Emperical Methods for Marketing Research and Analytics Using R

Prof. Dr. Martin Wetzels
Maastricht University

Maastricht University

About Me

- Name: Prof. Dr. Martin Wetzels
  Professor in Marketing and Supply Chain Research

- Address: Maastricht University
  School of Business and Economics
  Department of Marketing and SCM
  P.O. Box 616, 6200 MD Maastricht
  Tongersestraat 53, F1.05
  T: +31 43 388 3250/3839
  F: +31 43 388 4918
  E: m.wetzels@maastrichtuniversity.nl
  W: https://www.maastrichtuniversity.nl/m.wetzels

Using R for Regression Analysis

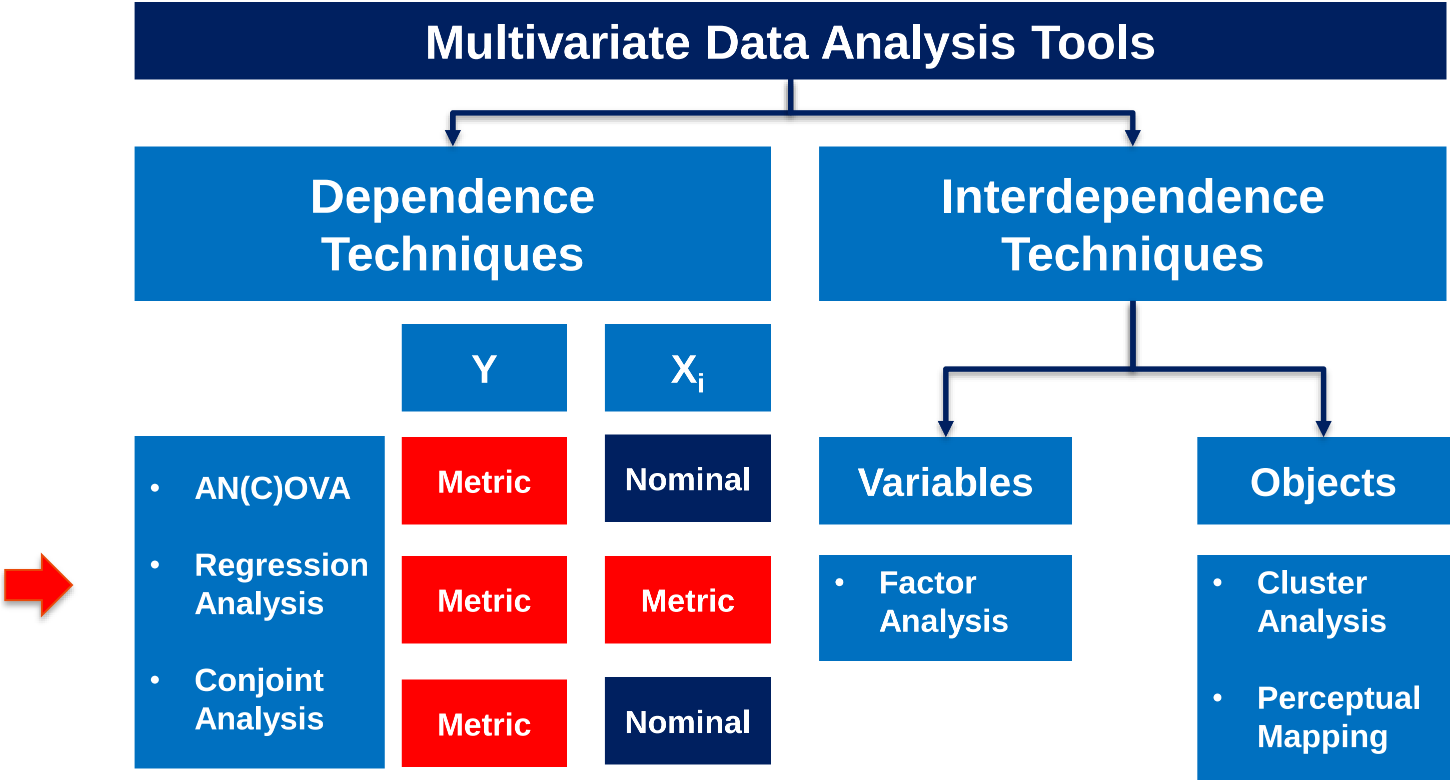
Prof. Dr. Martin Wetzels
Maastricht University

Course Outline

Session:	TOPIC:
DAY 1	
Session 1	INTRODUCING MULTIVARIATE ANALYSIS AND R
Session 2	USING R FOR BASIC ANALYSIS
DAY 2	
Session 3	USING R FOR AN(C)OVA
Session 4	USING R FOR REGRESSION ANALYSIS
DAY 3	
Session 5	USING R FOR SCALING AND FACTOR ANALYSIS
Session 6	USING R FOR SEM and PLS PATH MODELING

Positioning AN(C)OVA…

Hair et al. (2018); Malhotra (2010)



Correlation Analysis

Hair et al. (2018); Malhotra (2010)

- The **(Pearson) product moment correlation, ρ,** summarizes the strength of association between two **metric** (at least **interval**) variables [-1, 1].

- Assumptions
  - **Metric** (at least **interval**) variables
  - Independent, random sample
  - Linearity
  - (Bivariate) Normal distribution
  - Homoscedasticity
  - No causality!
  - Outliers!

Correlation Analysis

Hair et al. (2018); Malhotra (2010)

- Covariance **[-∞, + ∞]**

$$COV_{XY} = \sum_{i=1}^{n} \frac{(X_i - M_X) * (Y_i - M_Y)}{n-1}$$

- (Pearson Product-Moment) Correlation Coefficient **[-1, 1]**

$$r_{XY} = \frac{COV_{XY}}{S_X S_Y}$$

Correlation Analysis

Hair et al. (2018); Malhotra (2010)

- Hypotheses
  - H_0: correlation coefficient (ρ) = 0
  - H_1: correlation coefficient (ρ) ≠ 0

- Test statistic

$$t = r * \sqrt{\frac{n-2}{1-r^2}}$$

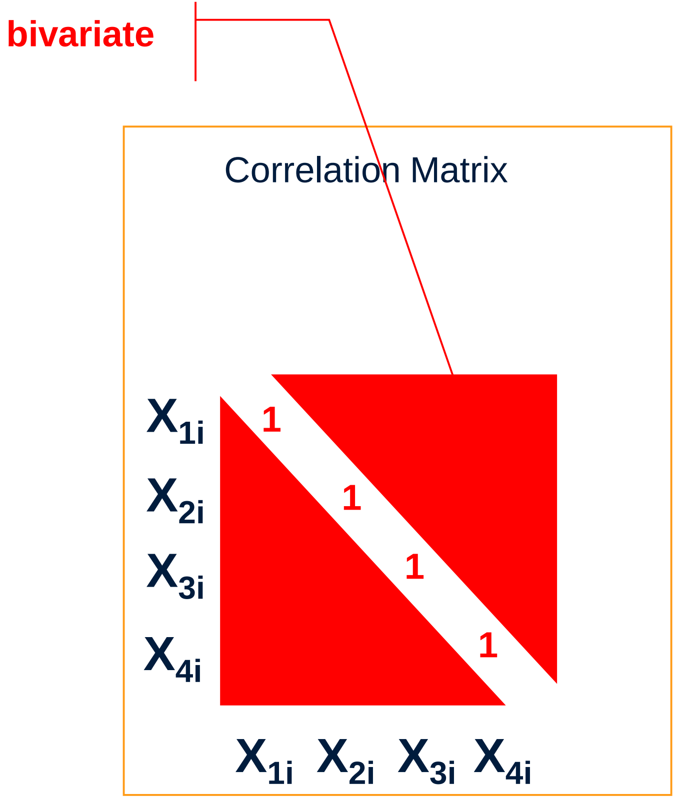
$$df = n - 2$$

- Effect Size (r) Effect Size (R^2)

Effect Size (r)	Effect Size (R^2)	
0.10	0.02	small
0.30	0.13	medium
0.50	0.26	large

Correlation Analysis

Hair et al. (2018); Malhotra (2010)



SPSS Application

TAM Data

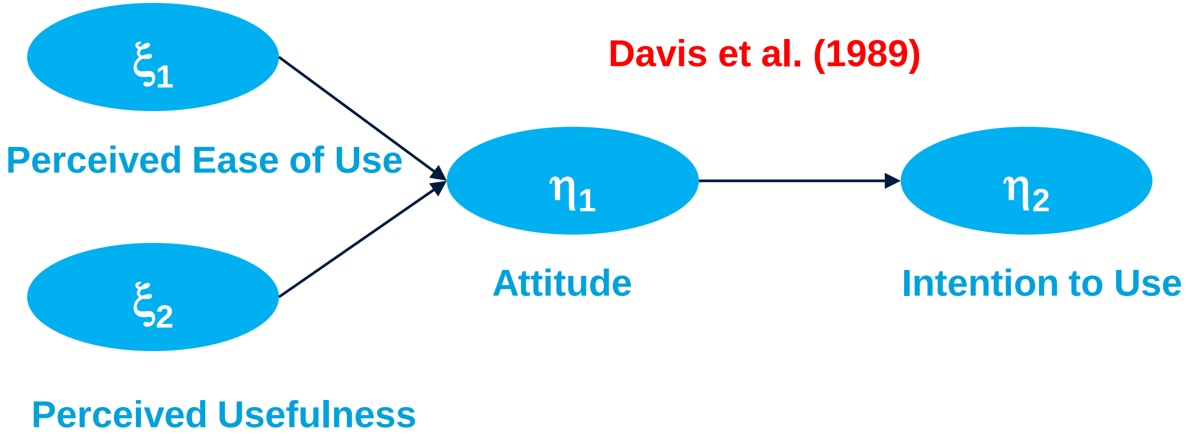
- Data Set *Data-01.sav* (n=439)

- Constructs and Items
  - Perceived Ease of Use (PEU, Q1-Q3)
  - Perceived Usefulness (PU, Q4-Q6)
  - Attitude (ATT, Q7-Q9)
  - Intention (INT, Q10-Q12)

- Reliability Analysis (Coefficient α)
  - PEU 0.93
  - PU 0.92
  - ATT 0.92
  - INT 0.87

SPSS Application

TAM Data



Construct	Items
PEU	Q1, Q2, Q3
PU	Q4, Q5, Q6
ATT	Q7, Q8, Q9
INT	Q10, Q11, Q12

Davis et al. (1989)

MANAGEMENT SCIENCE
Vol. 35, No. 8, August 1989
Printed in U.S.A.

USER ACCEPTANCE OF COMPUTER TECHNOLOGY: A COMPARISON OF TWO THEORETICAL MODELS*

FRED D. DAVIS, RICHARD P. BAGOZZI AND PAUL R. WARSHAW

School of Business Administration, University of Michigan, Ann Arbor, Michigan 48109-1234
School of Business Administration, University of Michigan, Ann Arbor, Michigan 48109-1234
School of Business Administration, California Polytechnic State University, San Luis Obispa, California 93407

Computer systems cannot improve organizational performance if they aren't used. Unfortunately, resistance to end-user systems by managers and professionals is a widespread problem. To better predict, explain, and increase user acceptance, we need to better understand why people accept or reject computers. This research addresses the ability to predict peoples' computer acceptance from a measure of their intentions, and the ability to explain their intentions in terms of their attitudes, subjective norms, perceived usefulness, perceived ease of use, and related variables. In a longitudinal study of 107 users, intentions to use a specific system, measured after a one-hour introduction to the system, were correlated 0.35 with system use 14 weeks later. The intention-usage correlation was 0.63 at the end of this time period. Perceived usefulness strongly influenced peoples' intentions, explaining more than half of the variance in intentions at the end of 14 weeks. Perceived ease of use had a small but significant effect on intentions as well, although this effect subsided over time. Attitudes only partially mediated the effects of these beliefs on intentions. Subjective norms had no effect on intentions. These results suggest the possibility of simple but powerful models of the determinants of user acceptance, with practical value for evaluating systems and guiding managerial interventions aimed at reducing the problem of underutilized computer technology.
(INFORMATION TECHNOLOGY; USER ACCEPTANCE; INTENTION MODELS)

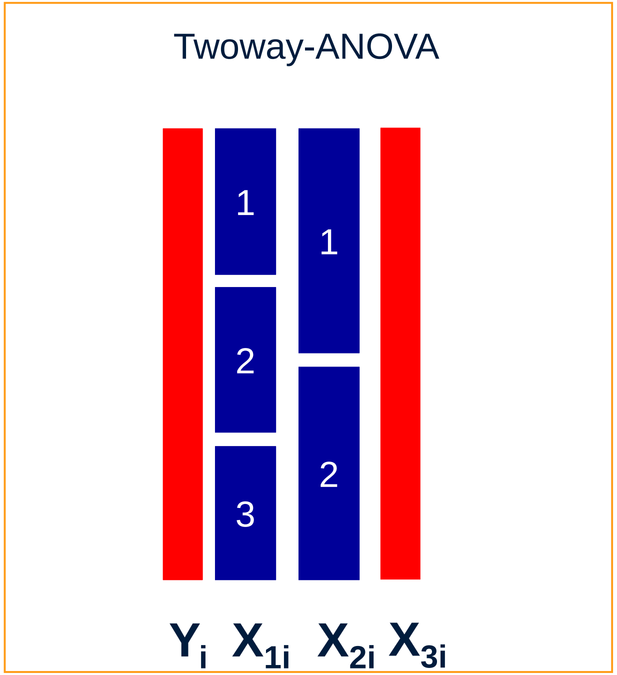
Conducting Regression Analysis

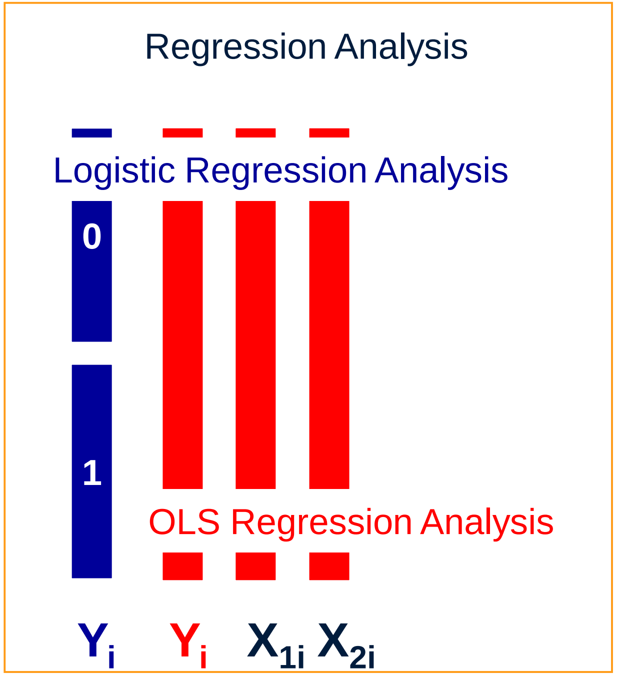
Hair et al. (2018); Malhotra (2010)

1. Regression Model
2. Regression Assumptions
3. Estimation Procedure
4. Significance Testing/Model Fit
5. Assessing Assumptions
6. Advanced Topics

1. Regression Model

GLM: General Linear Model





2. Regression Assumptions

Hair et al. (2018); Malhotra (2010)

- **Assumptions**
  - Independent and dependent variables are **metric** (at least **interval**)
  - For each combination of X_i values there is a normally distributed subpopulation of Y_i values (with equal variances)
  - All Y_i values are independent
  - The ERROR term has a mean of 0 (with equal variances)
  - There does not exist any exact linear relationship between the X_i's (assumption of **no multicollinearity**)
  - Model is correctly specified (no omitted an/or irrelevant variables)

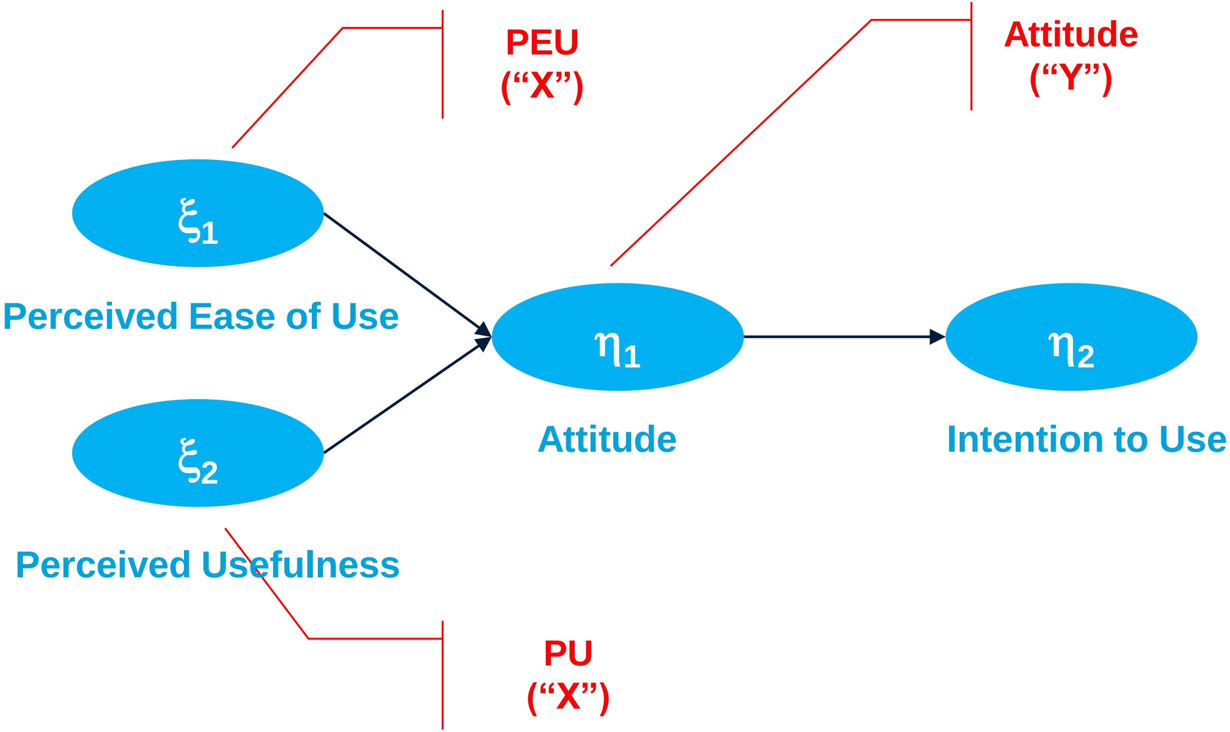
3. Estimation Procedure

Hair et al. (2018); Malhotra (2010)

- **Model**
  - $Y_i = a + b_1 * X_{1i} + b_2 * X_{2i} + ERROR_i$
- **Estimation of Parameters**
  - Ordinary Least Squares Estimation
  - Minimize Σ (Predicted Y-Observed Y)^2
- **Fit Measures**
  - R^2 (Adjusted R^2)

Regression Analysis

TAM Data



OLS Principle

TAM Data

PEU	PU	ATT	INT	PRE_1	RES_1
3.00	4.33	4.33	4.00	3.70533	.62800
3.00	2.33	2.00	4.33	3.34763	-1.34763
2.00	3.33	2.67	4.67	3.17929	-.51262
6.00	4.67	7.00	5.33	4.80653	2.19347
3.33	2.33	2.67	3.33	3.46336	-.79670
1.00	3.33	2.00	3.00	2.83210	-.83210
4.33	3.33	4.00	4.67	3.98941	.01059
3.00	3.33	3.33	2.67	3.52648	-.19315
4.67	2.33	4.67	4.67	3.92629	.74038
4.67	3.00	3.67	4.33	4.04552	-.37885
5.33	3.00	3.33	5.00	4.27698	-.94365
5.33	4.33	4.67	4.00	4.51545	.15122
4.00	3.67	5.00	3.67	3.93329	1.06671
4.33	4.67	5.33	4.00	4.22787	1.10546
4.33	4.00	4.67	4.67	4.10864	.55803

X X Y PRED RES

PRED_Attitude = 1.889 + 0.347*PEU + 0.179*PU

4. Significance Testing/Model Fit

Hair et al. (2018); Malhotra (2010)

- **Overall Significance of a Multiple Regression Model (F test)**
  H_0: $R^2 = 0$
  H_1: $R^2 > 0$

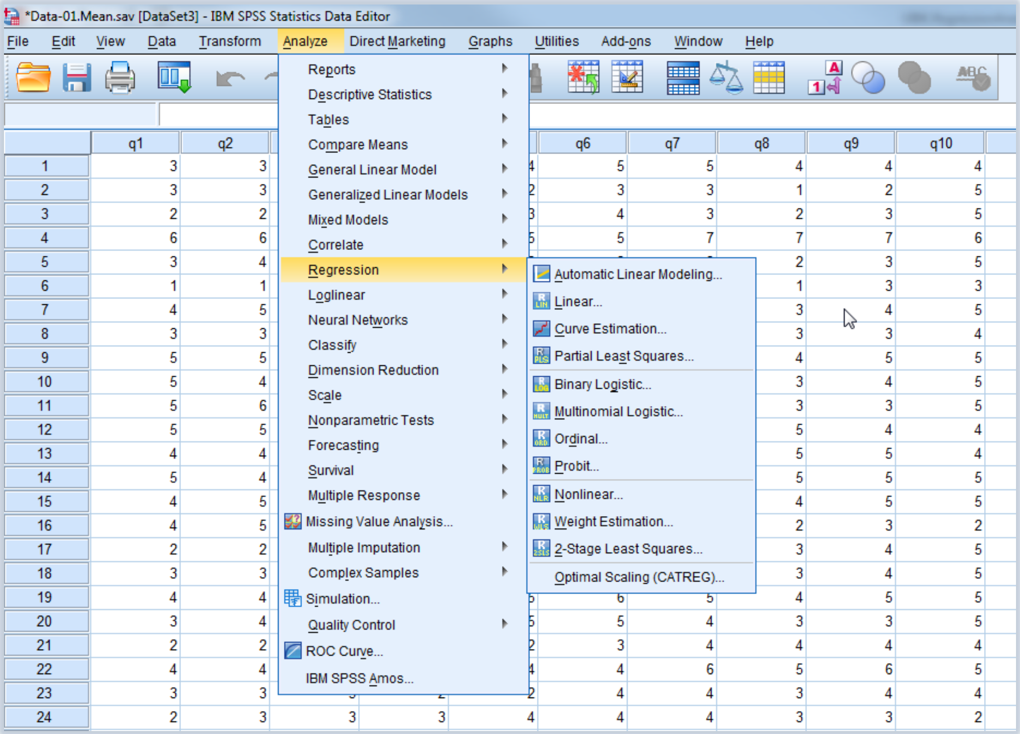
  $DF_1 = 1$, $DF_2 = n-k-1$

- **Significance of the Individual Regression Coefficients (t test)**
  H_0: $b_i = 0$
  H_1: $b_i \neq 0$

  $DF = n-k-1$

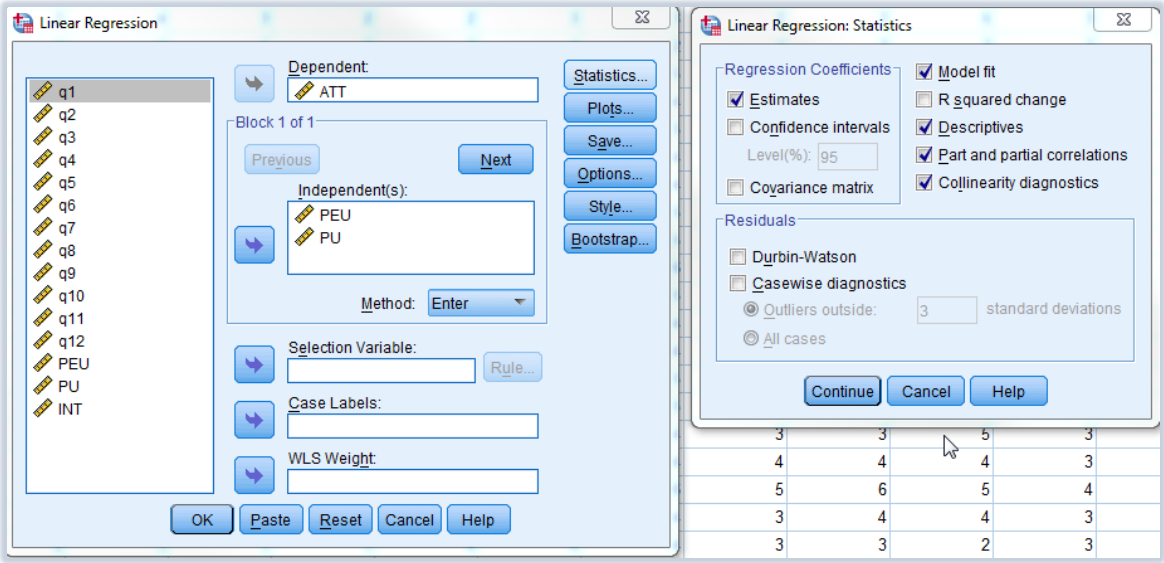
Regression Analysis

TAM Data



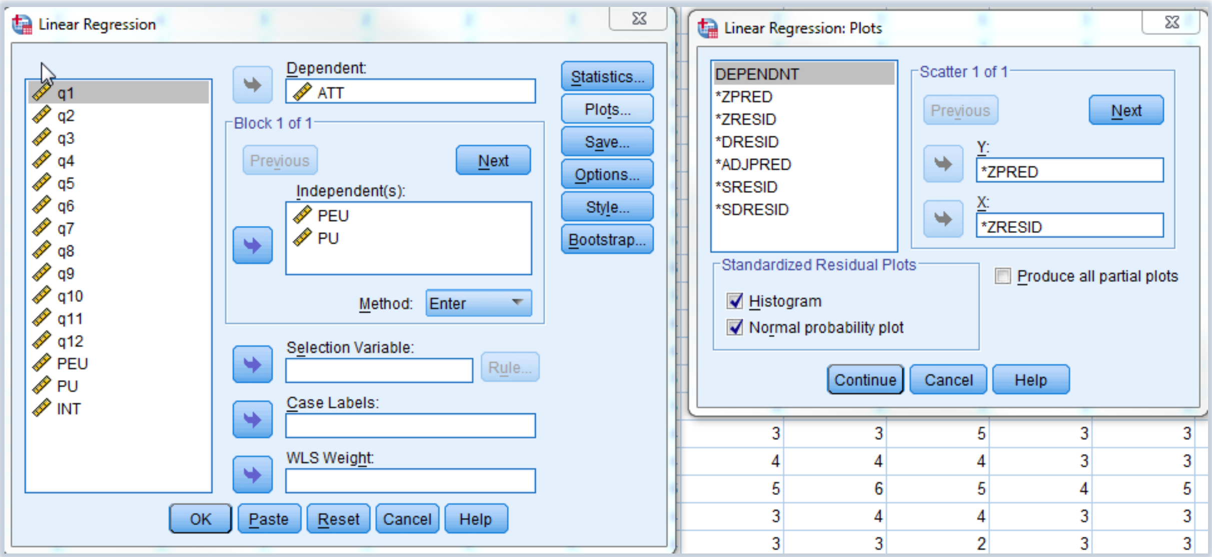
Regression Analysis

TAM Data



Regression Analysis

TAM Data



Regression Analysis

TAM Data

Model Fit

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.455[a]	.207	.204	.79880

a. Predictors: (Constant), PU, PEU

b. Dependent Variable: ATT

$R^2=72.751/350.958=0.207$

F test
(p value)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	72.751	2	36.376	57.007	.000[b]
	Residual	278.207	436	.638		
	Total	350.958	438			

Variance explained
Error variance
Total variance

a. Dependent Variable: ATT

b. Predictors: (Constant), PU, PEU

Regression Analysis

TAM Data

Coefficients
B, Beta

t test
(p value)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	1.889	.185		10.207	.000					
	PEU	.347	.045	.352	7.631	.000	.422	.343	.325	.855	1.170
	PU	.179	.045	.185	4.001	.000	.318	.188	.171	.855	1.170

a. Dependent Variable: ATT

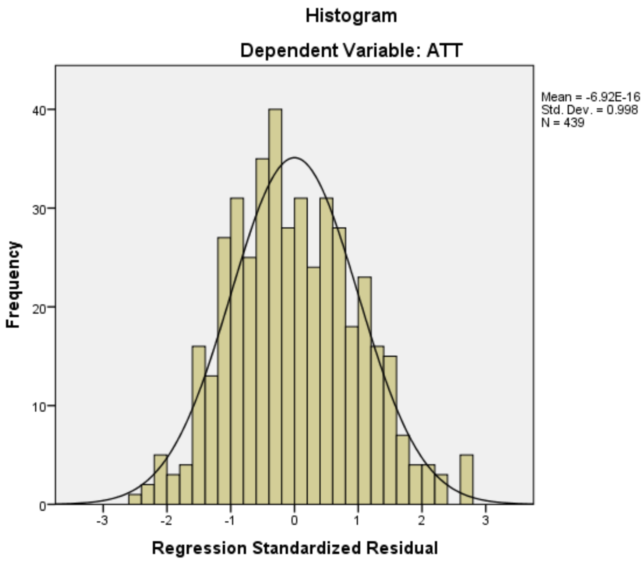
Correlations

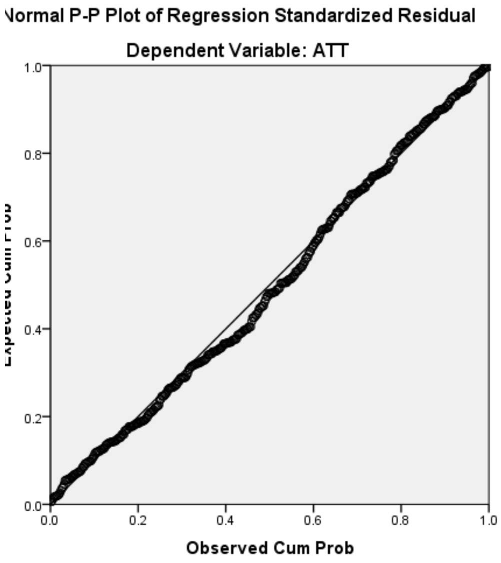
TOL= 1-R^2(All other X)
VIF = 1/TOL
VIF> 5, Multicollinearity!

5. Assessing Assumptions

Hair et al. (2018); Malhotra (2010)

Normality Assumption

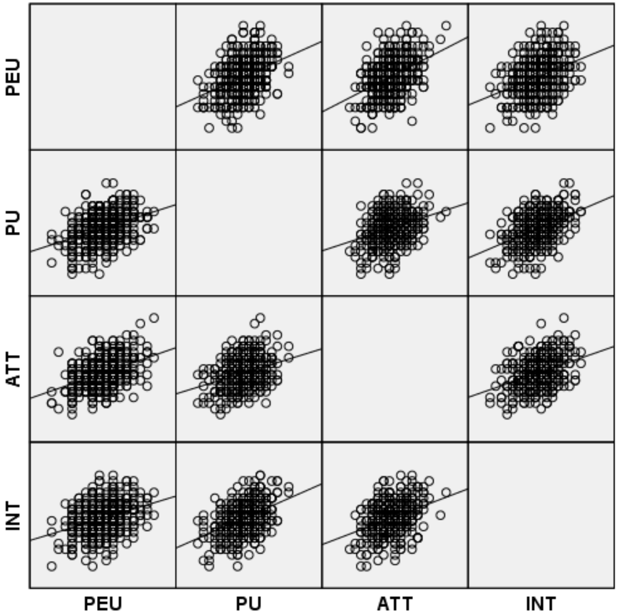




5. Assessing Assumptions

Hair et al. (2018); Malhotra (2010)

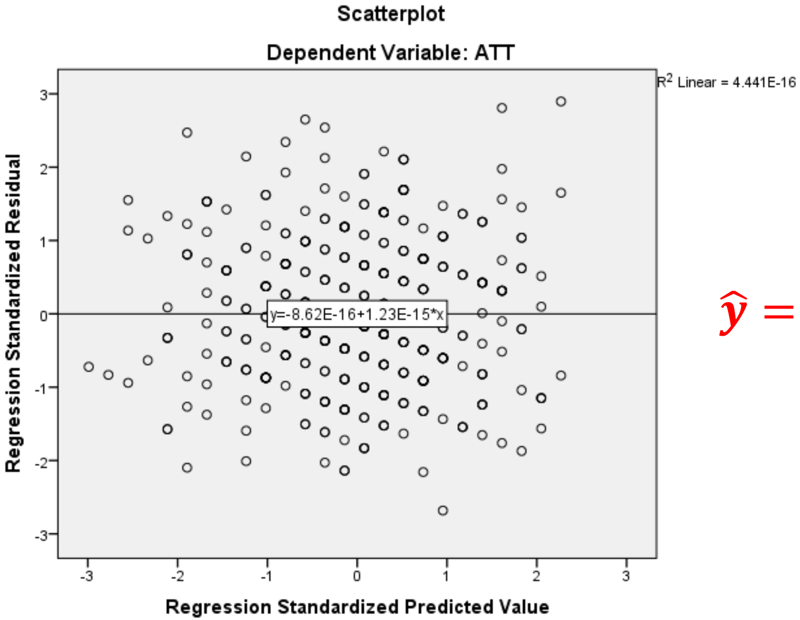
Linearity



5. Assessing Assumptions

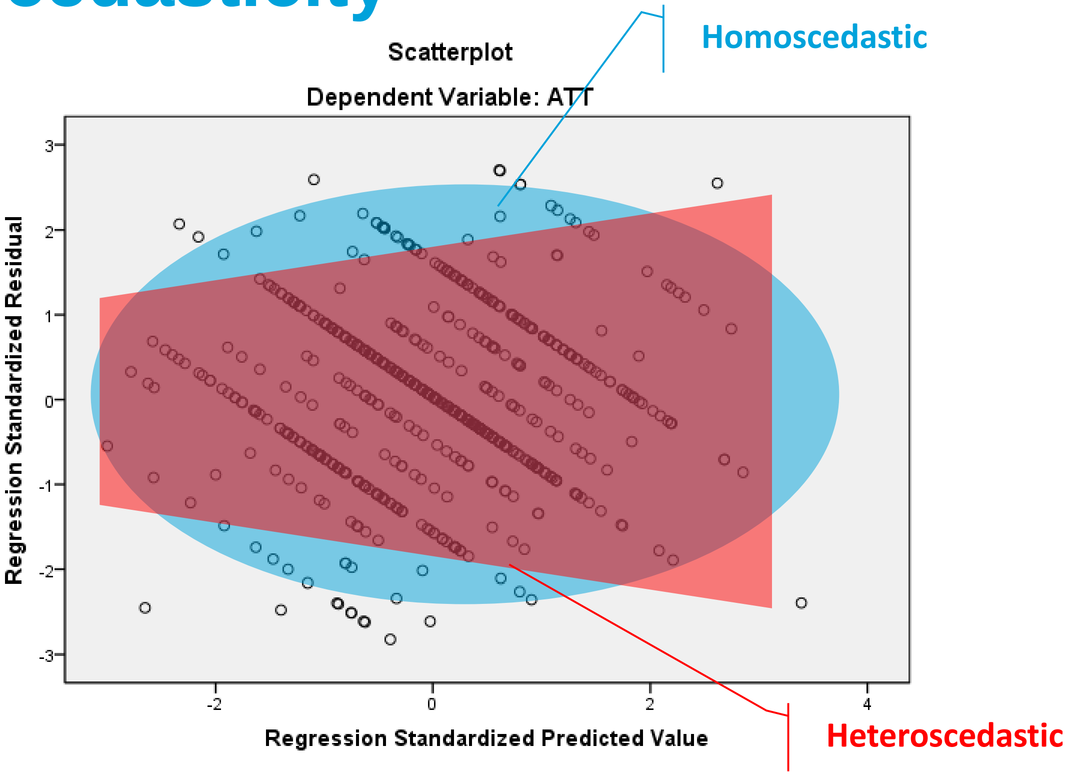
Hair et al. (2018); Malhotra (2010)

Outlier Detection: Residual Analysis



$$\hat{y} = 0$$

Regression Revisited… Homoscedasticity



Correlation Analysis

Editor

```r
# Correlation

cor(Data.01.PPT)

cor(Data.01.PPT[,13:16])

library(Hmisc)

rcorr(as.matrix(Data.01.PPT[,13:16]))

library(corrgram)

corrgram(Data.01.PPT)
```

Correlation Analysis

Console

	q1	q2	q3	q4	q5	q6
q1	1.0000000	0.8097975	0.8138406	0.2930891	0.3428762	0.3247076
q2	0.8097975	1.0000000	0.8166044	0.3060620	0.3243314	0.3472221
q3	0.8138406	0.8166044	1.0000000	0.3239042	0.3529852	0.3680337
q4	0.2930891	0.3060620	0.3239042	1.0000000	0.8009082	0.7970698
q5	0.3428762	0.3243314	0.3529852	0.8009082	1.0000000	0.8116305
q6	0.3247076	0.3472221	0.3680337	0.7970698	0.8116305	1.0000000
q7	0.3972089	0.3690113	0.3736317	0.2921024	0.2979363	0.2478376
q8	0.3653399	0.3518016	0.3532527	0.2462745	0.2597337	0.2426744
q9	0.3904451	0.3464314	0.3557382	0.3006136	0.3236039	0.2686053
q10	0.2589751	0.2588477	0.2689082	0.3145539	0.3231403	0.3321825
q11	0.3152825	0.3161738	0.3338670	0.4183882	0.3839118	0.3861757
q12	0.3138578	0.3141935	0.3201764	0.4015143	0.3765244	0.3707145
PEU	0.9328149	0.9346877	0.9396350	0.3291739	0.3636331	0.3709495
PU	0.3439444	0.3498104	0.3739597	0.9272086	0.9356141	0.9334255
ATT	0.4139689	0.3830650	0.3886151	0.3014059	0.3165713	0.2725093
INT	0.3257088	0.3260625	0.3384005	0.4142796	0.3981133	0.4008660

	q7	q8	q9	q10	q11	q12
q1	0.3972089	0.3653399	0.3904451	0.2589751	0.3152825	0.3138578
q2	0.3690113	0.3518016	0.3464314	0.2588477	0.3161738	0.3141935
q3	0.3736317	0.3532527	0.3557382	0.2689082	0.3338670	0.3201764
q4	0.2921024	0.2462745	0.3006136	0.3145539	0.4183882	0.4015143
q5	0.2979363	0.2597337	0.3236039	0.3231403	0.3839118	0.3765244
q6	0.2478376	0.2426744	0.2686053	0.3321825	0.3861757	0.3707145
q7	1.0000000	0.8026353	0.8059340	0.2570709	0.3164261	0.3369713
q8	0.8026353	1.0000000	0.7732401	0.2332786	0.2897454	0.3177918
q9	0.8059340	0.7732401	1.0000000	0.2726496	0.3677583	0.3916562
q10	0.2570709	0.2332786	0.2726496	1.0000000	0.6810155	0.6744540
q11	0.3164261	0.2897454	0.3677583	0.6810155	1.0000000	0.7859497
q12	0.3369713	0.3177918	0.3916562	0.6744540	0.7859497	1.0000000
PEU	0.4058197	0.3811711	0.3889120	0.2803795	0.3441139	0.3378509
PU	0.2995601	0.2678096	0.3193535	0.3469233	0.4246864	0.4105496
ATT	0.9369643	0.9230840	0.9262602	0.2740098	0.3498089	0.3757832
INT	0.3331362	0.3072647	0.3756716	0.8965442	0.8996885	0.8989238

	PEU	PU	ATT	INT
q1	0.9328149	0.3439444	0.4139689	0.3257088
q2	0.9346877	0.3498104	0.3830650	0.3260625
q3	0.9396350	0.3739597	0.3886151	0.3384005
q4	0.3291739	0.9272086	0.3014059	0.4142796
q5	0.3636331	0.9356141	0.3165713	0.3981133
q6	0.3709495	0.9334255	0.2725093	0.4008660
q7	0.4058197	0.2995601	0.9369643	0.3331362
q8	0.3811711	0.2678096	0.9230840	0.3072647
q9	0.3889120	0.3193535	0.9262602	0.3756716
q10	0.2803795	0.3469233	0.2740098	0.8965442
q11	0.3441139	0.4246864	0.3498089	0.8996885
q12	0.3378509	0.4105496	0.3757832	0.8989238
PEU	1.0000000	0.3807257	0.4221206	0.3528830
PU	0.3807257	1.0000000	0.3184720	0.4337164
ATT	0.4221206	0.3184720	1.0000000	0.3649072
INT	0.3528830	0.4337164	0.3649072	1.0000000

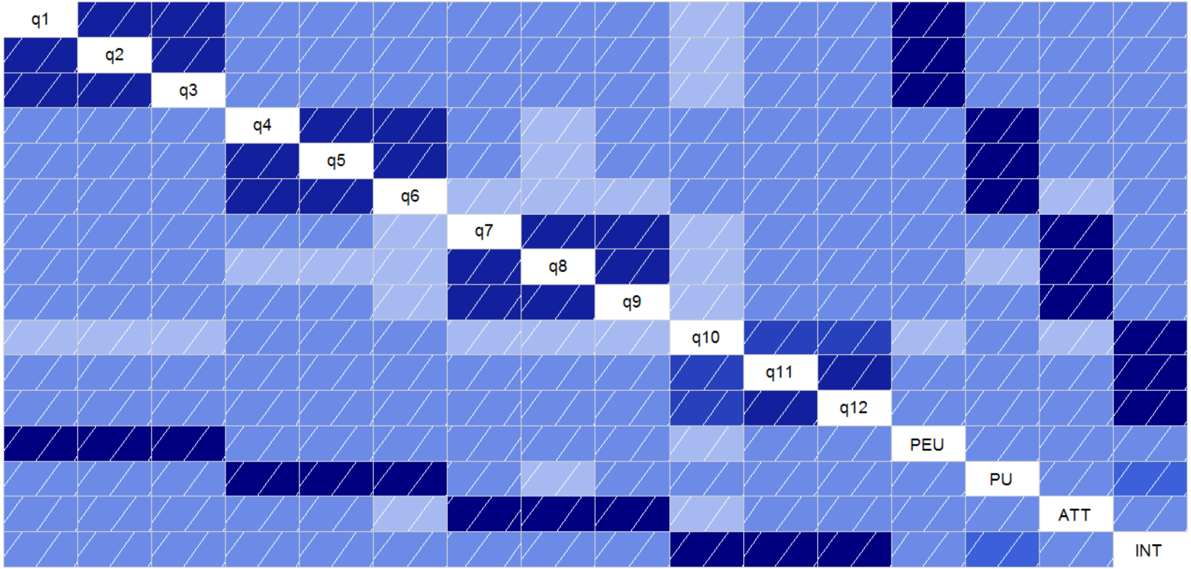
Correlation Analysis

Console

```
> cor(Data.01.PPT[,13:16])
          PEU        PU       ATT       INT
PEU 1.0000000 0.3807257 0.4221206 0.3528830
PU  0.3807257 1.0000000 0.3184720 0.4337164
ATT 0.4221206 0.3184720 1.0000000 0.3649072
INT 0.3528830 0.4337164 0.3649072 1.0000000
>
> library(Hmisc)
>
> rcorr(as.matrix(Data.01.PPT[,13:16]))
     PEU   PU  ATT  INT
PEU 1.00 0.38 0.42 0.35
PU  0.38 1.00 0.32 0.43
ATT 0.42 0.32 1.00 0.36
INT 0.35 0.43 0.36 1.00

n= 439


P
    PEU PU ATT INT
PEU      0  0   0
PU   0      0   0
ATT  0   0      0
INT  0   0  0
```

Correlation Analysis

Graphics



Regression Analysis

Editor

```r
# Regression Analsysis

library(car)

R00<-lm(ATT ~ PEU + PU, data=Data.01.PPT)

summary(R00)

vif(R00)

plot(R00)

```

Regression Analysis

Console

```
Call:
lm(formula = ATT ~ PEU + PU, data = Data.01.PPT)

Residuals:
     Min       1Q   Median       3Q      Max
-2.05603 -0.56699 -0.03866  0.54699  2.19347

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.8887     0.1850  10.207  < 2e-16 ***
PEU           0.3472     0.0455   7.631 1.48e-13 ***
PU            0.1789     0.0447   4.001 7.40e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7988 on 436 degrees of freedom
Multiple R-squared:  0.2073, Adjusted R-squared:  0.2037
F-statistic: 57.01 on 2 and 436 DF,  p-value: < 2.2e-16

>
> vif(R00)
     PEU       PU
1.169525 1.169525
```

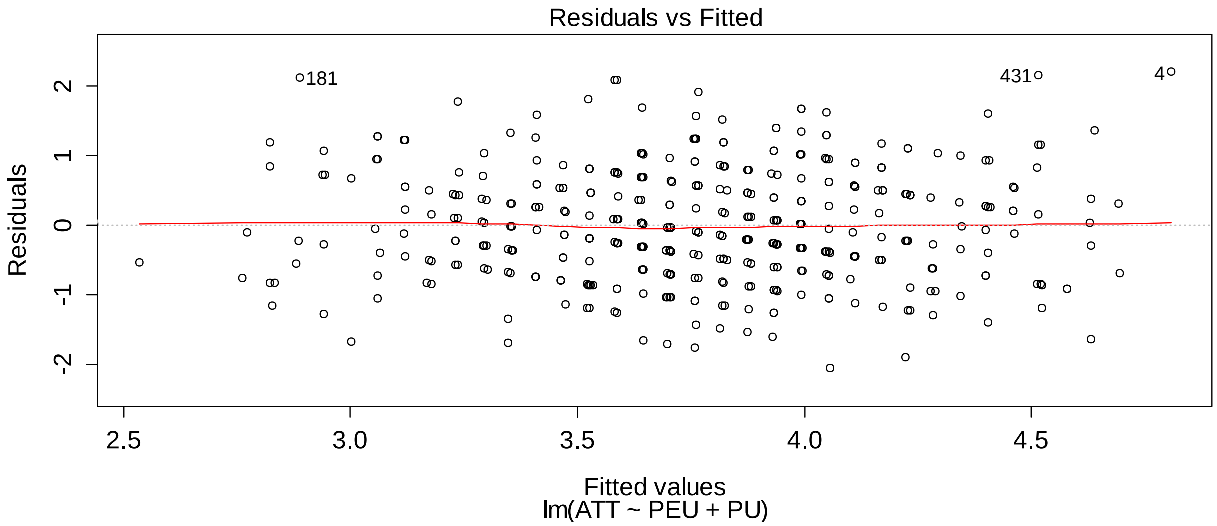
Regression Analysis

Console

```
> # Standardized Coefficients
>
> library(QuantPsyc)
> lm.beta(R00)
      PEU        PU
0.3518752 0.1845040
>
> # Alternatively
>
> R00$coefficients[["PEU"]]*(sd(Data.01.PPT$PEU)/sd(Data.01.PPT$ATT))
[1] 0.3518752
>
> R00$coefficients[["PU"]]*(sd(Data.01.PPT$PU)/sd(Data.01.PPT$ATT))
[1] 0.184504
```

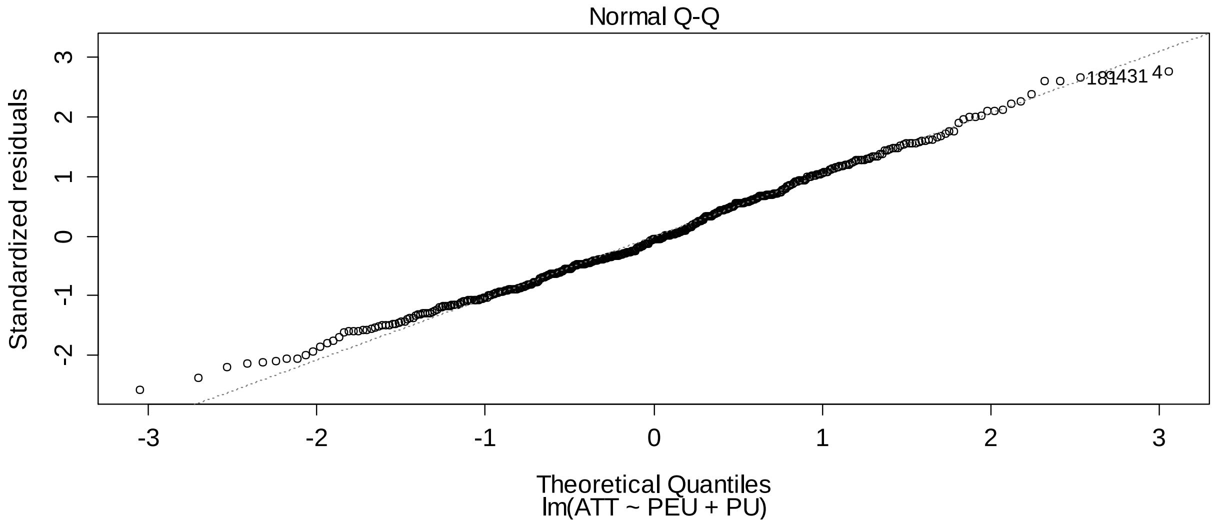
Regression Analysis

Graphics



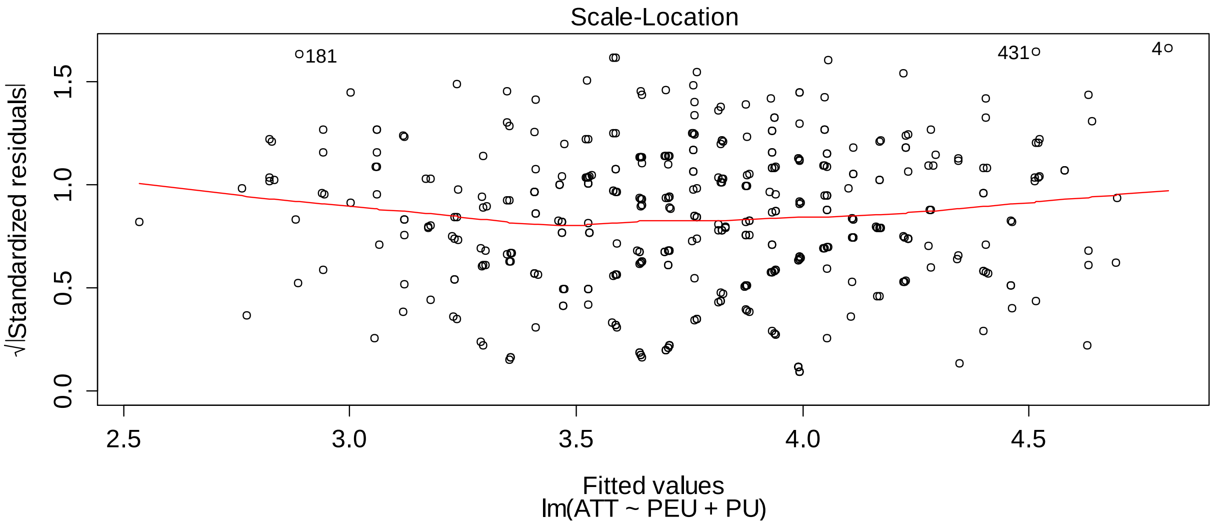
Regression Analysis

Graphics



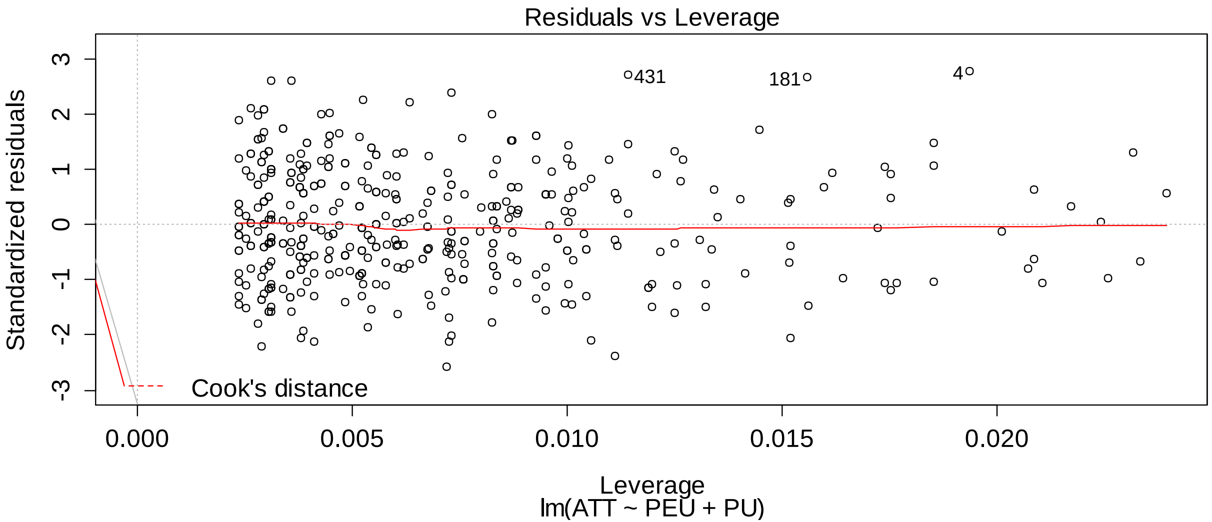
Regression Analysis

Graphics



Regression Analysis

Graphics



6. Advanced Topics: Nominal Variables

Malhotra (2010); Pallant (2016)

- Can nominal data be used in regression analysis?
  - Yes, how?
  - By using dummy variables (0,1)

R package: dummies

Two categories

	Dummy
A	1
B	0

Two categories

	Effect
A	1
B	-1

Three Categories

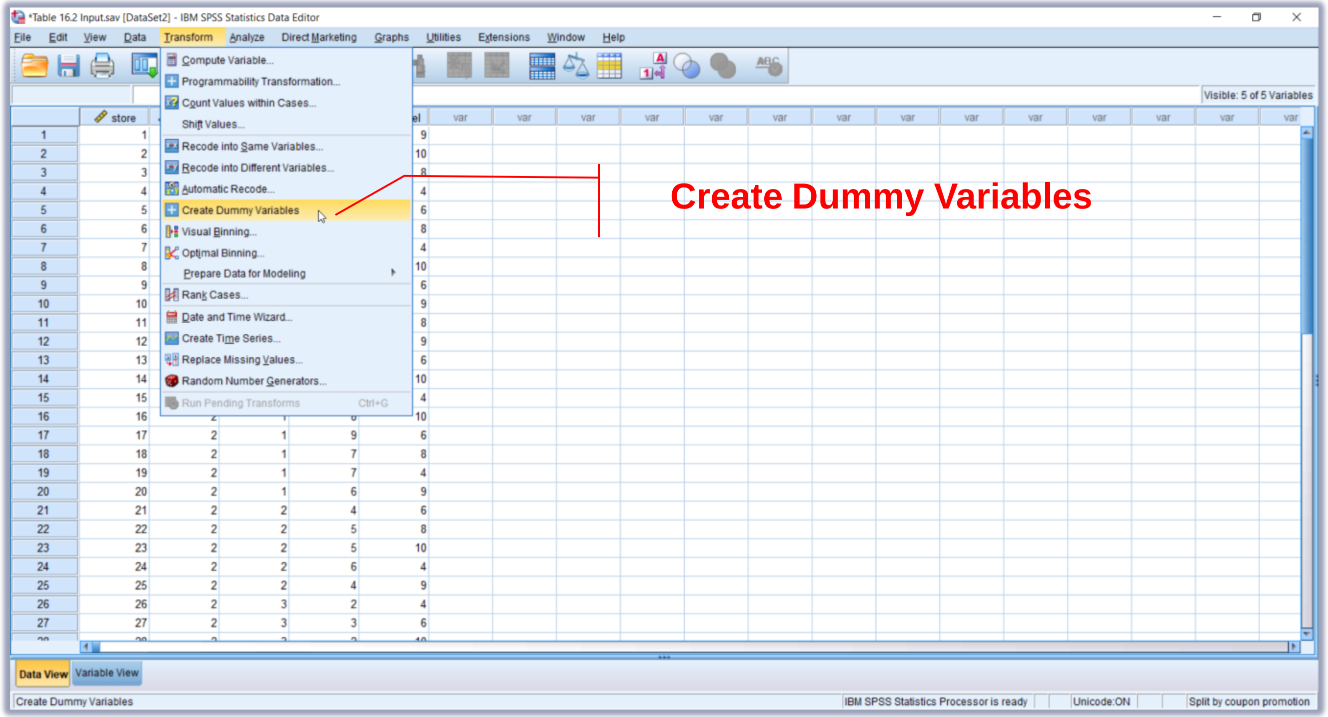
	Dummy1	Dummy2
A	1	0
B	0	1
C	0	0

Three Categories

	Effect1	Effect2
A	1	0
B	0	1
C	-1	-1

6. Advanced Topics: Nominal Variables

Malhotra (2010); Pallant (2016)



6. Advanced Topics: Nominal Variables

Malhotra (2010); Pallant (2016)

Department Store Data (Malhotra, 2010, Table 16.2)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	106.067	2	53.033	17.944	.000[b]
	Residual	79.800	27	2.956		
	Total	185.867	29			

a. Dependent Variable: sales Sales

b. Predictors: (Constant), D2, D1

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.700	.544		6.806	.000
	D1	4.600	8.30 .769	.871	5.983	.000
	D2	2.500	6.20 .769	.473	3.252	.003

a. Dependent Variable: sales Sales

Promotion	D1	D2
1	1	0
2	0	1
3	0	0

6. Advanced Topics: Nominal Variables

Editor

```r
# Dummy variables

A.01<-lm(sales ~ factor(promotion), data=Table16.2)
summary(A.01)

summary(lm(lm(sales ~ promotion, data=Table16.2)))

Table16.2$prom.1<-factor(Table16.2$promotion)
Table16.2$prom.2<-factor(Table16.2$promotion)

contrasts(Table16.2$prom.1)<-contr.treatment(3, base=3)
contrasts(Table16.2$prom.2)<-contr.sum(3, base=3)

summary(lm(lm(sales ~ prom.1, data=Table16.2)))

summary(lm(lm(sales ~ prom.2, data=Table16.2)))
```

6. Advanced Topics: Nominal Variables

Console

```
> summary(lm(sales ~ prom.1, data=Table16.2))

Call:
lm(formula = sales ~ prom.1, data = Table16.2)

Residuals:
   Min     1Q Median     3Q    Max 
  -2.7   -1.3   -0.2    1.3    3.3 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   3.7000     0.5437   6.806 2.61e-07 ***
prom.11       4.6000     0.7688   5.983 2.21e-06 ***
prom.12       2.5000     0.7688   3.252  0.00307 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.719 on 27 degrees of freedom
Multiple R-squared:  0.5707,	Adjusted R-squared:  0.5389 
F-statistic: 17.94 on 2 and 27 DF,  p-value: 1.104e-05
```

6. Advanced Topics: Nominal Variables

Console

```
> summary(lm(sales ~ prom.2, data=Table16.2))

Call:
lm(formula = sales ~ prom.2, data = Table16.2)

Residuals:
   Min     1Q Median     3Q    Max 
  -2.7   -1.3   -0.2    1.3    3.3 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   6.0667     0.3139  19.328  < 2e-16 ***
prom.21       2.2333     0.4439   5.031  2.8e-05 ***
prom.22       0.1333     0.4439   0.300    0.766    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.719 on 27 degrees of freedom
Multiple R-squared:  0.5707,	Adjusted R-squared:  0.5389 
F-statistic: 17.94 on 2 and 27 DF,  p-value: 1.104e-05
```

6. Advanced Topics: Equality of Coefficients

Testing Equality of Coefficients

General Approach

(1) $Y_i = a + b_1*X_{1i} + b_2*X_{2i} + b_3*X_{3i}$ + ERROR (unconstrained model)

Assume H_0: $b_1 = b_2$

(2) $Y_i = a + b_1*X_{1i} + b_1*X_{2i} + b_3*X_{3i}$ + ERROR (constrained model) ⇔

(3) $Y_i = a + b_1*(X_{1i} + X_{2i}) + b_3*X_{3i}$ + ERROR (constrained model) ⇔

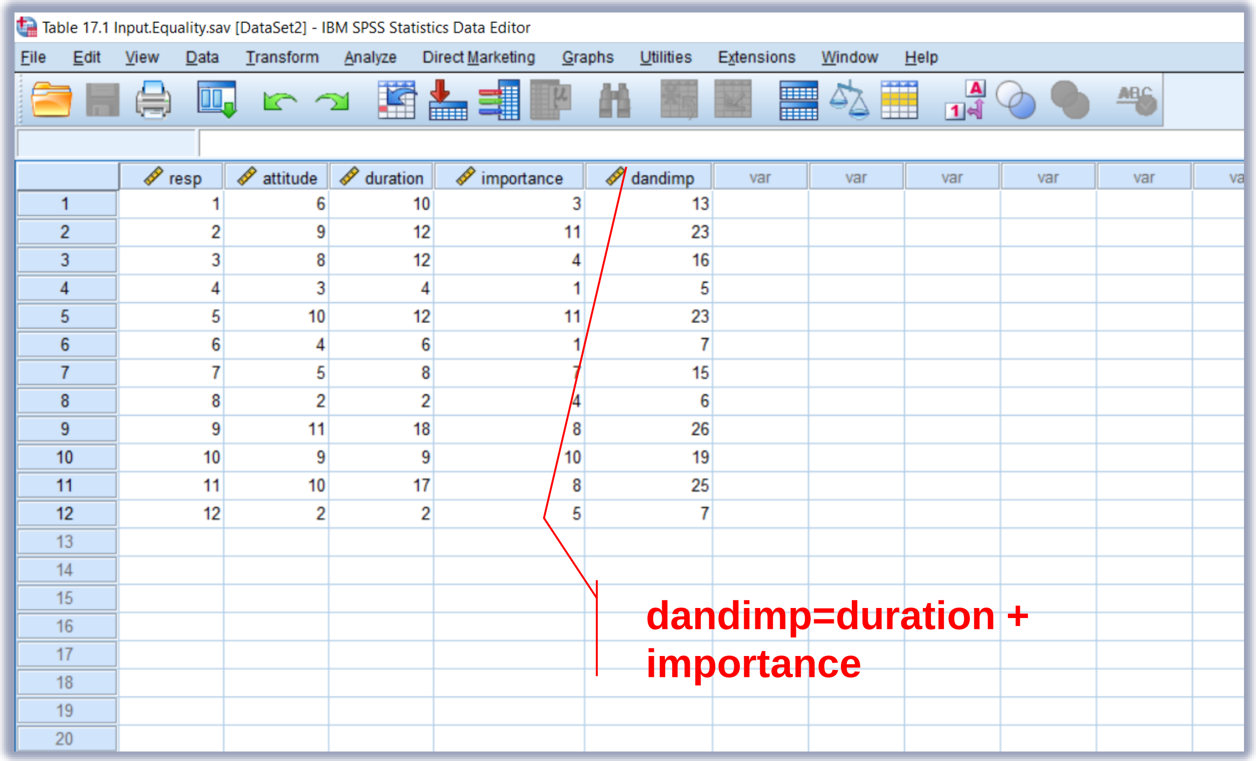
Compare (1) and (3) using:

$$F(j,n-k-1) = \frac{(R^2_U - R^2_C)*(n-k-1)}{(1-R^2_U)*j}$$

n=number of observations
k=number of variables in unconstrained model
j=number of constraints

6. Advanced Topics: Equality of Coefficients

Hayes (2013); Hayes and Matthes (2009)



	resp	attitude	duration	importance	dandimp
1	1	6	10	3	13
2	2	9	12	11	23
3	3	8	12	4	16
4	4	3	4	1	5
5	5	10	12	11	23
6	6	4	6	1	7
7	7	5	8	7	15
8	8	2	2	4	6
9	9	11	18	8	26
10	10	9	9	10	19
11	11	10	17	8	25
12	12	2	2	5	7

dandimp=duration + importance

6. Advanced Topics: Equality of Coefficients

UNconstrained model

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.972[a]	.945	.933	.860

a. Predictors: (Constant), Importance, Duration

$$F(j,n-k-1) = \frac{(R^2_U - R^2_C)*(n-k-1)}{(1-R^2_U)*j}$$

Constrained model ($b_1=b_2$)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.965[a]	.931	.924	.912

a. Predictors: (Constant), dandimp

$$F(1,9) = \frac{0.945-0.931*(12-2-1)}{(1-0.945)*1}$$

$F(1,9) = 2.29$ ($p=0.16$)

Coefficients[a]

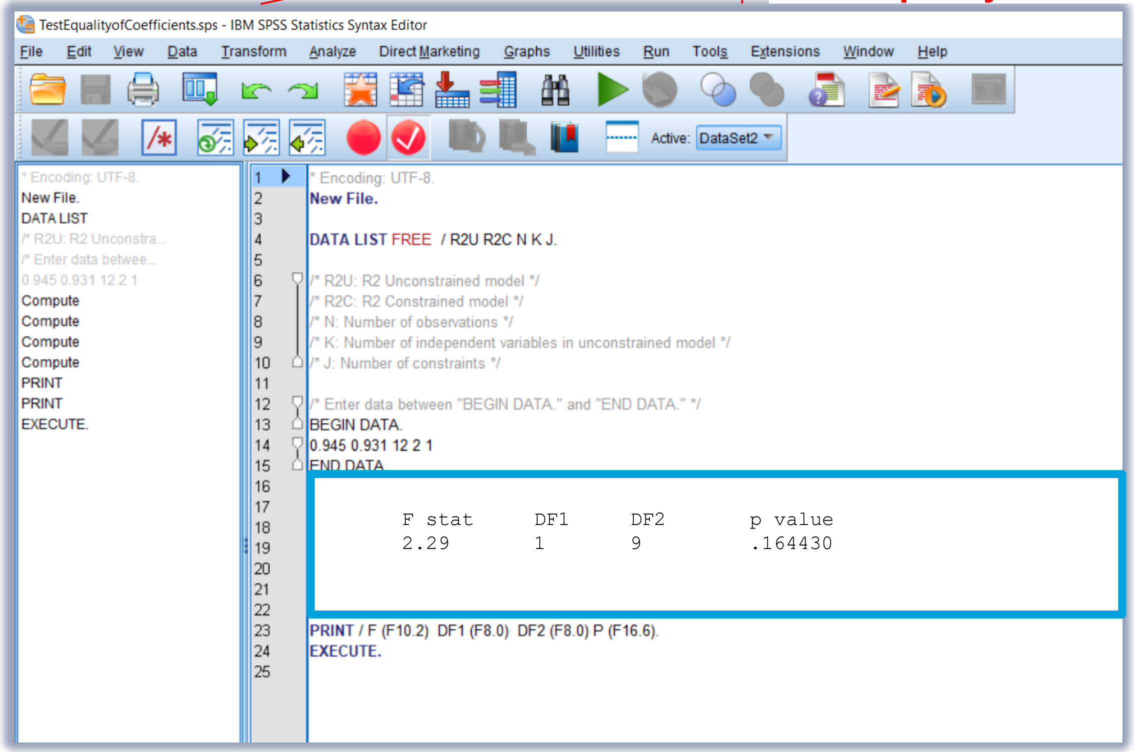
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.296	.601		.492	.633
	dandimp	.408	.035	.965	11.637	.000

$b_1=b_2$

a. Dependent Variable: Attitude

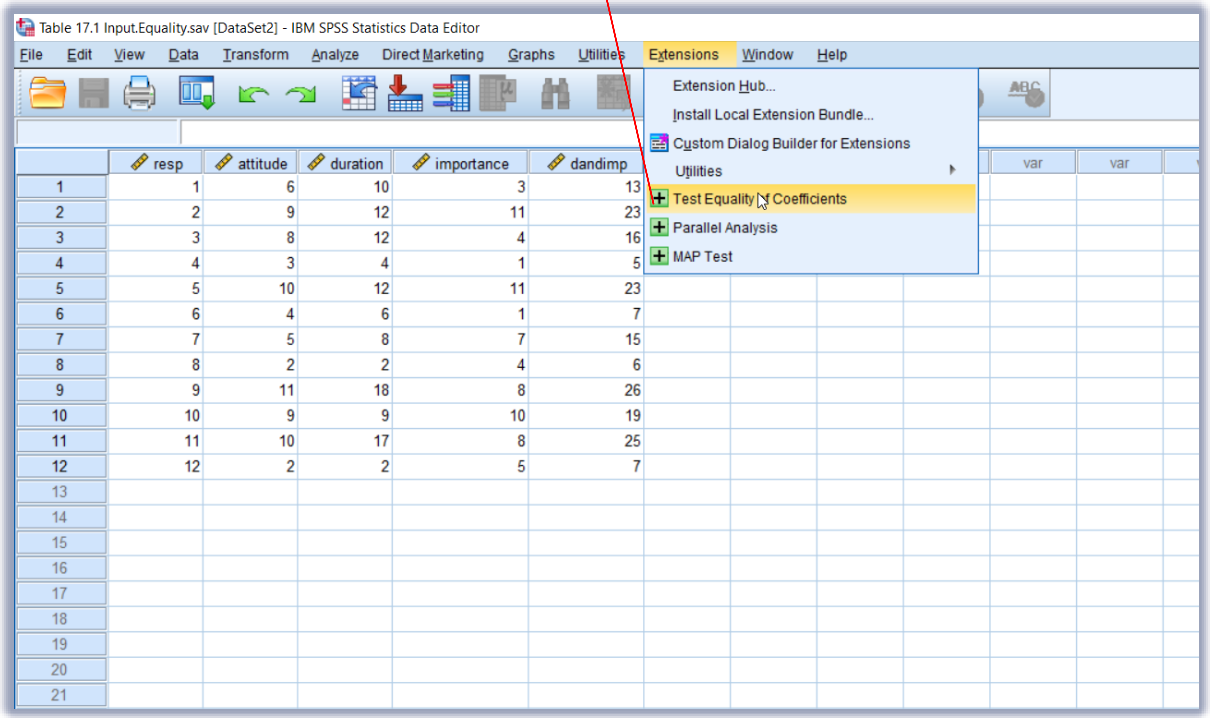
6. Advanced Topics: Equality of Coefficients

SPSS Macro
TestEqualityofCoefficients.sps



6. Advanced Topics: Equality of Coefficients

Extension (*.spe)
Equality of Coefficients



6. Advanced Topics: Equality of Coefficients

Hayes (2013); Hayes and Matthes (2009)

Integration

```
* Encoding: UTF-8.
* Test Equality of Coefficients ===.

BEGIN PROGRAM R.

# Load packages

install.packages("car", repos="http://cran.rstudio.com")
library(car)

# Transfer data to R data frame

testData<-spssdata.GetDataFromSPSS()

# Estimate regression model

M00<-lm(attitude ~ duration + importance, data=testData)

# Restriction

linearHypothesis(M00, "duration = importance")

END PROGRAM.
```

6. Advanced Topics: Equality of Coefficients

Hayes (2013); Hayes and Matthes (2009)

Integration

```
Hypothesis:
duration - importance = 0

Model 1: restricted model
Model 2: attitude ~ duration + importance

  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     10 8.3149
2      9 6.6524  1    1.6625 2.2492 0.1679
```

Alte R native…

```
> anova(R01,R00)
Analysis of Variance Table

Model 1: ATT ~ I(PEU + PU)
Model 2: ATT ~ PEU + PU
  Res.Df    RSS Df Sum of Sq      F  Pr(>F)
1    437 281.43
2    436 278.21  1    3.2195 5.0456 0.02519 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Baron and Kenny (1986)

Journal of Personality and Social Psychology
1986, Vol. 51, No. 6, 1173–1182

0022-3514/86/$00.75

The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations

Reuben M. Baron and David A. Kenny
University of Connecticut

In this article, we attempt to distinguish between the properties of moderator and mediator variables at a number of levels. First, we seek to make theorists and researchers aware of the importance of not using the terms *moderator* and *mediator* interchangeably by carefully elaborating, both conceptually and strategically, the many ways in which moderators and mediators differ. We then go beyond this largely pedagogical function and delineate the conceptual and strategic implications of making use of such distinctions with regard to a wide range of phenomena, including control and stress, attitudes, and personality traits. We also provide a specific compendium of analytic procedures appropriate for making the most effective use of the moderator and mediator distinction, both separately and in terms of a broader causal system that includes both moderators and mediators.

Preacher and Hayes (2004)

Behavior Research Methods, Instruments, & Computers
2004, 36 (4), 717-731

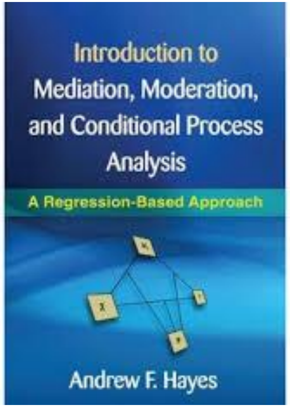
SPSS and SAS procedures for estimating indirect effects in simple mediation models

KRISTOPHER J. PREACHER
University of North Carolina, Chapel Hill, North Carolina

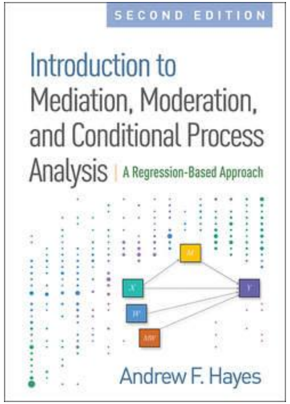
and

ANDREW F. HAYES
Ohio State University, Columbus, Ohio

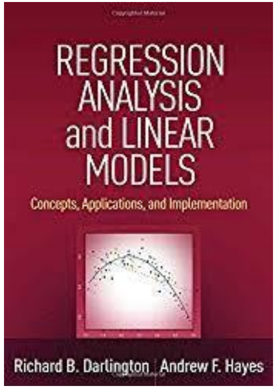
Researchers often conduct mediation analysis in order to indirectly assess the effect of a proposed cause on some outcome through a proposed mediator. The utility of mediation analysis stems from its ability to go beyond the merely descriptive to a more functional understanding of the relationships among variables. A necessary component of mediation is a statistically and practically significant indirect effect. Although mediation hypotheses are frequently explored in psychological research, formal significance tests of indirect effects are rarely conducted. After a brief overview of mediation, we argue the importance of directly testing the significance of indirect effects and provide SPSS and SAS macros that facilitate estimation of the indirect effect with a normal theory approach and a bootstrap approach to obtaining confidence intervals, as well as the traditional approach advocated by Baron and Kenny (1986). We hope that this discussion and the macros will enhance the frequency of formal mediation tests in the psychology literature. Electronic copies of these macros may be downloaded from the Psychonomic Society's Web archive at www.psychonomic.org/archive/.



Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach.* New York, NY: The Guilford Press.



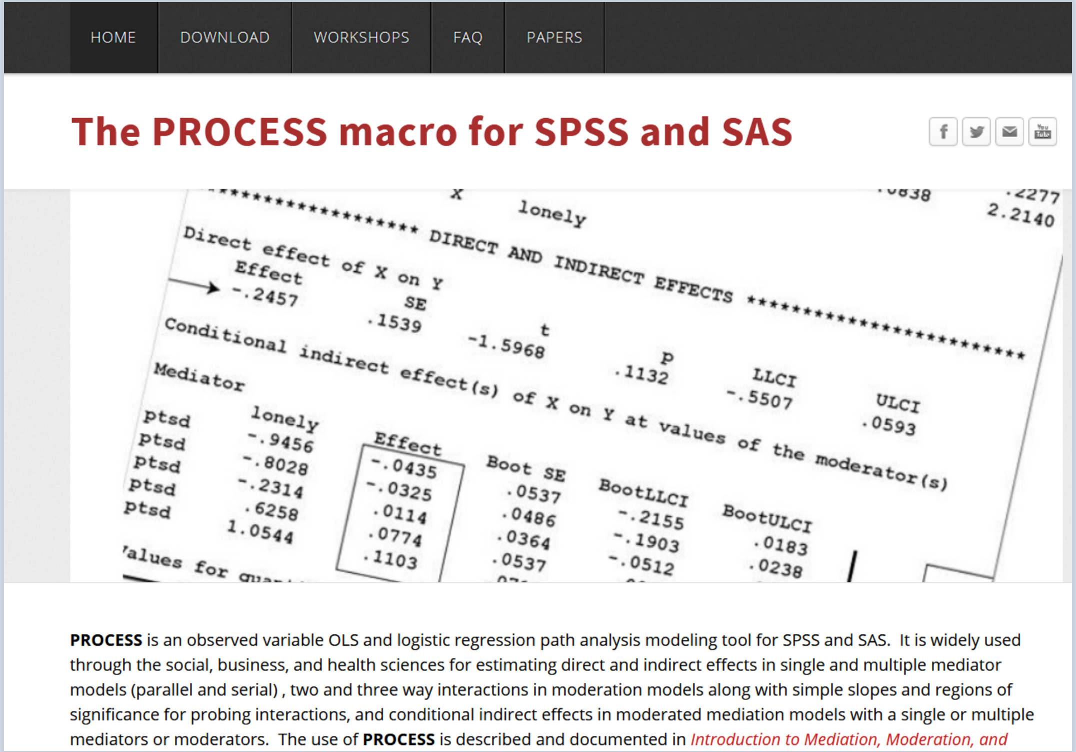
Hayes, A. F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach.* New York, NY: The Guilford Press.



Darlington, R. B., & Hayes, A. F. (2016). Regression Analysis and Linear models: Concepts, Applications, and Implementation. New York, NY: The Guilford Press.

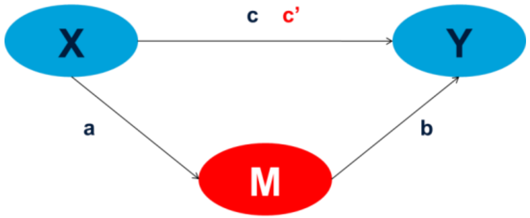
PROCESS Macro

http://www.processmacro.org/

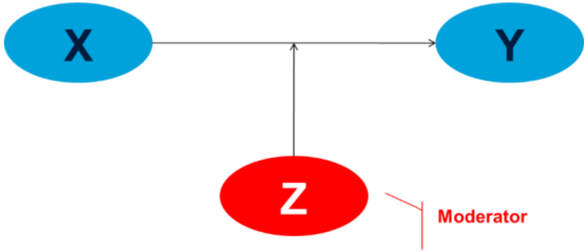


6. Advanced Topics

Mediation

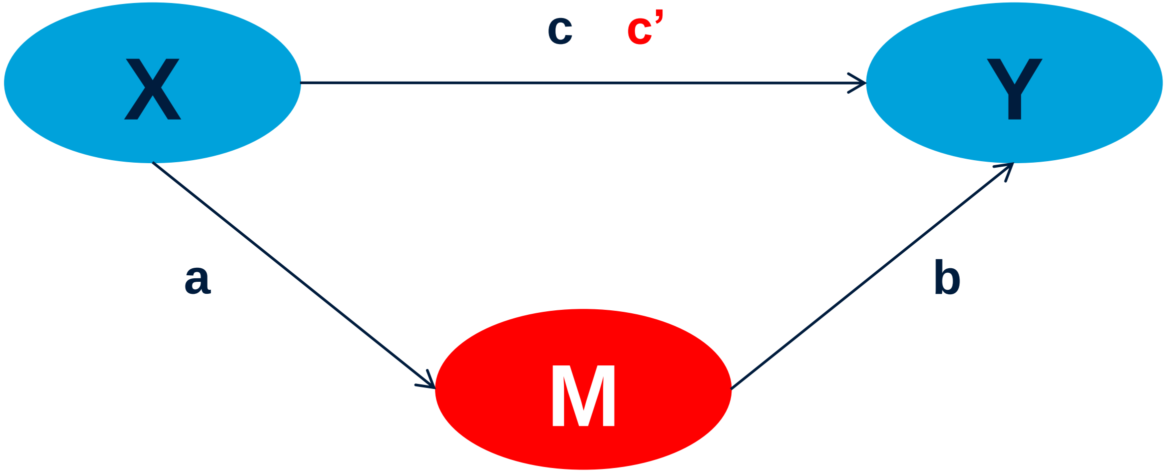


Moderation



Mediation

Preacher and Hayes (2004)



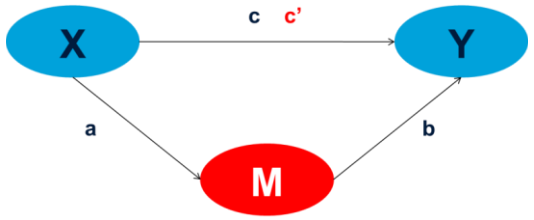
c’= 0 ⟶ COMPLETE (FULL) mediation

c’ ≠ 0 ⟶ PARTIAL mediation

Mediation

Preacher and Hayes (2004)

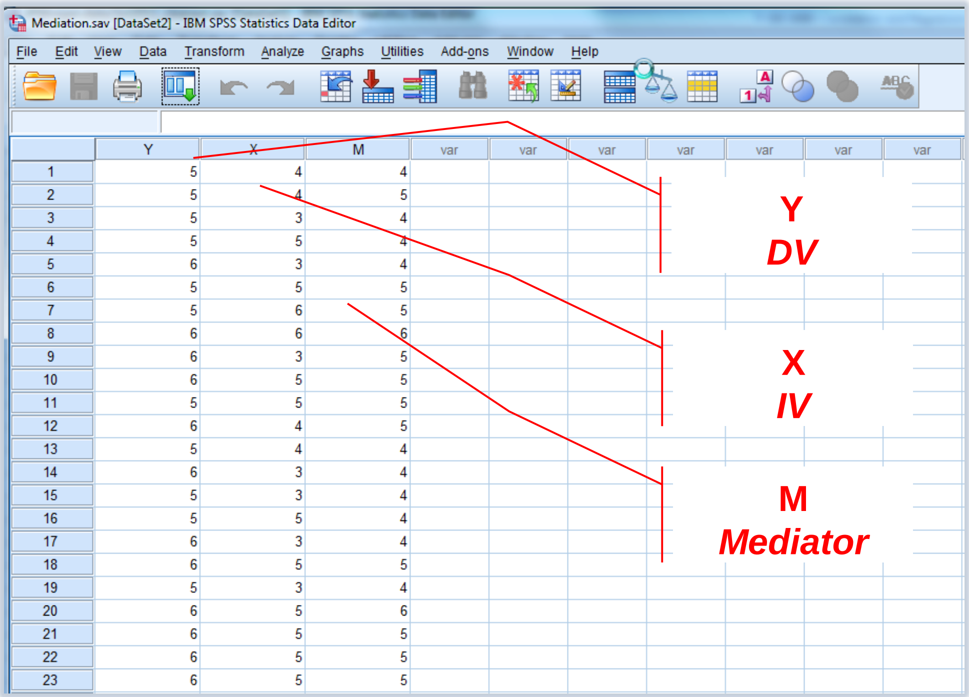
- Regression Approach to Mediation
  1. $Y = i_1 + c*X + ERROR_1$
  2. $M = i_2 + a*X + ERROR_2$
  3. $Y = i_3 + c'*X + b*M + ERROR_3$



- Sobel (1982) test
  - H_0: $a*b = 0$ $(c – c')$

Mediation (Preacher and Hayes, 2004)

Mediation Data

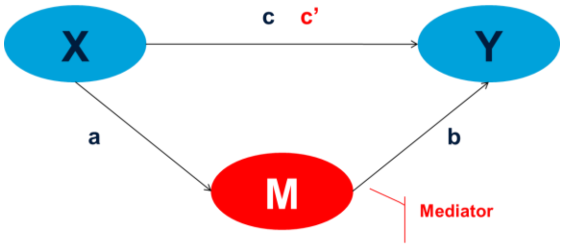


	Y	X	M
1	5	4	4
2	5	4	5
3	5	3	4
4	5	5	4
5	6	3	4
6	5	5	5
7	5	6	5
8	6	6	6
9	6	3	5
10	6	5	5
11	5	5	5
12	6	4	5
13	5	4	4
14	6	3	4
15	5	3	4
16	5	5	4
17	6	3	4
18	6	5	5
19	5	3	4
20	6	5	6
21	6	5	5
22	6	5	5
23	6	5	5

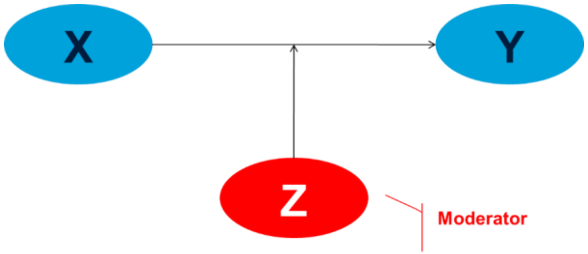
6. Advanced Topics: Mediation And Moderation

Hayes (2013); Hayes and Matthes (2009); Preacher and Hayes (2004)

Mediation

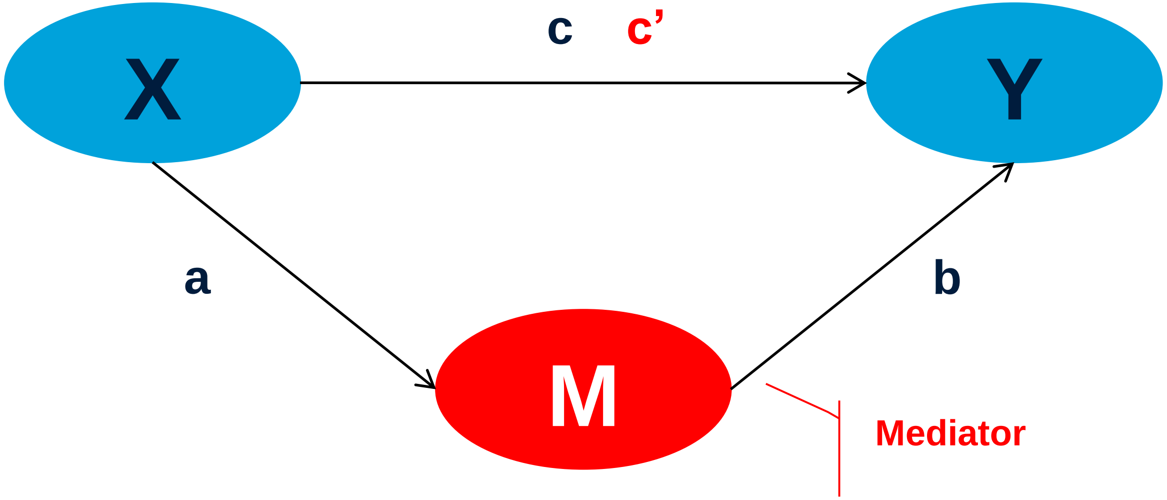


Moderation



6. Advanced Topics: Mediation

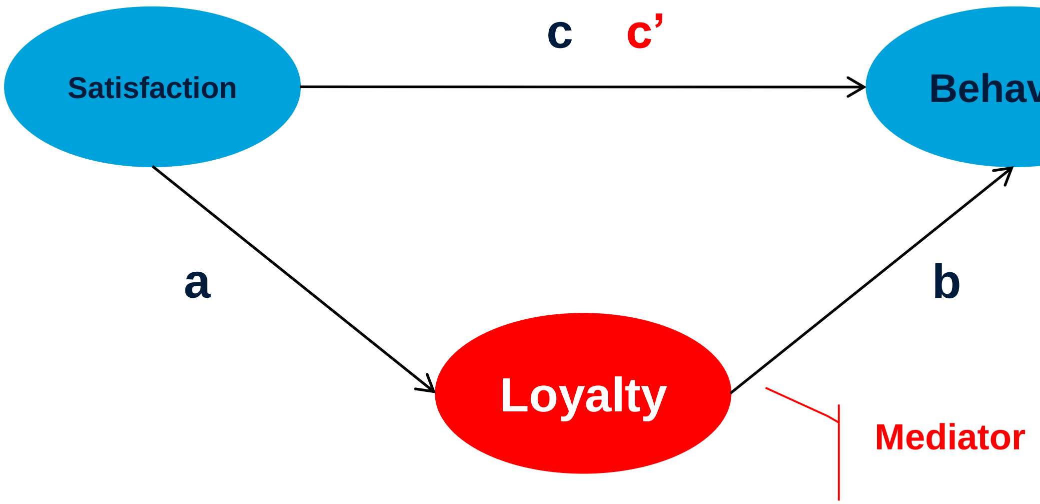
Hayes (2013); Preacher and Hayes (2004)



c’= 0 → COMPLETE (FULL) mediation
c’ ≠ 0 → PARTIAL mediation

6. Advanced Topics: Mediation

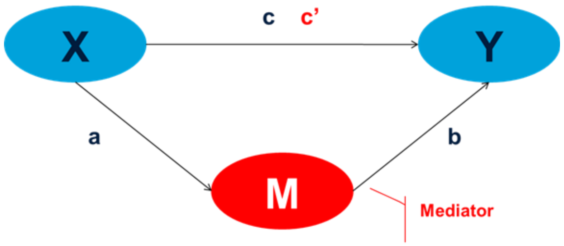
Hayes (2013); Preacher and Hayes (2004)



6. Advanced Topics: Mediation

Hayes (2013); Preacher and Hayes (2004)

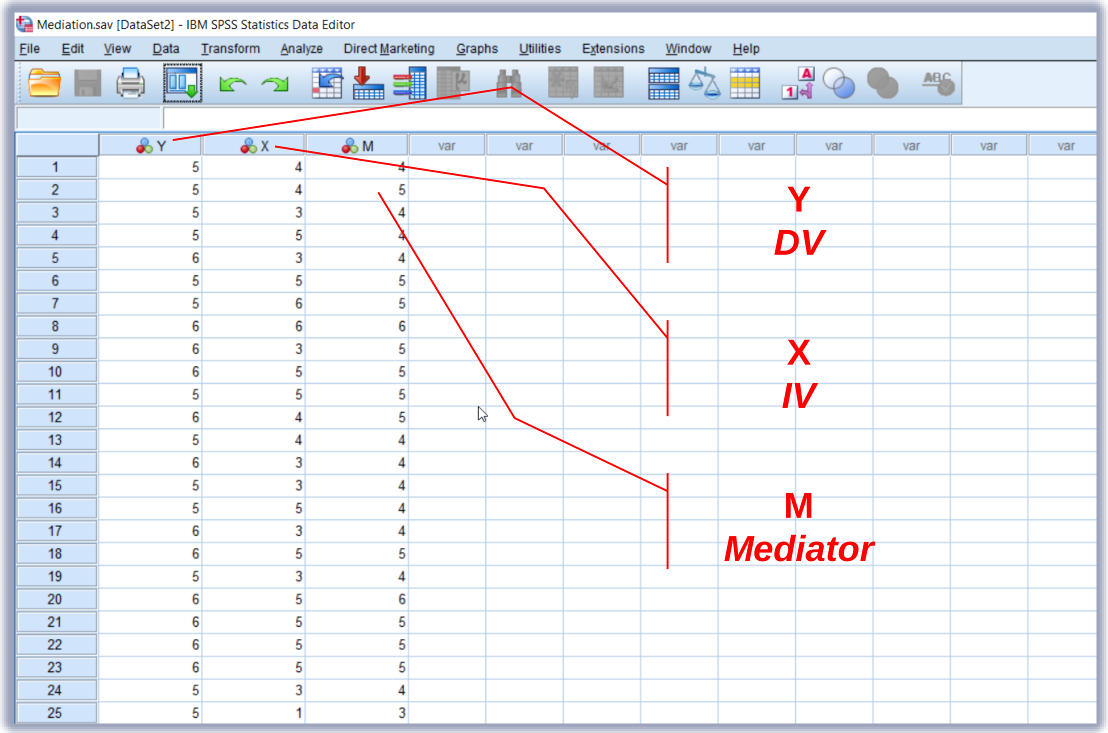
- Regression Approach to Mediation
  1. $Y = i_1 + c*X + ERROR_1$
  2. $M = i_2 + a*X + ERROR_2$
  3. $Y = i_3 + c'*X + b*M + ERROR_3$



- Sobel (1982) test
  - H_0: $a*b = 0$ $(c – c'=0)$

6. Advanced Topics: Mediation

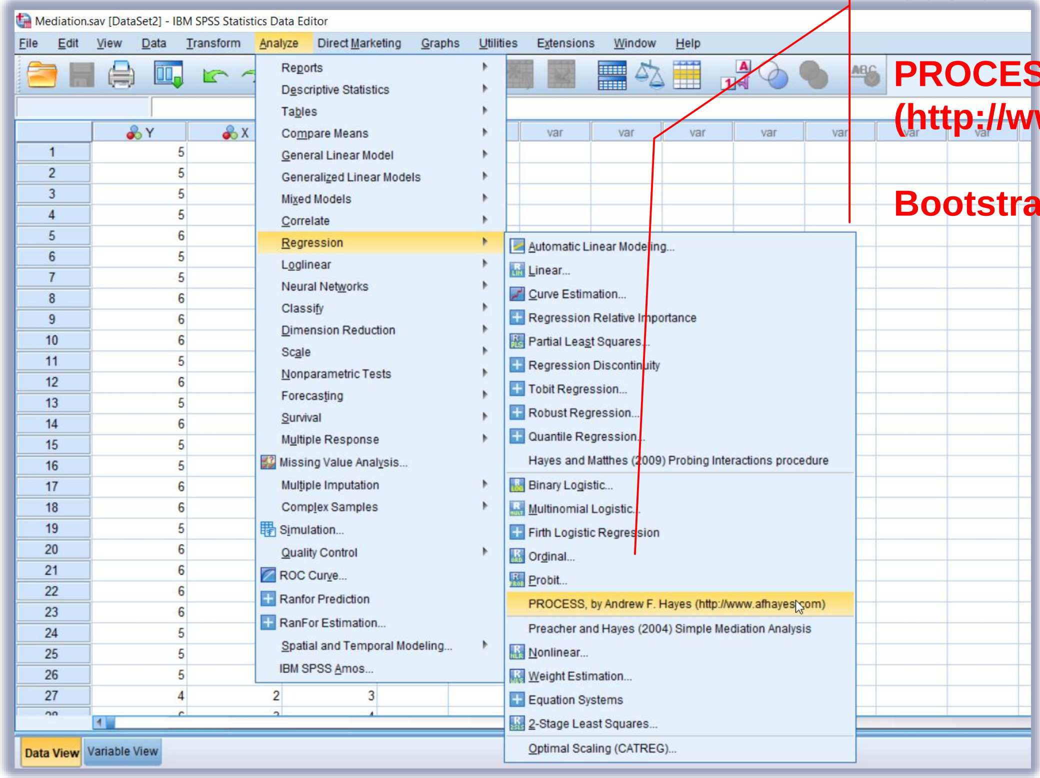
Hayes (2013); Preacher and Hayes (2004)



	Y	X	M
1	5	4	4
2	5	4	5
3	5	3	4
4	5	5	4
5	6	3	4
6	5	5	5
7	5	6	5
8	6	6	6
9	6	3	5
10	6	5	5
11	5	5	5
12	6	4	5
13	5	4	4
14	6	3	4
15	5	3	4
16	5	5	4
17	6	3	4
18	6	5	5
19	5	3	4
20	6	5	6
21	6	5	5
22	6	5	5
23	6	5	5
24	5	3	4
25	5	1	3

6. Advanced Topics: Mediation

Hayes (2013); Preacher and Hayes (2004)



Custom Dialogs (*.spd)

PROCESS Macro (http://www.afhayes.com/)

Bootstrapping!

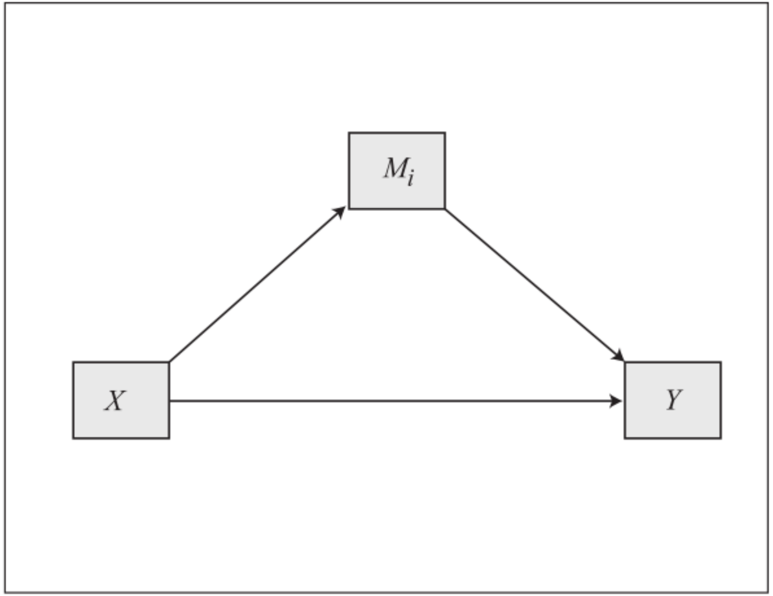
6. Advanced Topics: Mediation

Hayes (2013); Preacher and Hayes (2004)

templates.pdf

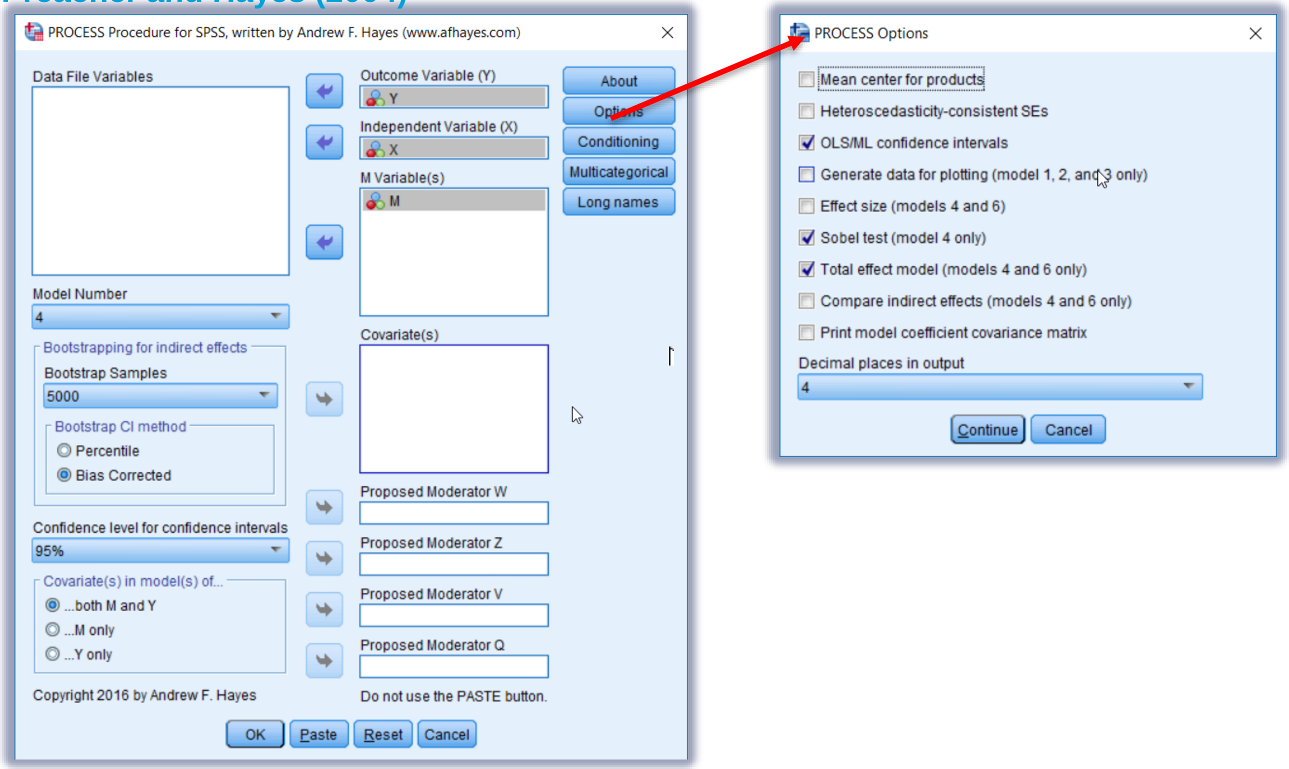
Model 4

Conceptual Diagram



6. Advanced Topics: Mediation

Hayes (2013); Preacher and Hayes (2004)



6. Advanced Topics: Mediation

Hayes (2013); Preacher and Hayes (2004)

```
**************************************************************************
Outcome: M

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .7607      .5787      .3098   271.9727     1.0000   198.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     2.5025      .1300    19.2568      .0000     2.2462     2.7588
X          a  .5132      .0311    16.4916      .0000      .4518      .5746

**************************************************************************
Outcome: Y

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5871      .3446      .3253    51.7991     2.0000   197.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     3.2430      .2257    14.3674      .0000     2.7979     3.6882
M          b  .4494      .0728     6.1707      .0000      .3058      .5930
X          c' .0275      .0491      .5598      .5763     -.0694      .1244
```

6. Advanced Topics: Mediation

Hayes (2013); Preacher and Hayes (2004)

```
************************* TOTAL EFFECT MODEL ****************************
Outcome: Y

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .4669      .2180      .3863    55.1864     1.0000   198.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     4.3677      .1451    30.0994      .0000     4.0815     4.6538
X             .2581      .0347     7.4288      .0000      .1896      .3266

***************** TOTAL, DIRECT, AND INDIRECT EFFECTS ********************

Total effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .2581      .0347     7.4288      .0000      .1896      .3266

Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
      .0275      .0491      .5598      .5763     -.0694      .1244

Indirect effect of X on Y
       Effect    Boot SE   BootLLCI   BootULCI
M       .2306      .0420      .1525      .3166

Normal theory tests for indirect effect
     Effect         se          Z          p
      .2306      .0400     5.7701      .0000
```

c

c

c’

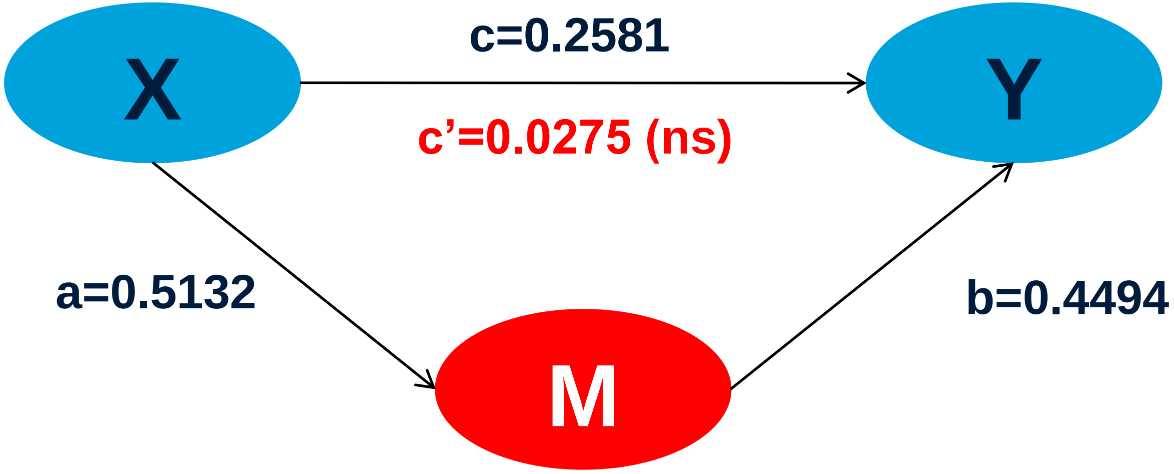
Bootstrap Results (5000 resamples [10000])

Sobel test

a*b (0.5132*0.4494) = c-c’ (0.2581-0.0275)=0.2306

6. Advanced Topics: Mediation

Hayes (2013); Preacher and Hayes (2004)



c'= 0 → COMPLETE (FULL) mediation
c' ≠ 0 → PARTIAL mediation

Regression Analysis

Editor

```r
# Mediation

library(mediation)

MM<-lm(M ~ X, data=Mediation)
MY<-lm(Y ~ X + M, data=Mediation)
MED1<-mediate(MM,MY, treat='X', mediator='M')
summary(MED1)

library(bda)

mediation.test(Mediation$M, Mediation$X, Mediation$Y)

library(QuantPsyc)

Data.Mediation<-Mediation
names(Data.Mediation)<-c('y','x','m')

proximal.med(Data.Mediation)

library(boot)
med.boot <- boot(Data.Mediation, proxInd.ef, R=1000)
sort(med.boot$t)[c(25,975)] #95% CI
plot(density(med.boot$t)) # Distribution of bootstapped indirect effect
summary(med.boot$t)
```

Regression Analysis

Console

```
> # Mediation
>
> options(scipen=999)
>
> library(mediation)
>
> MM<-lm(M ~ X, data=Mediation)
> MY<-lm(Y ~ X + M, data=Mediation)
> MED1<-mediate(MM,MY, treat='X', mediator='M')
> summary(MED1)

Causal Mediation Analysis

Quasi-Bayesian Confidence Intervals

               Estimate 95% CI Lower 95% CI Upper p-value
ACME             0.2317       0.1533       0.3122    0.00
ADE              0.0253      -0.0684       0.1301    0.63
Total Effect     0.2570       0.1914       0.3230    0.00
Prop. Mediated   0.9053       0.5620       1.3323    0.00

Sample Size Used: 200


Simulations: 1000
```

Regression Analysis

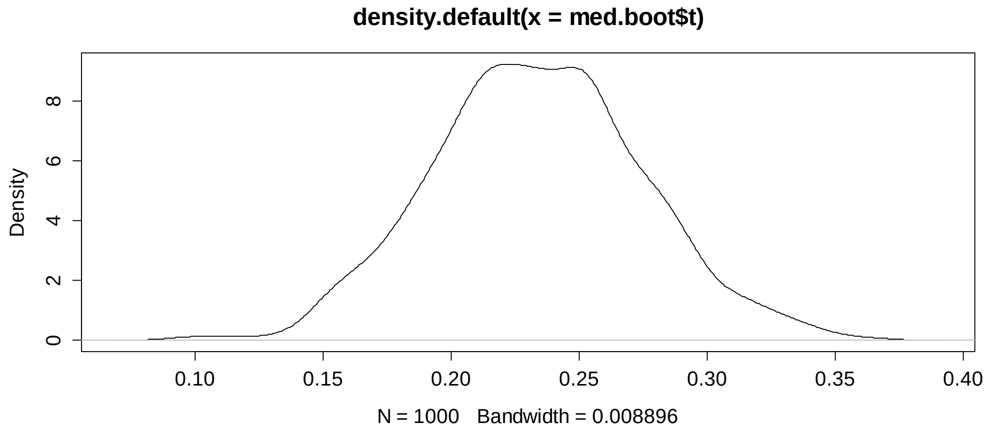
Console

```
> library(bda)
>
> mediation.test(Mediation$M, Mediation$X, Mediation$Y)
                       Sobel               Aroian              Goodman
z.value 5.779376771925725 5.770079189740873 5.788719444249967
p.value 0.000000007497785 0.000000007923429 0.000000007092505
>
> library(QuantPsyc)
>
> Data.Mediation<-Mediation
> names(Data.Mediation)<-c('y','x','m')
> proximal.med(Data.Mediation)
        Effect                SE                    t-ratio              Med.Ratio
a       "0.51319079769942 5"  "0.0311183270153198"  "16.4915934409577"   "--"
b       "0.449396851151261"   "0.072827539895 85"   "6.17069932327714"   "--"
t       "0.258127031757939"   "0.0347470100395941"  "7.42875520696038"   "--"
t'      "0.0275007032320145"  "0.0491301595620519"  "0.559751962484079"  "--"
ab      "0.230626328525925"   "0.0399050516391716"  "5.77937677192573"   "0.893460583943012"
Aroian  "--"                  "0.0399693524026456"  "--"                 "--"
Goodman "--"                  "0.0398406470976945"  "--"                 "--"
```

Regression Analysis

Console

```
> library(boot)
> med.boot <- boot(Data.Mediation, proxInd.ef, R=1000)
> sort(med.boot$t)[c(25,975)] #95% CI
[1] 0.1541803 0.3175103
> plot(density(med.boot$t)) # Distribution of bootstapped indirect effect
```



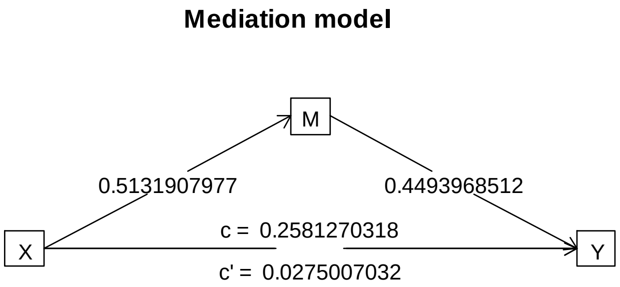
Regression Analysis

Console

```
> library(psych)
> MED.1<-mediate(Y ~ X + (M), data=Mediation, n.iter=50)
> mediate.diagram(MED.1, digits=10)
> print(MED.1, digits=10, short=FALSE)

Mediation/Moderation Analysis
Call: mediate(y = Y ~ X + (M), data = Mediation, n.iter = 50)

The DV (Y) was  Y . The IV (X) was  X . The mediating variable(s) =  M .

Total effect(c) of  X  on  Y  =  0.258127   S.E. =  0.03474701  t  =  7.428755  df=  197   with p =  3.25063e-12
Direct effect (c') of  X  on  Y  removing  M  =  0.0275007   S.E. =  0.04913016  t  =  0.559752  df=  197   with p =  0.5762845
Indirect effect (ab) of  X  on  Y  through  M   =  0.2306263
Mean bootstrapped indirect effect =  0.2328739  with standard error =  0.0433557  Lower CI =  0.1449465    Upper CI =  0.3094526
R = 0.5870605 R2 = 0.34464   F = 51.79908 on 2 and 197 DF   p-value:  8.380536e-19


 Full output

 Total effect estimates (c)
         Y         se        t  df        Prob
X 0.258127 0.03474701 7.428755 197 3.25063e-12

Direct effect estimates     (c')
          Y         se        t  df         Prob
X 0.0275007 0.04913016 0.559752 197 5.762845e-01
M 0.4493969 0.07282754 6.170699 197 3.796747e-09

 'a'  effect estimates
          M         se        t  df        Prob
X 0.5131908 0.03111833 16.49159 198 5.06913e-39

 'b'  effect estimates
          Y         se        t  df         Prob
M 0.4493969 0.07282754 6.170699 197 3.796747e-09

 'ab'  effect estimates
          Y      boot        sd     lower     upper
X 0.2306263 0.2328739 0.0433557 0.1449465 0.3094526
>
```

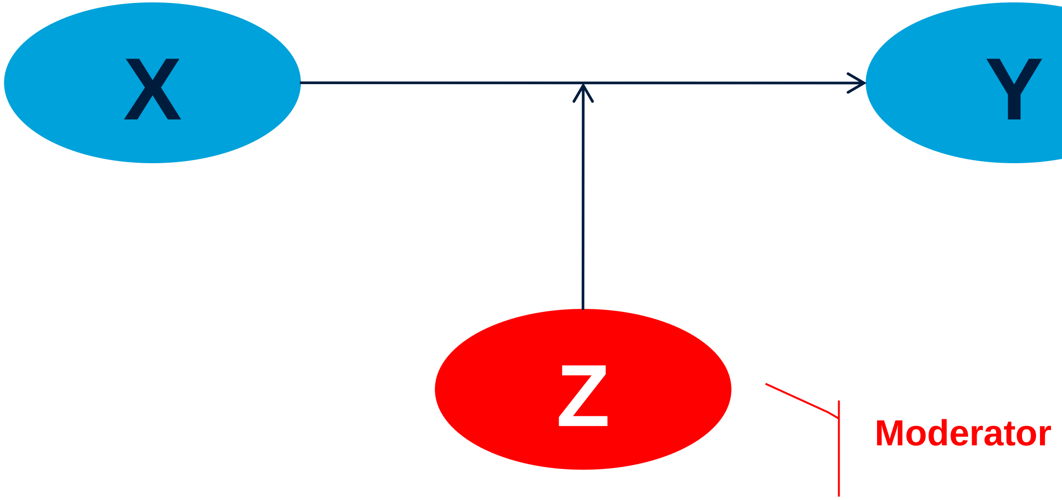
Regression Analysis

Graphics



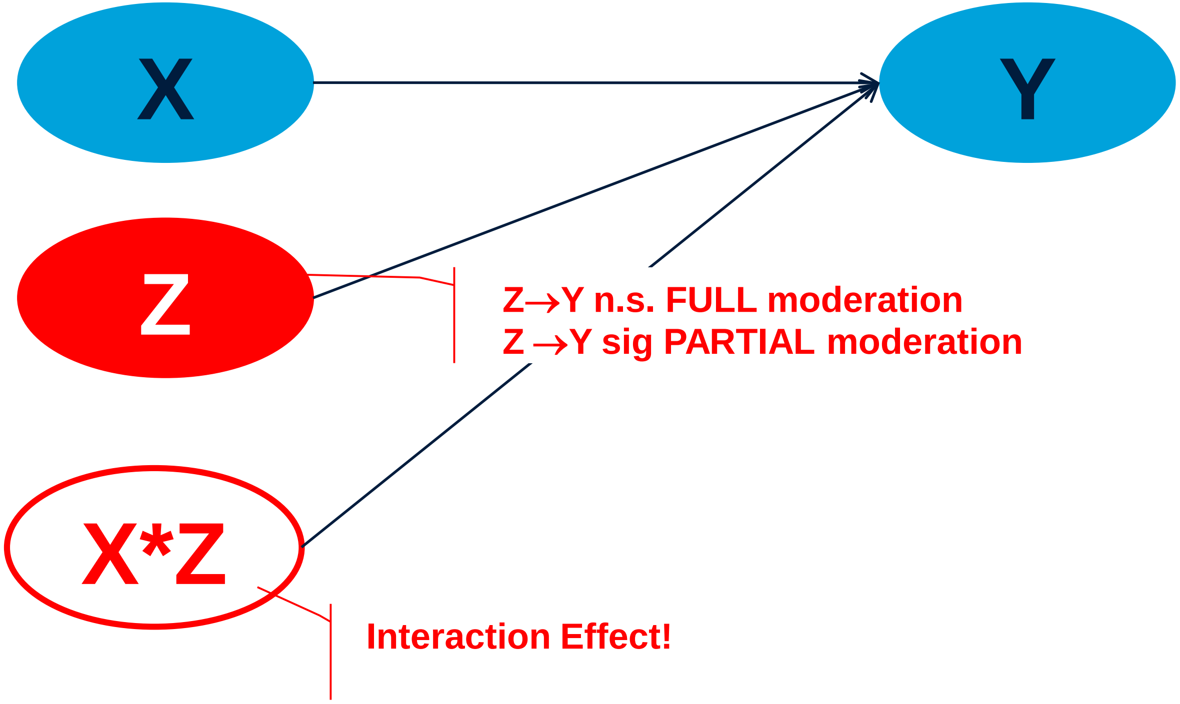
Moderation

Hayes and Matthes (2009)



Moderation

Hayes and Matthes (2009)



6. Advanced Topics: Moderation

Hayes (2013); Hayes and Matthes (2009)

- Equations:

$$Y_i = a + b_1 * X_i + b_2 * Z_i + b_3 * Z_i * X_i + \text{ERROR}$$

$$Y_i = a + (b_1 + b_3 * Z) * X_i + b_2 * Z_i + \text{ERROR}$$

What happens if $b_3 = 0$?

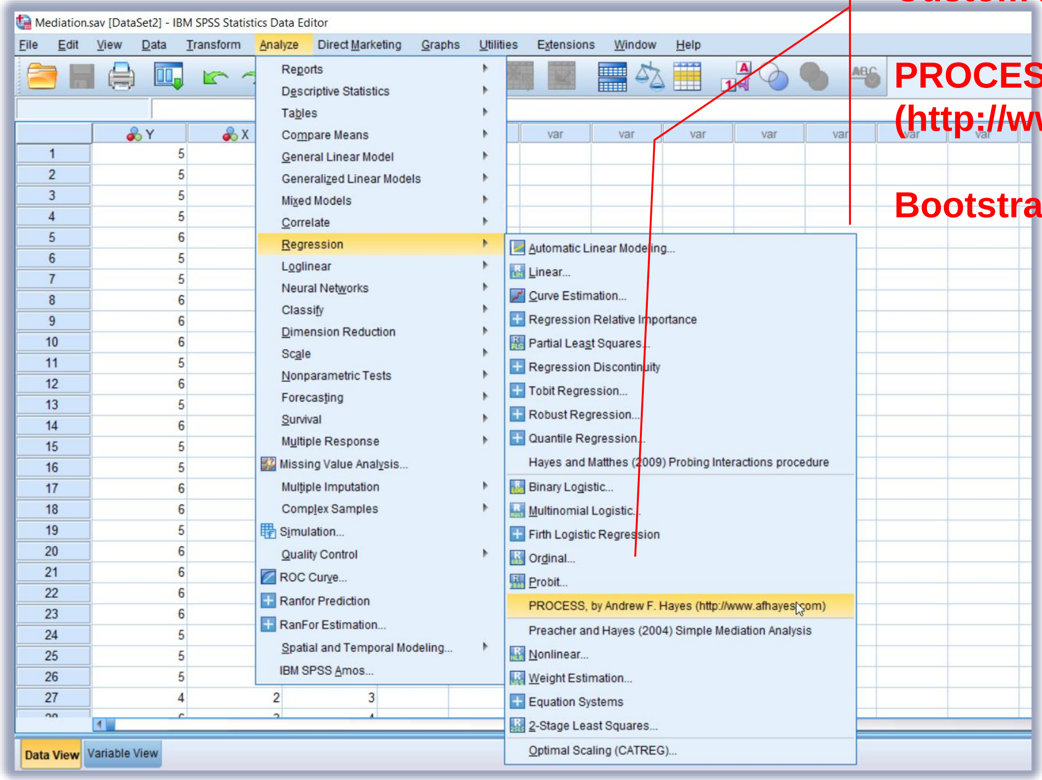
$$Y_i = a + (b_1 + 0 * Z) * X_i + b_2 * Z_i + \text{ERROR}$$

$$Y_i = a + b_1 * X_i + b_2 * Z_i + \text{ERROR}$$

No Moderation!

6. Advanced Topics: Moderation

Hayes (2013); Hayes and Matthes (2009)



Custom Dialogs (*.spd)

PROCESS Macro (http://www.afhayes.com/)

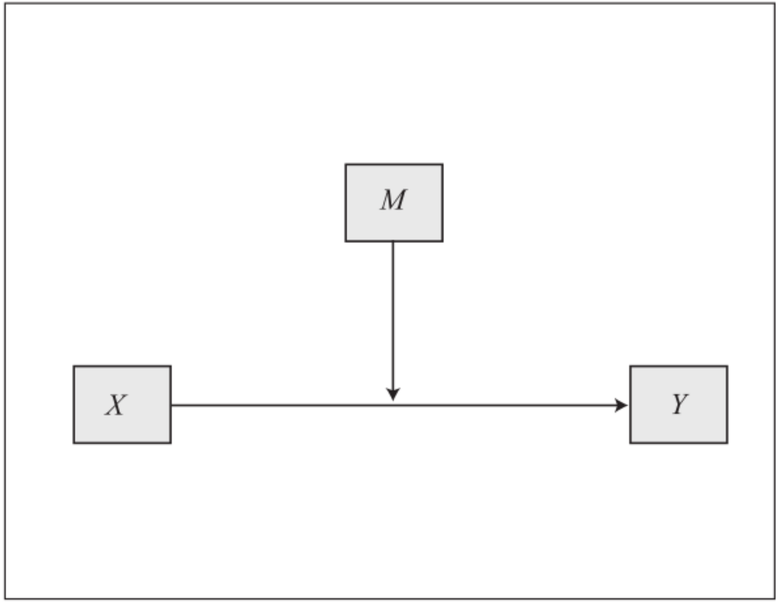
Bootstrapping!

6. Advanced Topics: Moderation

Hayes (2013); Preacher and Hayes (2004)

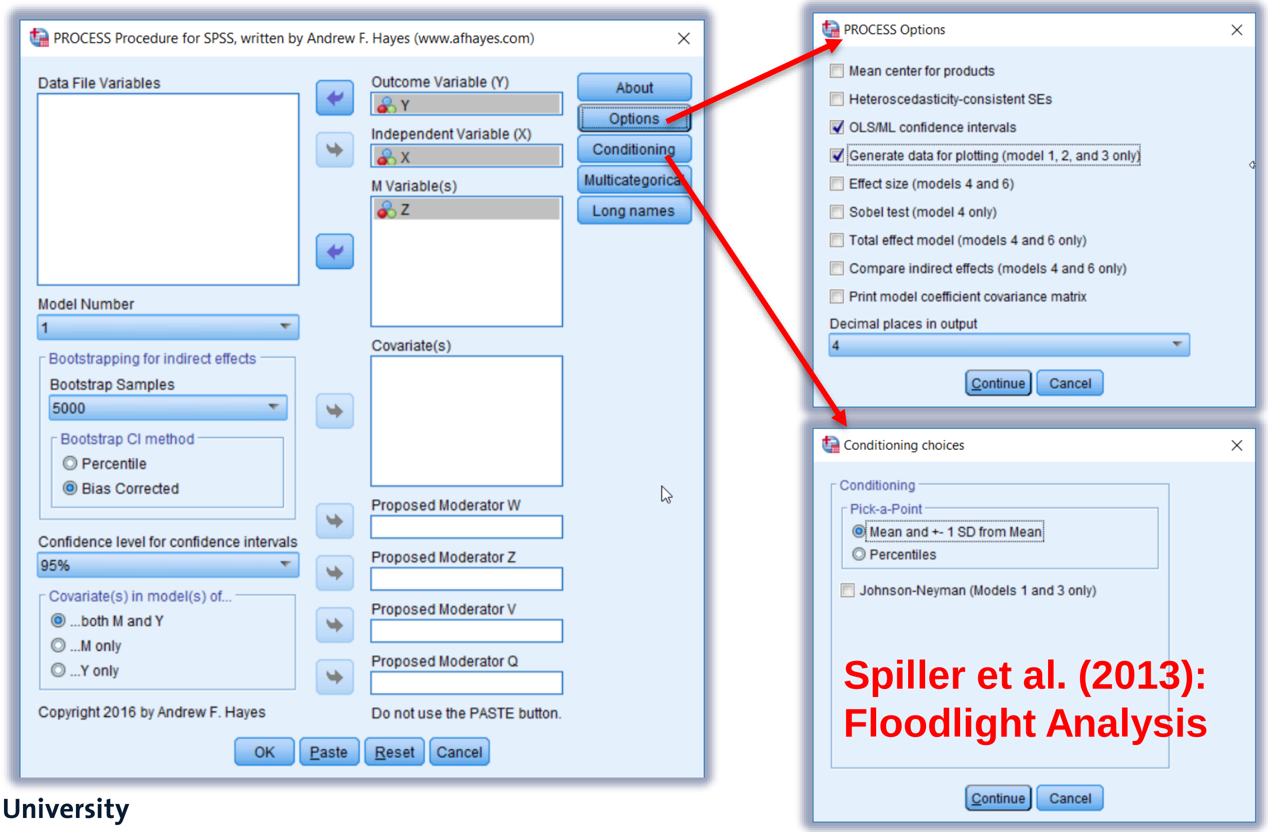
Model 1 — **templates.pdf**

Conceptual Diagram



6. Advanced Topics: Moderation

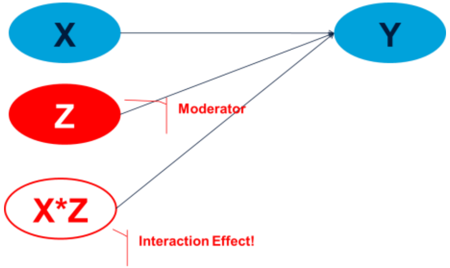
Hayes (2013); Hayes and Matthes (2009)



6. Advanced Topics: Moderation

Hayes (2013); Hayes and Matthes (2009)

```
**************************************************************************
Outcome: Y

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .8687      .7546      .4623   200.8704     3.0000   196.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     -.1238      .5556     -.2227      .8240    -1.2196      .9720
Z             .0956      .1315      .7272      .4680     -.1637      .3549
X             .4689      .1296     3.6188      .0004      .2134      .7245
int_1         .0786      .0300     2.6159      .0096      .0193      .1379

Product terms key:

 int_1    X          X      Z

R-square increase due to interaction(s):
        R2-chng          F        df1        df2          p
int_1     .0086     6.8430     1.0000   196.0000      .0096

**************************************************************************

Conditional effect of X on Y at values of the moderator(s):
          Z     Effect         se          t          p       LLCI       ULCI
     2.9473      .7006      .0528    13.2648      .0000      .5964      .8048
     4.2250      .8010      .0399    20.0643      .0000      .7223      .8798
     5.5027      .9015      .0578    15.5837      .0000      .7874     1.0156

Values for quantitative moderators are the mean and plus/minus one SD from mean.
Values for dichotomous moderators are the two values of the moderator.
```



Incremental F test

Conditional Effects

6. Advanced Topics: Moderation

Hayes (2013); Hayes and Matthes (2009)

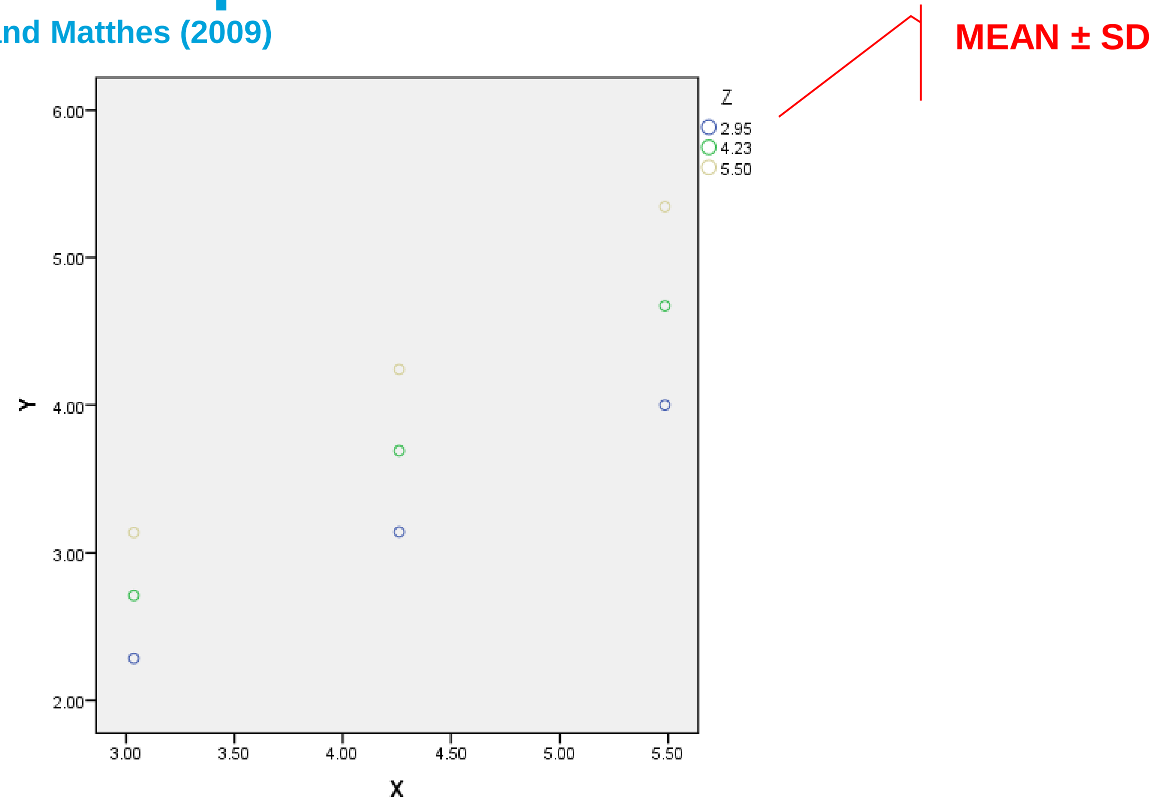
SPSS.PROCESS.Moderation.Plot.sps

```
* Encoding: UTF-8.
DATA LIST FREE/X Z Y.
BEGIN DATA.

   3.0353   2.9473   2.2845
   4.2600   2.9473   3.1426
   5.4847   2.9473   4.0006
   3.0353   4.2250   2.7116
   4.2600   4.2250   3.6926
   5.4847   4.2250   4.6736
   3.0353   5.5027   3.1386
   4.2600   5.5027   4.2426
   5.4847   5.5027   5.3466

END DATA.

GRAPH /SCATTERPLOT=X WITH Y BY Z.

GRAPH /LINE(MULTIPLE)=mean(Y) BY X BY Z.
```

```
END DATA.
GRAPH/SCATTERPLOT=X WITH Y BY Z.
```

6. Advanced Topics: Moderation

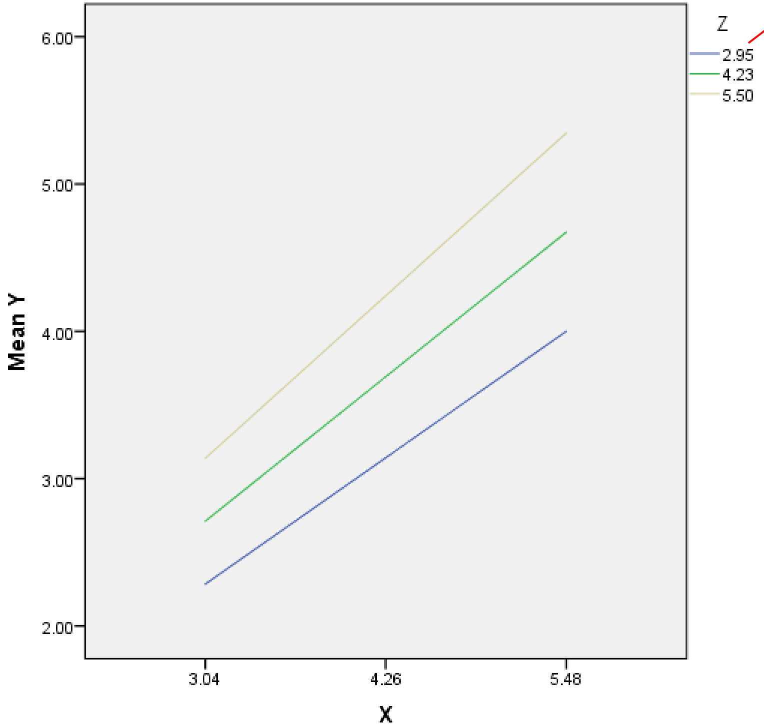
Hayes (2013); Hayes and Matthes (2009)



6. Advanced Topics: Moderation

Hayes (2013); Hayes and Matthes (2009)

MEAN ± SD



Spiller et al. (2013)

STEPHEN A. SPILLER, GAVAN J. FITZSIMONS, JOHN G. LYNCH JR., and GARY H. McCLELLAND*

It is common for researchers discovering a significant interaction of a measured variable X with a manipulated variable Z to examine simple effects of Z at different levels of X. These "spotlight" tests are often misunderstood even in the simplest cases, and it appears that consumer researchers are unsure how to extend them to more complex designs. The authors explain the general principles of spotlight tests, show that they rely on familiar regression techniques, and provide a tutorial demonstrating how to apply these tests across an array of experimental designs. Rather than following the common practice of reporting spotlight tests at one standard deviation above and below the mean of X, it is recommended that when X has focal values, researchers should report spotlight tests at those focal values. When X does not have focal values, it is recommended that researchers report ranges of significance using a version of Johnson and Neyman's test the authors term a "floodlight."

Keywords: moderated regression, spotlight analysis, simple effects tests

Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression

Regression Analysis

Editor

```
# Moderation

library(QuantPsyc)

MOD1<-moderate.lm(X, Z, Y, Moderation, mc=FALSE)
summary(MOD1)

GMOD1<-sim.slopes(MOD1, Moderation$Z, zsd=2, mcz=FALSE)

graph.mod(GMOD1, X, Y, Moderation)

```

Regression Analysis

Console

```
Call:
lm(formula = y ~ mcx * mcz, na.action = na.omit)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.68679 -0.39206  0.01796  0.42799  2.19791 

Coefficients:
            Estimate Std. Error t value             Pr(>|t|)    
(Intercept)  3.69259    0.04854  76.080 < 0.0000000000000002 ***
mcx          0.80104    0.03992  20.064 < 0.0000000000000002 ***
mcz          0.43046    0.03817  11.277 < 0.0000000000000002 ***
mcx:mcz      0.07861    0.03005   2.616              0.00959 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6799 on 196 degrees of freedom
Multiple R-squared:  0.7546, Adjusted R-squared:  0.7508 
F-statistic: 200.9 on 3 and 196 DF,  p-value: < 0.00000000000000022

>
```

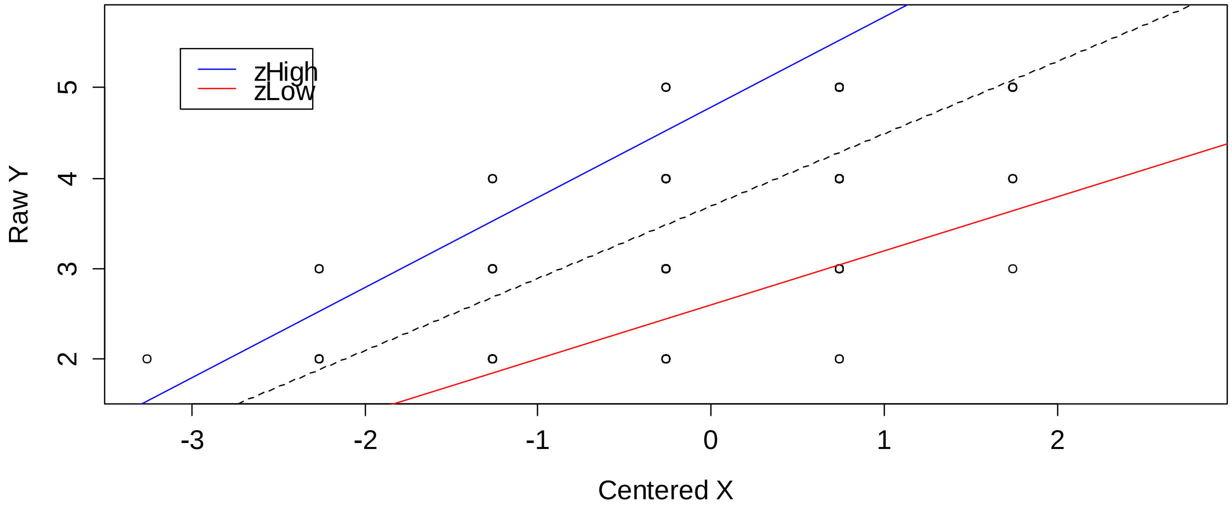
Regression Analysis

Console

```
Call:
lm(formula = y ~ mcx * mcz, na.action = na.omit)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.68679 -0.39206  0.01796  0.42799  2.19791 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.12376    0.55564  -0.223 0.823970    
mcx          0.46893    0.12958   3.619 0.000377 ***
mcz          0.09560    0.13146   0.727 0.467960    
mcx:mcz      0.07861    0.03005   2.616 0.009592 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6799 on 196 degrees of freedom
Multiple R-squared:  0.7546, Adjusted R-squared:  0.7508 
F-statistic: 200.9 on 3 and 196 DF,  p-value: < 0.00000000000000022
```

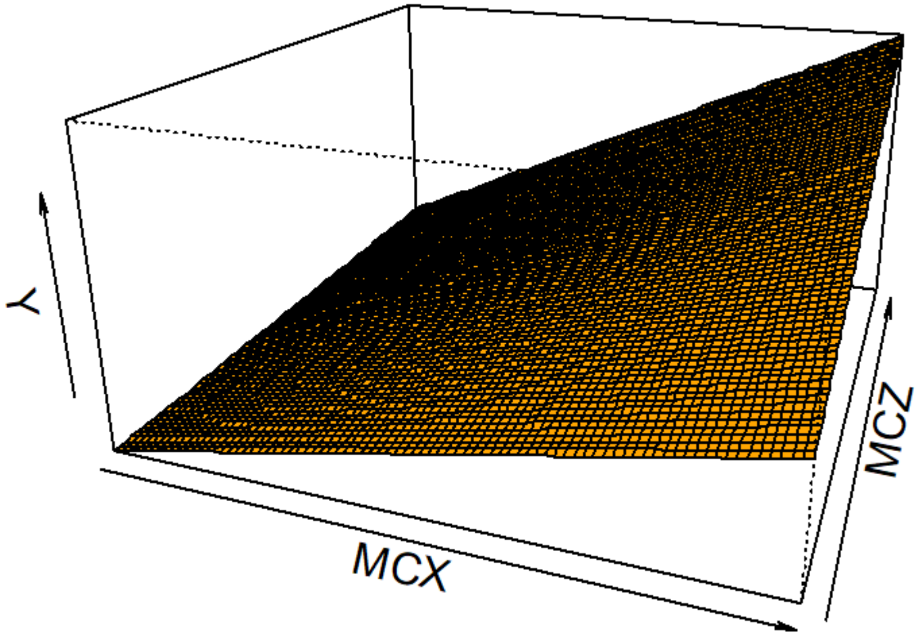
Regression Analysis

Graphics



Regression Analysis

Graphics



Regression Analysis

Console

```
> library(psych)
> MOD2<-mediate(Y ~ X*Z, data=Moderation)
> print(MOD2, digits=10)

Mediation/Moderation Analysis
Call: mediate(y = Y ~ X * Z, data = Moderation)

The DV (Y) was  Y . The IV (X) was  X Z X*Z . The mediating variable(s) =  .
 DV =  Y
        slope         se         t            p
X   0.8010410 0.03992372 20.064290 2.107443e-49
Z   0.4304616 0.03817242 11.276767 4.698036e-23
X*Z 0.0786056 0.03004897  2.615917 9.591751e-03

With R2 =  0.754574
R = 0.8686622 R2 = 0.754574   F = 200.8704 on 3 and 196 DF   p-value:  1.590401e-59
```

Regression Analysis

Editor

```r
# Rockchalk

library(rockchalk)

R.rc<-lm(Y ~ X*Z, data=Moderation)
PS.rc<-plotSlopes(R.rc, plotx="X", modx="Z", modxVals="std.dev")
TS.rc<-testSlopes(PS.rc)
plot(TS.rc)
```

6. Advanced Topics: Moderation

Hayes (2013); Hayes and Matthes (2009); Spiller et al. (2013)

- Equations:

$$Y_i = a + b_1 * X_i + b_2 * Z_i + b_3 * Z_i * X_i + \text{ERROR}$$

$$Y_i = a + (b_1 + b_3 * Z) * X_i + b_2 * Z_i + \text{ERROR}$$

What happens if $b_3 = 0$?

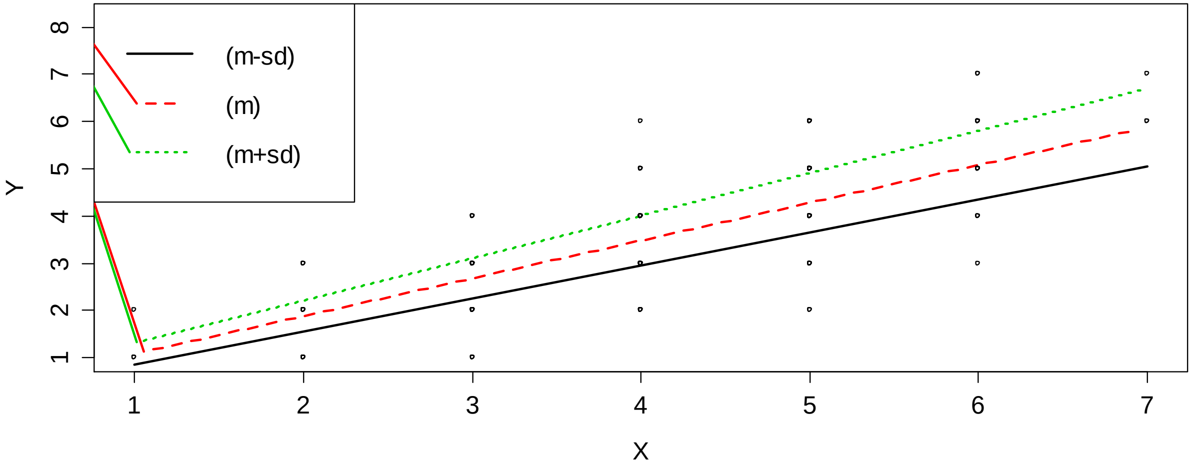
$$Y_i = a + (b_1 + 0 * Z) * X_i + b_2 * Z_i + \text{ERROR}$$

$$Y_i = a + b_1 * X_i + b_2 * Z_i + \text{ERROR}$$

No Moderation!

Regression Analysis

Graphics



Moderation

Hayes (2013); Hayes and Matthes (2009)

Z (M=4.23; SD=1.27)

Y.MEAN=-0.1238+0.4689*X+0.0956*4.23+0.0786*4.23*X

Y.MEAN=0.2806+0.8014*X

Y+1*SD=-0.1238+0.4689*X+0.0956*5.50+0.0786*5.50*X

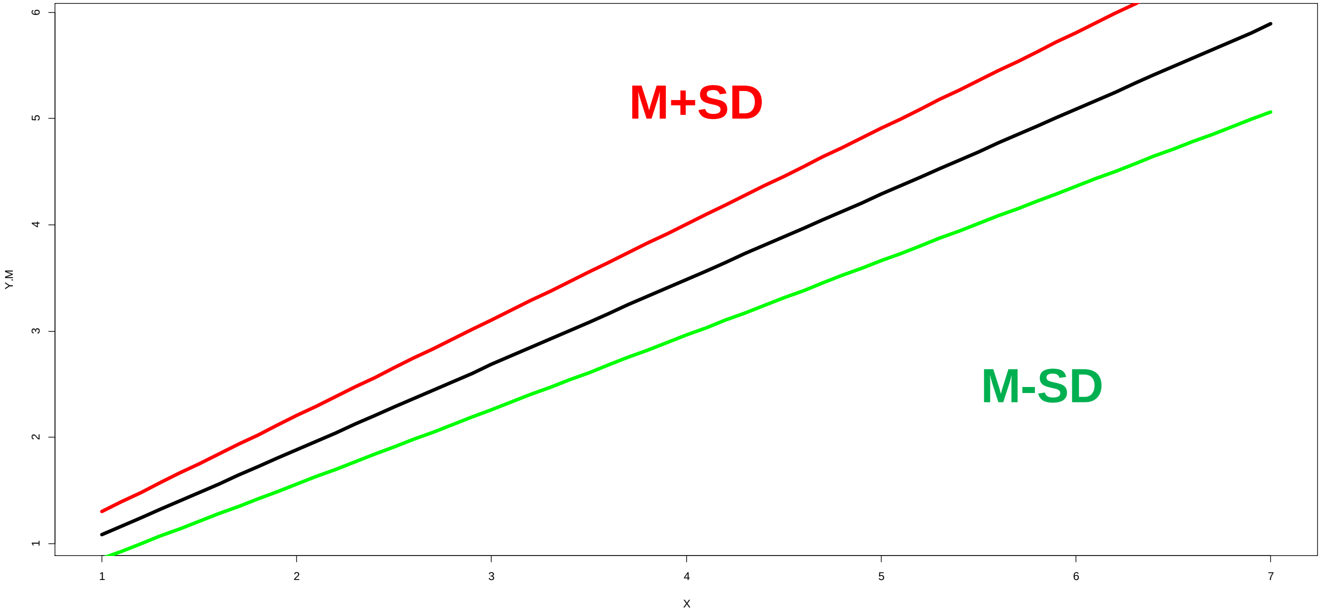
Y+1*SD=0.4020+0.9012*X

Y-1*SD=-0.1238+0.4689*X+0.0956*2.95+0.0786*2.95*X

Y-1*SD=0.1582+0.7008*X

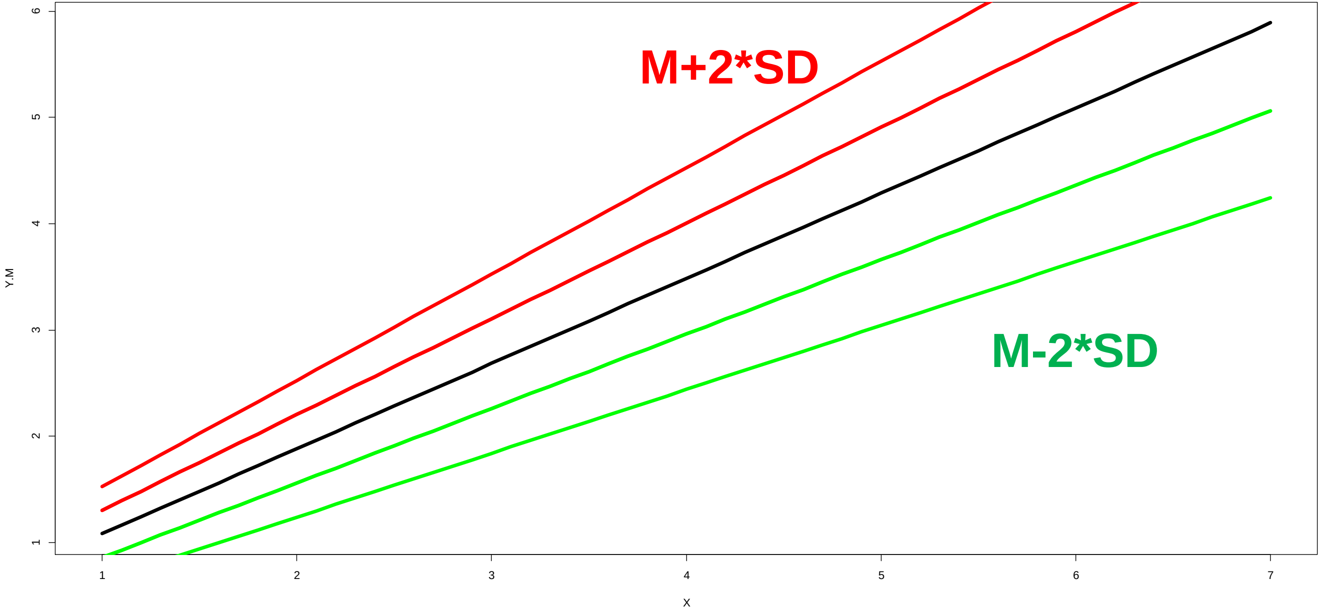
Moderation

Hayes (2013); Hayes and Matthes (2009)



Moderation

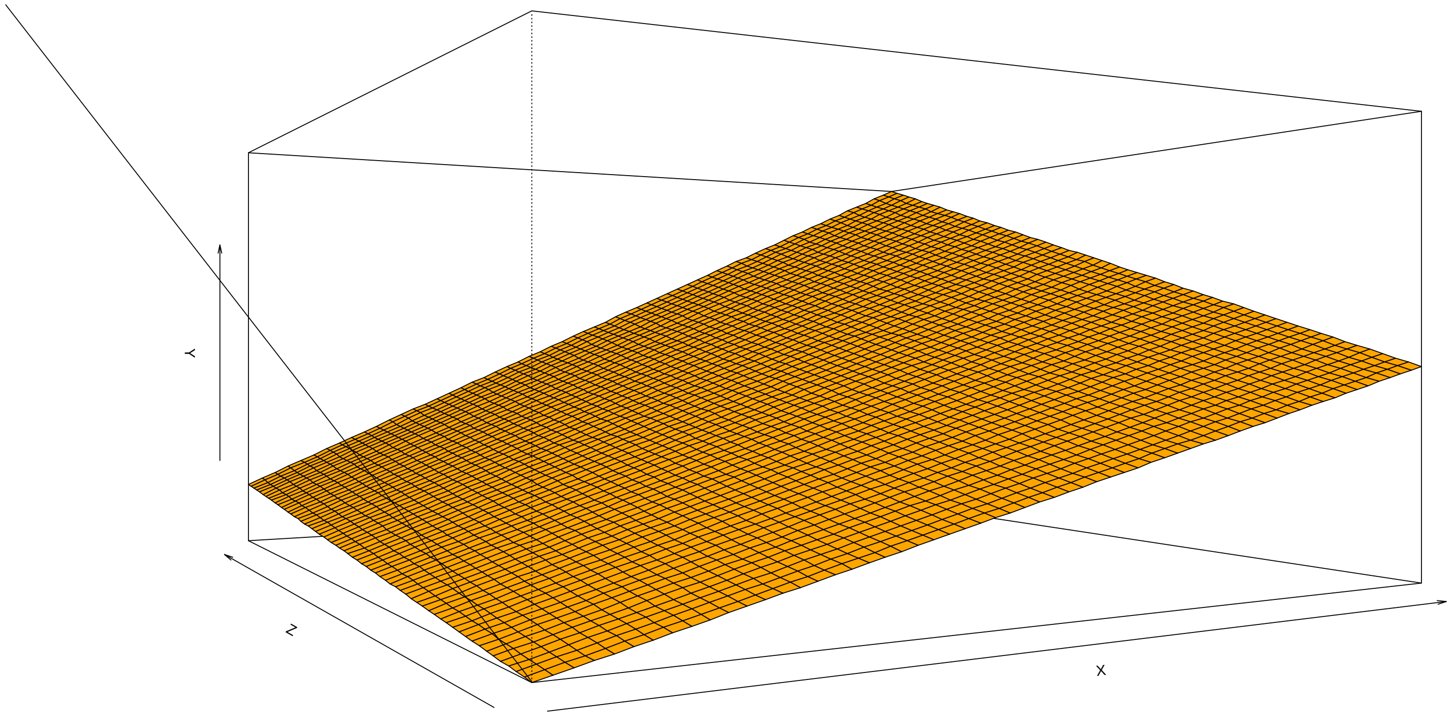
Hayes (2013); Hayes and Matthes (2009)



Moderation

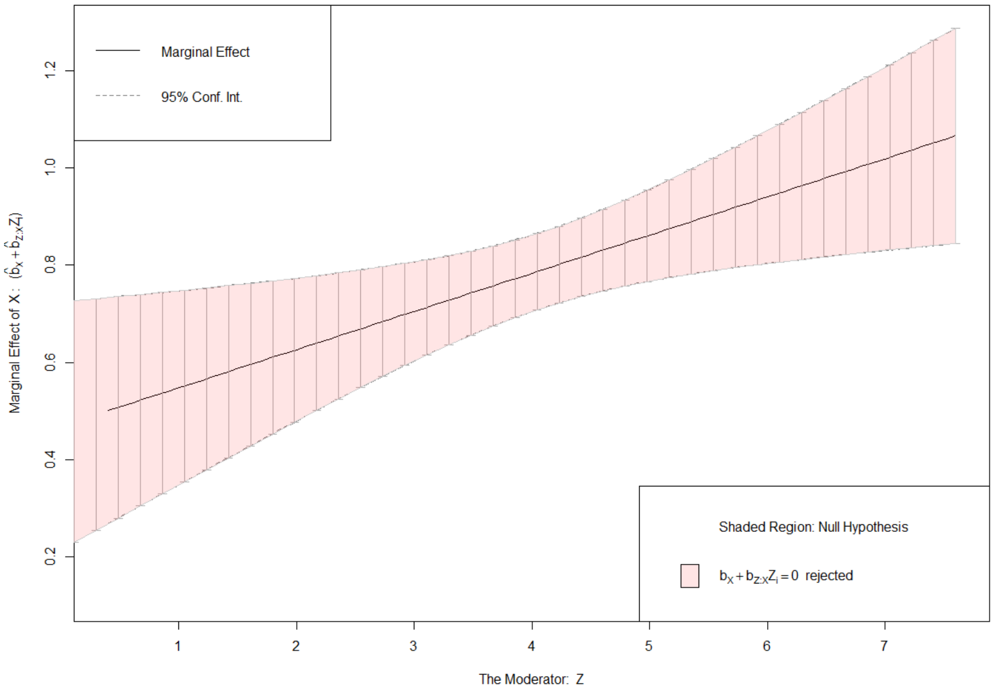
Hayes (2013); Hayes and Matthes (2009)

Surface Plot



Regression Analysis

Graphics



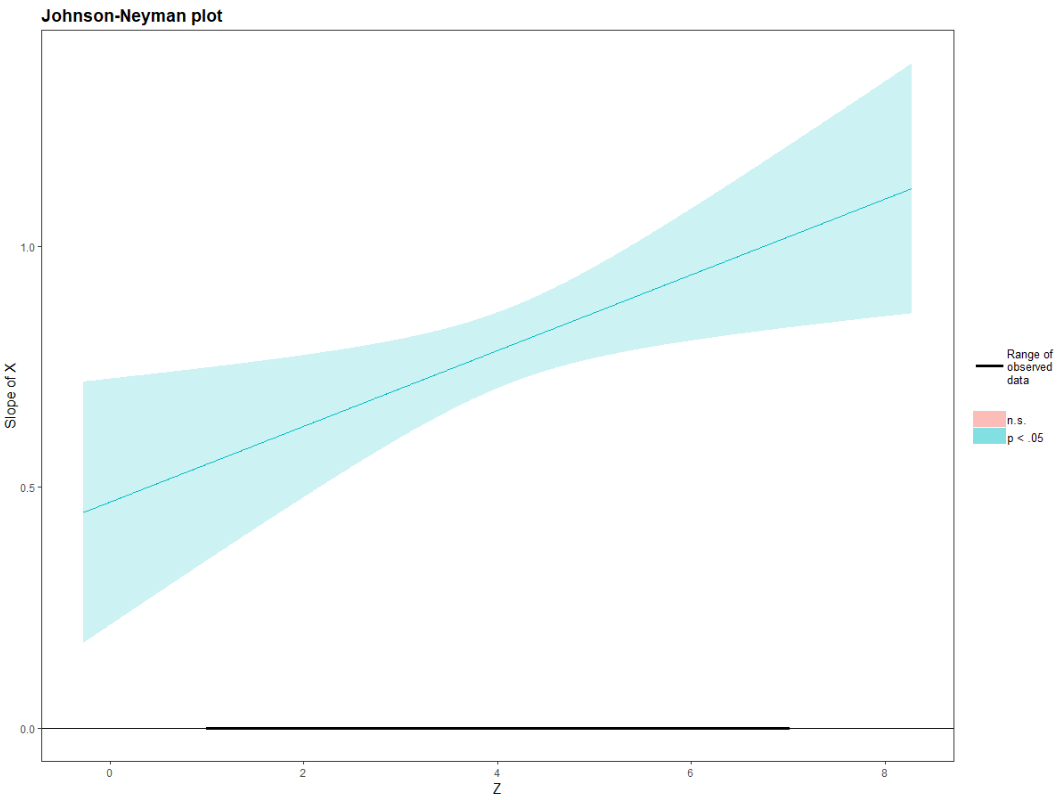
Regression Analysis

R Editor R

```
> # jtools
>
> library(jtools)
>
> MOD3<-lm(Y ~ X*Z, data=Moderation)
>
> johnson_neyman(model = MOD3, pred = X, modx = Z)
JOHNSON-NEYMAN INTERVAL

When Z is OUTSIDE the interval [-36.88, -1.57], the slope of X is p < .05.

Note: The range of observed values of Z is [1.00, 7.00]
```

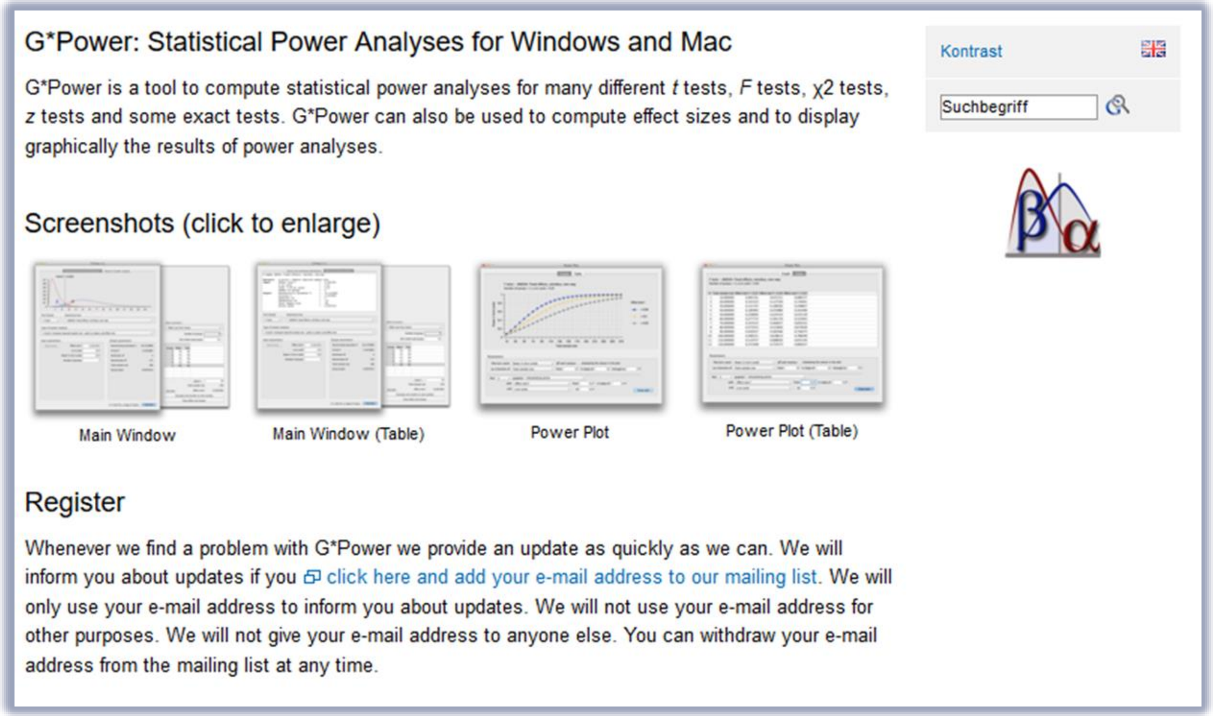
Regression Analysis

Graphics



G*Power

http://www.gpower.hhu.de/



G*Power: Statistical Power Analyses for Windows and Mac

G*Power is a tool to compute statistical power analyses for many different *t* tests, *F* tests, χ2 tests, *z* tests and some exact tests. G*Power can also be used to compute effect sizes and to display graphically the results of power analyses.

Screenshots (click to enlarge)

Main Window | Main Window (Table) | Power Plot | Power Plot (Table)

Register

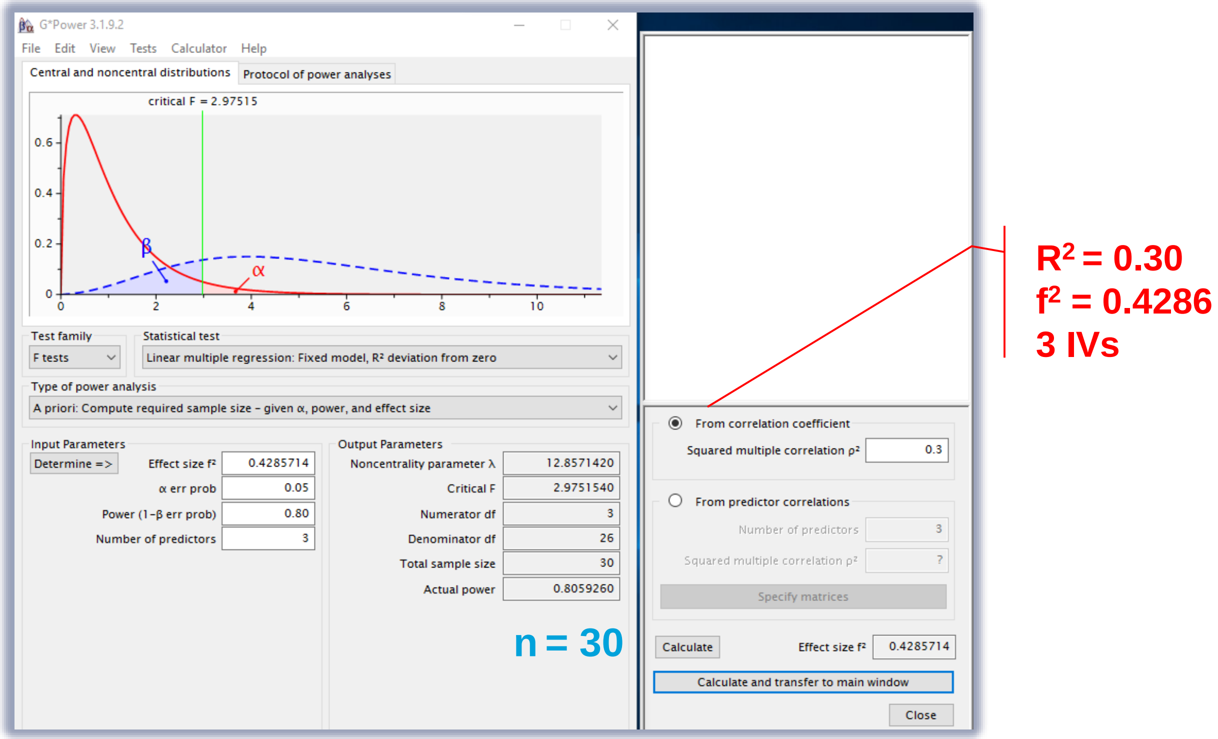
Whenever we find a problem with G*Power we provide an update as quickly as we can. We will inform you about updates if you click here and add your e-mail address to our mailing list. We will only use your e-mail address to inform you about updates. We will not use your e-mail address for other purposes. We will not give your e-mail address to anyone else. You can withdraw your e-mail address from the mailing list at any time.

Kontrast

Suchbegriff

G*Power Sample Size Determination

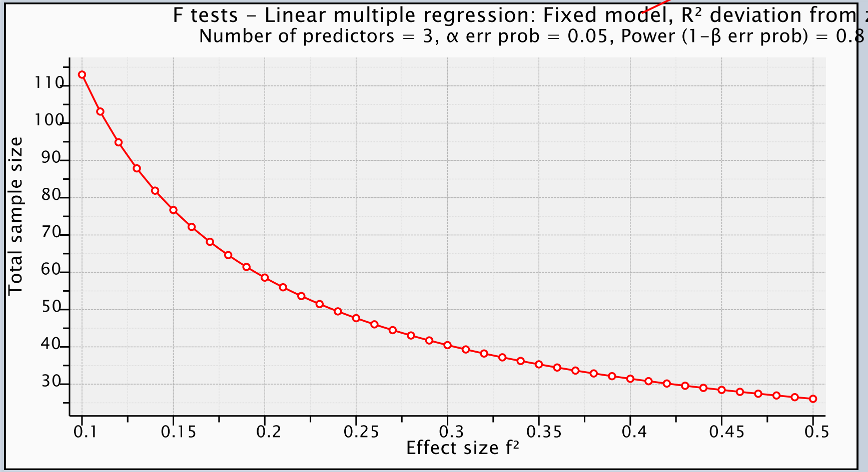
http://www.gpower.hhu.de/



G*Power Sample Size Determination

http://www.gpower.hhu.de/

X-Y plot for range of values



IBM SPSS and R

```
* Encoding: UTF-8.
* pwr

BEGIN PROGRAM R.

# pwr Regression

install.packages("pwr", repos="http://cran.rstudio.com")

library(pwr)

# R2=0.30
# u=3 (independent variables)
# n=v+u+1

pwr.f2.test(u=3, f2=0.30/(1-0.30), sig.level=0.05, power=0.80)


END PROGRAM.
```

IBM SPSS and R

```
     Multiple regression power calculation

              u = 3
              v = 25.65333
             f2 = 0.4285714
      sig.level = 0.05
          power = 0.8
```

n = v + u + 1 = 26 + 3 + 1 = 30

References

Baron, R. M., & Kenny, D. A. (1986). The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical considerations. *Journal of Personality and Social Psychology, 51* (6), 1173.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Cohen, J. (1992). “A Power Primer,” *Psychological Bulletin*, 112 (1), 155-159. [condensed version].

Darlington, R. B., & Hayes, A. F. (2016). *Regression Analysis and Linear Models: Concepts, Applications, and Implementation*. Guilford Publications.

Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Management Science, 35*, 982-1003.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). “G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences,” *Behavior Research Methods, 39* (2), 175-191.

Hair, J.F., Jr., Black, W.C., Babin, B.J, and Anderson, R.E. (2018). *Multivariate Data Analysis.* Cengage.

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. Guilford Press.

Hayes, A.F. and Matthes, J (2009). Computational Procedures for Probing Interactions in OLS and Logistic regression: SPSS and SAS Implementations. *Behavior Research Methods, 41* (3), 924-936.

References

Howell, D. C. (2012). *Statistical Methods for Psychology*. Cengage Learning.

Malhotra, N. (2010). *Marketing Research: An Applied Orientation*. Upper Saddle River: Pearson/Prentice-Hall.

Preacher, K.J. and Hayes, A.F. (2004). SPSS and SAS Procedures for Estimating Effects in Simple Mediation Models. *Behavior Research Methods, Instruments and Computers, 36* (4), 717-731.

Spiller, S. A., Fitzsimons, G. J., Lynch Jr, J. G., & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero: Simple effects tests in moderated regression. *Journal of Marketing Research, 50*(2), 277-288.

References

Chapman, C., & Feit, E. M. (2015). *R for Marketing Research and Analytics.* New York, NY: Springer.

Crawley, MJ (2013). *The R Book.* Chichester, UK: John Wiley and Sons.

Everitt, BS and Hothorn, T (2006). *A Handbook of Statistical Analysis Using R.* Boca Raton, FL: Chapman and Hall/CRC.

Faraday, J.A. (2015). *Linear Models with R*. Boca Raton, FL: CRC Press.

Field, A, Miles, J and Field, Z (2012). *Discovering Statistics Using R.* Los Angeles, CA: Sage Publications.

Kabacoff, RI (2011). *R in Action*. Shelter Island, NY: Manning.

Muenchen, RA (2009). *R for SAS and SPSS Users.* New York, NY; Springer Science and Business Media.

Revelle, W, Psychometric Theory. [http://www.personality-project.org/r/book/]

Thank you for Your Attention!