$$y = a + bx$$

where:

y equals the predicted value of the dependent variable;

a equals the intercept;

b equals the slope or the regression coefficient;

x equals an observed value of the independent variable.

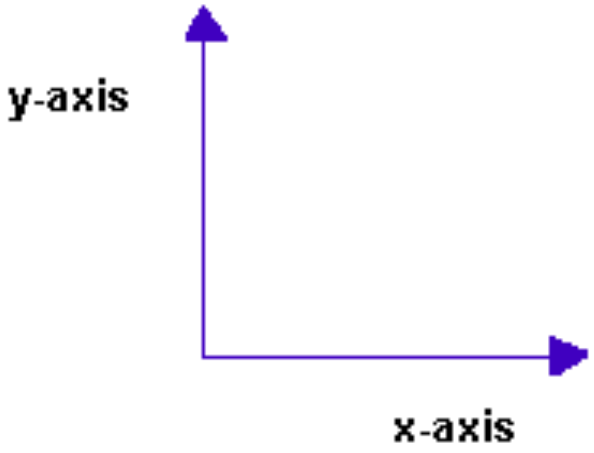
$$Y' = a + bX$$

$$Y = a + b(X).$$

$$Y = a + b(X).$$

$$\hat{Y} = a + b\ (x).$$

$$y = a + bx$$


y-axis
x-axis

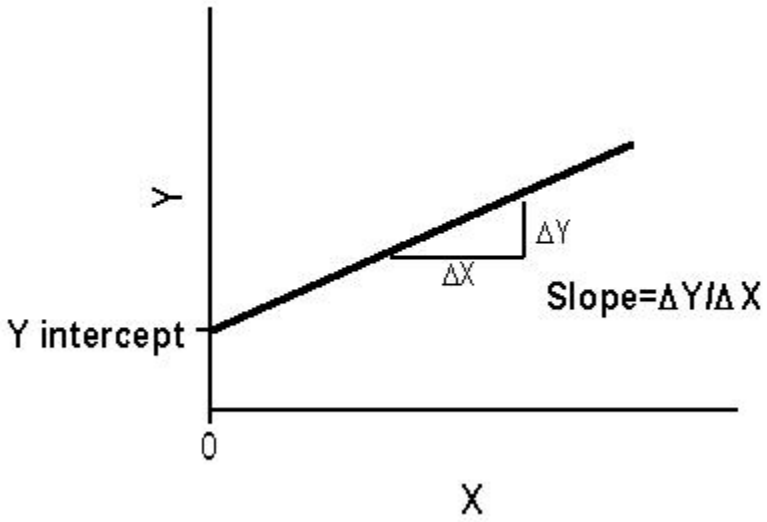

Y
ΔY
ΔX
Slope=ΔY/ΔX
Y intercept
0
X

Exhibit 7.1

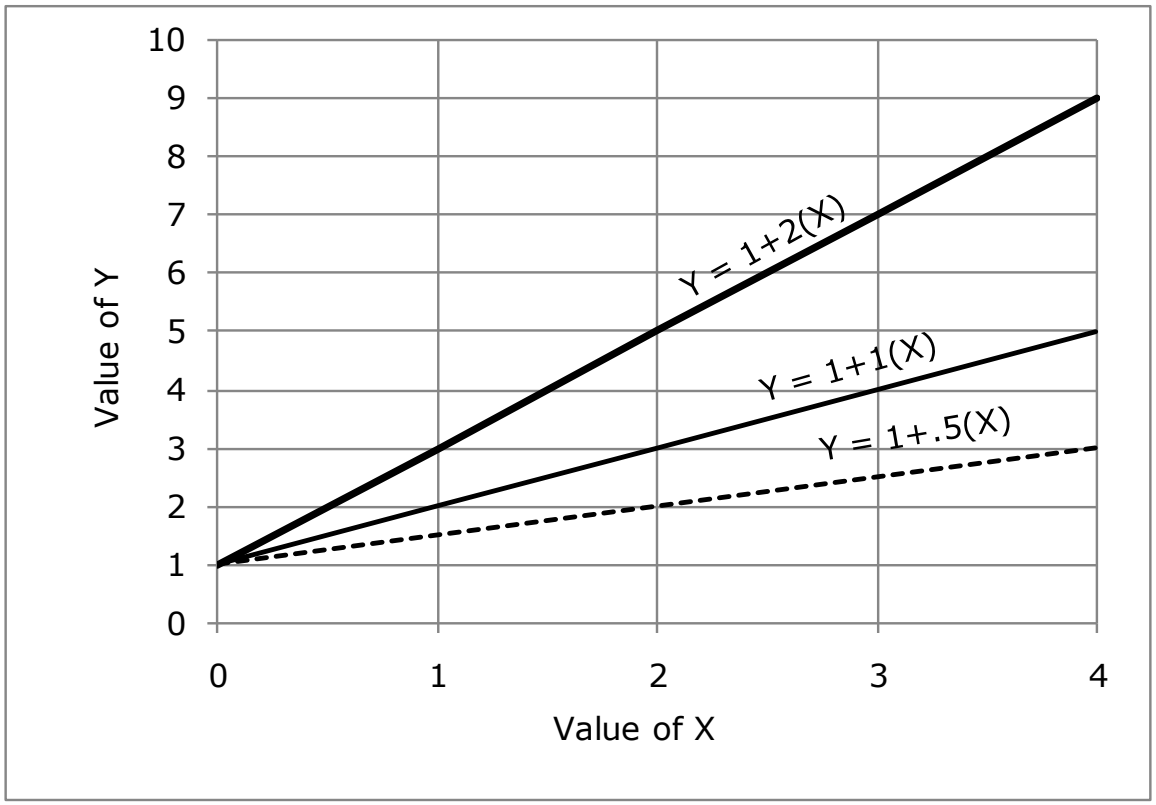
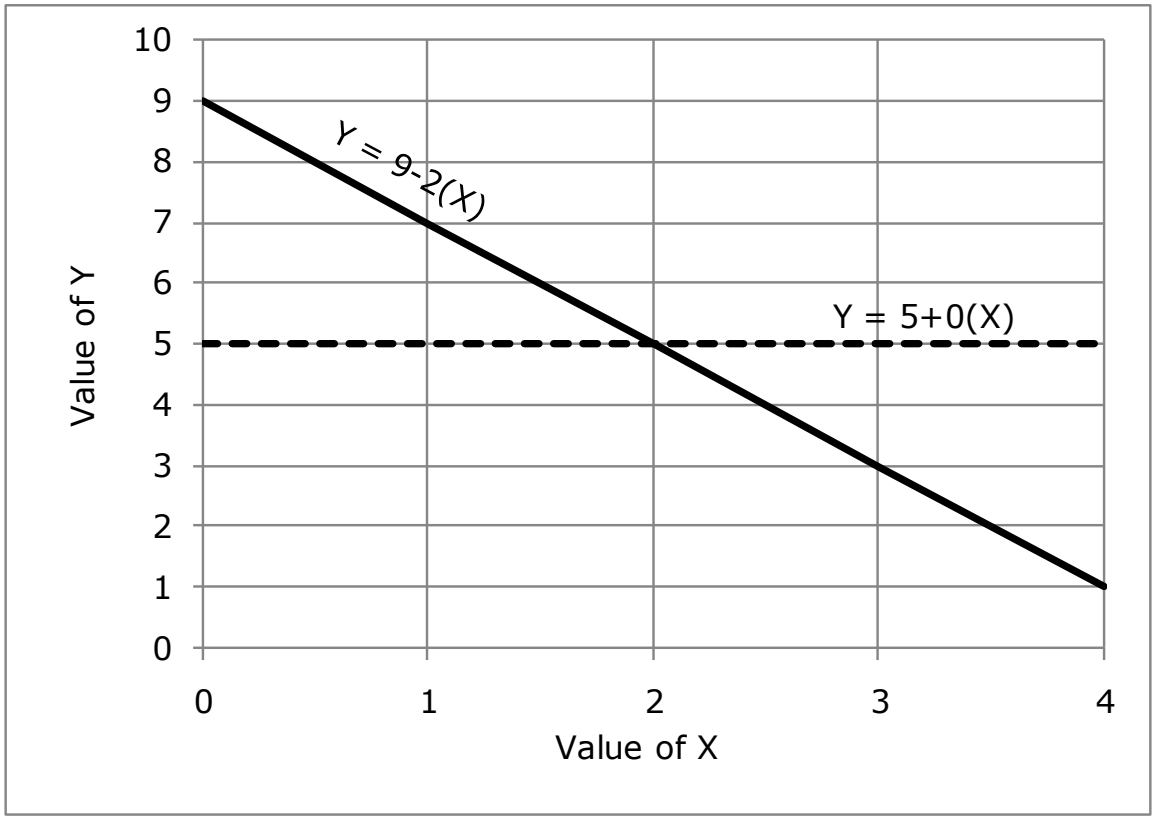


Exhibit 7.2



Total Variation

– Unexplained Variation

Explained Variation.

$$SDy = \sqrt{\frac{\sum \left(y_i - \bar{Y} \right)^2}{N}}$$

$$SD^2 = \frac{\sum (y_i - \bar{Y})^2}{N}$$

$$Variance = \frac{\sum (y_i - \bar{Y})^2}{N}$$

Total Variation
– Unexplained Variation
Explained Variation.

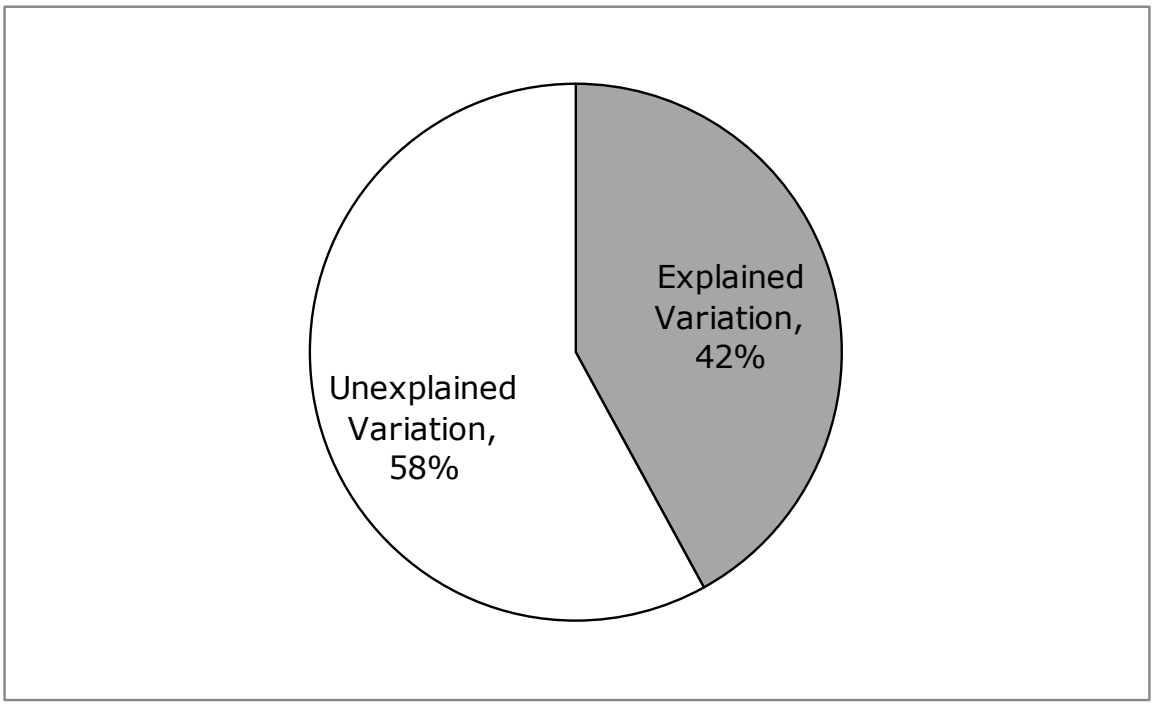
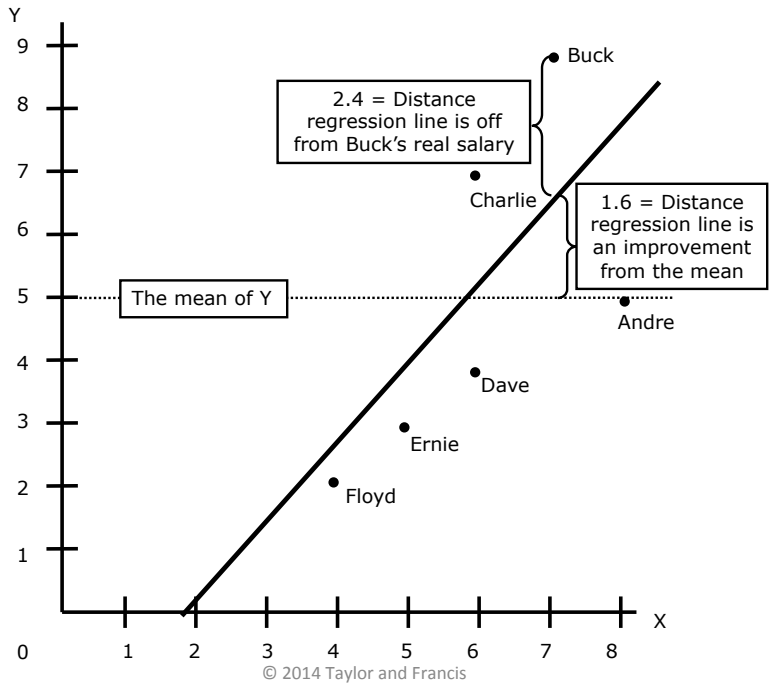


Exhibit 7.19: Illustrating Explained and Unexplained Variation



Total Variation

– **<u>Unexplained Variation</u>**

Explained Variation.

$$r = \sqrt{\exp lained\ \mathrm{var} iance}$$

r^2 = explained variance

My terms	SPSS terms
Total Variation	Total
– Unexplained Variation	-Residual
Explained Variation.	Regression

```
*Weighting the Data*.
weight by weight.
*Recoding MJ Index Items*.
recode q21 (1=1) (2=0) into MJPropD.
value labels MJPropD 1 'yes' 0 'no'.
recode q36 (1=1) (2=0) into MJLegalD.
value labels MJLegalD 1 'yes' 0 'no'.
recode q36a (1=1) (2=.5) (3=.0) into MJTry.
value labels MJTry 1 'recent' .5 'not recent' 0 'no'.

*Constructing an Index with alpha = .777*.
compute RawMJ3 = (MJPropD + MJLegalD + MJTry).

*Creating IV Indicators of Party Identification & Ideology*.
recode q40c (1=0) (3=.5) (2=1) into Democrat.
value labels Democrat 1 'Democ' .5 'Indep' 0 'Repub'.
*Democrat5 (adapted from from lab 7)*.
if (q40c = 1) and (q40e =1) Democrat5 =0.
if (q40c = 1) and (q40e =2) Democrat5 =.25.
if (q40c = 3) Democrat5 =.5.
if (q40c =2) and (q40d =2) Democrat5 = .75.
if (q40c =2) and (q40d=1) Democrat5 =1.
value labels Democrat5 0 'strRep' .25 'Rep' .5 'Indep' .75 'Dem'
   1 'strDem'.
recode q37 (1,2=1) (3=.5) (4,5= 0) into liberal.
value labels liberal 1 'liberal' .5 'middle' 0 'conserv'.
recode q37 (1=1) (2=.75) (3= .5 ) (4=.25) (5= 0) into liberal5.
value labels liberal5 1 'vlib' .75 'liberal'.5 'middle' .25 'conserv' 0 'vcons'.

regression variables=RawMJ3 Democrat5
  /dependent =  RawMJ3
  /method = enter.
```

```
regression variables=RawMJ3 Democrat5
 /dependent =  RawMJ3
 /method = enter.
```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.209[a]	.044	.043	1.12629

a. Predictors: (Constant), Democrat5

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	55.490	1	55.490	43.744	.000[b]
	Residual	1218.702	961	1.269		
	Total	1274.191	962			

a. Dependent Variable: RawMJ3

b. Predictors: (Constant), Democrat5

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	1.108	.072		15.355	.000
	Democrat5	.734	.111	.209	6.614	.000

a. Dependent Variable: RawMJ3

y= a+ bx

RawMJ3 = 1.108 + .734 (Democrat5)

r^2= .044

Total Variation – Unexplained Variation = Explained Variation

1274.2 – 1218.7 =55.5;

Ratio of Explained to Total Variation = 55.5/1274.2 = .044 = r^2

```
regression variables=RawMJ3 liberal5
  /dependent = RawMJ3
  /method = enter.
```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.361[a]	.130	.129	1.07650

a. Predictors: (Constant), liberal5

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	169.650	1	169.650	146.395	.000[b]
	Residual	1134.606	979	1.159		
	Total	1304.256	980			

a. Dependent Variable: RawMJ3

b. Predictors: (Constant), liberal5

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.832	.067		12.476	.000
	liberal5	1.351	.112	.361	12.099	.000

a. Dependent Variable: RawMJ3

y= a+ bx

RawMJ3 = .832 + 1.351 (liberal5)

r^2 = .130

Total Variation – Unexplained Variation = Explained Variation

1304.3 – 1134.6 =169.7.

Ratio of Explained to Total Variation = 169.7/1304.3 = .130 = r^2

Comparing H1 & H2

	b	**r^2**
H1 Partisanship (Democrat5)	.73	.04
H2 Ideology (Liberal5)	1.35	.13

Observation:
Both b value and R^2 are greater for Ideology than Partisanship

Conclusion:
Ideology is a better predictor that Party of attitudes on Rec Marijuana

Calculating the predicted data points

y= a+ bx

RawMJ3 = .832 + 1.351 (liberal5)

liberal5 scored as:

1= very liberal
.75 = liberal
.50 = middle
.25 = conservative
.00 = very conservative

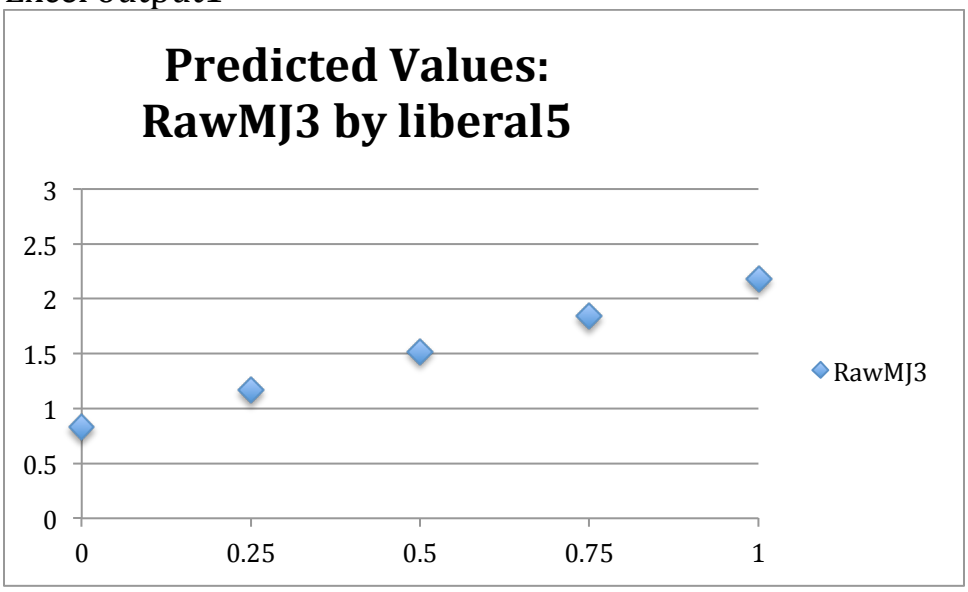
y(vlib) = .832 + 1.351 (1) = 2.183
y(lib) = .832 + 1.351 (.75) = 1.013
y(mid) = .832 + 1.351 (.50) = 1.507
y(con) = .832 + 1.351 (.25) = 1.013
y(vcon) = .832 + 1.351 (.00) = .832

Data entered in Excel

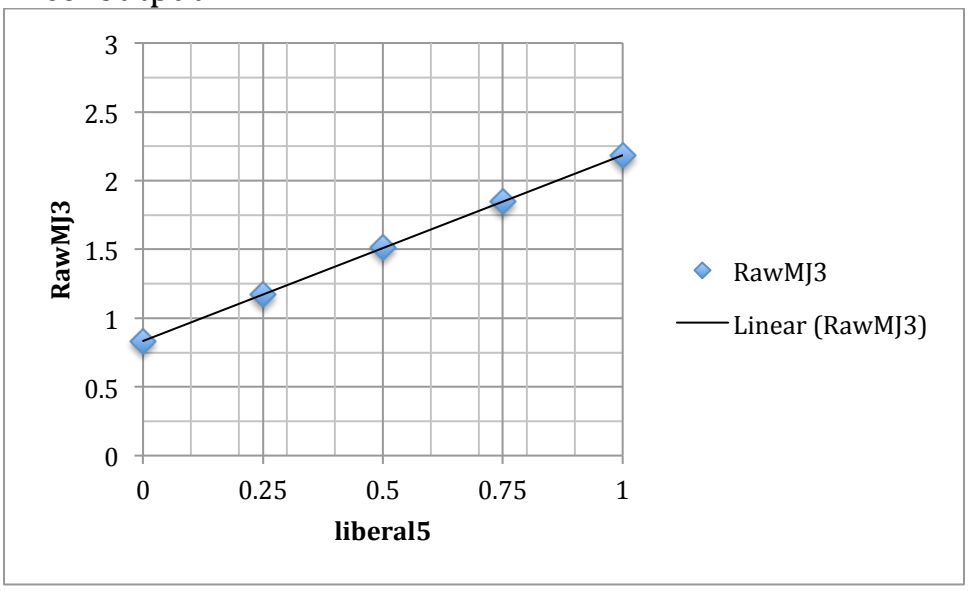
liberal5	RawMJ3
0	0.832
0.25	1.169
0.5	1.517
0.75	1.845
1	2.183

Visualizing Regression in Excel

Excel output1

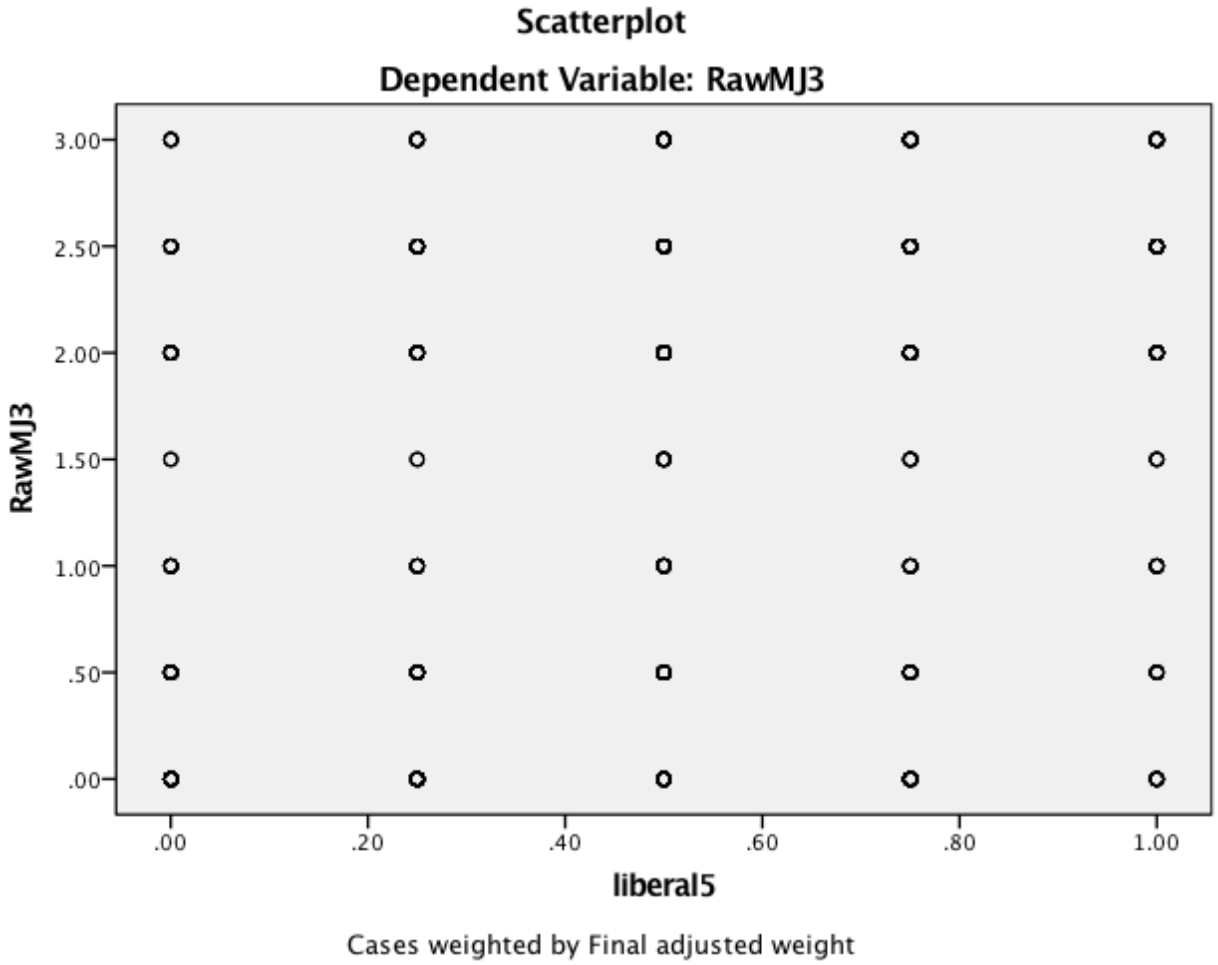


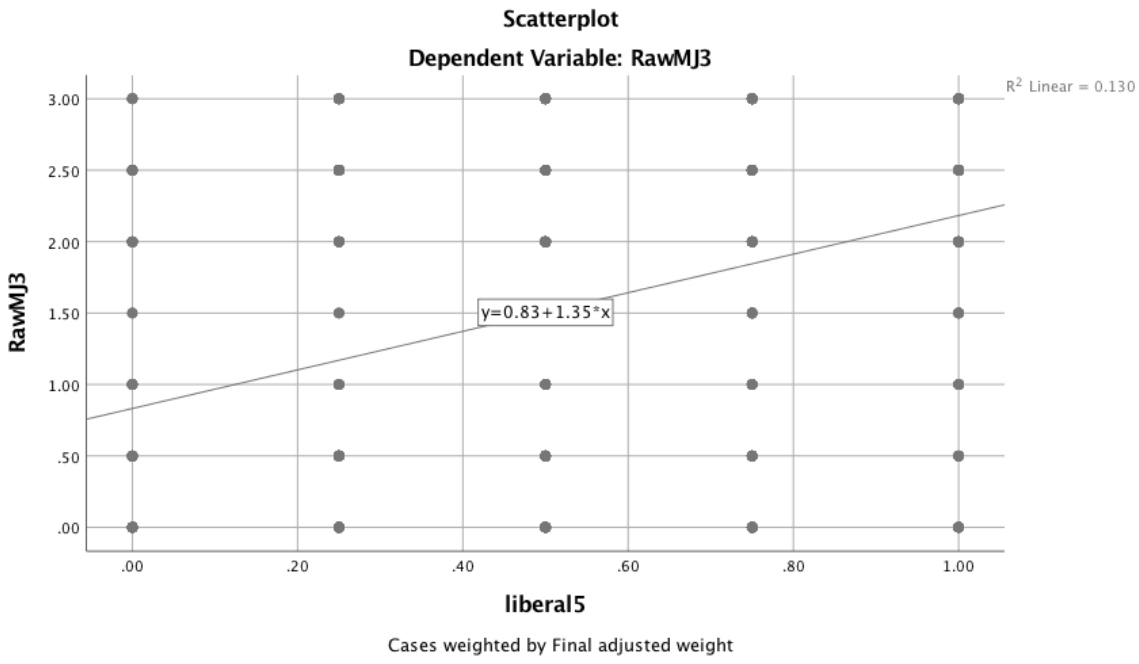
Excel output 2



Visualizing Regression results in SPSS

```
regression variables=RawMJ3 liberal5
  /dependent = RawMJ3
  /method = enter
 /scatterplot = (RawMJ3 liberal5).
```



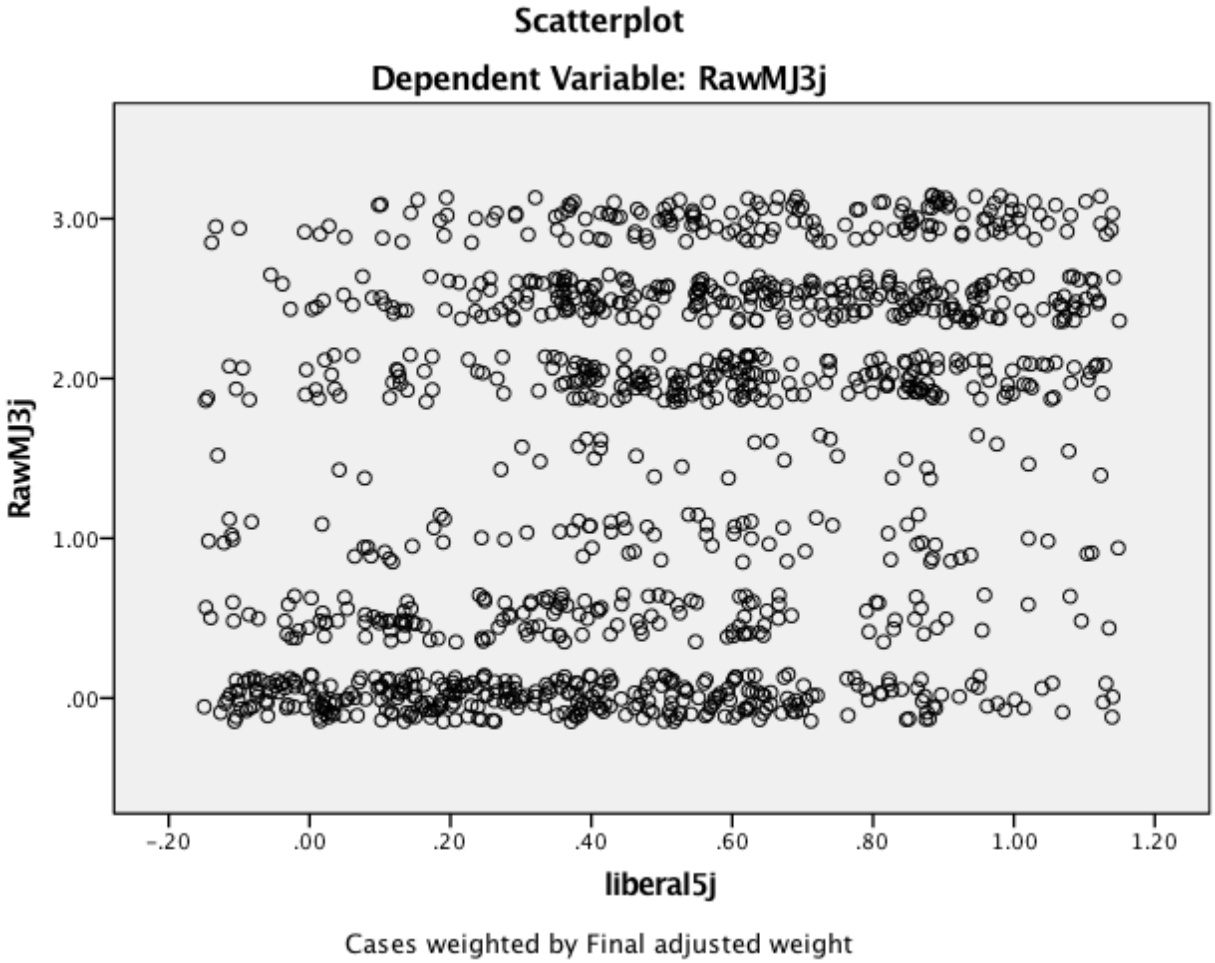


Cases weighted by Final adjusted weight

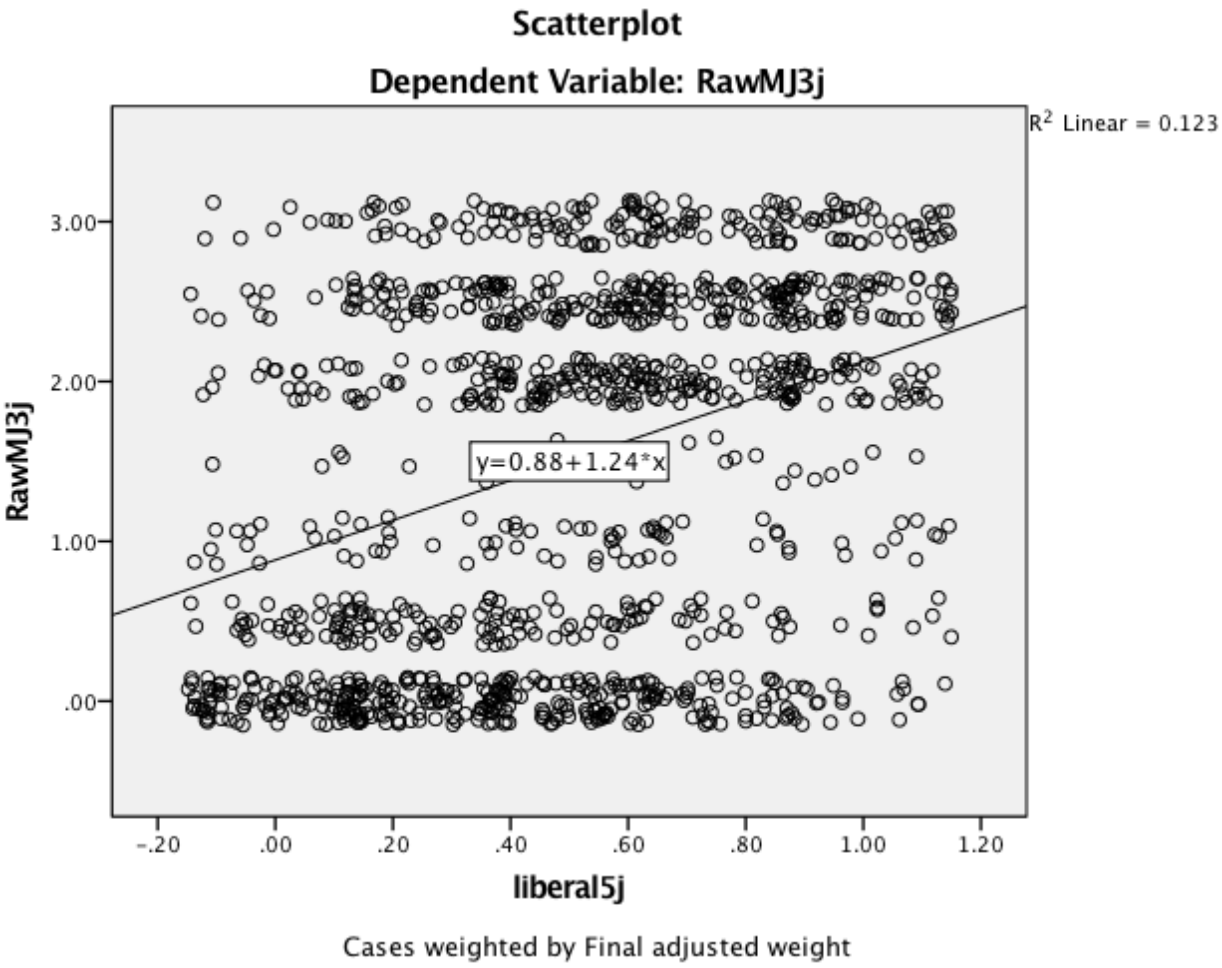
Improving the Vizualization through Jittering

```
COMPUTE liberal5j = liberal5 + RV.UNIFORM(-0.15, +0.15).
COMPUTE RawMJ3j = RawMJ3 + RV.UNIFORM(-0.15, +0.15).

regression variables=RawMJ3j liberal5j
  /dependent = RawMJ3j
  /method = enter
 /scatterplot = (RawMJ3j liberal5j).
```



Further improving Vizualization by adding a regression line



Comparing Original and Jittered Models for H2

	a	**b**	**r^2**
H2 Ideology model	.83	1.35	.13
H2 Jittered Ideology model	.88	1.24	.12

```
REGRESSION variables = IncomeShareTop10 civilliberties
 /DEPENDENT IncomeShareTop10
 /METHOD=ENTER
 /scatterplot = (IncomeShareTop10 civilliberties).
```

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.197[a]	.039	.031	6.60503

a. Predictors: (Constant), Freedom House score 2012, out of 60

b. Dependent Variable: Income share held by highest 10%. From GapMinder, most recent data 1997-2007

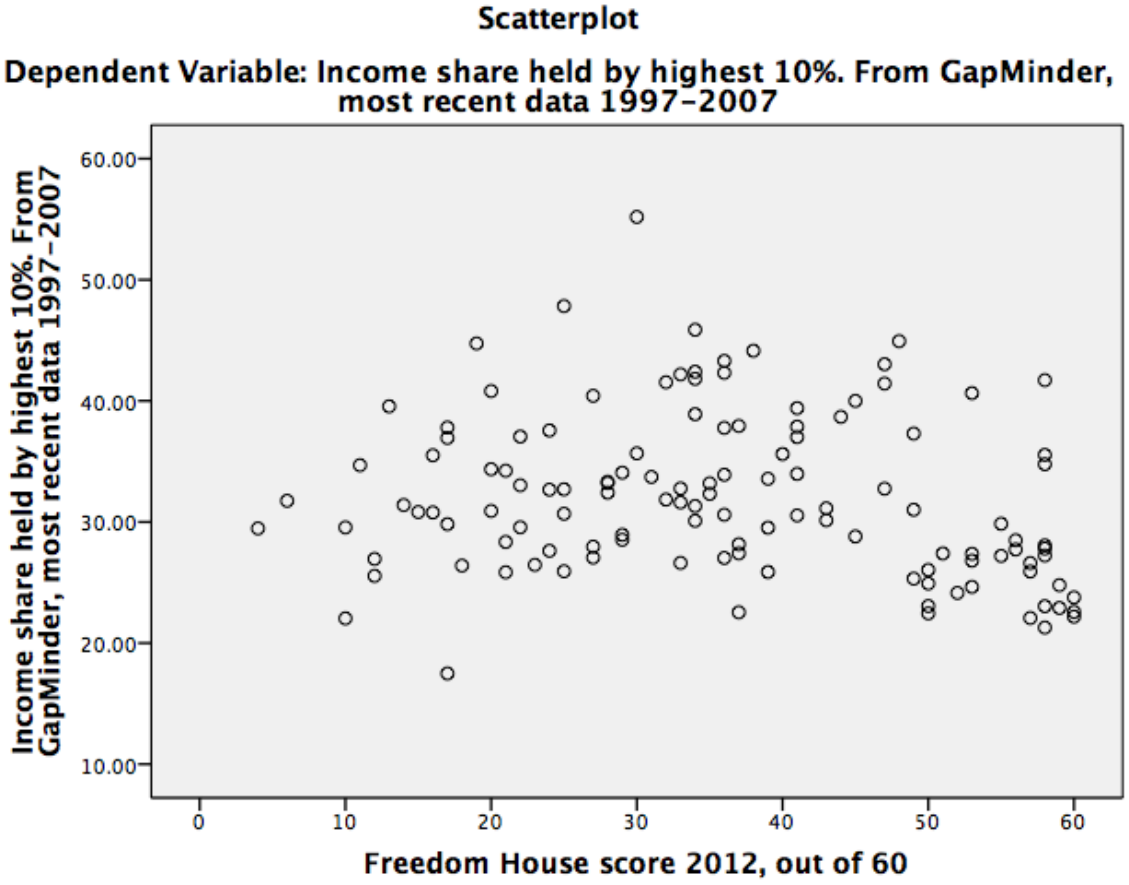
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	35.283	1.573		22.424	.000
	Freedom House score 2012, out of 60	-.089	.040	-.197	-2.210	.029

a. Dependent Variable: Income share held by highest 10%. From GapMinder, most recent data 1997-2007

$$y = a + bx$$

$$y = 35.3 + (-.089)$$


Scatterplot
Dependent Variable: Income share held by highest 10%. From GapMinder, most recent data 1997-2007
Income share held by highest 10%. From GapMinder, most recent data 1997-2007
60.00
50.00
40.00
30.00
20.00
10.00
0
10
20
30
40
50
60
Freedom House score 2012, out of 60

GRAPH

/SCATTERPLOT=Civilliberties WITH IncomeShareTop10 BY Country(name).