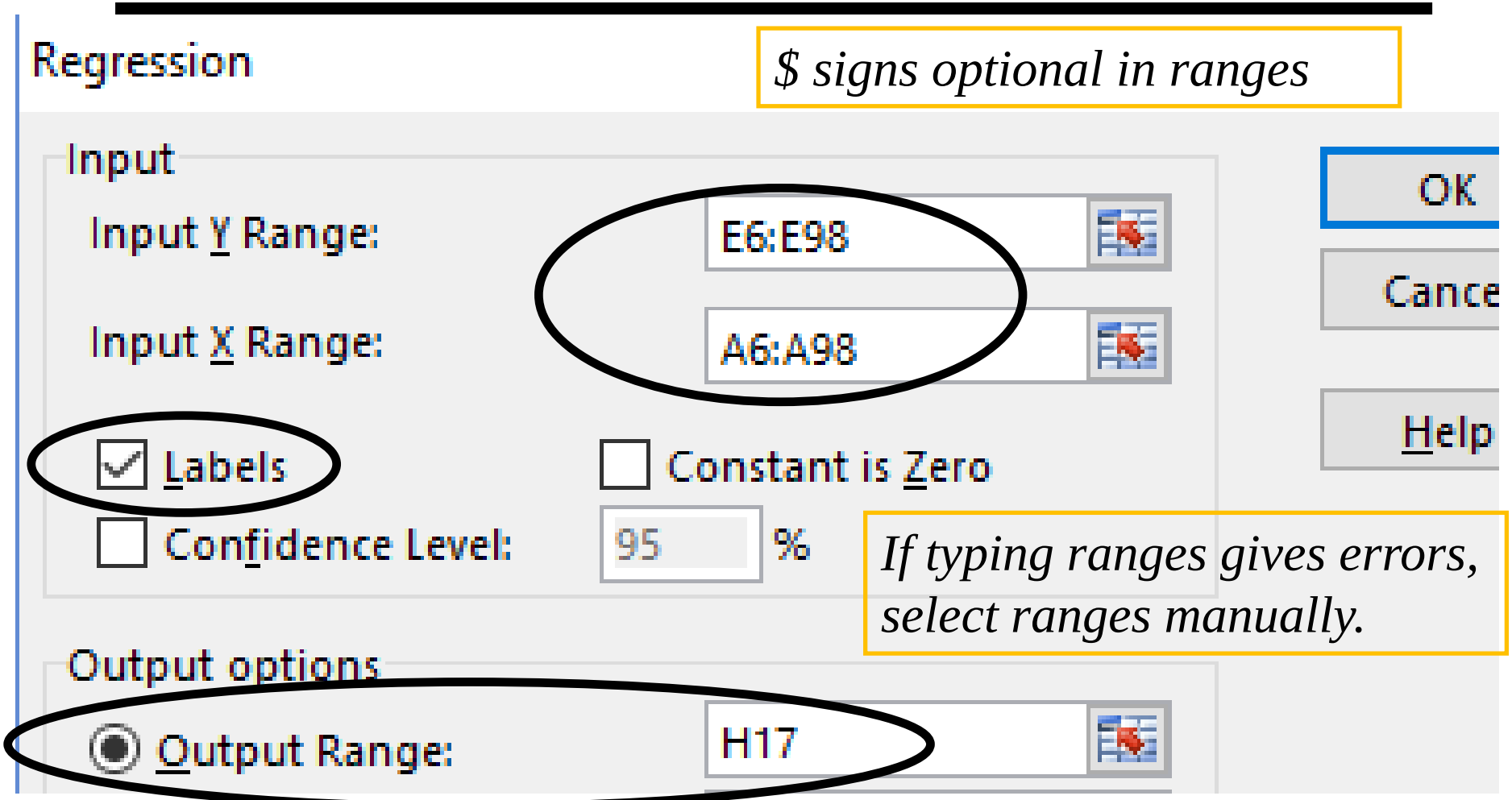
	A	B	C	D	E	F	G
	Predict chance of being male given weight. Regres						
	C7	=IF(B7=0, 0.001, 0.999)			E7	=LN(D7)	
	D7	=C7/(1-C7)					
	A	B	C	D	E	F	
	Weight	Male	Male1	Odds	LN(Odds)	yPred	6
	95	0	0.001	0.001	-6.91		7
	102	0					8
	108	0					9

A) From Data Bar, Select Data Analysis; Regression





B) Select Data, Labels, Output Range. Press OK

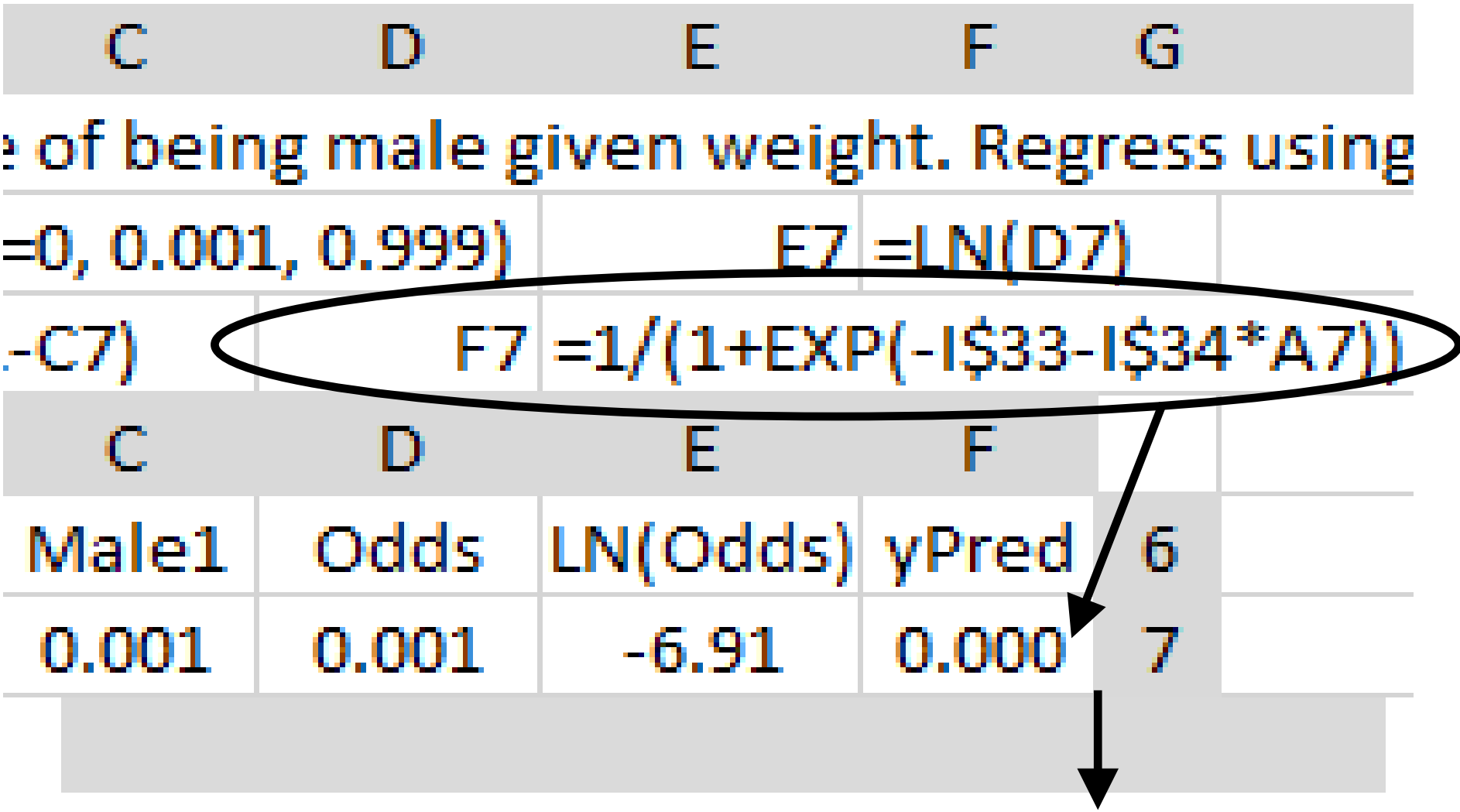


C) Logistic Regression: Results Using OLS1

19	*Regression Statistics*						
20	Multiple R	0.708667					
21	R Square	0.5022089					
22	Adjusted R Square	0.4966779					
23	Standard Error	4.7839195					
24	Observations	92					
25							
26	ANOVA						
27		*df*	*SS*	*MS*	*F*	*ignificance F*	
28	Regression	1	2078.01	2078.01	90.79874	2.73E-15	
29	Residual	90	2059.73	22.88589			
30	Total	91	4137.739				
31							
32		*Coefficients*	*ndard Err*	*t Stat*	*P-value*	*Lower 95%*	*Jpper 95%*
33	Intercept	-27.56682	3.106615	-8.87359	6.31E-14	-33.7387	-21.395
34	Weight	0.2012952	0.021125	9.528837	2.73E-15	0.159327	0.24326

Check to see that you get the same results.

D) Generate F7; Pull F7 down to F98



E) Insert XY-Plot. Add Two Series: Male vs Weight; yPred vs Weight

Edit Series

Series name:

='Male|Wt'!$B$6

Series X values:

='Male|Wt'!$A$7:$A$98

Series Y values:

='Male|Wt'!$B$7:$B$98

Edit Series

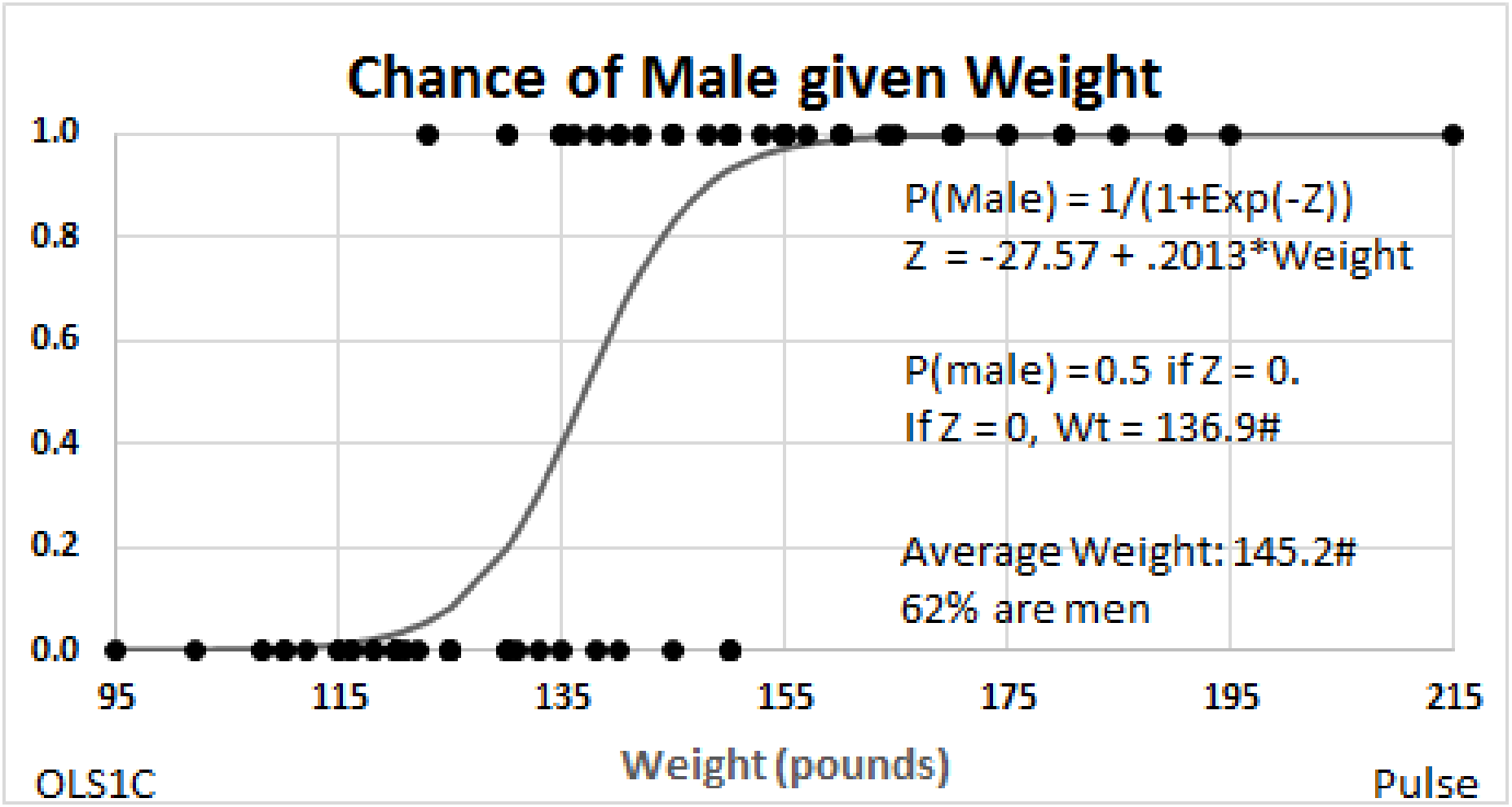
Series name:

='Male|Wt'!$F$6

Series X values:

='Male|Wt'!$A$7:$A$98

Series Y values:

='Male|Wt'!$F$7:$F$98

F) Add Titles and Textboxes. Format yPred with Solid Line