Multiple OLS Regression

Richard Lee Rogers

Last Update: February 14, 2016

Log of Violent Crime Rate on Poverty Rate

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.391[a]	.152	.135	.36130

a. Predictors: (Constant), Poverty Rate

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.127	1	1.127	8.637	.005[b]
	Residual	6.266	48	.131		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Poverty Rate

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.097	.243		20.981	.000
	Poverty Rate	.046	.016	.391	2.939	.005

a. Dependent Variable: Log of Violent Crime Rate

Previous Regression Results

- Equation of the line: $\hat{y} = 5.10 + 0.5$(poverty rate)
- Moderate positive relationship that was statistically significant (standardized B=.39, $p<.01$)
- Model explains 15% of the variance ($R^2=.15$, $p<.01$)

Log of Violent Crime Rate on Temperature

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.415[a]	.172	.155	.35712

a. Predictors: (Constant), Temperature

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.272	1	1.272	9.972	.003[b]
	Residual	6.121	48	.128		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.834	.309		15.668	.000
	Temperature	.019	.006	.415	3.158	.003

a. Dependent Variable: Log of Violent Crime Rate

Log of Violent Crime Rate on Temperature

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.415[a]	.172	.155	.35712

a. Predictors: (Constant), Temperature

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.272	1	1.272	9.972	.003[b]
	Residual	6.121	48	.128		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature

$\hat{y} = 4.83 + .02(\text{temperature})$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.834	.309		15.668	.000
	Temperature	.019	.006	.415	3.158	.003

a. Dependent Variable: Log of Violent Crime Rate

Log of Violent Crime Rate on Temperature

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.415[a]	.172	.155	.35712

a. Predictors: (Constant), Temperature

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.272	1	1.272	9.972	.003[b]
	Residual	6.121	48	.128		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature

$\hat{y} = 4.83 + .02(\text{temperature})$

moderate + positive

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.834	.309		15.668	.000
	Temperature	.019	.006	.415	3.158	.003

a. Dependent Variable: Log of Violent Crime Rate

Log of Violent Crime Rate on Temperature

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.415[a]	.172	.155	.35712

a. Predictors: (Constant), Temperature

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.272	1	1.272	9.972	.003[b]
	Residual	6.121	48	.128		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature

$\hat{y} = 4.83 + .02(\text{temperature})$

moderate + positive

relationship is significant

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.834	.309		15.668	.000
	Temperature	.019	.006	.415	3.158	.003

a. Dependent Variable: Log of Violent Crime Rate

Log of Violent Crime Rate on Temperature

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.415[a]	.172	.155	.35712

a. Predictors: (Constant), Temperature

R^2

model is significant

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.272	1	1.272	9.972	.003[b]
	Residual	6.121	48	.128		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature

$\hat{y} = 4.83 + .02(\text{temperature})$

moderate + positive

relationship is significant

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.834	.309		15.668	.000
	Temperature	.019	.006	.415	3.158	.003

a. Dependent Variable: Log of Violent Crime Rate

A Common Mistake

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.05*
Temperature		.02**	.02**
Constant	5.10***	4.83***	9.93***
R^2	.15**	.17**	.32**

Note: t p<.10 *p<.05 **p<.01 ***p<.001

A Common Mistake

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.05*
Temperature		.02**	.02**
Constant	5.10***	4.83***	9.93***
R^2	.15**	.17**	.32**

Note: t p<.10 *p<.05 **p<.01 ***p<.001

A Common Mistake

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.05*
Temperature		.02**	.02**
Constant	5.10***	4.83***	9.93***
R^2	.15**	.17**	.32**

Note: t p<.10 *p<.05 **p<.01 ***p<.001

A Common Mistake

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.05*
Temperature		.02**	.02**
Constant	5.10***	4.83***	9.93***
R^2	.15**	.17**	.32**

Note: t p<.10 *p<.05 **p<.01 ***p<.001

A Common Mistake

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)		(2)		(3)
% poverty rate	.05**	→			.05*
Temperature			.02**	→	.02**
Constant	5.10***	+	4.83***	=	9.93***
R^2	.15**	+	.17**	=	.32**

Note: t p<.10 * p<.05 **p<.01 ***p<.001

A Common Mistake

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)		(2)		(3)
% poverty rate	.05**	→			.05*
Temperature			.02**	→	.0[illegible]**
Constant	5.10***	+	4.83***	=	9.93***
R^2	.15**	+	.17**		.32**

Note: t p<.10 * p<.05 **p<.01 ***p<.001

Multiple Regression

One dependent and two or more independent variables:

$$\hat{y} = \beta_0 + \beta_1 x + \beta_2 x$$

$$y = \beta_0 + \beta_1 x + \beta_2 x + e$$

Correlation Matrix

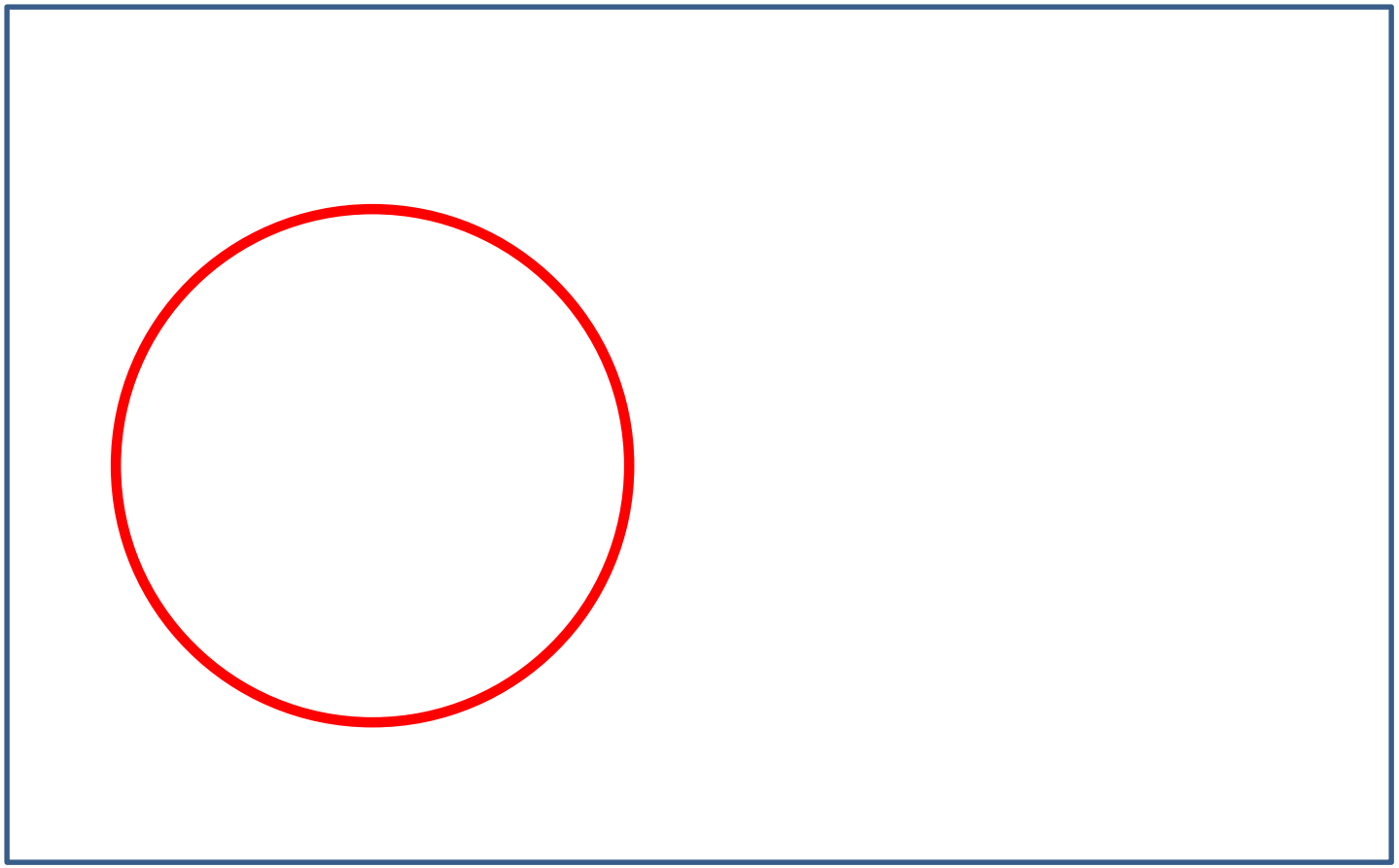
Correlations

		Log of Violent Crime Rate	Poverty Rate	Temperature
Log of Violent Crime Rate	Pearson Correlation	1	.391**	.415**
	Sig. (2-tailed)		.005	.003
	N	50	50	50
Poverty Rate	Pearson Correlation	.391**	1	.569**
	Sig. (2-tailed)	.005		.000
	N	50	50	50
Temperature	Pearson Correlation	.415**	.569**	1
	Sig. (2-tailed)	.003	.000	
	N	50	50	50

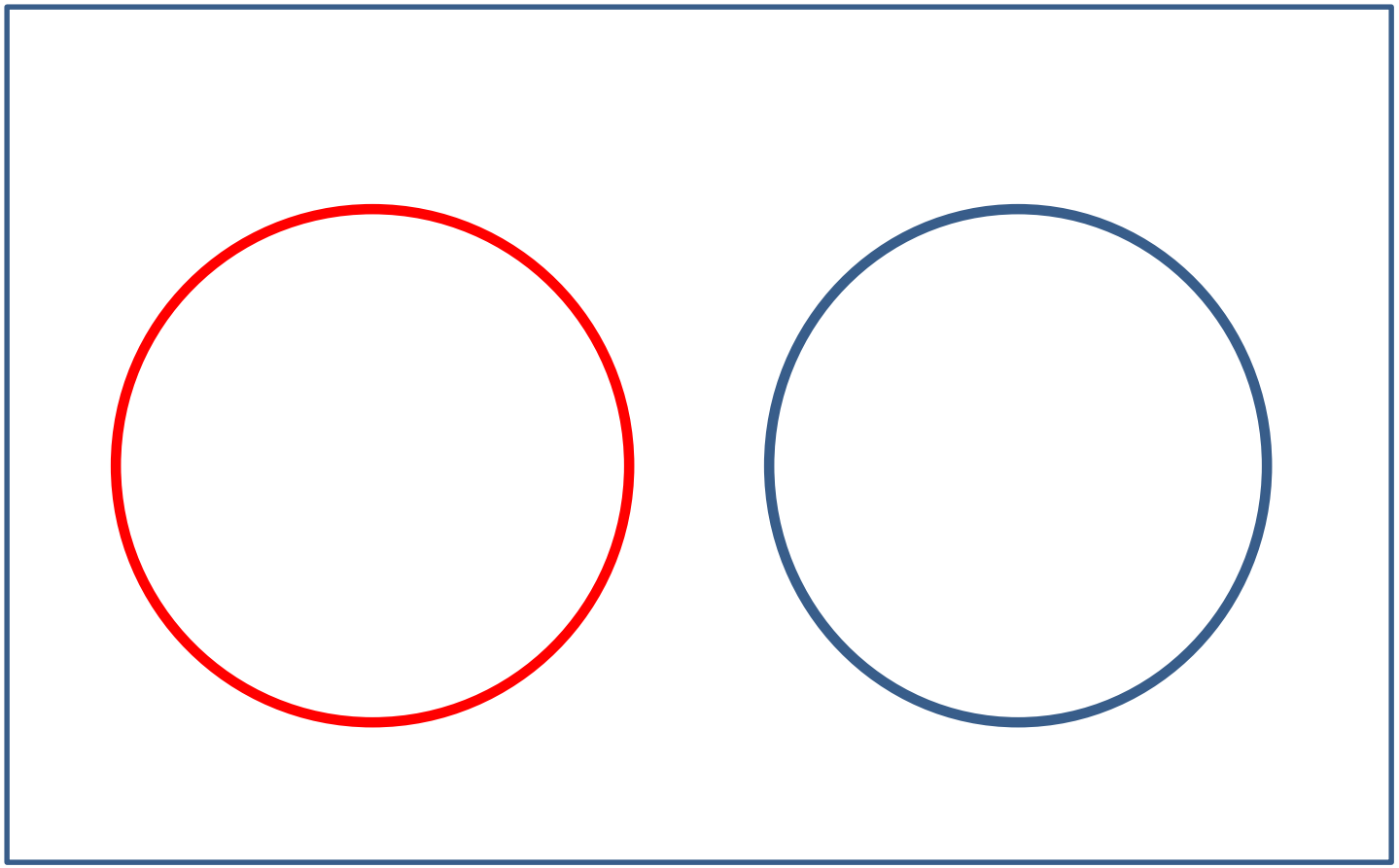
**. Correlation is significant at the 0.01 level (2-tailed).

Explaining Multiple Regression

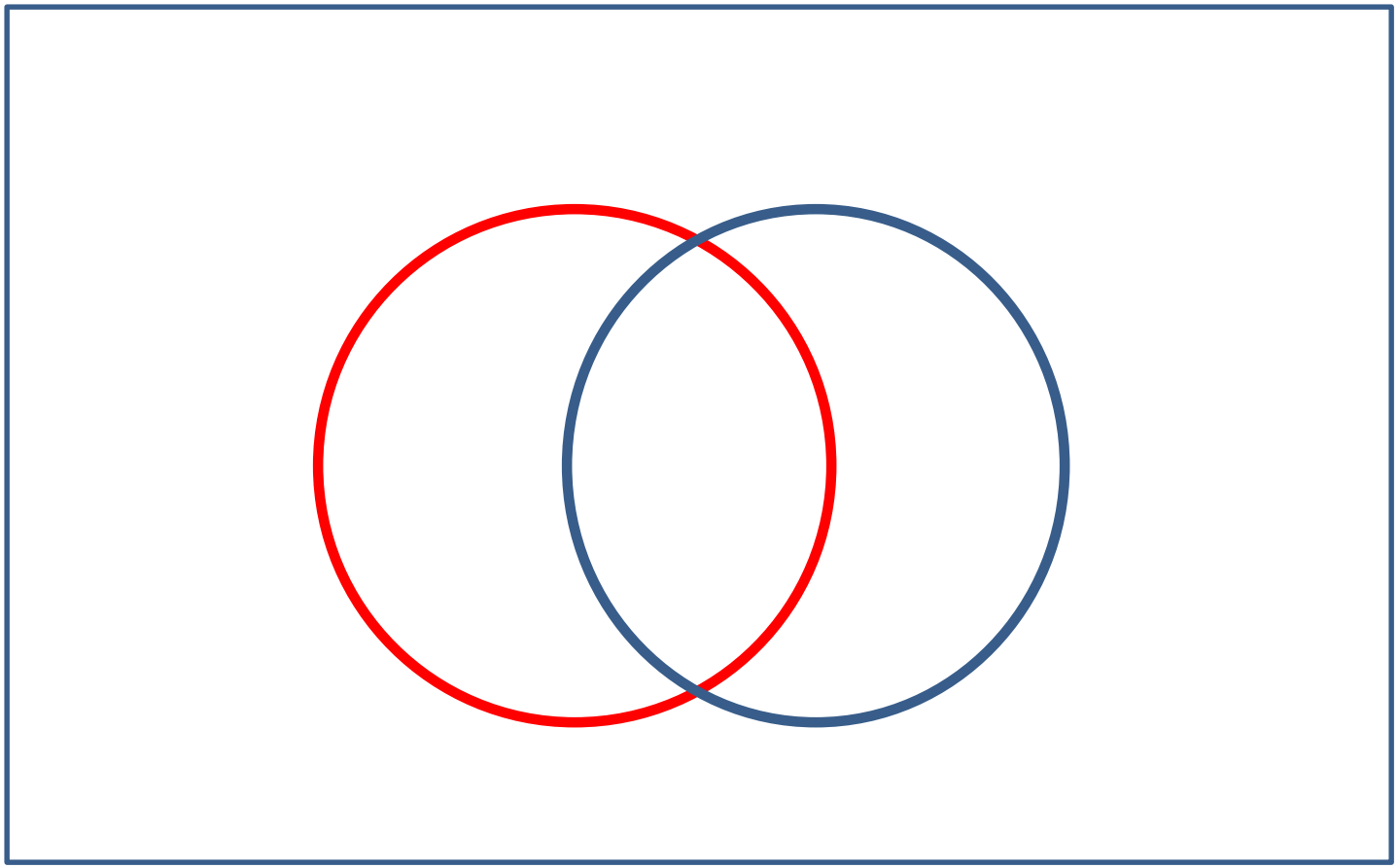
Explaining Multiple Regression



Explaining Multiple Regression



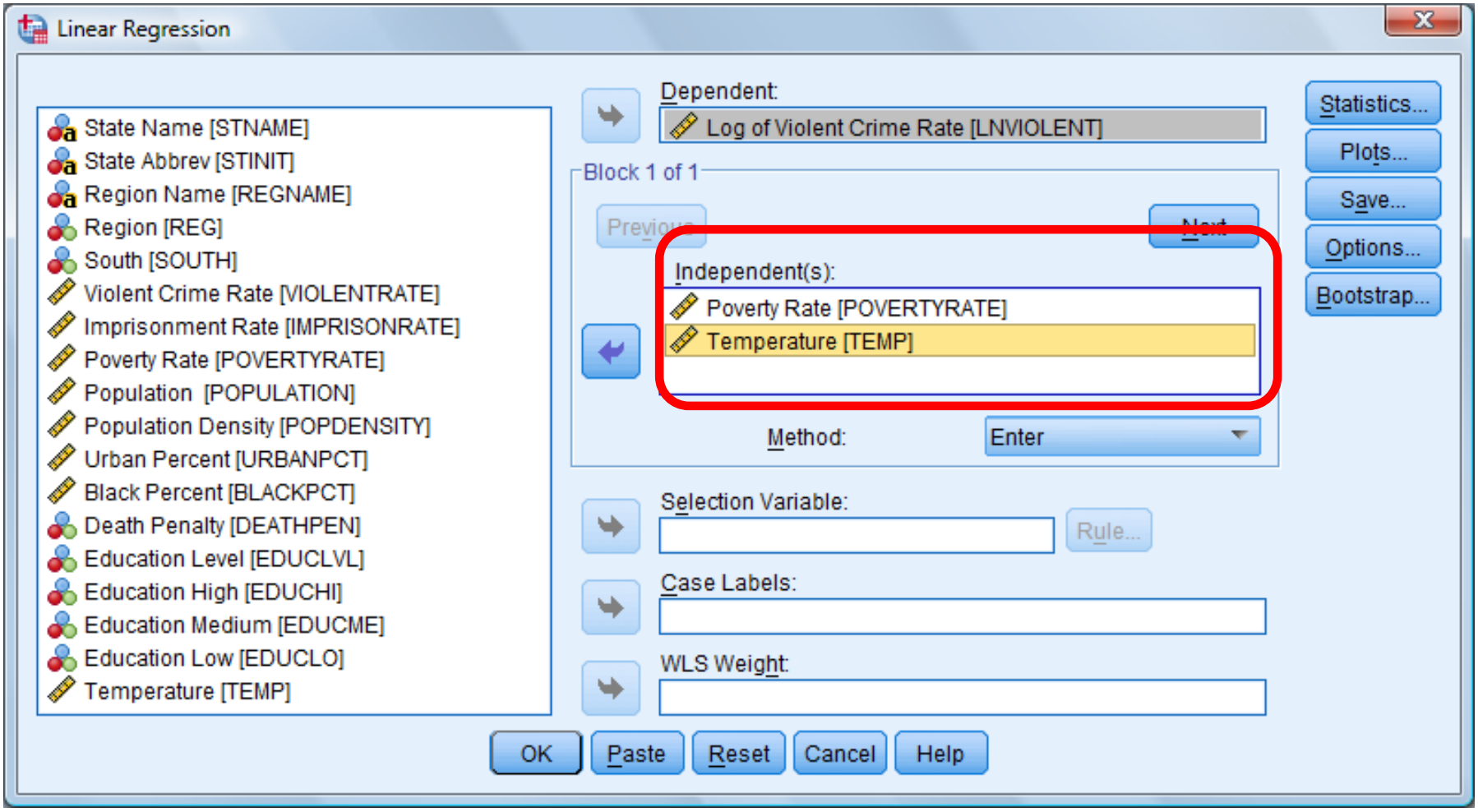
Explaining Multiple Regression



Adjusting for the Overlap

$$b_1 = \frac{(\sum x_2^2)(\sum x_1 y) - (\sum x_1 x_2)(\sum x_2 y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

Analyze > Regression > Linear



Multiple Regression Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.455[a]	.207	.174	.35312

a. Predictors: (Constant), Temperature, Poverty Rate

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.533	2	.766	6.146	.004[b]
	Residual	5.860	47	.125		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature, Poverty Rate

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.727	.314		15.057	.000
	Poverty Rate	.027	.019	.228	1.447	.155
	Temperature	.013	.007	.285	1.803	.078

a. Dependent Variable: Log of Violent Crime Rate

X_1 controlling for X_2

- An effect can stay the same
- An effect can increase in size
- An effect can decrease in size
- An effect can disappear
- An effect can change direction

"Control"

- Multiple regression always has one variable controlling for the effect of another, and vice versa.

- What we call the control variable depends on our intent
  - An independent variable is of theoretical interest (specified in our hypothesis).
  - A control variable is included in the model but does not contribute to the theory.

- The computer does not distinguish between an independent and control variable.

"Control"

- Multiple regression always has one variable controlling for the effect of another, and vice versa.

- What we call the control variable depends on our intent
  - An independent variable is of theoretical interest (specified in our hypothesis).
  - A control variable is included in the model but does not contribute to the theory.

- The computer does not distinguish between an independent and control variable.

Interpreting the Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.455[a]	.207	.174	.35312

a. Predictors: (Constant), Temperature, Poverty Rate

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.533	2	.766	6.146	.004[b]
	Residual	5.860	47	.125		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature, Poverty Rate

$\hat{y} = 4.73 + .03(\text{poverty rate}) + .01(\text{temperature})$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.727	.314		15.057	.000
	Poverty Rate	.027	.019	.228	1.447	.155
	Temperature	.013	.007	.285	1.803	.078

a. Dependent Variable: Log of Violent Crime Rate

Interpreting the Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.455[a]	.207	.174	.35312

a. Predictors: (Constant), Temperature, Poverty Rate

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.533	2	.766	6.146	.004[b]
	Residual	5.860	47	.125		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature, Poverty Rate

ŷ = 4.73 + .03(poverty rate) + .01(temperature)

weak + positive

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.727	.314		15.057	.000
	Poverty Rate	.027	.019	.228	1.447	.155
	Temperature	.013	.007	.285	1.803	.078

a. Dependent Variable: Log of Violent Crime Rate

Multiple Regression Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.455[a]	.207	.174	.35312

a. Predictors: (Constant), Temperature, Poverty Rate

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.533	2	.766	6.146	.004[b]
	Residual	5.860	47	.125		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature, Poverty Rate

ŷ = 4.73 + .03(poverty rate) + .01(temperature)

weak + positive

not significant

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.727	.314		15.057	.000
	Poverty Rate	.027	.019	.228	1.447	.155
	Temperature	.013	.007	.285	1.803	.078

a. Dependent Variable: Log of Violent Crime Rate

Interpreting the Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.455[a]	.207	.174	.35312

a. Predictors: (Constant), Temperature, Poverty Rate

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.533	2	.766	6.146	.004[b]
	Residual	5.860	47	.125		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature, Poverty Rate

ŷ = 4.73 + .03(poverty rate) + .01(temperature)

weak + positive

not significant

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.727	.314		15.057	.000
	Poverty Rate	.027	.019	.228	1.447	.155
	Temperature	.013	.007	.285	1.803	.078

a. Dependent Variable: Log of Violent Crime Rate

Interpreting the Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.455[a]	.207	.174	.35312

a. Predictors: (Constant), Temperature, Poverty Rate

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.533	2	.766	6.146	.004[b]
	Residual	5.860	47	.125		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature, Poverty Rate

ŷ = 4.73 + .03(poverty rate) + .01(temperature)

weak + positive

not significant

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.727	.314		15.057	.000
	Poverty Rate	.027	.019	.228	1.447	.155
	Temperature	.013	.007	.285	1.803	.078

a. Dependent Variable: Log of Violent Crime Rate

Interpreting the Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.455[a]	.207	.174	.35312

a. Predictors: (Constant), Temperature, Poverty Rate

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.533	2	.766	6.146	.004[b]
	Residual	5.860	47	.125		
	Total	7.393	49			

a. Dependent Variable: Log of Violent Crime Rate

b. Predictors: (Constant), Temperature, Poverty Rate

$\hat{y} = 4.73 + .03(\text{poverty rate}) + .01(\text{temperature})$

weak + positive

not significant

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	4.727	.314		15.057	.000
	Poverty Rate	.027	.019	.228	1.447	.155
	Temperature	.013	.007	.285	1.803	.078

a. Dependent Variable: Log of Violent Crime Rate

APA Table

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.03
Temperature		.02**	.01t
Constant	5.10***	4.83***	4.73***
R^2	.15**	.17**	.21**

Note: t p<.10 *p<.05 **p<.01 ***p<.001

APA Table

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.03
Temperature		.02**	.01t
Constant	5.10***	4.83***	4.73***
R^2	.15**	.17**	.21**

Note: t p<.10 * p<.05 **p<.01 ***p<.001

APA Table

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.03
Temperature		.02**	.01t
Constant	5.10***	4.83***	4.73***
R^2	.15**	.17**	.21**

Note: t p<.10 * p<.05 **p<.01 ***p<.001

APA Table

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.03
Temperature		.02**	.01t
Constant	5.10***	4.83***	4.73***
R^2	.15**	.17**	.21**

Note: t p<.10 * p<.05 **p<.01 ***p<.001

APA Table

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.03
Temperature		.02**	.01t
Constant	5.10***	4.83***	4.73***
R^2	.15**	.17**	.21**

Note: t p<.10 *p<.05 **p<.01 ***p<.001

APA Table

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.03
Temperature		.02**	.01t
Constant	5.10***	4.83***	4.73***
R^2	.15**	.17**	.21**

Note: t p<.10 *p<.05 **p<.01 ***p<.001

APA Table

Table 3

OLS Regression Results for Log of the Violent Crime Rate (N=50)

	(1)	(2)	(3)
% poverty rate	.05**		.03
Temperature		.02**	.01t
Constant	5.10***	4.83***	4.73***
R^2	.15**	.17**	.21**

Note: t p<.10 * p<.05 **p<.01 ***p<.001

A Complex Model

Table 5

OLS Regressions (n=1,069)

	Prison Programs			Prison Safety			
Dependent Variable	% Work	# Education	# Treatment	Disciplinary Reports	Major Events	Assaults on Inmates	Assaults on Staff
Prison ownership (ref=Private)							
Federal	3.83***	0.47***	0.73***	-7.80**	-0.49**	0.37*	-0.02
State	-0.43	-0.02	0.04	1.55	-0.25**	0.06	-0.03
Log of inmates	-0.24*	0.17***	0.12***	1.35*	-0.06*	0.23***	0.12***
Security level (ref=Minimum)							
Maximum	-1.53***	0.08*	0.10t	7.29***	-0.13t	0.52***	0.67***
Medium	-1.09***	0.14***	0.20***	5.39***	-0.14*	0.46***	0.30***
Gender assignment (ref=Female)							
Male	0.26	-0.20***	-0.40***	-2.62t	-0.05	-0.17t	0.16*
Mixed	1.28**	-0.10	-0.32***	-2.32	0.32**	0.34*	0.18
Construction (ref=Unknown)							
Before 1945	-2.50***	0.53***	0.61***	-0.38	-0.52**	-0.18	-0.37*
1945 to 1964	-1.98***	0.78***	0.74***	-3.70	-0.40*	-0.26	-0.22
1965 to 1979	-1.30*	0.58***	0.48***	-6.09*	-0.62***	0.00	-0.31t
1980 or later	-2.18***	0.52***	0.48***	-0.91	-0.53***	-0.10	0.42**
% in work programs (transformed)				-0.39*	0.02t	-0.03**	0.01
# education programs (transformed)				-1.97t	-0.08	-0.08	-0.16**
# treatment programs (transformed)				-5.42***	-0.07t	-0.15**	-0.02
Constant	0.61	-3.54***	.21***	2.34	1.04***	-0.62t	-0.73**
Adjusted R^2	.26***	.34***	.20***.	.21***	.08***	.17***	.16***

t p<.10 * p<.05 ** p<.01 *** p<.001

How Many Variables?

The minimum of the following:

$$k = \frac{N - 50}{8}$$

$$k = N - 104$$

where N=number of observations and k=number of variables.

Derived from Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson.

The End

- Multiple regression builds on the principles of simple regression to allow the creation of models with two or more independent variables.

- The need to remove correlational effects among independent variables creates results that are different from simple regression.

- “Once you can do three, you can do them all.”