

SAS Example One: The math data

```
tuzo.utm.utoronto.ca:~% mkdir 431s13
tuzo.utm.utoronto.ca:~% cd 431s13
tuzo.utm.utoronto.ca:~/431s13% ls
tuzo.utm.utoronto.ca:~/431s13% curl
http://www.utstat.toronto.edu/~brunner/431s13/code_n_data/lecture/calculus.data >
calculus.data
```

```
  % Total      % Received % Xferd  Average Speed   Time    Time       Time  Current
                             Dload  Upload  Total   Spent    Left   Speed
100 19296  100 19296    0     0  32439      0  --:--:--  --:--:--  --:--:-- 1177k
```

```
tuzo.utm.utoronto.ca:~/431s13% ls
```

```
calculus.data
```

```
tuzo.utm.utoronto.ca:~/431s13% less calculus.data
```

	hscalc	hsengl	hsgpa	test	grade
1	65	80	78.0	2	39
2	54	75	66.0	8	57
3	77	70	80.2	8	62
4	80	67	81.7	7	76
5	87	80	86.8	8	86
6	53	75	76.7	4	60
7	86	81	85.8	5	54
8	75	77	73.0	7	17
9	87	88	90.3	14	76

```
. . .
```

278	90	91	90.5	14	90
279	91	96	93.7	6	67
280	85	90	85.2	10	57
281	88	68	84.8	9	80
282	83	84	78.3	7	56
283	81	90	88.3	9	70
284	73	79	77.0	14	60
285	80	70	80.7	7	50
286	56	81	80.7	6	50
287	80	86	82.2	7	61

```
tuzo.utm.utoronto.ca:~/431s13% emacs calculus1.sas
```

```
/* calculus1.sas */
options linesize=79 noovp formdlim='_';
title 'Calculus 1: Basic descriptive stats, correlation and regression';
title2 'On a subset of the math data';

data math;
  infile 'calculus.data' firstobs=2;
  input id hscalc hsengl hsgpa test grade;

proc means;
  var hscalc hsengl hsgpa test grade;

proc corr;
  var hscalc hsengl hsgpa test grade;

proc reg;
  title3 'Predict grade, test HS Calc and English at the same time';
  model grade = hscalc hsengl hsgpa test;
  HS_Calc_and_English: test hscalc = hsengl = 0;
```

```
tuzo.utm.utoronto.ca:~/431s13% ls
calculus1.sas  calculus.data
tuzo.utm.utoronto.ca:~/431s13% sas calculus1
tuzo.utm.utoronto.ca:~/431s13% ls
calculus1.log  calculus1.lst  calculus1.sas  calculus.data
tuzo.utm.utoronto.ca:~/431s13% less calculus1.log
```

1

The SAS System

21:04 Thursday, January 10, 2013

NOTE: Copyright (c) 2002-2003 by SAS Institute Inc., Cary, NC, USA.

NOTE: SAS (r) 9.1 (TS1M3)

Licensed to UNIVERSITY OF TORONTO/COMPUTING & COMMUNICATIONS, Site 0008987
002.

NOTE: This session is executing on the Linux 2.6.9-67.ELsmp platform.

NOTE: SAS 9.1.3 Service Pack 3

You are running SAS 9. Some SAS 8 files will be automatically converted by the V9 engine; others are incompatible. Please see <http://support.sas.com/rnd/migration/planning/platform/64bit.html>

PROC MIGRATE will preserve current SAS file attributes and is recommended for converting all your SAS libraries from any SAS 8 release to SAS 9. For details and examples, please see <http://support.sas.com/rnd/migration/index.html>

This message is contained in the SAS news file, and is presented upon initialization. Edit the file "news" in the "misc/base" directory to display site-specific news and information in the program log. The command line option "-nonews" will prevent this display.

NOTE: SAS initialization used:
real time 0.01 seconds
cpu time 0.01 seconds

```
1      /* calculus1.sas */
2      options linesize=79 noovp formdlm='_';
3      title 'Calculus 1: Basic descriptive stats, correlation and
3      ! regression';
4      title2 'On a subset of the math data';
5
6      data math;
7          infile 'calculus.data' firstobs=2;
8          input id hscal hsenl hsgpa test grade;
9
```

NOTE: The infile 'calculus.data' is:
File Name=/student/cslec/brunnerj/431s13/calculus.data,
Owner Name=brunnerj,Group Name=lecturers,
Access Permission=rw-----,
File Size (bytes)=19296

NOTE: 287 records were read from the infile 'calculus.data'.
The minimum record length was 66.
The maximum record length was 66.

NOTE: The data set WORK.MATH has 287 observations and 6 variables.

NOTE: DATA statement used (Total process time):
 real time 0.00 seconds
 cpu time 0.00 seconds

10 proc means;
 ^L2 The SAS System

11 var hscal hsenl hsgpa test grade;
 12

NOTE: There were 287 observations read from the data set WORK.MATH.
 NOTE: The PROCEDURE MEANS printed page 1.
 NOTE: PROCEDURE MEANS used (Total process time):
 real time 0.45 seconds
 cpu time 0.02 seconds

13 proc corr;
 14 var hscal hsenl hsgpa test grade;
 15

NOTE: The PROCEDURE CORR printed page 2.
 NOTE: PROCEDURE CORR used (Total process time):
 real time 0.05 seconds
 cpu time 0.01 seconds

16 proc reg;
 17 title3 'Predict grade, test HS Calc and English at the same
 17 ! time';
 18 model grade = hscal hsenl hsgpa test;
 19 HS_Calc_and_English: test hscal = hsenl = 0;

NOTE: The PROCEDURE REG printed pages 3-4.
 NOTE: PROCEDURE REG used (Total process time):
 real time 0.11 seconds
 cpu time 0.00 seconds

NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414
 NOTE: The SAS System used:
 real time 0.68 seconds
 cpu time 0.04 seconds

m.utoronto.ca/~431s13/less/calculus1.lst

Calculus 1: Basic descriptive stats, correlation and regression 1
 On a subset of the math data

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
hscal	287	78.5191638	11.3402009	50.0000000	100.0000000
hsenl	287	77.0662021	8.4761579	50.0000000	96.0000000
hsgpa	287	80.9829268	5.9706293	66.0000000	96.2000000
test	287	8.8153310	3.5691043	1.0000000	20.0000000
grade	287	60.9895470	17.7335533	12.0000000	99.0000000

Calculus 1: Basic descriptive stats, correlation and regression 2
 On a subset of the math data

The CORR Procedure

5 Variables: hscalc hsenl hsgpa test grade

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
hscalc	287	78.51916	11.34020	22535	50.00000	100.00000
hsenl	287	77.06620	8.47616	22118	50.00000	96.00000
hsgpa	287	80.98293	5.97063	23242	66.00000	96.20000
test	287	8.81533	3.56910	2530	1.00000	20.00000
grade	287	60.98955	17.73355	17504	12.00000	99.00000

Pearson Correlation Coefficients, N = 287
 Prob > |r| under H0: Rho=0

	hscalc	hsenl	hsgpa	test	grade
hscalc	1.00000	0.11102 0.0603	0.63061 <.0001	0.43501 <.0001	0.55240 <.0001
hsenl	0.11102 0.0603	1.00000	0.54148 <.0001	0.06236 0.2924	0.11815 0.0455
hsgpa	0.63061 <.0001	0.54148 <.0001	1.00000	0.34019 <.0001	0.58667 <.0001
test	0.43501 <.0001	0.06236 0.2924	0.34019 <.0001	1.00000	0.43976 <.0001
grade	0.55240 <.0001	0.11815 0.0455	0.58667 <.0001	0.43976 <.0001	1.00000

Calculus 1: Basic descriptive stats, correlation and regression 3
 On a subset of the math data

Predict grade, test HS Calc and English at the same time

The REG Procedure

Model: MODEL1

Dependent Variable: grade

Number of Observations Read 287
 Number of Observations Used 287

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	41449	10362	60.26	<.0001
Error	282	48492	171.95883		
Corrected Total	286	89941			

Root MSE	13.11331	R-Square	0.4608
Dependent Mean	60.98955	Adj R-Sq	0.4532
Coeff Var	21.50091		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-61.76855	10.80433	-5.72	<.0001
hscalc	1	0.23744	0.09775	2.43	0.0158
hsengl	1	-0.41251	0.11647	-3.54	0.0005
hsgpa	1	1.56646	0.21266	7.37	<.0001
test	1	1.02647	0.24270	4.23	<.0001

Calculus 1: Basic descriptive stats, correlation and regression 4

On a subset of the math data

Predict grade, test HS Calc and English at the same time

The REG Procedure

Model: MODEL1

Test HS_Calc_and_English Results
for Dependent Variable grade

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	2312.45704	13.45	<.0001
Denominator	282	171.95883		