Salary and Publications Example
(Cohen, Cohen, West, & Aiken Table 3.2.1)

SPSS Syntax

```
get file='c:\jason\spsswin\da2\ccwa3_2_1.sav'.

correlations vars=pubs time salary.

regression vars=salary time pubs
    /descriptives=mean stdev
    /statistics=anova coeff ses r ci
    /dependent=salary
    /method=enter pubs time.
```

Correlations

		number of publications	years since PhD	annual salary in dollars
number of publications	Pearson Correlation	1	.657	.588
	Sig. (2-tailed)		.008	.021
	N	15	15	15
years since PhD	Pearson Correlation	.657	1	.710
	Sig. (2-tailed)	.008		.003
	N	15	15	15
annual salary in dollars	Pearson Correlation	.588	.710	1
	Sig. (2-tailed)	.021	.003	
	N	15	15	15

Note: this table was produced by the first, separate `correlations` *procedure because it has two-tailed tests.*

Descriptive Statistics

	Mean	Std. Deviation	N
annual salary in dollars	53045.6000	7889.76815	15
years since PhD	7.667	4.5774	15
number of publications	19.933	13.8227	15

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	PUBS number of publications, TIME years since PhD[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: SALARY annual salary in dollars

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.728[a]	.530	.452	5839.23054

a. Predictors: (Constant), number of publications, years since PhD

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.623E8	2	2.312E8	6.780	.011[a]
	Residual	4.092E8	12	3.410E7		
	Total	8.715E8	14			

a. Predictors: (Constant), number of publications, years since PhD

b. Dependent Variable: annual salary in dollars

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta	Std. Error			Lower Bound	Upper Bound
1	(Constant)	43082.394	3099.493			13.900	.000	36329.178	49835.610
	years since PhD	982.867	452.057	.570	.262	2.174	.050	-2.081	1967.815
	number of publications	121.801	149.699	.213	.262	.814	.432	-204.364	447.966

a. Dependent Variable: annual salary in dollars

R Code

```
> corrvars = subset(d,select=c(PUBS,TIME,SALARY))
>
> library(correlation)
> correlation(corrvars, p_adjust="none")
# Correlation Matrix (pearson-method)

Parameter1 | Parameter2 |    r |       95% CI | t(13) |       p
-----------------------------------------------------------------
PUBS       |       TIME | 0.66 | [0.22, 0.87] |  3.14 | 0.008**
PUBS       |     SALARY | 0.59 | [0.11, 0.85] |  2.62 | 0.021*
TIME       |     SALARY | 0.71 | [0.31, 0.90] |  3.64 | 0.003**

p-value adjustment method: none
Observations: 15
```

lessR

```
> #clear active frame from previous analyses
> rm(d)
>
> library(haven)
> d = read_sav("c:/jason/spsswin/da2/ccwa3_2_1.sav")

> library(lessR)
> Regression(SALARY ~ PUBS + TIME, brief=TRUE)


Number of cases (rows) of data:  15
Number of cases retained for analysis:  15


  BASIC ANALYSIS

-- Estimated Model for SALARY

             Estimate     Std Err  t-value  p-value   Lower 95%   Upper 95%
(Intercept) 43082.394   3099.493   13.900    0.000   36329.178   49835.610
       PUBS   121.801    149.699    0.814    0.432    -204.364     447.966
       TIME   982.867    452.057    2.174    0.050      -2.081    1967.815

-- Model Fit

Standard deviation of SALARY: 7,889.768

Standard deviation of residuals:  5,839.231 for 12 degrees of freedom
95% range of residual variation:  25,445.181 = 2 * (2.179 * 5,839.231)

R-squared:  0.530    Adjusted R-squared:  0.452    PRESS R-squared:  0.325

Null hypothesis of all 0 population slope coefficients:
  F-statistic: 6.780     df: 2 and 12     p-value:  0.011

-- Analysis of Variance

            df         Sum Sq        Mean Sq   F-value   p-value
```

```
     PUBS     1  301137778.671  301137778.671     8.832     0.012
     TIME     1  161181041.670  161181041.670     4.727     0.050

Model         2  462318820.340  231159410.170     6.780     0.011
Residuals    12  409159359.260   34096613.272
SALARY       14  871478179.600   62248441.400
```

```
> #note that lessR standardized coefficients--ignore SE and significance tests
> Regression(SALARY ~ PUBS + TIME, brief=TRUE, new_scale="z",scale_response=TRUE)

-- Estimated Model for zsalary

             Estimate    Std Err  t-value  p-value   Lower 95%   Upper 95%
(Intercept)    0.0002     0.1911    0.001    0.999     -0.4162      0.4166
      zpubs    0.2135     0.2623    0.814    0.432     -0.3581      0.7851
      ztime    0.5701     0.2623    2.174    0.050     -0.0014      1.1416
```

Base R

```
> #multiple regression using base R function lm
> mod = lm(SALARY ~ PUBS + TIME, data=d)
> summary(mod)
```

Output omitted

```
> library(lm.beta)
> modcoef <- lm.beta(mod)
> coef(modcoef,standardized=TRUE)
(Intercept)        PUBS        TIME
         NA   0.2133919   0.5702261
```

Example Write-up

To examine whether the number of publications was independently associated with annual academic salary after controlling for the number of years since receiving the doctorate, a simultaneous multiple regression analysis was conducted. Results indicated that the number of publications did not independently predict salary, B = 121.80, SE_B = 149.70, 95% CI [-204.36,447.97], β = .21, p = .43. Years since receiving a doctorate was marginally significantly related to salary, B = 982.87, SE_B = 452.06, 95% CI [-2.08,1967.81], β = .57, p = .05, indicating that salary increased by approximately $982 for each additional year since receiving the doctorate. Overall, the number of publications and years since finishing a doctorate accounted for over 50% of the variance in salary, R^2 = .53, $F(2,12)$ = 6.78, p = .01.

Note: Current APA format is to use b for unstandardized and b^* for standardized. B and β are still widely used, however, even in APA journals.