Appendix 2: SAS Code and Output

This appendix contains SAS code and output for all of the analyses described in this article. Readers who have access to SAS can generate that output by running the SAS programs that are available via the second author's website (http://core.ecu.edu/psyc/wuenschk/W&W/W&W-SAS.htm). The data file used by two of the programs, *lung.sas7bdat*, can be downloaded from the UCLA Academic Technology Services website (http://www.ats.ucla.edu/stat/spss/examples/cama4/default.htm). The data are also available at https://sites.google.com/a/lakeheadu.ca/bweaver/Home/statistics/spss/my-spss-page/weaver_wuensch and the second author's website, linked above .

```
/* File name:   1_Pearson_r.txt
Written by:  Karl L. Wuensch, WuenschK@ECU.edu
Date:        8-September-2012.
=====================================================

Test the null hypothesis that rho = a specified value
and compute a confidence interval for rho using
a confidence level specified by the user.

NOTE: When the null hypothesis (H0) states that rho = 0,
  a t-test is used; but when H0 states that rho = a non-zero
  value, a z-test on r-prime is used (where r-prime is obtained
  via Fisher's r-to-z transformation).

Regardless of the value of rho, the confidence interval
on rho is obtained by computing a CI on rho-prime, and
then using the inverse of the r-to-z transformation.

Use correlations between father's height and father's weight
from the 4 areas to test the null hypothesis that rho = .65.


In the INPUT statement below:
r = observed value of Pearson r
rho = population correlation according to the null hypothesis
n = sample size
alpha = value used to set the confidence level for the CI,
with Confidence Level = (1-alpha)*100
Note = a brief description of the data in that row.

Users can replace the data lines (between CARDS and PROC)
with their own data.

**********************************************************/

data rho;
input r rho n alpha Note $30.;
rprime=0.5*log(abs((1+r)/(1-r)));
rhoprime=0.5*log(abs((1+rho)/(1-rho)));
SE=SQRT(1/(n-3));
*Compute confidence interval for rho;
cump=1-alpha/2;
*CV = critical value of z;
CV = PROBIT(cump);
LLprime = rprime - CV*SE; ULprime = rprime + CV*SE;
*LL=(exp(L)-exp(LLprime*-1))/(exp(L)+exp(LLprime*-1));
*UL=(exp(U)-exp(ULprime*-1))/(exp(U)+exp(ULprime*-1));
CI_Lower=(exp(2*LLprime)-1 )/(exp(2*LLprime)+ 1);
CI_Upper=(exp(2*ULprime)-1 )/(exp(2*ULprime)+ 1);
z=(rprime-rhoprime)/SE; zneg = 0-abs(z);
p_z=2*PROBNORM(zneg);
If rho=0 then Do;
```

```
  df=n-2; t=r*SQRT(df)/SQRT(1-r**2);
  tneg=0-abs(t); p_t=2*probt(tneg,df); end;

CARDS;
.628 .000 24 .05 Burbank
.418 .000 49 .05 Lancaster
.438 .000 19 .05 Long Beach
.589 .000 58 .05 Glendora
.628 .650 24 .05 Burbank
.628 .650 24 .01 Burbank, .01
.418 .650 49 .05 Lancaster
.418 .650 49 .01 Lancaster, .01
.438 .650 19 .05 Long Beach
.438 .650 19 .01 Long Beach, .01
.589 .650 58 .05 Glendora
.589 .650 58 .01 Glendora, .01
;
PROC print; var r rho n t df p_t z p_z alpha CI_Lower CI_Upper Note; ID;
title 'Test H0: rho = value, with CI'; run;

/* If rho = 0, TestVal = t, otherwise TestVal = z (and df is missing).
* Confidence level for CI = (1-alpha)*100.
* Notice that when p > alpha, the (1-alpha)*100% CI
* includes the value of rho specified by the null hypothesis.
* But when p < alpha, the (1-alpha)*100% CI does not include
* the value of rho specified by H0. */
```

Test H0: rho = value, with CI

r	rho	n	t	df	p_t	z	p_z	alpha	CI_Lower	CI_Upper	Note
0.628	0.00	24	3.78506	22	0.001017	3.38243	0.00072	0.05	0.30081	0.82292	Burbank
0.418	0.00	49	3.15446	47	0.002802	3.01994	0.00253	0.05	0.15503	0.62566	Lancaster
0.438	0.00	19	2.00887	17	0.060706	1.87901	0.06024	0.05	-0.02023	0.74416	Long Beach
0.589	0.00	58	5.45414	56	0.000001	5.01434	0.00000	0.05	0.39004	0.73541	Glendora
0.628	0.65	24	.	.	.	-0.17043	0.86467	0.05	0.30081	0.82292	Burbank
0.628	0.65	24	.	.	.	-0.17043	0.86467	0.01	0.17422	0.86177	Burbank, .01
0.418	0.65	49	.	.	.	-2.23839	0.02520	0.05	0.15503	0.62566	Lancaster
0.418	0.65	49	.	.	.	-2.23839	0.02520	0.01	0.06539	0.67781	Lancaster, .01
0.438	0.65	19	.	.	.	-1.22218	0.22164	0.05	-0.02023	0.74416	Long Beach
0.438	0.65	19	.	.	.	-1.22218	0.22164	0.01	-0.17246	0.80537	Long Beach, .01
0.589	0.65	58	.	.	.	-0.73543	0.46208	0.05	0.39004	0.73541	Glendora
0.589	0.65	58	.	.	.	-0.73543	0.46208	0.01	0.31745	0.77127	Glendora, .01

```
/*=========================================================
File name:   2_single_regression_coefficient.txt
Written by:  Karl L. Wuensch, WuenschK@ECU.edu
Date:        28-August-2012.
=====================================================

Use intercepts and slopes from Table 2.
First test H0: intercept = 0.
Then test HO: slope = 0.
Then test H0: intercept = 145.
Then test H0: slope = 3.5.

In the INPUT statement below:
  b = the value of an OLS regression coefficient
  bstar = population value of b according to the null hypothesis
  se = the standard error of b
  df = degrees of freedom associated with b
  alpha = value used to set the confidence level for the CI,
  with Confidence Level = (1-alpha)*100
*  Note = a brief description of the data in that row.

Users can replace the data lines (between CARDS and PROC)
  with their own data. */
data ab;
input b bstar se df alpha Note $35.;
t=(b-bstar)/se;
tneg=0-abs(t);
p=2*probt(tneg,df);
cump=1-alpha/2;
CV=TINV(cump,df);
CI_Lower=b-CV*se; CI_Upper=b+CV*se;

CARDS;
142.011   0   10.664  22 .05  Intercept, Burbank
148.053   0   11.142  47 .05  Intercept, Lancaster
144.038   0   18.250  17 .05  Intercept, Long Beach
130.445   0   10.228  56 .05  Intercept, Glendora
  4.179   0    1.105  22 .05  Slope, Burbank
  3.709   0    1.177  47 .05  Slope, Lancaster
  3.749   0    1.866  17 .05  Slope, Long Beach
  5.689   0    1.044  56 .05  Slope, Glendora
142.011 145.0 10.664  22 .05  Intercept, Burbank
148.053 145.0 11.142  47 .05  Intercept, Lancaster
144.038 145.0 18.250  17 .05  Intercept, Long Beach
130.445 145.0 10.228  56 .05  Intercept, Glendora
  4.179   3.5  1.105  22 .05  Slope, Burbank
  3.709   3.5  1.177  47 .05  Slope, Lancaster
  3.749   3.5  1.866  17 .05  Slope, Long Beach
  5.689   3.5  1.044  56 .05  Slope, Glendora
;
PROC print; var b bstar se t df p CI_Lower CI_Upper Note; ID;
title 'Test HO: regression parameter = value with CI’; run;
```

Test HO: regression parameter = value with CI

b	bstar	se	t	df	p	CI_Lower	CI_Upper	Note
142.011	0.0	10.664	13.3169	22	0.00000	119.895	164.127	Intercept, Burbank
148.053	0.0	11.142	13.2878	47	0.00000	125.638	170.468	Intercept, Lancaster
144.038	0.0	18.250	7.8925	17	0.00000	105.534	182.542	Intercept, Long Beach
130.445	0.0	10.228	12.7537	56	0.00000	109.956	150.934	Intercept, Glendora
4.179	0.0	1.105	3.7819	22	0.00103	1.887	6.471	Slope, Burbank
3.709	0.0	1.177	3.1512	47	0.00283	1.341	6.077	Slope, Lancaster
3.749	0.0	1.866	2.0091	17	0.06068	-0.188	7.686	Slope, Long Beach
5.689	0.0	1.044	5.4492	56	0.00000	3.598	7.780	Slope, Glendora
142.011	145.0	10.664	-0.2803	22	0.78187	119.895	164.127	Intercept, Burbank
148.053	145.0	11.142	0.2740	47	0.78528	125.638	170.468	Intercept, Lancaster
144.038	145.0	18.250	-0.0527	17	0.95858	105.534	182.542	Intercept, Long Beach
130.445	145.0	10.228	-1.4231	56	0.16027	109.956	150.934	Intercept, Glendora
4.179	3.5	1.105	0.6145	22	0.54520	1.887	6.471	Slope, Burbank
3.709	3.5	1.177	0.1776	47	0.85982	1.341	6.077	Slope, Lancaster
3.749	3.5	1.866	0.1334	17	0.89541	-0.188	7.686	Slope, Long Beach
5.689	3.5	1.044	2.0967	56	0.04055	3.598	7.780	Slope, Glendora

---

```
/*======================================================
File name:    3_two_independent_correlations.txt
Written by:   Karl L. Wuensch, WuenschK@ECU.edu
Date:         10-September-2012.
======================================================


z-test for comparing two independent correlations
Confidence intervals for rho1 and rho2
Confidence interval for difference between rho1 and rho2 (Zou (2007)

This syntax works only for the case where the null
hypothesis specifies that the difference between the
population correlations = 0.

Use correlation between FHEIGHT and FWEIGHT.
Compare one city with another.

In the INPUT statement below:
r1 = first correlation
r2 = second correlation (r1 and r2 must be independent of each other)
n1 = first sample size
n2 = second sample size
alpha = value used to set the confidence level for the CI on b1-b2,
  with Confidence Level = (1-alpha)*100
Note = a brief description of the data in that row.

Users can replace the data lines (between CARDS and PROC)
with their own data. */
```

```
data rho1I2;
input r1 r2 n1 n2 alpha Note $40.;
rprime1=0.5*log(abs((1+r1)/(1-r1))); rprime2=0.5*log(abs((1+r2)/(1-r2)));
se1=SQRT(1/(n1-3)); se2=SQRT(1/(n2-3));
rprimediff=rprime1-rprime2; sediff= SQRT(se1**2+se2**2);
z=rprimediff/sediff;
zneg=0-abs(z); p=2*PROBNORM(zneg);
cump=1-alpha/2;
*CV = critical value of z;
CV = PROBIT(cump);
*Confidence intervals for rho 1 and rho 2;
LPrime1=rprime1-CV*se1; UPrime1=rprime1+CV*se1;
LPrime2=rprime2-CV*se2; UPrime2=rprime2+CV*se2;
CI_1_Lower=(exp(2*LPrime1)-1)/(exp(2*LPrime1)+1);
CI_1_Upper=(exp(2*UPrime1)-1)/(exp(2*UPrime1)+1);
CI_2_Lower=(exp(2*LPrime2)-1)/(exp(2*LPrime2)+1);
CI_2_Upper=(exp(2*UPrime2)-1)/(exp(2*UPrime2)+1);
*Use method of Zou (2007) to compute CI for rho1-rho2 difference;
CI_Lower=r1-r2-sqrt((r1-CI_1_Lower)**2+(CI_2_Upper-r2)**2);
CI_Upper=r1-r2+SQRT((CI_1_Upper-r1)**2+(r2-CI_2_Lower)**2);

CARDS;
.4180 .5890 49 58 .05 r(FHT,FWT), Lancaster v Glendora
 .0400 .3640 49 58 .05 r(MHT,MWT), Lancaster v Glendora
 .1980 .3660 49 58 .05 r(FHT,MHT), Lancaster v Glendora
 .2990 .2090 49 58 .05 r(FWT,MWT), Lancaster v Glendora
-.1810 .3300 49 58 .05 r(FWT,MHT), Lancaster v Glendora
-.1810 .3300 49 58 .01 r(FWT,MHT), Lancaster v Glendora, .01
 .0650 .0710 49 58 .05 r(FHT,MWT), Lancaster v Glendora
.49 .36 145 87 .05 Zou (2007) Example 1
;
PROC print; var r1 r2 z p Note; ID;
title 'Test of H0: rho1 = rho2, independent samples'; run;
proc print; var r1 CI_1_Lower CI_1_Upper r2 CI_2_Lower CI_2_Upper alpha note;
title 'Confidence intervals for rho 1 and rho 2'; ID;
proc print; var r1 r2 alpha CI_Lower CI_Upper Note; ID;
title 'Confidence interval for difference between rho1 and rho2'; run;

/* Zou, G. Y. (2007).  Toward using confidence intervals to compare correlations.
Psychological Bulletin, 12, 399-413. */
```

Test of H0: rho1 = rho2, independent samples

r1	r2	z	p	Note
0.418	0.589	-1.15548	0.24789	r(FHT,FWT), Lancaster v Glendora
0.040	0.364	-1.70903	0.08745	r(MHT,MWT), Lancaster v Glendora
0.198	0.366	-0.91664	0.35933	r(FHT,MHT), Lancaster v Glendora
0.299	0.209	0.48195	0.62984	r(FWT,MWT), Lancaster v Glendora
-0.181	0.330	-2.63183	0.00849	r(FWT,MHT), Lancaster v Glendora
-0.181	0.330	-2.63183	0.00849	r(FWT,MHT), Lancaster v Glendora, .01
0.065	0.071	-0.03017	0.97593	r(FHT,MWT), Lancaster v Glendora
0.490	0.360	1.15639	0.24752	Zou (2007) Example 1

Confidence intervals for rho 1 and rho 2							
r1	**CI_1_Lower**	**CI_1_Upper**	**r2**	**CI_2_Lower**	**CI_2_Upper**	**alpha**	**Note**
0.418	0.15503	0.62566	0.589	0.39004	0.73541	0.05	r(FHT,FWT), Lancaster v Glendora
0.040	-0.24394	0.31762	0.364	0.11667	0.56882	0.05	r(MHT,MWT), Lancaster v Glendora
0.198	-0.08810	0.45392	0.366	0.11895	0.57037	0.05	r(FHT,MHT), Lancaster v Glendora
0.299	0.01944	0.53520	0.209	-0.05211	0.44336	0.05	r(FWT,MWT), Lancaster v Glendora
-0.181	-0.43981	0.10557	0.330	0.07839	0.54209	0.05	r(FWT,MHT), Lancaster v Glendora
-0.181	-0.51005	0.19427	0.330	-0.00450	0.59808	0.01	r(FWT,MHT), Lancaster v Glendora, .01
0.065	-0.22022	0.33998	0.071	-0.19079	0.32337	0.05	r(FHT,MWT), Lancaster v Glendora
0.490	0.35538	0.60471	0.360	0.16161	0.53042	0.05	Zou (2007) Example 1

Confidence interval for difference between rho1 and rho2					
r1	**r2**	**alpha**	**CI_Lower**	**CI_Upper**	**Note**
0.418	0.589	0.05	-0.47199	0.11658	r(FHT,FWT), Lancaster v Glendora
0.040	0.364	0.05	-0.67410	0.04781	r(MHT,MWT), Lancaster v Glendora
0.198	0.366	0.05	-0.51960	0.18771	r(FHT,MHT), Lancaster v Glendora
0.299	0.209	0.05	-0.27480	0.44209	r(FWT,MWT), Lancaster v Glendora
-0.181	0.330	0.05	-0.84561	-0.12964	r(FWT,MHT), Lancaster v Glendora
-0.181	0.330	0.01	-0.93543	-0.00829	r(FWT,MHT), Lancaster v Glendora, .01
0.065	0.071	0.05	-0.38684	0.37367	r(FHT,MWT), Lancaster v Glendora
0.490	0.360	0.05	-0.08718	0.35917	Zou (2007) Example 1

```
/*==========================================================
File name:   4_two_independent_regression_coefficients.txt
Written by:  Karl L. Wuensch, WuenschK@ECU.edu
Date:        7-February-2013.
==========================================================

* Use data from Table 2.  Compare regression coefficients for Lancaster and Glendora.
* "K&K" uses data from Kleinbaum & Kupper <Applied regression analysis and other
multivariable
   methods, 1978, Boston: Duxbury>, Table 8.1.
* m is the number of predictors in the regression model.
* Assign to "Pool" value 1 if you wish the error terms and degrees of freedom to be pooled
between groups;
* Assign value 1 if you wish them not to be pooled (df with Satterthwaite adjustment);
* Users can replace the data lines (between CARDS and PROC)
  with their own data. */
```

```
data b1Ib2;
input Pool m b1 b2 se1 se2 MSE1 MSE2 n1 n2 alpha Note $20.;
bdiff=(b1-b2);
df1 = n1-m-1; df2 = n2-m-1;
If pool = 0 then goto Satter;
df=df1+df2;
MSE=((n1-m-1)*MSE1+(n2-m-1)*MSE2)/df;
sediff = SQRT(MSE*(se1**2/MSE1+se2**2/MSE2));
GOTO T;
Satter: v1=se1*se1; v2=se2*se2;
df=(v1+v2)**2/(v1**2/df1+v2**2/df2);
sediff=SQRT(v1+v2);
T: t=bdiff/sediff;
tneg=0-abs(t); p=2*probt(tneg,df);
cump=1-alpha/2;
CV=TINV(cump,df);
CI_Lower=bdiff-CV*sediff; CI_Upper=bdiff+CV*sediff;
CARDS;
1 1 148.053 130.445 11.142 10.228 457.956 407.826 49 58 .05 Int,   Lan v Glen
0 1 148.053 130.445 11.142 10.228 457.956 407.826 49 58 .05 Int,   Lan v Glen
1 1 3.709   5.689   1.177  1.044  457.956 407.826 49 58 .05 Slope, Lan v Glen
0 1 3.709   5.689   1.177  1.044  457.956 407.826 49 58 .05 Slope, Lan v Glen
1 1 0.9493 0.96135 0.116145 0.0913 91.457 71.897 29 40 .05 Slope,  K&K
0 1 0.9493 0.96135 0.116145 0.0913 91.457 71.897 29 40 .05 Slope,  K&K
;
PROC print; var Pool b1 b2 bdiff sediff t df p Note; id;
title 'Difference between two independent slopes or intercepts';
proc print; var Pool b1 b2 bdiff alpha CI_Lower CI_Upper Note;
title 'CI for Difference between two independent slopes or intercepts'; run;
```

Difference between two independent slopes or intercepts

Pool	b1	b2	bdiff	sediff	t	df	p	Note
1	148.053	130.445	17.6080	15.0743	1.16808	103.000	0.24547	Int, Lan v Glen
0	148.053	130.445	17.6080	15.1247	1.16419	99.993	0.24712	Int, Lan v Glen
1	3.709	5.689	-1.9800	1.5665	-1.26396	103.000	0.20910	Slope, Lan v Glen
0	3.709	5.689	-1.9800	1.5733	-1.25850	98.748	0.21118	Slope, Lan v Glen
1	0.949	0.961	-0.0121	0.1452	-0.08299	65.000	0.93411	Slope, K&K
0	0.949	0.961	-0.0121	0.1477	-0.08157	55.595	0.93529	Slope, K&K

CI for Difference between two independent slopes or intercepts

Obs	Pool	b1	b2	bdiff	alpha	CI_Lower	CI_Upper	Note
1	1	148.053	130.445	17.6080	0.05	-12.2884	47.5044	Int, Lan v Glen
2	0	148.053	130.445	17.6080	0.05	-12.3990	47.6150	Int, Lan v Glen
3	1	3.709	5.689	-1.9800	0.05	-5.0868	1.1268	Slope, Lan v Glen
4	0	3.709	5.689	-1.9800	0.05	-5.1019	1.1419	Slope, Lan v Glen
5	1	0.949	0.961	-0.0121	0.05	-0.3020	0.2779	Slope, K&K
6	0	0.949	0.961	-0.0121	0.05	-0.3080	0.2839	Slope, K&K

```
/*=============================================================
File name:   5_k_independent_parameters.txt
Written by:  Karl L. Wuensch, WuenschK@ECU.edu
Date:        28-August-2012.
=============================================================

 To test the null hypothesis that k independent
 parameters all have the same value, one can use
 the chi-square test of heterogenity that is commonly
 used in meta-analysis.  See Fleiss's article on
 meta-analysis for details.

 -----------------------------------------------------------
 Fleiss, JL. The statistical basis of meta-analysis.
   Statistical Methods in Medical Research, 1993, 2: 121-145.
 -----------------------------------------------------------

 For this method, one needs k independent effect size estimates
 (called Y) along with their variances (i.e., the square of their
 standard errors).  Each one of those estimates appears on a separate
 row in the data file.  The INPUT data are:

   Group - consecutive integers starting with 1
   Y   - the measure of effect size
   seY - the SE of Y
   n   - the sample size
   Note - a string variable with a note.

 In some cases, Y and se.Y may have to be computed
 from other variables -- e.g., when testing the
 heterogeneity of correlation coefficients, we compute
 Y = the Fisher r-to-z transformed value of Pearson r.

Users can replace the data lines (between "CARDS" and "%MACRO ComputeQ") with their own
data.

*Create macro for computing Q. */

%MACRO ComputeQ;
proc means noprint; var Group w wy;
 output out=sums sum=x1 sumW sumWY max=S x2 x3;
data sums2; set sums;
Do Group=1 to S; sw=sumw; swy=sumWY; output; end; keep group S sw swy;
data Qgroup; merge xyzzy sums2; by Group;
Ybar=swy/sw; Qgroup=w*(Y-Ybar)**2;
proc means noprint; var Group Qgroup; output out=Q sum=nix1 Q max=S nix2;
data p; set Q;
df=S-1; p=1-probchi(Q,df);
proc print; var Q df p; ID;
%MEND ComputeQ;

* Test the heterogeneity of the INTERCEPTS in Table 2;

data xyzzy;
input Group Y seY n Note $35.;
varY=seY*seY; w=1/varY; wy=w*y;

CARDS;
1 142.011   10.664  22  Intercept, Burbank
```

```
2 148.053   11.142  47  Intercept, Lancaster
3 144.038   18.250  17  Intercept, Long Beach
4 130.445   10.228  56  Intercept, Glendora
;
%ComputeQ
title 'Comparing Intercepts, k Independent Groups'; run;

* Now test the heterogeneity of the SLOPES in Table 2;

data xyzzy;
input Group Y seY n Note $35.;
varY=seY*seY; w=1/varY; wy=w*y;
*GROUPS MUST BE CONSECUTIVE INTEGERS STARTING WITH 1;
CARDS;
1  4.179   1.105  22  Slope, Burbank
2  3.709   1.177  47  Slope, Lancaster
3  3.749   1.866  17  Slope, Long Beach
4  5.689   1.044  56  Slope, Glendora
;
%ComputeQ
title 'Comparing Slopes, k Independent Groups'; run;

/**** Test equivalence of correlations ***

 In the case where one has k independent Pearson
 correlations, Y = r-prime, the Fisher r-to-z transformed
 correlation.  The variance of Y = 1/(n-3).

 We illustrate using correlations between FHEIGHT and FWEIGHT
 in 4 different areas -- see Table 1 in the article. */

data xyzzy;
input Group r n;
Y = 0.5*log(abs((1+r)/(1-r)));
varY=1/(n-3);
w=1/varY;
wy=w*y;
*GROUPS MUST BE CONSECUTIVE INTEGERS STARTING WITH 1;
CARDS;
1 .628 24
2 .418 49
3 .438 19
4 .589 58
;
%ComputeQ
title 'Comparing Correlation Coefficients, k Independent Groups'; run;

/* Below, we run this code on the 2 independent correlations
* between FWT and MHT in Lancaster and in Glendora.
* The Q-value we obtain below should be equal to the square of
* the z-value (2.63183**2 = 6.9265) obtained earlier and the one-tailed p from Q
* should be identical to the two-tailed p (.00849) from z. */

data xyzzy;
input Group r n;
Y = 0.5*log(abs((1+r)/(1-r)));
varY=1/(n-3);
w=1/varY;
wy=w*y;
CARDS;
1 -.181 49
2  .330 58
;
```

```
%ComputeQ
title 'Comparing Correlation Coefficients, k=2 Independent Groups'; run;
```

Comparing Intercepts, k Independent Groups

Q	df	p
1.47925	3	0.68707

Comparing Slopes, k Independent Groups

Q	df	p
1.99446	3	0.57356

Comparing Correlation Coefficients, k Independent Groups

Q	df	p
2.05974	3	0.56010

Comparing Correlation Coefficients, k=2 Independent Groups

Q	df	p
6.92650	1	.008492759

```
/*===========================================================
File name:   6_Williams_test.txt
Written by:  Karl L. Wuensch, WuenschK@ECU.edu
Date:        10-September-2012.
=============================================================

The user wants the difference r12 - r13.

 Let r12 = correlation between father's height and mother's height.
 Let r13 = correlation between father's height and mother's weight.
 Therefore, r23 = correlation between mother's height and mother's weight.
 Compare r12 and r13 in each of the 4 areas. */

data williams; input r12 r13 r23 n alpha Note $25.;
R=(1-r12**2-r13**2-r23**2)+2*r12*r13*r23;
diff=r12-r13;
t=diff*SQRT(((n-1)*(1+r23))/(2*((n-1)/(n-3))*R+((r12+r13)**2/4)*(1-r23)**3));
SEdiff=1/SQRT(((n-1)*(1+r23))/(2*((n-1)/(n-3))*R+((r12+r13)**2/4)*(1-r23)**3));
df=n-3;
tneg=0-abs(t);
p=2*probt(tneg,df);
*Use method of Zou (2007, p. 409) to compute CI for rho12 - rho13;
num=(r23-.5*r12*r13)*(1-r12**2-r13**2-r23**2)+r23**3;
denom=(1-r12**2)*(1-r13**2);
corr12_13=num/denom;
cump=1-alpha/2;
CV=PROBIT(cump);
*CV = critical value of z;
SE=SQRT(1/(n-3));
rprime12=0.5*log(abs((1+r12)/(1-r12)));
```

```
L12p=rprime12 - CV*SE; * lower limit of 95% CI for rho-prime;
U12p=rprime12 + CV*SE; * upper limit of 95% CI for rho-prime;
L12=(exp(2*L12p)-1)/(exp(2*L12p)+1);
U12=(exp(2*U12p)-1)/(exp(2*U12p)+1);
rprime13=0.5*log(abs((1+r13)/(1-r13)));
L13p=rprime13-CV*SE; * lower limit of 95% CI for rho-prime;
U13p=rprime13+CV*SE; * upper limit of 95% CI for rho-prime;
L13=(exp(2*L13p)-1)/(exp(2*L13p)+1);
U13=(exp(2*U13p)-1)/(exp(2*U13p)+1);
Label L12='Lower CI rho12'; Label L13='Lower CI rho13';
Label U12='Upper CI rho12'; Label U13='Upper CI rho13';
CI_Lower=r12-r13-SQRT((r12-L12)**2+(U13-r13)**2-2*corr12_13*(r12-L12)*(U13-r13));
CI_Upper=r12-r13+SQRT((U12-r12)**2+(r13-L13)**2-2*corr12_13*(U12-r12)*(r13-L13));

CARDS;
.164 -.189 .624  24 .05 Burbank
.198  .065 .040  49 .05 Lancaster
.412  .114 .487  19 .05 Long Beach
.366  .071 .364  58 .05 Glendora
.366  .071 .364  58 .01 Glendora
.396  .179 .088 66 .05 Zou (2007) Example 2
;
proc print Label; var r12 L12 U12 r13 L13 U13 alpha Note; ID;
 Title 'Confidence Intervals for rho-12 and rho-13'; run;
proc print; var r12 r13 t df p CI_Lower CI_Upper alpha Note; ID; title 'Williams'' test';
run;
```

Confidence Intervals for rho-12 and rho-13

r12	Lower CI rho12	Upper CI rho12	r13	Lower CI rho13	Upper CI rho13	alpha	Note
0.164	-0.25636	0.53219	-0.189	-0.55043	0.23209	0.05	Burbank
0.198	-0.08810	0.45392	0.065	-0.22022	0.33998	0.05	Lancaster
0.412	-0.05193	0.72966	0.114	-0.35879	0.54024	0.05	Long Beach
0.366	0.11895	0.57037	0.071	-0.19079	0.32337	0.05	Glendora
0.366	0.03646	0.62375	0.071	-0.26939	0.39562	0.01	Glendora
0.396	0.17029	0.58223	0.179	-0.06589	0.40355	0.05	Zou (2007) Example 2

Williams' test

r12	r13	t	df	p	CI_Lower	CI_Upper	alpha	Note
0.164	-0.189	2.04251	21	0.05386	-0.00774	0.66586	0.05	Burbank
0.198	0.065	0.66274	46	0.51080	-0.25711	0.50976	0.05	Lancaster
0.412	0.114	1.29494	16	0.21372	-0.16178	0.72642	0.05	Long Beach
0.366	0.071	2.08160	55	0.04205	0.01085	0.56437	0.05	Glendora
0.366	0.071	2.08160	55	0.04205	-0.07717	0.64200	0.01	Glendora
0.396	0.179	1.38095	63	0.17217	-0.09290	0.51676	0.05	Zou (2007) Example 2

```
/*=========================================================
File name:   7_ZPF.txt
Written by:  Karl L. Wuensch, WuenschK@ECU.edu
Date:        7. February 2013
===========================================================
 For the following correlation matrix, the correlations to be compared are r12 and r34

      X1   X2   X3   X4
  X1  --  r12  r13  r14
  X2  --   --  r23  r24
  X3  --   --   --  r34

 Users can replace the data lines (between CARDS and PROC) with their own data.
 Each data row represents the matrix for a single group.

 "Steiger" is whether to apply (1) or not apply (0) Steiger's (1980) modification,
 which is to set r12 and r34 to the mean of r12 and r34 when estimating the
 standard errors of the difference between r12 and r34 under the null hypothesis
 of no difference.  This modification is not appropriate and is not used when
 constructing the confidence interval for the difference.

 "N" is the sample size, and "Note" is a descriptive string.

 The correlation coefficients comprise the remainder of the input data.

*****************************************************************************/

data ZPF; input Steiger N r12 r13 r14 r23 r24 r34 alpha Note $15.;
If Steiger = 0 then do; r12S = r12; r34S = r34; end;
Else If Steiger = 1 then do; r12S=(r12+r34)/2; r34S=r12S; end;
k=(r13-r23*r12S)*(r24-r23*r34S)+(r14-r13*r34S)*(r23-r13*r12S)+
 (r13-r14*r34S)*(r24-r14*r12S)+(r14-r12S*r24)*(r23-r24*r34S);
Z12=.5*log((1+r12)/(1-r12)); Z34=.5*log((1+r34)/(1-r34));
ZPF = sqrt((n-3)/2)*(Z12-Z34)/sqrt(1-(k/(2*(1-r12S**2)*(1-r34S**2))));
*Remove leading asterisk on next two lines if you want PF;
*PF=(r12-r34)*SQRT(n)/SQRT((1-r12**2)**2+(1-r34**2)**2-k);
*p_PF=2*PROBNORM(-1*ABS(PF));
p_ZPF=2*PROBNORM(-1*ABS(ZPF));

*Use method of Zou (2007, p. 409-410) to compute CI for rho12 - rho34;
num=.5*r12*r34*(r13**2 + r14**2 + r23**2 + r24**2) + r13*r24 + r14*r23 -
      (r12*r13*r14 + r12*r23*r24 + r13*r23*r34 + r14*r24*r34);
denom=(1 - r12**2)*(1 - r34**2);
corr12_34=num/denom;
cump=1-alpha/2;
CV=PROBIT(cump);
*CV = critical value of z;
SE=SQRT(1/(n-3));
rprime12=0.5*log(abs((1+r12)/(1-r12)));
L12p=rprime12 - CV*SE; * lower limit of 95% CI for rho-prime;
U12p=rprime12 + CV*SE; * upper limit of 95% CI for rho-prime;
L12=(exp(2*L12p)-1)/(exp(2*L12p)+1);
U12=(exp(2*U12p)-1)/(exp(2*U12p)+1);
rprime34=0.5*log(abs((1+r34)/(1-r34)));
L34p=rprime34-CV*SE; * lower limit of 95% CI for rho-prime;
U34p=rprime34+CV*SE; * upper limit of 95% CI for rho-prime;
L34=(exp(2*L34p)-1)/(exp(2*L34p)+1);
U34=(exp(2*U34p)-1)/(exp(2*U34p)+1);
Label L12='Lower CI rho12'; Label L34='Lower CI rho34';
Label U12='Upper CI rho12'; Label U34='Upper CI rho34';
CI_Lower=r12-r34-SQRT((r12-L12)**2+(U34-r34)**2-2*corr12_34*(r12-L12)*(U34-r34));
CI_Upper=r12-r34+SQRT((U12-r12)**2+(r34-L34)**2-2*corr12_34*(U12-r12)*(r34-L34));
```

```
CARDS;
0 24 .628 .164 -.189 -.145 -.201 .624 .05 Burbank
1 24 .628 .164 -.189 -.145 -.201 .624 .05 Burbank
0 49 .418 .198  .065 -.181  .299 .040 .05 Lancaster
1 49 .418 .198  .065 -.181  .299 .040 .05 Lancaster
0 19 .438 .412  .114 -.032  .230 .487 .05 Long Beach
1 19 .438 .412  .114 -.032  .230 .487 .05 Long Beach
0 58 .589 .366  .071  .330  .209 .364 .05 Glendora
1 58 .589 .366  .071  .330  .209 .364 .05 Glendora
0 150  .521 .281  .043  .042  .199 .316 .05 All Areas
1 150  .521 .281  .043  .042  .199 .316 .05 All Areas
0 66  .396 .208  .143  .023  .423 .189 .05 Zou Example 3
1 66  .396 .208  .143  .023  .423 .189 .05 Zou Example 3
0 603 .38 .45   .53   .31   .55  .25  .05 Wuensch example
1 603 .38 .45   .53   .31   .55  .25  .05 Wuensch example
0 103 .50 .70 .50 .50 .80 .60 .05 Steiger B
1 103 .50 .70 .50 .50 .80 .60 .05 Steiger B
;
/* The Wuensch example is from the document found at
http://core.ecu.edu/psyc/wuenschk/stathelp/ZPF.docx.
"Steiger B" is from Steiger <Tests for comparing elements of a correlation matrix,
Psychological Bulletin, 1980, 245-251>, Case B. */
proc print Label; var Steiger r12 L12 U12 r34 L34 U34 alpha Note; ID;
 Title 'Confidence Intervals for rho-12 and rho-34'; run;
PROC print; var Steiger r12 r34 ZPF -- p_ZPF CI_Lower CI_Upper alpha Note; ID;
title 'ZPF test with confidence intervals'; run;
```

Confidence Intervals for rho-12 and rho-34

Steiger	r12	Lower CI rho12	Upper CI rho12	r34	Lower CI rho34	Upper CI rho34	alpha	Note
0	0.628	0.30081	0.82292	0.624	0.29481	0.82079	0.05	Burbank
1	0.628	0.30081	0.82292	0.624	0.29481	0.82079	0.05	Burbank
0	0.418	0.15503	0.62566	0.040	-0.24394	0.31762	0.05	Lancaster
1	0.418	0.15503	0.62566	0.040	-0.24394	0.31762	0.05	Lancaster
0	0.438	-0.02023	0.74416	0.487	0.04210	0.77072	0.05	Long Beach
1	0.438	-0.02023	0.74416	0.487	0.04210	0.77072	0.05	Long Beach
0	0.589	0.39004	0.73541	0.364	0.11667	0.56882	0.05	Glendora
1	0.589	0.39004	0.73541	0.364	0.11667	0.56882	0.05	Glendora
0	0.521	0.39360	0.62876	0.316	0.16405	0.45330	0.05	All Areas
1	0.521	0.39360	0.62876	0.316	0.16405	0.45330	0.05	All Areas
0	0.396	0.17029	0.58223	0.189	-0.05558	0.41218	0.05	Zou Example 3
1	0.396	0.17029	0.58223	0.189	-0.05558	0.41218	0.05	Zou Example 3
0	0.380	0.30955	0.44630	0.250	0.17362	0.32339	0.05	Wuensch example
1	0.380	0.30955	0.44630	0.250	0.17362	0.32339	0.05	Wuensch example
0	0.500	0.33931	0.63234	0.600	0.45987	0.71097	0.05	Steiger B
1	0.500	0.33931	0.63234	0.600	0.45987	0.71097	0.05	Steiger B

ZPF test with confidence intervals

Steiger	r12	r34	ZPF	p_ZPF	CI_Lower	CI_Upper	alpha	Note
0	0.628	0.624	0.02164	0.98274	-0.37274	0.38153	0.05	Burbank
1	0.628	0.624	0.02164	0.98274	-0.37274	0.38153	0.05	Burbank
0	0.418	0.040	2.02730	0.04263	0.01138	0.71593	0.05	Lancaster
1	0.418	0.040	2.01788	0.04360	0.01138	0.71593	0.05	Lancaster
0	0.438	0.487	-0.19137	0.84824	-0.55040	0.45172	0.05	Long Beach
1	0.438	0.487	-0.19122	0.84835	-0.55040	0.45172	0.05	Long Beach
0	0.589	0.364	1.58168	0.11372	-0.05394	0.50660	0.05	Glendora
1	0.589	0.364	1.57597	0.11503	-0.05394	0.50660	0.05	Glendora
0	0.521	0.316	2.23588	0.02536	0.02506	0.38436	0.05	All Areas
1	0.521	0.316	2.23365	0.02551	0.02506	0.38436	0.05	All Areas
0	0.396	0.189	1.33836	0.18078	-0.09596	0.50092	0.05	Zou Example 3
1	0.396	0.189	1.33536	0.18176	-0.09596	0.50092	0.05	Zou Example 3
0	0.380	0.250	2.86917	0.00412	0.04116	0.21845	0.05	Wuensch example
1	0.380	0.250	2.86508	0.00417	0.04116	0.21845	0.05	Wuensch example
0	0.500	0.600	-1.40794	0.15915	-0.24521	0.03934	0.05	Steiger B
1	0.500	0.600	-1.40497	0.16003	-0.24521	0.03934	0.05	Steiger B

```
/*============================================================
File name:    8_Potthoff.txt
Written by:   Karl L. Wuensch, WuenschK@ECU.edu
Date:         28-August-2012.
============================================================

This is an example of how to test conincidence, intercepts, and slopes
with k independent groups.

Edit the LIBNAME and SET statements to point to the location of the SAS data set.
*********************************************************/
data Areas2_4; LIBNAME Sol 'c:\Users\Vati\Documents\SAS\SASdata';
SET Sol.Lung; if area = 2 or area = 4;
fheight=fheight-60; interaction=area*fheight;
proc reg; model fweight=area fheight interaction;
  test area=0, interaction=0;
title 'Compare Area 2 with Area 4'; run;

data Areas1to4; set Sol.Lung;
fheight=fheight-60;
*Create dummy variables for area;
If area=1 then Area1=1; else Area1=0;
If area=2 then Area2=1; else Area2=0;
If area=3 then Area3=1; else Area3=0;
*Create dummy variables for interaction;
I1=Area1*fheight; I2=Area2*fheight;I3=Area3*fheight;
proc reg; model fweight=Area1 Area2 Area3 fheight I1 I2 I3;
  test Area1=0, Area2=0, Area3=0, I1=0, I2=0, I3=0;
  test I1=0, I2=0, I3=0;
  test Area1=0, Area2=0, Area3=0;
```

```
title 'Test coincidence, slopes, and intercepts across four areas'; run;

/********* If you can do without the test of coincidence, Proc GLM will do the dummy
coding for you and give you the tests of slopes and of intercepts *******/

proc GLM; CLASS area; model fweight=area|fheight / ss3; run;
```

Compare Area 2 with Area 4

The REG Procedure
Model: MODEL1
Dependent Variable: FWEIGHT weight of father in pounds

Number of Observations Read	107
Number of Observations Used	107

Analysis of Variance

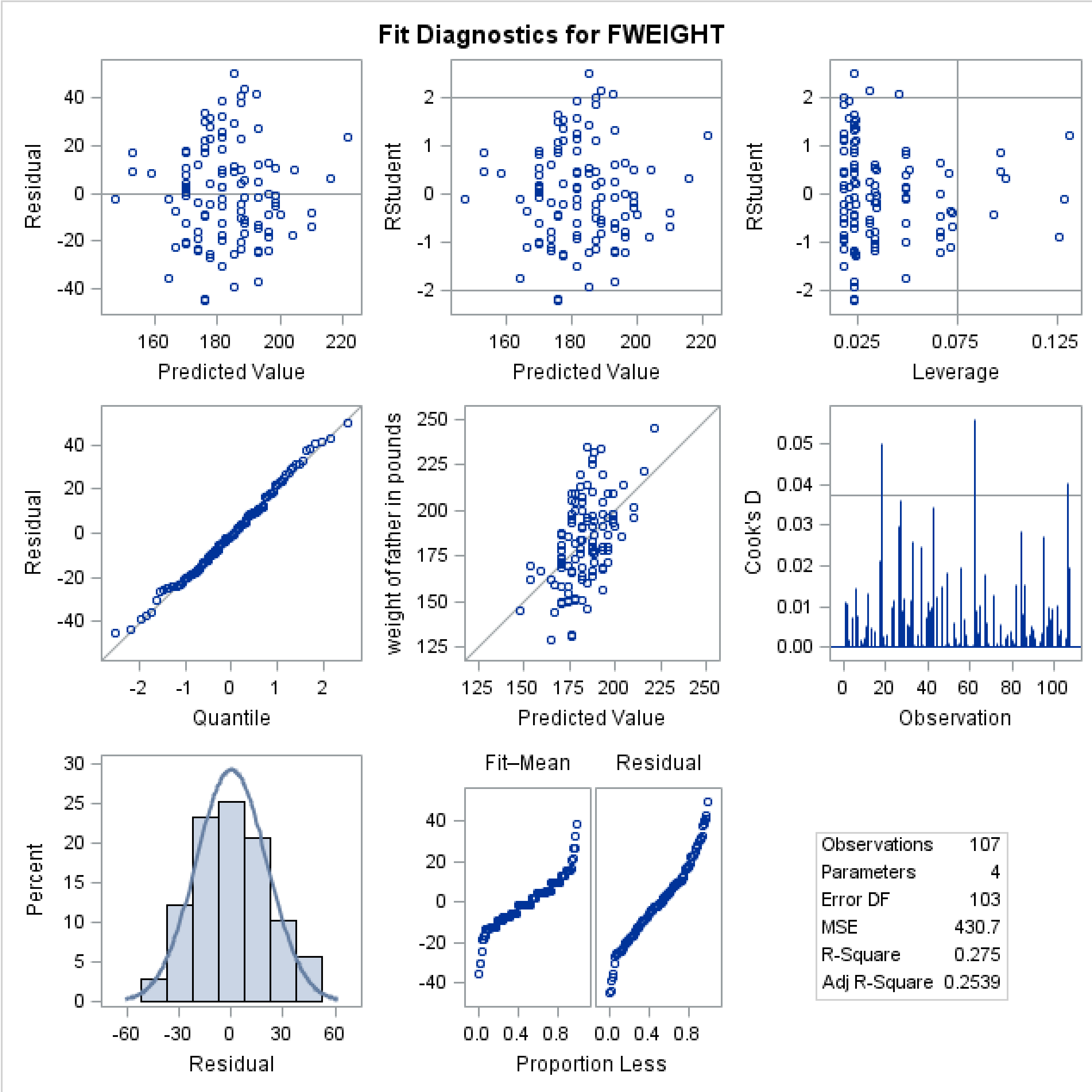
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	16828	5609.37320	13.02	<.0001
Error	103	44362	430.70077		
Corrected Total	106	61190			

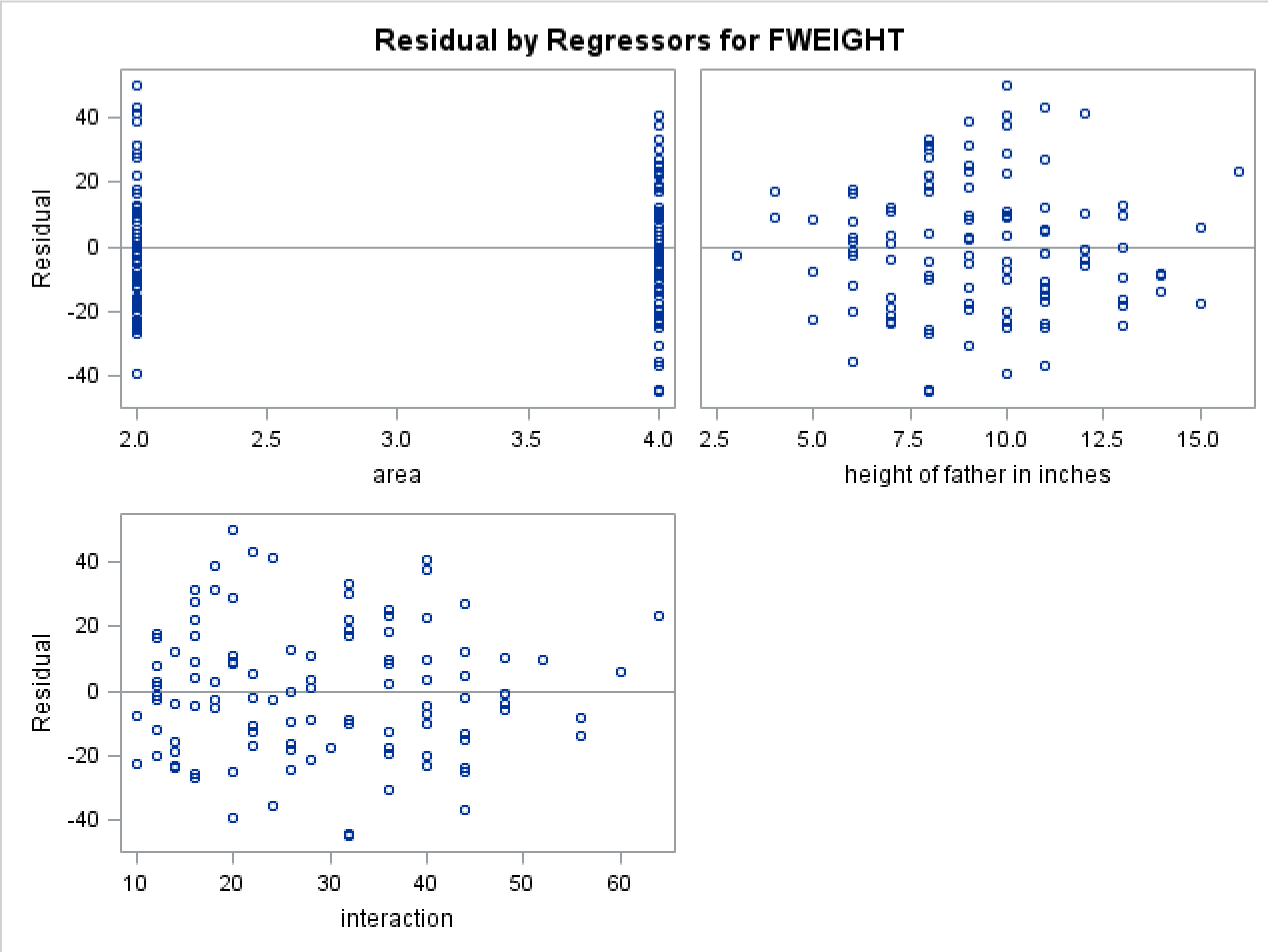
Root MSE	20.75333	**R-Square**	0.2750
Dependent Mean	183.15888	**Adj R-Sq**	0.2539
Coeff Var	11.33078		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	165.66137	24.03160	6.89	<.0001
AREA	area	1	-8.80407	7.53731	-1.17	0.2455
FHEIGHT	height of father in inches	1	1.72994	2.52253	0.69	0.4944
interaction		1	0.98973	0.78318	1.26	0.2092

Area tests differences in intercepts.
The interaction tests differences in slopes.


Fit Diagnostics for FWEIGHT
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
weight of father in pounds
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 107
Parameters 4
Error DF 103
MSE 430.7
R-Square 0.275
Adj R-Square 0.2539



The REG Procedure
Model: MODEL1

Test 1 Results for Dependent Variable FWEIGHT				
Source	**DF**	**Mean Square**	**F Value**	**Pr > F**
Numerator	2	351.27807	0.82	0.4452
Denominator	103	430.70077		

This is the test of coincidence.

Test coincidence, slopes, and intercepts across four areas

The REG Procedure
Model: MODEL1
Dependent Variable: FWEIGHT weight of father in pounds

Number of Observations Read	150
Number of Observations Used	150

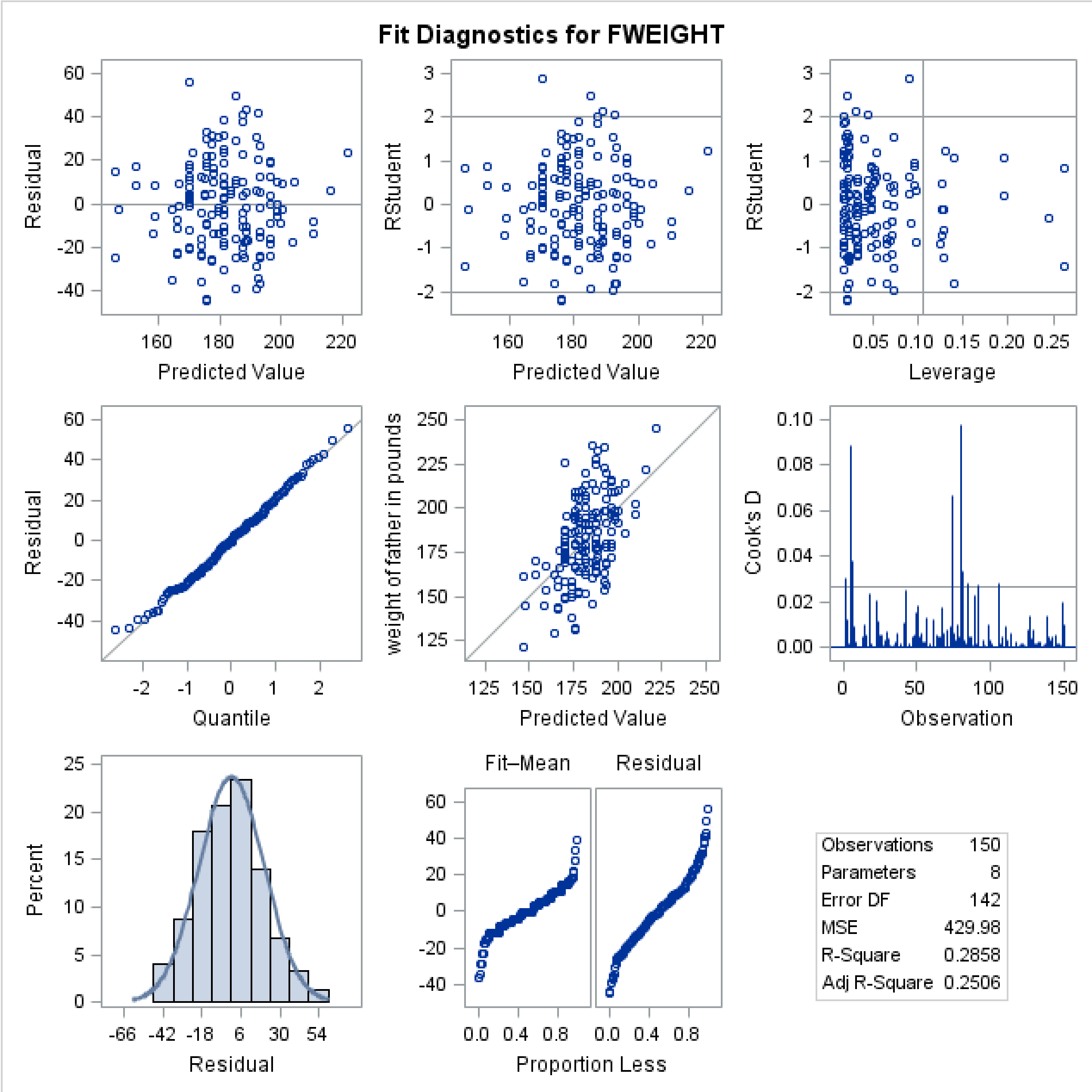
Analysis of Variance

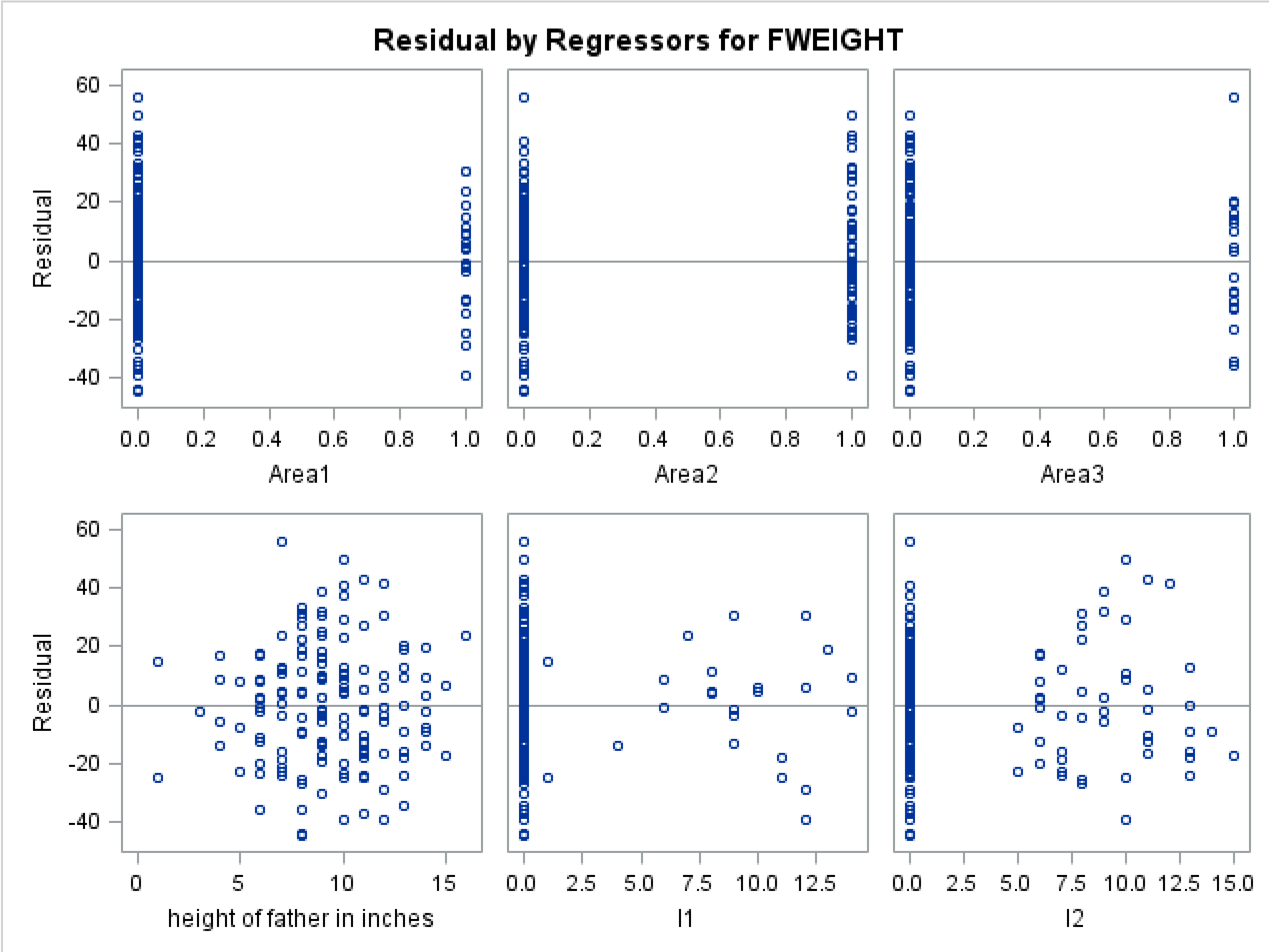
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	24439	3491.25448	8.12	<.0001
Error	142	61057	429.97952		
Corrected Total	149	85496			

Root MSE	20.73595	**R-Square**	0.2858
Dependent Mean	182.08667	**Adj R-Sq**	0.2506
Coeff Var	11.38795		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	**1**	130.44509	10.50245	12.42	<.0001
Area1		**1**	11.56612	15.75801	0.73	0.4642
Area2		**1**	17.60814	15.06200	1.17	0.2443
Area3		**1**	13.59270	19.59121	0.69	0.4889
FHEIGHT	height of father in inches	**1**	5.68886	1.07161	5.31	<.0001
I1		**1**	-1.50955	1.62205	-0.93	0.3536
I2		**1**	-1.97946	1.56505	-1.26	0.2080
I3		**1**	-1.94009	2.00171	-0.97	0.3341


Fit Diagnostics for FWEIGHT
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
weight of father in pounds
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 150
Parameters 8
Error DF 142
MSE 429.98
R-Square 0.2858
Adj R-Square 0.2506



Residual by Regressors for FWEIGHT
Residual
Area1
Area2
Area3
height of father in inches
I1
I2

The REG Procedure
Model: MODEL1

Test 1 Results for Dependent Variable FWEIGHT				
Source	**DF**	**Mean Square**	**F Value**	**Pr > F**
Numerator	6	202.84034	0.47	0.8284
Denominator	142	429.97952		

Test 1 is the test of coincidence.

The REG Procedure
Model: MODEL1

Test 2 Results for Dependent Variable FWEIGHT				
Source	**DF**	**Mean Square**	**F Value**	**Pr > F**
Numerator	3	283.35831	0.66	0.5786
Denominator	142	429.97952		

Test 2 is the test of equality of slopes.

The REG Procedure
Model: MODEL1

Test 3 Results for Dependent Variable FWEIGHT				
Source	**DF**	**Mean Square**	**F Value**	**Pr > F**
Numerator	3	210.42221	0.49	0.6902
Denominator	142	429.97952		

Test 3 is the test of equality of intercepts.

The GLM Procedure

Class Level Information		
Class	**Levels**	**Values**
AREA	4	1 2 3 4

Number of Observations Read	150
Number of Observations Used	150

The GLM Procedure

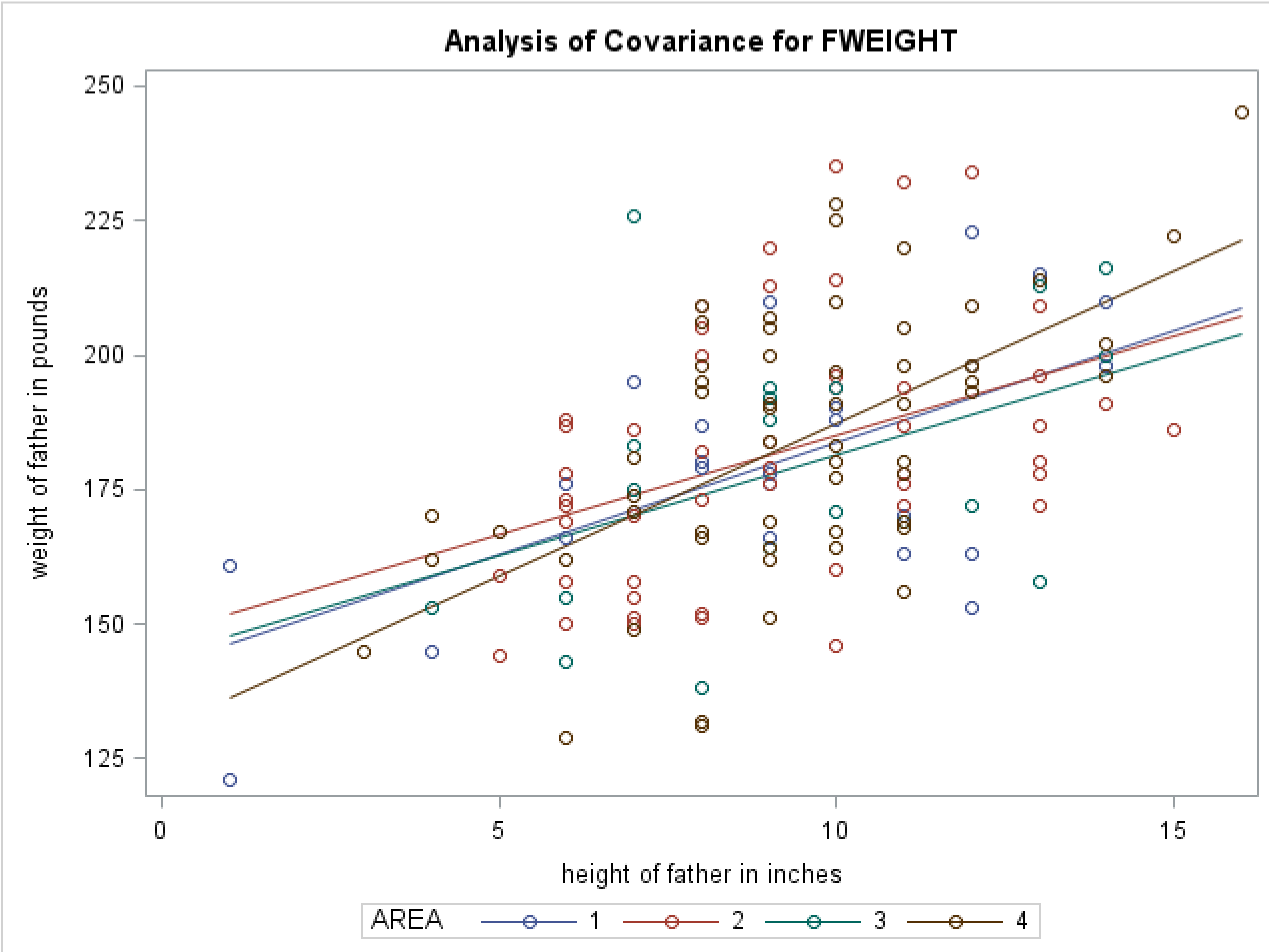
Dependent Variable: FWEIGHT weight of father in pounds

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	24438.78138	3491.25448	8.12	<.0001
Error	142	61057.09195	429.97952		
Corrected Total	149	85495.87333			

R-Square	Coeff Var	Root MSE	FWEIGHT Mean
0.285847	11.38795	20.73595	182.0867

Source	DF	Type III SS	Mean Square	F Value	Pr > F
AREA	3	631.26664	210.42221	0.49	0.6902
FHEIGHT	1	19008.78980	19008.78980	44.21	<.0001
FHEIGHT*AREA	3	850.07493	283.35831	0.66	0.5786

AREA is the test of intercepts, the interaction is the test of slopes.


Analysis of Covariance for FWEIGHT
weight of father in pounds
250
225
200
175
150
125
0
5
10
15
height of father in inches
AREA
1
2
3
4