@Coefficient of variation
 (R^2)

① Sum of Squares

③ Multiple regression & various
   aspects

Sum of Squares

Y

regression in
the middle
- least distance
between each do
& the line

X

the same
number of obs
are above the
below the line.
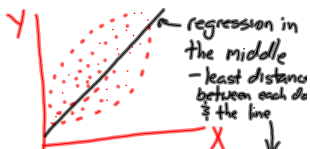
Each distance between
the ith observation and
the regression line is the
residual error (ε)

$$\varepsilon = \sum_i (\hat{y}_i - y_i)^2$$

predicted   actual value
value of the   of the ith
i^{th} observation   observation

(based on the
regression)   - Why do we
                     square the differ
                        ↓
                     To make ε p

Summing all our residuals is
called sum of squares error/re

Y   Y

X   X
Low SSR   High SSR

Same
Slope
Same
Intercept
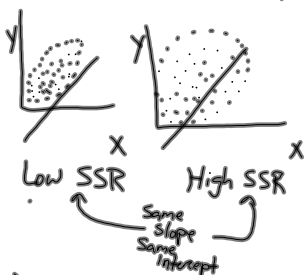
→ Total amount of variation
explained by the independent

Sum of squares expl
→ Total amount of varia
explained by the independent

actual

hypothesized

Sum of Squares To

Sum of Sq

*Output2 [Document2] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

Output
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients
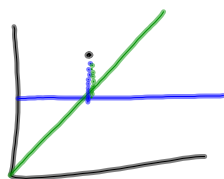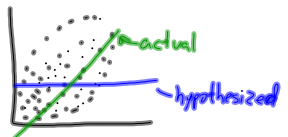
| 1 | DENS[a] | . | Enter |

a. All requested variables entered.
b. Dependent Variable: PVT

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.374[a]	.140	.122	7.3391431

a. Predictors: (Constant), DENS

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	419.680	1	419.680	7.792	.008[a]
	Residual	2585.425	48	53.863		
	Total	3005.105	49			

a. Predictors: (Constant), DENS
b. Dependent Variable: PVT

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	7.575	3.209		2.361	.022
	DENS	.254	.091	.374	2.791	.008

a. Dependent Variable: PVT

IBM SPSS Statistics Processor is ready

Handwritten annotations:

1 to 1

β

For every 1% increase in DENS, PVT will increase 0.254 percent.

less than 0.05, so it is significant.

% of public sector unionized

When you have multiple ind. variables, R^2 is no longer valid.

*Output2 [Document2] - IBM SPSS Statistics Viewer

File　Edit　View　Data　Transform　Insert　Format　Analyze　Graphs　Utilities　Add-ons　Window　Help

Output
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients

a. All requested variables entered.
b. Dependent Variable: PVT

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.598[a]	.358	.330	6.4088783

a. Predictors: (Constant), RTW, DENS

(handwritten note with arrow pointing to "Adjusted R Square"): accounts for the number of variables used in regression.

(handwritten note with arrow pointing to the .330 value): Valid for regressions w/ multiple independent variables (tends to be lower than R^2)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1074.640	2	537.320	13.082	.000[a]
	Residual	1930.465	47	41.074		
	Total	3005.105	49			

a. Predictors: (Constant), RTW, DENS
b. Dependent Variable: PVT

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17.565	3.756		4.676	.000
	DENS	.059	.093	.086	.628	.533
	RTW	-8.676	2.173	-.548	-3.993	.000

a. Dependent Variable: PVT

IBM SPSS Statistics Processor is ready

*Output2 [Document2] - IBM SPSS Statistics Viewer

File   Edit   View   Data   Transform   Insert   Format   Analyze   Graphs   Utilities   Add-ons   Window   Help

Output
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients

a. All requested variables entered.
b. Dependent Variable: PVT

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.598[a]	.358	.330	6.4088783

a. Predictors: (Constant), RTW, DENS

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1074.640	2	537.320	13.082	.000[a]
	Residual	1930.465	47	41.074		
	Total	3005.105	49			

a. Predictors: (Constant), RTW, DENS
b. Dependent Variable: PVT

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17.565	3.756		4.676	.000
	DENS	.059	.093	.086	.628	.533
	RTW	-8.676	2.173	-.548	-3.993	.000

a. Dependent Variable: PVT

Handwritten annotations:

Multiple regressors

.330 (Adj. R^2)

RTW Only

.352 (R^2)

IBM SPSS Statistics Processor is ready

*Output2 [Document2] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

Output
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients
- Log
- Regression
  - Title
  - Notes
  - Active Dataset
  - Variables Entered
  - Model Summary
  - ANOVA
  - Coefficients

a. All requested variables entered.
b. Dependent Variable: PVT

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.598[a]	.358	.330	6.4088783

a. Predictors: (Constant), RTW, DENS

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1074.640	2	537.320	13.082	.000[a]
	Residual	1930.465	47	41.074		
	Total	3005.105	49			

a. Predictors: (Constant), RTW, DENS
b. Dependent Variable: PVT

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	17.565	3.756		4.676	.000
	DENS	.059	.093	.086	.628	.533
	RTW	-8.676	2.173	-.548	-3.993	.000

a. Dependent Variable: PVT

IBM SPSS Statistics Processor is ready

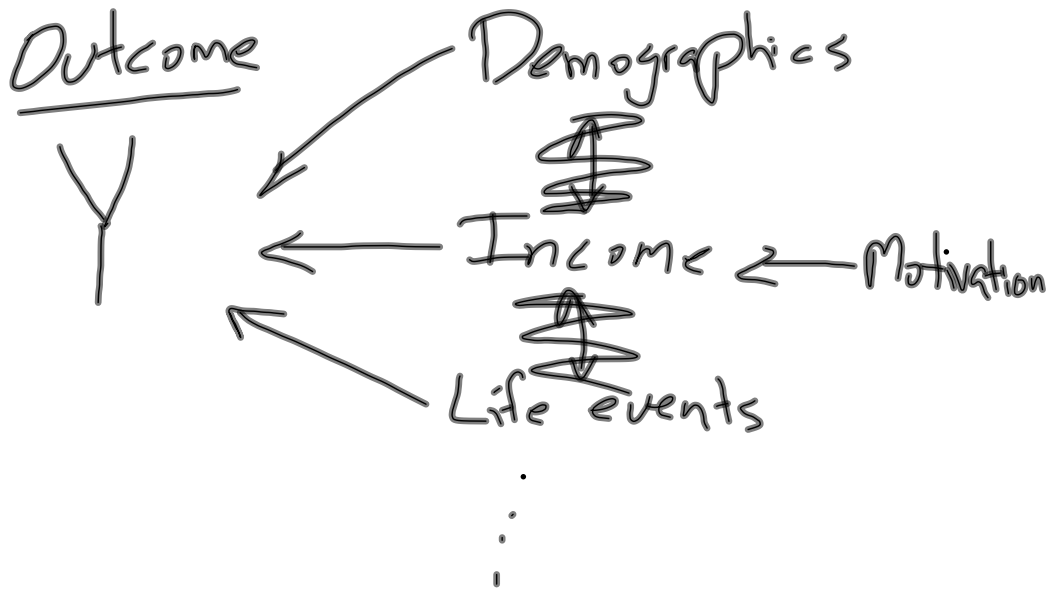*Handwritten annotations:*

DENS – for every 1% increase in DENS, PVT will rise 0.059 percent.

RTW states have a value of PVT that is 8.676 percent less than non-RTW.

Is statistically significant

7

Outcome

Demographics

Y

Income — Motivation

Life events

$\vdots$

Correlation is not sufficient for causation, but it is necessary.
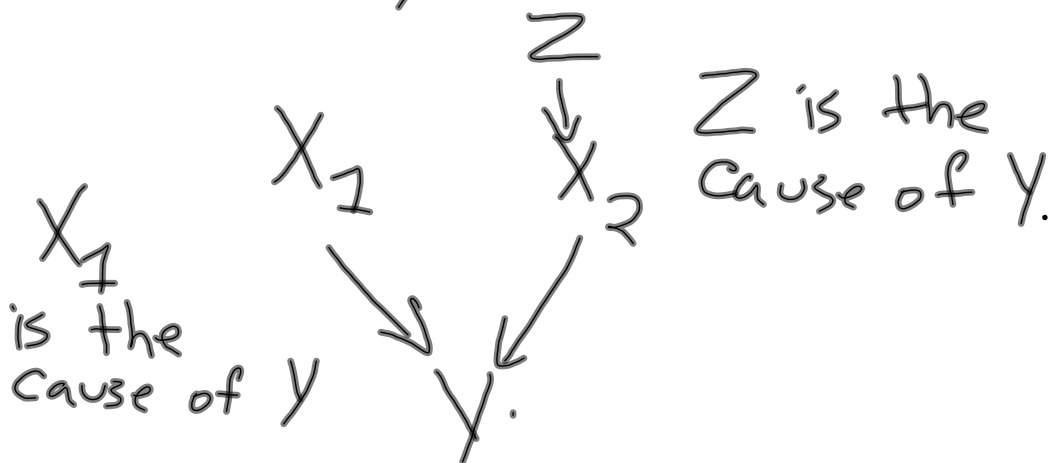
Z

X_1

X_2

Z is the cause of Y.

X_1 is the cause of Y

Y

*Output2 [Document2] - IBM SPSS Statistics Viewer

File   Edit   View   Data   Transform   Insert   Format   Analyze   Graphs   Utilities   Add-ons   Window   Help

- Output
  - Log
  - Regression
    - Title
    - Notes
    - Active Dataset
    - Variables Entered
    - Model Summary
    - ANOVA
    - Coefficients

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	RTW[a]	.	Enter

a. All requested variables entered.
b. Dependent Variable: PVT

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.593[a]	.352	.339	6.3683203

a. Predictors: (Constant), RTW

$$.593 \times .593 = .352 = R^2$$

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1058.441	1	1058.441	26.099	.000[a]
	Residual	1946.664	48	40.556		
	Total	3005.105	49			

a. Predictors: (Constant), RTW
b. Dependent Variable: PVT

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	19.807	1.163		17.035	.000

IBM SPSS Statistics Processor is ready

*Output2 [Document2] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

- Output
  - Log
  - Regression
    - Title
    - Notes
    - Active Dataset
    - Variables Entered
    - Model Summary
    - ANOVA
    - Coefficients

Model	Variables Entered	Variables Removed	Method
1	RTW[a]	.	Enter

a. All requested variables entered.
b. Dependent Variable: PVT

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.593[a]	.352	.339	6.3683203

a. Predictors: (Constant), RTW

Explained

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1058.441	1	1058.441	26.099	.000[a]
	Residual	1946.664	48	40.556		
	Total	3005.105	49			

a. Predictors: (Constant), RTW
b. Dependent Variable: PVT

$$\frac{\text{Sum of Sq. Regression}}{\text{Sum of Sq. Total}} = R^2$$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	19.807	1.163		17.035	.000
	RTW	-9.392	1.838	-.593	-5.109	.000

a. Dependent Variable: PVT

IBM SPSS Statistics Processor is ready

35.2% of the variation in PVT is explained by RTW.