Handout 3: Interactions with Regression using the Math Data

```
/* mathread.sas    Just read the data and do basic transformations */
options nodate linesize=79 noovp formdlim=' ';
title 'Gender, Ethnicity and Math performance';

proc format;
     value ynfmt  0 = 'No'  1 = 'Yes';
     value crsfmt   4 = 'No Resp';
     value nfmt
               1  = 'Chinese'
               2  = 'Japanese'
               3  = 'Korean'
               4  = 'Vietnamese'
               5  = 'Other Asian'
               6  = 'Eastern European'
               7  = 'Hispanic'
               8  = 'English-speaking'
               9  = 'French'
              10   = 'Italian'
              11   = 'Greek'
              12   = 'Germanic'
              13   = 'Other European'
              14   = 'Middle-Eastern'
              15   = 'Pakistani'
              16   = 'East Indian'
              17   = 'Sub-Saharan'
              18  = 'OTHER or DK';

      value ncfmt
                1 = 'Asian'
                2 = 'Eastern European'
                3 = 'European not Eastern'
                4 = 'Middle-Eastern  and Pakistani'
                5 = 'East Indian'
                6 = 'Other    and DK' ;

data math;
     infile 'mathexplore.data';
     input id course precalc calc gpa calculus english mark lang $ sex $
          nation1 nation2 sample;

/* Computed Variables: totscore, passed, grade, hsgpa, hscalc, hsengl,
                       tongue, cat1, cat2, ethnic */

     totscore = precalc+calc;
     if (50<=mark<=100) then passed=1; else passed=0;
     if mark=0 then grade=.;
        else if mark > 100 then grade=.;
        else grade=mark;
     if 65 le gpa le 100 then hsgpa = gpa; /* Else missing is automatic */
     if 0 < calculus < 101 then hscalc = calculus;
     if 0 < english < 101 then hsengl = english;
     if lang='French' then tongue='Other  '; else tongue=lang;
     label tongue = 'Mother Tongue (Eng or Other)';
```

```
      /********* Nationality According to the 2 raters **********/
       if         1 <= nation1 <= 5 then cat1 = 1;
       else if    nation1 = 6 then cat1 = 2;
       else if    7 <= nation1 <= 13 then cat1 = 3;
       else if    14 <= nation1 <= 15 then cat1 = 4;
       else if    nation1 = 16 then cat1 = 5;
       else       cat1 = 6;
       if         1 <= nation2 <= 5 then cat2 = 1;
       else if    nation2 = 6 then cat2 = 2;
       else if    7 <= nation2 <= 13 then cat2 = 3;
       else if    14 <= nation2 <= 15 then cat2 = 4;
       else if    nation2 = 16 then cat2 = 5;
       else       cat2 = 6;
       if cat1=4 then ethnic=cat1; else ethnic=cat2;
       /* Rater 1 knows Middle Eastern names -- otherwise believe Rater 2 */
       /********************************/

       label
             precalc  = 'Number precalculus correct'
             calc  = 'Number calculus correct'
             totscore = 'Total # right on diagnostic test'
             passed = 'Passed the course'
             grade = 'Final mark (if any)'
             hsgpa = 'High School GPA'
             hscalc = 'HS Calculus'
             hsengl = 'HS English'
             lang = 'Mother Tongue'
             nation1 = 'Nationality of name acc to rater1'
             nation2 = 'Nationality of name acc to rater2'
             tongue = 'Mother Tongue (Eng or Other)'
             cat1 = 'Nationality Acc to Rater 1'
             cat2 = 'Nationality Acc to Rater 2'
             ethnic = 'Judged Nationality of name'
             ;

       format course crsfmt.;
       format passed ynfmt.;
       format nation1 nation2 nfmt.;
       format cat1 cat2 ethnic ncfmt.;

/* Dummy variables and Interaction Terms */

/* Indicator dummy variables for course; Use them all only if no intercept */
       if course = . then c1=.;
          else if course = 1 then c1 = 1;
          else c1 = 0;
       if course = . then c2=.;
          else if course = 2 then c2 = 1;
          else c2 = 0;
       if course = . then c3=.;
          else if course = 3 then c3 = 1;
          else c3 = 0;
/* Sex and tongue are character variables. Make numeric versions. */
       if sex = 'Female' then gender=1;
          else if sex = 'Male' then gender=0;
       if tongue='English' then lang1=1;
          else if tongue='Other' then lang1=0;
```

```
/* Indicators for ethnic background. Leave out e3 to make European non-eastern
the reference category. */

    if ethnic = . then e1 = .;
        else if ethnic = 1 then e1 = 1;
        else e1 = 0;
    if ethnic = . then e2 = .;
        else if ethnic = 2 then e2 = 1;
        else e2 = 0;
    if ethnic = . then e3 = .;
        else if ethnic = 3 then e3 = 1;
        else e3 = 0;
    if ethnic = . then e4 = .;
        else if ethnic = 4 then e4 = 1;
        else e4 = 0;
    if ethnic = . then e5 = .;
        else if ethnic = 5 then e5 = 1;
        else e5 = 0;
    if ethnic = . then e6 = .;
        else if ethnic = 6 then e6 = 1;
        else e6 = 0;

/* Interaction Terms */
    hscalcc1 = hscalc*c1 ; hscalcc2 = hscalc*c2 ; hscalcc3 = hscalc*c3;
    scorec1 = totscore*c1 ; scorec2 = totscore*c2 ; scorec3 = totscore*c3;
    gencalc = gender*hscalc; genscore = gender*score;

/* Effect coding for course and sex */

    if course = . then crs1=.;
        else if course = 1 then crs1 = 1;
        else if course = 3 then crs1 = -1;
        else crs1 = 0;
    if course = . then crs2=.;
        else if course = 2 then crs2 = 1;
        else if course = 3 then crs2 = -1;
        else crs2 = 0;
    gndr = 2*gender-1;  /* Using algebra instead of if */
    /* Interaction: Categorical by Categorical  */
    gcrs1 = gndr*crs1;  gcrs2 = gndr*crs2;
```

```
/*********************** mathreg1.sas **************************
* Illustrate interactions, not necessarily trying to answer the   *
* most important questions yet.                                   *
******************************************************************/

%include 'mathread.sas';
/* Quantitative Vars: precalc calc totscore hsgpa hscalc hsengl grade
   Categorical Vars:  course sex ethnic (6 cats)  tongue passed   */

/* The data step continues ... */

if 1 le course le 3; /* Discard other cases */


proc reg;
     title2 'Predict Diagnostic test score from HSCALC and Course';
     model totscore = hscalc c1 c3 ;
     course: test c1=c3=0;

proc reg simple;
     title2 'Test course by HSCALC interaction';
     model totscore = hscalc c1 c3 hscalcc1 hscalcc3;
     Interac: test hscalcc1=hscalcc3=0;
     Slop1vs3: test hscalcc1=hscalcc3;

/* The following is easier than using a calculator. Regression coefficients
for the different groups will be okay, but not standard errors and tests. */

proc sort;       /* Sort data before using "by" statement; */
     by course;

proc reg;
     title2 'Separate regressions for each course';
     model totscore = hscalc;
     by course;

proc means;
     title2 'Check Range of data in courses for plotting';
     var hscalc;
     by course;

proc reg simple;
     title2 'Test course differences at specific points';
     model totscore = hscalc c1 c3 hscalcc1 hscalcc3;
     Atmean: test c1 + 76*hscalcc1 = 0, c3 + 76*hscalcc3 = 0;
     At80:   test c1 + 80*hscalcc1 = 0, c3 + 80*hscalcc3 = 0;
     At90:   test c1 + 90*hscalcc1 = 0, c3 + 90*hscalcc3 = 0;
     At200:  test c1 + 200*hscalcc1 = 0, c3 + 200*hscalcc3 = 0;

proc reg;
     title2 'No Intercept: Cell means coding. ANCOVA F = 2.64';
     model totscore = hscalc c1 c2 c3 / noint;
     course: test c1=c2=c3;
```

```
/* With an intercept, you MUST have a variable in the model if it is in
   an interaction, or else the meaning of the model is obscure. But
   without an intercept, this rule does not apply. */

proc reg;
     title2 'Test Interaction with No Intercept: F = 2.69, p = 0.0690';
     model totscore = c1-c3 hscalcc1-hscalcc3 / noint;
     Interac: test hscalcc1=hscalcc2=hscalcc3 ;

/* Two categorical IVs */

proc freq;
     tables course*sex / nocol nopercent;

proc reg;
     title2 'Course by Sex Regression with Effect Coding';
     model totscore = crs1 crs2 gndr gcrs1 gcrs2;
     course: test crs1=crs2=0;
     sex: test gndr=0;
     SexByCrs: test gcrs1=gcrs2=0;
proc glm;
     title2 'Course by Sex with proc glm: Much easier';
     class sex course;
     model totscore = sex course sex*course;
     means sex|course;

proc glm;
     title2 'Quick peek at Final Grade';
     class sex course;
     model grade = sex course sex*course;

proc reg;
     title2 'Course by Sex ANCOVA with effect coding';
     model totscore = hscalc crs1 crs2 gndr gcrs1 gcrs2;
     sex: test gndr = 0;
     Course: test crs1 = crs2 = 0;
     SexByCrs: test gcrs1 = gcrs2 = 0;

proc glm;
     title2 'Course by Sex ANCOVA with proc glm';
     class sex course;
     model totscore = hscalc sex course sex*course;
     lsmeans sex|course;

proc glm;
     title2 'Parallel slopes?';
     class sex course;
     model totscore = hscalc sex course sex*course
                      hscalc*sex hscalc*course hscalc*sex*course;

/* I'm satisfied with a parallel slopes model. */
```

mathreg1.lst

Predict Diagnostic test score from HSCALC and Course

The REG Procedure
Model: MODEL1
Dependent Variable: totscore Total # right on diagnostic test

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1161.98625	387.32875	38.70	<.0001
Error	380	3803.32364	10.00875		
Corrected Total	383	4965.30990			

Root MSE	3.16366	R-Square	0.2340
Dependent Mean	8.17448	Adj R-Sq	0.2280
Coeff Var	38.70167		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error
Intercept	Intercept	1	-1.61265	1.06398
hscalc	HS Calculus	1	0.12845	0.01373
c1		1	-1.10455	0.74199
c3		1	1.03430	0.60758

Parameter Estimates

Variable	Label	DF	t Value	Pr > |t|
Intercept	Intercept	1	-1.52	0.1304
hscalc	HS Calculus	1	9.35	<.0001
c1		1	-1.49	0.1374
c3		1	1.70	0.0895

Predict Diagnostic test score from HSCALC and Course

The REG Procedure
Model: MODEL1

Test course Results for Dependent Variable totscore

Mean

Source	DF	Square	F Value	Pr > F
Numerator	2	26.43889	2.64	0.0726
Denominator	380	10.00875		

The REG Procedure

Descriptive Statistics

Variable	Sum	Mean	Uncorrected SS	Variance	Standard Deviation
Intercept	384.00000	1.00000	384.00000	0	0
hscalc	29197	76.03385	2278733	153.45316	12.38762
c1	21.00000	0.05469	21.00000	0.05183	0.22767
c3	30.00000	0.07813	30.00000	0.07221	0.26872
hscalcc1	1285.00000	3.34635	80453	198.83273	14.10081
hscalcc3	2455.00000	6.39323	203239	489.67003	22.12849
totscore	3139.00000	8.17448	30625	12.96426	3.60059

Descriptive Statistics

Variable	Label
Intercept	Intercept
hscalc	HS Calculus
c1	
c3	
hscalcc1	
hscalcc3	
totscore	Total # right on diagnostic test

The REG Procedure
Model: MODEL1
Dependent Variable: totscore Total # right on diagnostic test

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1215.41746	243.08349	24.50	<.0001
Error	378	3749.89244	9.92035		
Corrected Total	383	4965.30990			

Root MSE		3.14966	R-Square	0.2448
Dependent Mean		8.17448	Adj R-Sq	0.2348

```
          Coeff Var               38.53039


                      Parameter Estimates

                                               Parameter        Standard
Variable       Label                      DF    Estimate          Error

Intercept      Intercept                   1    -1.51021         1.10225
hscalc         HS Calculus                 1     0.12711         0.01424
c1                                         1     5.59828         4.69683
c3                                         1    -8.45918         5.47342
hscalcc1                                   1    -0.10987         0.07513
hscalcc3                                   1     0.11610         0.06668

                      Parameter Estimates

   Variable       Label                    DF   t Value    Pr > |t|

   Intercept      Intercept                 1     -1.37      0.1715
   hscalc         HS Calculus               1      8.93      <.0001
   c1                                        1      1.19      0.2340
   c3                                        1     -1.55      0.1231
   hscalcc1                                  1     -1.46      0.1444
   hscalcc3                                  1      1.74      0.0825
```

              Test course by HSCALC interaction

                    The REG Procedure
                      Model: MODEL1

       Test Interac Results for Dependent Variable totscore

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	26.71560	2.69	0.0690
Denominator	378	9.92035		

              Test course by HSCALC interaction

                    The REG Procedure
                      Model: MODEL1

       Test Slop1vs3 Results for Dependent Variable totscore

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	52.31074	5.27	0.0222
Denominator	378	9.92035		

The interaction is not significant, but let's pause and figure out what's going on, to illustrate the process.

$$\hat{Y} = b_0 + b_1 x + b_2 c_1 + b_3 c_3 + b_4 c_1 x + b_5 c_3 x$$
$$= -1.51021 + 0.12711x + 0.12711c_1 + -8.45918c_3 + -0.10987c_1 x + 0.11610c_3 x$$

Class	C1	C3	$\hat{Y} = b_0 + b_1 x + b_2 c_1 + b_3 c_3 + b_4 c_1 x + b_5 c_3 x$
1 = Remedial	1	0	$(b_0 + b_2) + (b_1 + b_4)x$
2 = Standard	0	0	$b_0 + b_1 x$
3 = Advanced	0	1	$(b_0 + b_3) + (b_1 + b_5)x$

Class	C1	C3	$\hat{Y} = b_0 + b_1 x + b_2 c_1 + b_3 c_3 + b_4 c_1 x + b_5 c_3 x$
1 = Remedial	1	0	`(-1.51021 + 5.59828) + (0.12711 + -0.10987)*hscalc`
2 = Standard	0	0	`-1.51021+ 0.12711*hscalc`
3 = Advanced	0	1	`(-1.51021 + -8.45918) + (0.12711 + 0.11610)*hscalc`

Class	C1	C3	$\hat{Y} = b_0 + b_1 x + b_2 c_1 + b_3 c_3 + b_4 c_1 x + b_5 c_3 x$
1 = Remedial	1	0	`4.08807 + 0.01724*hscalc`
2 = Standard	0	0	`-1.51021 + 0.12711*hscalc`
3 = Advanced	0	1	`-9.96939 + 0.24321*hscalc`

```
/* The following is easier than using a calculator. Regression coefficients
for the different groups will be okay, but not standard errors and tests. */

proc sort;        /* Sort data before using "by" statement; */
     by course;
proc reg;
     title2 'Separate regressions for each course';
     model totscore = hscalc;
     by course;

proc means;
     title2 'Check Range of data in courses for plotting';
     var hscalc;
     by course;
```

-------------------------------- course=1 ----------------------------------

The REG Procedure
Model: MODEL1
Dependent Variable: totscore Total # right on diagnostic test

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.54176	0.54176	0.10	0.7565
Error	19	104.02967	5.47525		
Corrected Total	20	104.57143			

Root MSE	2.33992	R-Square	0.0052	
Dependent Mean	5.14286	Adj R-Sq	-0.0472	
Coeff Var	45.49853			

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error
Intercept	Intercept	1	4.08807	3.39189
hscalc	HS Calculus	1	0.01724	0.05480

Parameter Estimates

Variable	Label	DF	t Value	Pr > |t|
Intercept	Intercept	1	1.21	0.2429
hscalc	HS Calculus	1	0.31	0.7565

Page 10 of 28

-------------------------------- course=2 ---------------------------------

The REG Procedure
Model: MODEL1
Dependent Variable: totscore Total # right on diagnostic test

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	790.40090	790.40090	78.70	<.0001
Error	331	3324.30180	10.04321		
Corrected Total	332	4114.70270			

Root MSE	3.16910	R-Square	0.1921
Dependent Mean	8.20721	Adj R-Sq	0.1897
Coeff Var	38.61365		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error
Intercept	Intercept	1	-1.51021	1.10906
hscalc	HS Calculus	1	0.12711	0.01433

Parameter Estimates

Variable	Label	DF	t Value	Pr > |t|
Intercept	Intercept	1	-1.36	0.1742
hscalc	HS Calculus	1	8.87	<.0001

-------------------------------- course=3 ----------------------------------

The REG Procedure
Model: MODEL1
Dependent Variable: totscore Total # right on diagnostic test

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	138.30570	138.30570	12.04	0.0017
Error	28	321.56097	11.48432		
Corrected Total	29	459.86667			

Root MSE	3.38885	R-Square	0.3008	
Dependent Mean	9.93333	Adj R-Sq	0.2758	
Coeff Var	34.11596			

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error
Intercept	Intercept	1	−9.96939	5.76843
hscalc	HS Calculus	1	0.24321	0.07008

Parameter Estimates

Variable	Label	DF	t Value	Pr > |t|
Intercept	Intercept	1	−1.73	0.0950
hscalc	HS Calculus	1	3.47	0.0017

```
                Gender, Ethnicity and Math performance                  10
                Check Range of data in courses for plotting
```

```
------------------------------ course=1 ------------------------------

                            The MEANS Procedure

                    Analysis Variable : hscalc HS Calculus

      N            Mean          Std Dev          Minimum          Maximum
    ----------------------------------------------------------------------
     21       61.1904762        9.5478744       50.0000000       80.0000000
    ----------------------------------------------------------------------


------------------------------ course=2 ------------------------------

                    Analysis Variable : hscalc HS Calculus

      N            Mean          Std Dev          Minimum          Maximum
    ----------------------------------------------------------------------
    333       76.4474474       12.1385489       50.0000000       99.0000000
    ----------------------------------------------------------------------


------------------------------ course=3 ------------------------------

                    Analysis Variable : hscalc HS Calculus

      N            Mean          Std Dev          Minimum          Maximum
    ----------------------------------------------------------------------
     30       81.8333333        8.9792225       61.0000000      100.0000000
    ----------------------------------------------------------------------
```

Class	C1	C3	$\hat{Y} = b_0 + b_1 x + b_2 c_1 + b_3 c_3 + b_4 c_1 x + b_5 c_3 x$
1 = Remedial	1	0	4.08807 + 0.01724*hscalc
2 = Standard	0	0	-1.51021 + 0.12711*hscalc
3 = Advanced	0	1	-9.96939 + 0.24321*hscalc

HS Calculus by Course

HS Calculus by Course
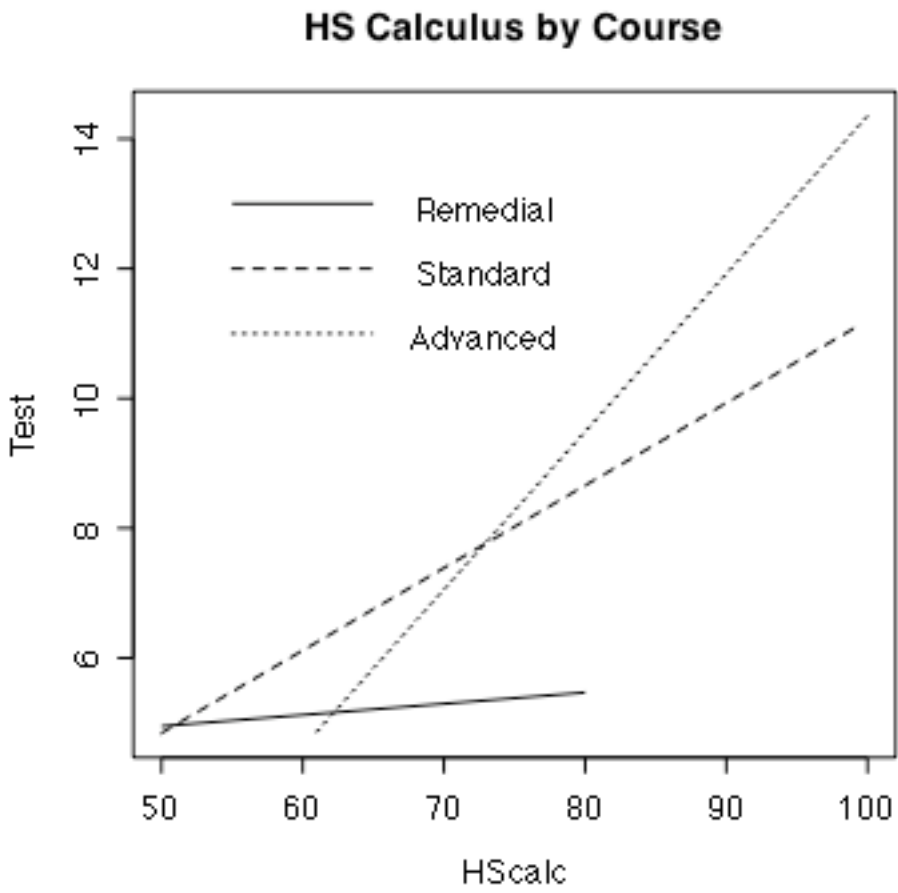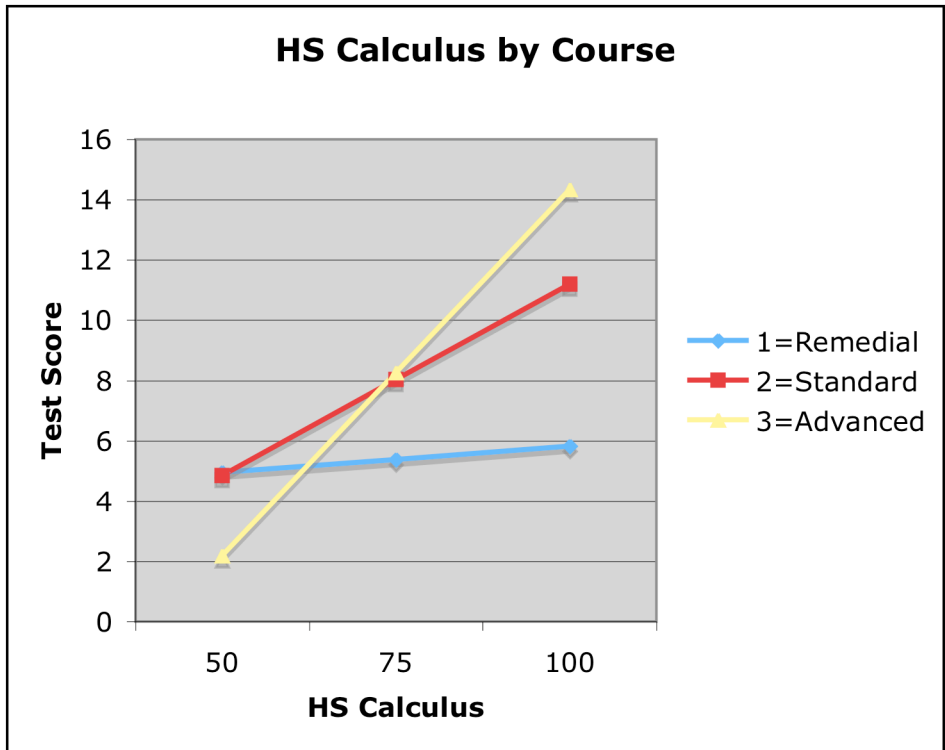
It was easier with Excel, but the picture from R is better. I could not figure out how to make lines of different lengths in Excel. We have not gotten to R yet, but for the record, here is the R code:

```
HScalc <- c(50,50,61,99,80,100)
Test <- c(4.84529,4.95007,4.86642,11.07368,5.46727,14.35161)
xOne1 <- 50 ; yOne1 <- 4.95007 ; xOne2 <- 80 ; yOne2 <- 5.46727
xTwo1 <- 50; yTwo1 <- 4.84529; xTwo2 <- 99; yTwo2 <- 11.07368
xThree1 <- 61; yThree1 <- 4.86642; xThree2 <- 100; yThree2 <- 14.35161
plot(HScalc,Test,pch=' ')
lines(c(xOne1,xOne2),c(yOne1,yOne2),lty=1)
lines(c(xTwo1,xTwo2),c(yTwo1,yTwo2),lty=2)
lines(c(xThree1,xThree2),c(yThree1,yThree2),lty=3)
title("HS Calculus by Course")
x1 <- c(55,65) ; y1 <- c(13,13) ; lines(x1,y1,lty=1) ; text(73,13,"Remedial")
x2 <- c(55,65) ; y2 <- c(12,12) ; lines(x2,y2,lty=2) ; text(73,12,"Standard")
x3 <- c(55,65) ; y3 <- c(11,11) ; lines(x3,y3,lty=3) ; text(73,11,"Advanced")
```

Next we test differences among the levels of the three lines at selected points: HScalc = mean = 76, HScalc=80, and a deliberately crazy HScalc=200. Skipping right to the tests,

The REG Procedure
Model: MODEL1

Test At80 Results for Dependent Variable totscore

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	30.78440	3.10	0.0461
Denominator	378	9.92035		

Test At90 Results for Dependent Variable totscore

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	48.62123	4.90	0.0079
Denominator	378	9.92035		

Test At200 Results for Dependent Variable totscore

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	30.65004	3.09	0.0467
Denominator	378	9.92035		

No intercept (and no interactions yet)

```
Gender, Ethnicity and Math performance                    17
            No Intercept: Cell means coding. ANCOVA F = 2.64

                        The REG Procedure
                          Model: MODEL1
         Dependent Variable: totscore Total # right on diagnostic test

          NOTE: No intercept in model. R-Square is redefined.

                        Analysis of Variance

                            Sum of          Mean
Source                  DF   Squares        Square     F Value   Pr > F

Model                    4    26822      6705.41909    669.96    <.0001
Error                  380  3803.32364     10.00875
Uncorrected Total      384    30625


            Root MSE              3.16366   R-Square     0.8758
            Dependent Mean       8.17448   Adj R-Sq     0.8745
            Coeff Var           38.70167


                        Parameter Estimates

                                          Parameter        Standard
Variable   Label                    DF     Estimate           Error

hscalc     HS Calculus               1      0.12845         0.01373
c1                                   1     -2.71721         1.08749
c2                                   1     -1.61265         1.06398
c3                                   1     -0.57836         1.26347

                        Parameter Estimates

   Variable    Label                DF    t Value   Pr > |t|

   hscalc      HS Calculus           1       9.35    <.0001
   c1                                1      -2.50    0.0129
   c2                                1      -1.52    0.1304
   c3                                1      -0.46    0.6474


        Test course Results for Dependent Variable totscore

                            Mean
        Source          DF  Square     F Value   Pr > F

        Numerator        2  26.43889     2.64    0.0726
        Denominator    380  10.00875
```

Test **Interaction** with No Intercept: F = 2.69, p = 0.0690

The REG Procedure
Model: MODEL1
Dependent Variable: totscore Total # right on diagnostic test

NOTE: No intercept in model. R-Square is redefined.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	26875	4479.18459	451.51	<.0001
Error	378	3749.89244	9.92035		
Uncorrected Total	384	30625			

Root MSE	3.14966	R-Square	0.8776	
Dependent Mean	8.17448	Adj R-Sq	0.8756	
Coeff Var	38.53039			

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error
c1		1	4.08807	4.56566
c2		1	-1.51021	1.10225
c3		1	-9.96939	5.36128
hscalcc1		1	0.01724	0.07376
hscalcc2		1	0.12711	0.01424
hscalcc3		1	0.24321	0.06514

Parameter Estimates

Variable	Label	DF	t Value	Pr > |t|
c1		1	0.90	0.3711
c2		1	-1.37	0.1715
c3		1	-1.86	0.0637
hscalcc1		1	0.23	0.8154
hscalcc2		1	8.93	<.0001
hscalcc3		1	3.73	0.0002

Test Interac Results for Dependent Variable totscore

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	26.71560	2.69	0.0690
Denominator	378	9.92035		

Sex by Course

            Test Interaction with No Intercept: F = 2.69, p = 0.0690

                        The FREQ Procedure

                    Table of course by sex

               course   sex

               Frequency|
               Row Pct |Female |Male | Total
               ---------+--------+--------+
                     1 | 28 | 28 | 56
                        | 50.00 | 50.00 |
               ---------+--------+--------+
                     2 | 198 | 173 | 371
                        | 53.37 | 46.63 |
               ---------+--------+--------+
                     3 | 8 | 29 | 37
                        | 21.62 | 78.38 |
               ---------+--------+--------+
               Total   234   230   464

                        Frequency Missing = 7

                Course by Sex Regression with Effect Coding

                        The REG Procedure
                         Model: MODEL1
            Dependent Variable: totscore Total # right on diagnostic test

                        Analysis of Variance

                              Sum of   Mean
Source   DF   Squares   Square   F Value   Pr > F

Model   5   1009.57812   201.91562   17.82   <.0001
Error   458   5190.20636   11.33233
Corrected Total   463   6199.78448

            Root MSE   3.36635   R-Square   0.1628
            Dependent Mean   7.77155   Adj R-Sq   0.1537
            Coeff Var   43.31635

Parameter Estimates

                                                       Parameter   Standard
Variable   Label   DF   Estimate   Error

Intercept   Intercept   1   7.09726   0.27586
crs1   1   -2.56154   0.37888
crs2   1   0.97130   0.29381
gndr   1   -0.98284   0.27586
gcrs1   1   0.80427   0.37888
gcrs2   1   0.52539   0.29381

                    Parameter Estimates

   Variable   Label   DF   t Value   Pr > |t|

   Intercept   Intercept   1   25.73   <.0001
   crs1   1   -6.76   <.0001
   crs2   1   3.31   0.0010
   gndr   1   -3.56   0.0004
   gcrs1   1   2.12   0.0343
   gcrs2   1   1.79   0.0744

                    The REG Procedure
                      Model: MODEL1

     Test course Results for Dependent Variable totscore

                              Mean
       Source   DF   Square   F Value   Pr > F

       Numerator   2   317.37256   28.01   <.0001
       Denominator   458   11.33233

                    The REG Procedure
                      Model: MODEL1

      Test sex Results for Dependent Variable totscore

                              Mean
       Source   DF   Square   F Value   Pr > F

       Numerator   1   143.85210   12.69   0.0004
       Denominator   458   11.33233

The REG Procedure
Model: MODEL1

Test SexByCrs Results for Dependent Variable totscore

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	44.27167	3.91	0.0208
Denominator	458	11.33233		

The GLM Procedure

Class Level Information

Class	Levels	Values
sex	2	Female Male
course	3	1 2 3

Number of observations   471

NOTE: Due to missing values, only 464 observations can be used in this
      analysis.

The GLM Procedure

Dependent Variable: totscore   Total # right on diagnostic test

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1009.578122	201.915624	17.82	<.0001
Error	458	5190.206361	11.332328		
Corrected Total	463	6199.784483			

R-Square	Coeff Var	Root MSE	totscore Mean
0.162841	43.31635	3.366352	7.771552

Source	DF	Type I SS	Mean Square	F Value	Pr > F
sex	1	165.4794660	165.4794660	14.60	0.0002
course	2	755.5553159	377.7776580	33.34	<.0001
sex*course	2	88.5433401	44.2716701	3.91	0.0208

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sex	1	143.8520968	143.8520968	12.69	0.0004
course	2	634.7451233	317.3725616	28.01	<.0001
sex*course	2	88.5433401	44.2716701	3.91	0.0208

The GLM Procedure

Level of sex	N	----------totscore---------- Mean	Std Dev
Female	234	7.17948718	3.29810814
Male	230	8.37391304	3.90936815

Level of course	N	----------totscore---------- Mean	Std Dev
1	56	4.5357143	2.44178985
2	371	8.0377358	3.47946551
3	37	10.0000000	4.06885187

Level of sex	Level of course	N	----------totscore---------- Mean	Std Dev
Female	1	28	4.3571429	2.19788258
Female	2	198	7.6111111	3.26421573
Female	3	8	6.3750000	2.38671921
Male	1	28	4.7142857	2.69233676
Male	2	173	8.5260116	3.65927696
Male	3	29	11.0000000	3.88219378

Repeating the ANOVA summary table,

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sex	1	143.8520968	143.8520968	12.69	0.0004
course	2	634.7451233	317.3725616	28.01	<.0001
sex*course	2	88.5433401	44.2716701	3.91	0.0208

Sex by Course: Means
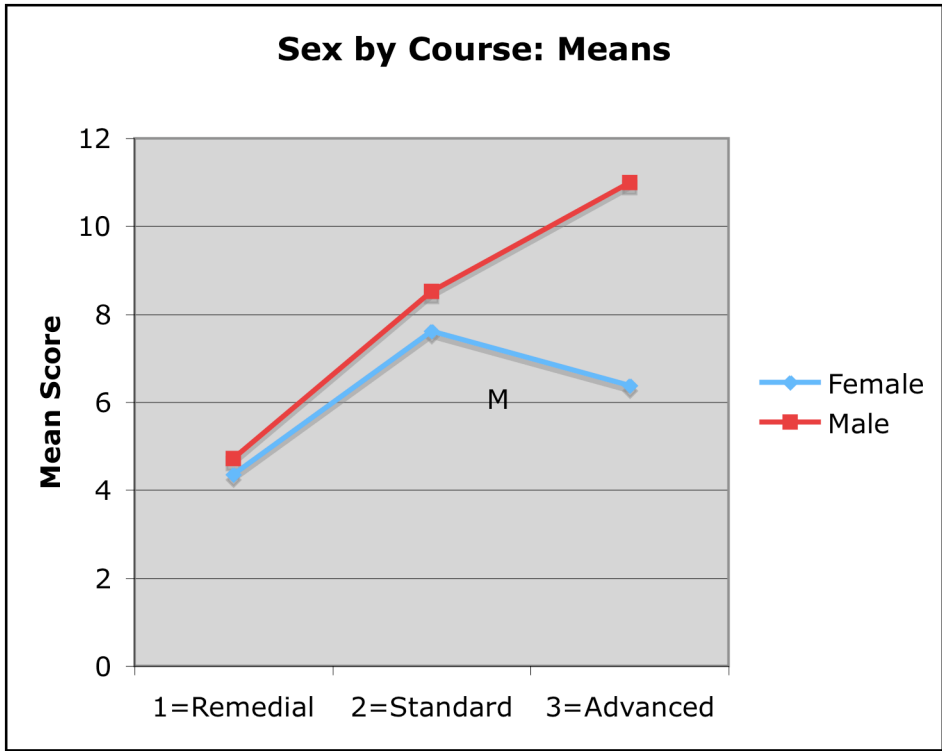
Quick peek at Final Grade

Source	DF	Type III SS	Mean Square	F Value	Pr > F
sex	1	5.6850661	5.6850661	0.02	0.9013
course	2	623.1656390	311.5828195	0.84	0.4307
sex*course	2	178.3222847	89.1611424	0.24	0.7855

Course by Sex ANCOVA with effect coding

The REG Procedure
Model: MODEL1
Dependent Variable: totscore Total # right on diagnostic test

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	1315.63856	219.27309	22.43	<.0001
Error	372	3636.68334	9.77603		
Corrected Total	378	4952.32190			

Root MSE	3.12666	R-Square	0.2657	
Dependent Mean	8.18997	Adj R-Sq	0.2538	
Coeff Var	38.17673			

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error
Intercept	Intercept	1	-1.71679	1.05852
hscalc	HS Calculus	1	0.12497	0.01366
crs1		1	-0.81540	0.55729
crs2		1	0.41176	0.35973
gndr		1	-0.65399	0.34280
gcrs1		1	1.17230	0.53712
gcrs2		1	0.23339	0.35742

Parameter Estimates

Variable	Label	DF	t Value	Pr > |t|
Intercept	Intercept	1	-1.62	0.1057
hscalc	HS Calculus	1	9.15	<.0001
crs1		1	-1.46	0.1443
crs2		1	1.14	0.2531
gndr		1	-1.91	0.0572
gcrs1		1	2.18	0.0297
gcrs2		1	0.65	0.5142

                         The REG Procedure
                          Model: MODEL1

        Test sex Results for Dependent Variable totscore

                                  Mean
     Source   DF   Square   F Value   Pr > F

     Numerator   1   35.58209   3.64   0.0572
     Denominator   372   9.77603

                         The REG Procedure
                          Model: MODEL1

       Test Course Results for Dependent Variable totscore

                                  Mean
     Source   DF   Square   F Value   Pr > F

     Numerator   2   12.74225   1.30   0.2728
     Denominator   372   9.77603

                         The REG Procedure
                          Model: MODEL1

      Test SexByCrs Results for Dependent Variable totscore

                                  Mean
     Source   DF   Square   F Value   Pr > F

     Numerator   2   33.51968   3.43   0.0335
     Denominator   372   9.77603

                            The GLM Procedure

                        Class Level Information

                Class   Levels   Values

                sex   2   Female Male

                course   3   1 2 3

                    Number of observations   471

NOTE: Due to missing values, only 379 observations can be used in this
      analysis.

                            The GLM Procedure

Dependent Variable: totscore   Total # right on diagnostic test

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	1315.638562	219.273094	22.43	<.0001
Error	372	3636.683338	9.776030		
Corrected Total	378	4952.321900			

R-Square	Coeff Var	Root MSE	totscore Mean
0.265661	38.17673	3.126664	8.189974

Source	DF	Type I SS	Mean Square	F Value	Pr > F
hscalc	1	1108.454933	1108.454933	113.38	<.0001
sex	1	83.392225	83.392225	8.53	0.0037
course	2	56.752050	28.376025	2.90	0.0561
sex*course	2	67.039354	33.519677	3.43	0.0335

Source	DF	Type III SS	Mean Square	F Value	Pr > F
hscalc	1	817.6922866	817.6922866	83.64	<.0001
sex	1	35.5820931	35.5820931	3.64	0.0572
course	2	25.4845063	12.7422532	1.30	0.2728
sex*course	2	67.0393536	33.5196768	3.43	0.0335

```
        Gender, Ethnicity and Math performance                37
           Course by Sex ANCOVA with proc glm
```

```
                The GLM Procedure
                Least Squares Means

                           totscore
            sex            LSMEAN

            Female      7.13805968
            Male        8.44603872


                           totscore
            course         LSMEAN

            1           6.97665373
            2           8.20381366
            3           8.19568020


                                totscore
       sex         course        LSMEAN

     Female        1          7.4949641
     Female        2          7.7832168
     Female        3          6.1359981
     Male          1          6.4583433
     Male          2          8.6244105
     Male          3         10.2553623
```
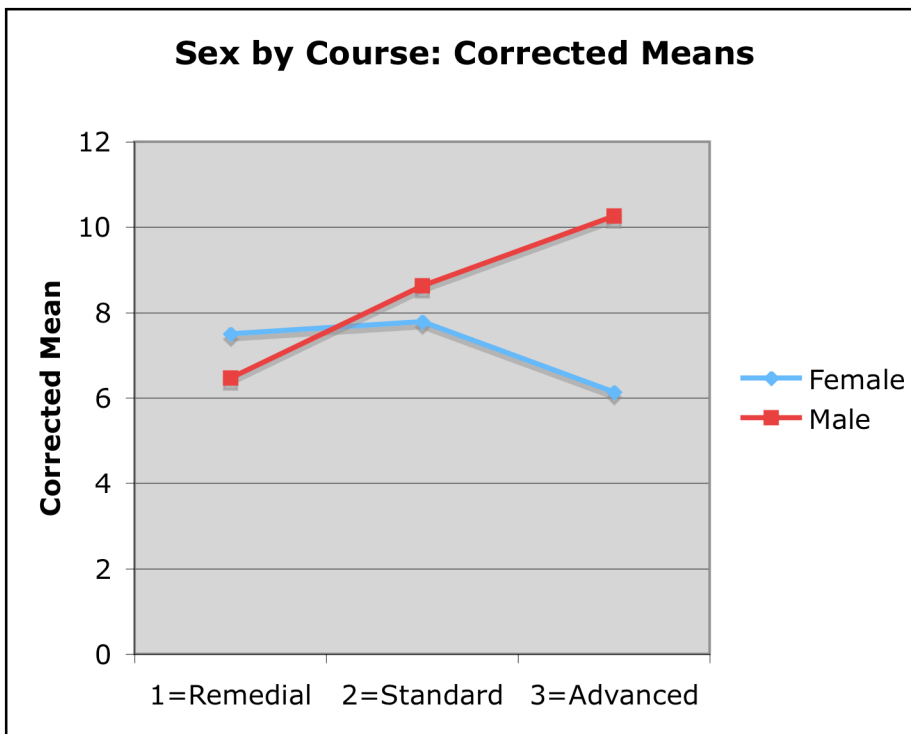

Sex by Course: Corrected Means

                              The GLM Procedure

                           Class Level Information

                    Class   Levels   Values

                    sex   2   Female Male

                    course   3   1 2 3

                    Number of observations   471

NOTE: Due to missing values, only 379 observations can be used in this
      analysis.

Dependent Variable: totscore   Total # right on diagnostic test

                                     Sum of
Source   DF   Squares   Mean Square   F Value   Pr > F

Model   11   1447.714992   131.610454   13.78   <.0001

Error   367   3504.606908   9.549338

Corrected Total   378   4952.321900

            R-Square   Coeff Var   Root MSE   totscore Mean

            0.292331   37.73150   3.090200   8.189974

Source   DF   Type I SS   Mean Square   F Value   Pr > F

hscalc   1   1108.454933   1108.454933   116.08   <.0001
sex   1   83.392225   83.392225   8.73   0.0033
course   2   56.752050   28.376025   2.97   0.0525
sex*course   2   67.039354   33.519677   3.51   0.0309
hscalc*sex   1   35.898621   35.898621   3.76   0.0533
hscalc*course   2   49.921878   24.960939   2.61   0.0746
hscalc*sex*course   2   46.255931   23.127965   2.42   0.0902

Source	DF	Type III SS	Mean Square	F Value	Pr > F
hscalc	1	81.48477752	81.48477752	8.53	0.0037
sex	1	1.11720492	1.11720492	0.12	0.7325
course	2	18.87293847	9.43646924	0.99	0.3732
sex*course	2	35.95964019	17.97982009	1.88	0.1536
hscalc*sex	1	1.65655207	1.65655207	0.17	0.6773
hscalc*course	2	27.47843107	13.73921554	1.44	0.2386
hscalc*sex*course	2	46.25593076	23.12796538	2.42	0.0902