

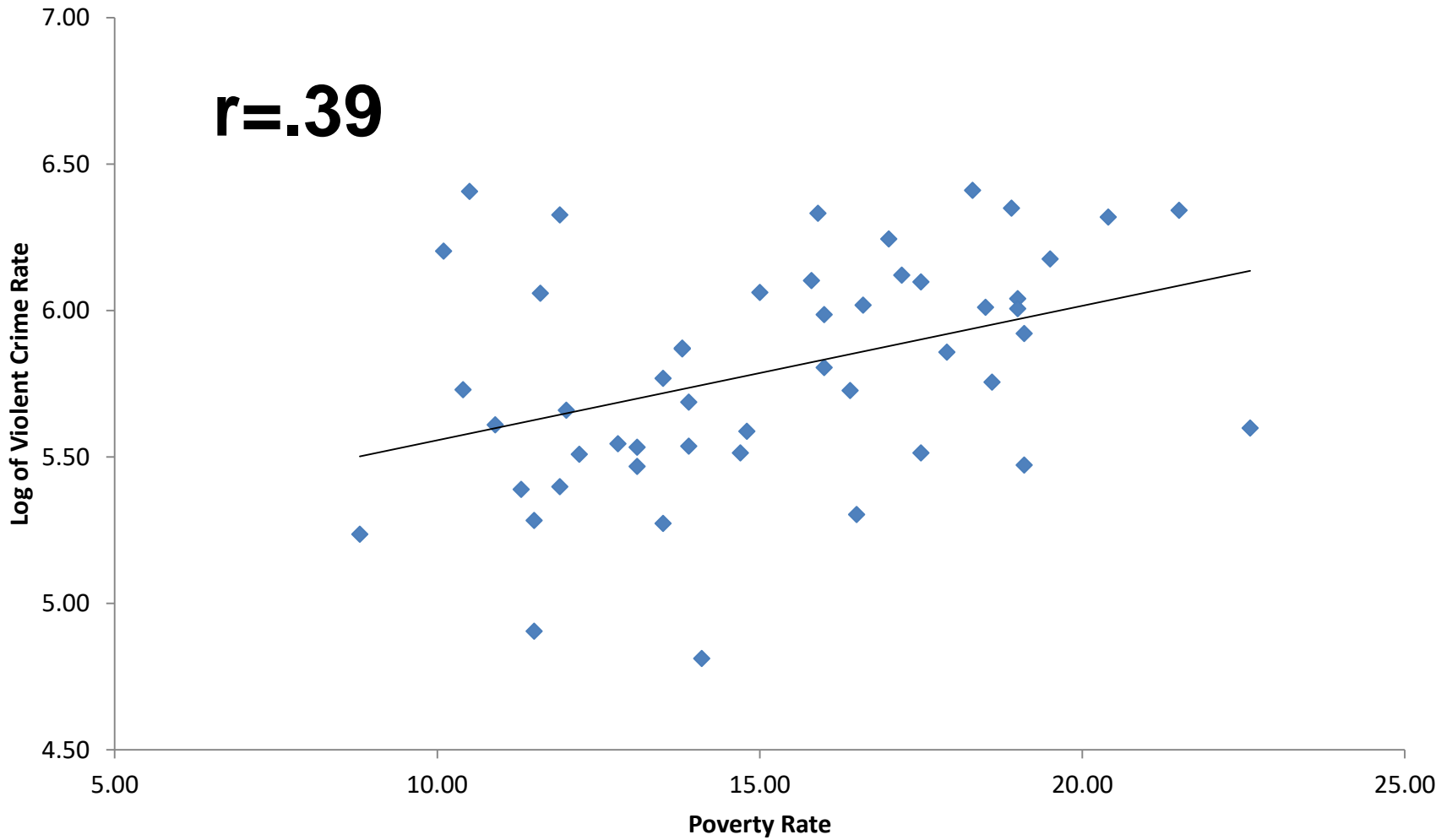
SIMPLE OLS REGRESSION, PART I: THE EQUATION OF A LINE

Richard Lee Rogers

Last Update: February 14, 2016



Example



Analyze > Regression > Linear

Transform **Analyze** Direct Marketing Graphs Utilities Add-ons Window Help

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response

Visible:

EG	SOUTH	VIOLENT RATE	IM
3	1	420	
4	0	606	
4	0	406	
2	1	481	
		411	
		320	
		273	
		559	
		515	
		373	
		287	
		201	
		429	
		332	

Automatic Linear Modeling...
Linear...
Curve Estimation...
Partial Least Squares...
Binary Logistic...
Multinomial Logistic...
Ordinal...
Probit...
Nonlinear



Analyze > Regression > Linear

The screenshot shows the SPSS software interface. The 'Analyze' menu is open, and the 'Regression' sub-menu is also open, highlighting the 'Linear...' option. The background shows a data table with columns: STATE, REGION, TRA, IMPRISONR, POVERTYRA, and POPU.

STATE	REGION	TRA	IMPRISONR	POVERTYRA	POPU		
AL	South	420	648.0	19.0	4		
AK	West	606	340.0	10.5			
AZ	West	406	572.0	19.0	6		
AR	South			9.5	2		
CA	West			6.6	37		
CO	West	320	445.0	13.5	5		
CT	Northeast	1	0	273	376.0	10.9	3
DE	South	3	1	559	443.0	11.9	



Command Dialog Box

Linear Regression

State Name [STNAME]
State Abbrev [STINIT]
Region Name [REGNAME]
Region [REG]
South [SOUTH]
Violent Crime Rate [VIOLENTRATE]
Imprisonment Rate [IMPRISONRATE]
Poverty Rate [POVERTYRATE]
Population [POPULATION]
Population Density [POPDENSITY]
Urban Percent [URBANPCT]
Black Percent [BLACKPCT]
Death Penalty [DEATHPEN]
Education Level [EDUCLVL]
Education High [EDUCHI]
Education Medium [EDUCME]
Education Low [EDUCLO]
Temperature [TEMP]

Dependent:
Log of Violent Crime Rate [LNVIOLENT]

Block 1 of 1
Previous Next

Independent(s):
Poverty Rate [POVERTYRATE]

Method: Enter


Selection Variable:
Rule...

Case Labels:

WLS Weight:

OK Paste Reset Cancel Help

Statistics...
Plots...
Save...
Options...
Bootstrap...



Command Dialog Box

Linear Regression

State Name [STNAME]
State Abbrev [STINIT]
Region Name [REGNAME]
Region [REG]
South [SOUTH]
Violent Crime Rate [VIOLENTRATE]
Imprisonment Rate [IMPRISONRATE]
Poverty Rate [POVERTYRATE]
Population [POPULATION]
Population Density [POPDENSITY]
Urban Percent [URBANPCT]
Black Percent [BLACKPCT]
Death Penalty [DEATHPEN]
Education Level [EDUCLVL]
Education High [EDUCHI]
Education Medium [EDUCME]
Education Low [EDUCLO]
Temperature [TEMP]

Dependent:
Log of Violent Crime Rate [LNVIOLENT]

Block 1 of 1

Previous Next

Independent(s):
Poverty Rate [POVERTYRATE]

Method: Enter

Selection Variable:
Rule...

Case Labels:

WLS Weight:

OK Paste Reset Cancel Help

Statistics...
Plots...
Save...
Options...
Bootstrap...



Command Dialog Box

Linear Regression

State Name [STNAME]
State Abbrev [STINIT]
Region Name [REGNAME]
Region [REG]
South [SOUTH]
Violent Crime Rate [VIOLENTRATE]
Imprisonment Rate [IMPRISONRATE]
Poverty Rate [POVERTYRATE]
Population [POPULATION]
Population Density [POPDENSITY]
Urban Percent [URBANPCT]
Black Percent [BLACKPCT]
Death Penalty [DEATHPEN]
Education Level [EDUCLVL]
Education High [EDUCHI]
Education Medium [EDUCME]
Education Low [EDUCLO]
Temperature [TEMP]

Dependent:
Log of Violent Crime Rate [LNVIOLENT]

Block 1 of 1
Previous Next

Independent(s):
Poverty Rate [POVERTYRATE]

Method: Enter

Selection Variable:
Rule...

Case Labels:

WLS Weight:

OK Paste Reset Cancel Help

Statistics...
Plots...
Save...
Options...
Bootstrap...



Output

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	Poverty Rate ^b	.	Enter

a. Dependent Variable: Log of Violent Crime Rate

b. All requested variables entered.

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



Parameters of the Line

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	.005

a. Dependent Variable: Log of Violent Crime Rate



Parameters of the Line: B_0

b_0

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



Parameters of the Line: B_1

b_1

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



The Equation of the Line

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



The Equation of the Line

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

$$\hat{y} = b_0 + b_1x_1 = 5.10 + .05(\text{poverty rate})$$



The Equation of the Line

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

$$\hat{y} = b_0 + b_1x_1 = 5.10 + .05(\text{poverty rate})$$

$$b_1 = \frac{\text{covariance of } x \text{ and } y}{\text{variance of } x} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$



The Equation of the Line

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

$$\hat{y} = b_0 + b_1x_1 = 5.10 + .05(\text{poverty rate})$$

$$b_1 = \frac{\text{covariance of } x \text{ and } y}{\text{variance of } x} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = r_{xy} \frac{s_y}{s_x}$$



What the Slope Tells Us

- Sign of b_1 is the direction of the relationship.
- The magnitude of b_1 is unstandardized
- In this analysis
 - The relationship is positive
 - Magnitude: The log of the violent crime rate increases .05 for every percent increase in the poverty rate



Standard Error of the Beta Coefficient

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	.005

a. Dependent Variable: Log of Violent Crime Rate



Standardized Beta

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	2.939	.005

a. Dependent Variable: Log of Violent Crime Rate



Statistical Significance

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

Null Hypotheses:

$$b_0 = 0$$

$$b_1 = 0$$



t

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	.005

a. Dependent Variable: Log of Violent Crime Rate



Elements of An Inferential Statistic

Elements in the parameter of a line are associated with B_1

- Direction: sign
- Unstandardized magnitude: beta (slope)
- Standardized magnitude: standardized beta
- Statistical significance

Constant usually has no substantive bearing on the relationship

Example results

- The relationship is positive
- The relationship is moderate (standardized magnitude: $b_1=.39$)
- Statistically significant
- The log of violent crime rate increases .05 for each percent increase in the poverty rate (unstandardized beta: $b_0=.05$)

