

SAS

Your One Stop Shop for IR

Greg Gengo & Tracy Stegmair

Institutional Research & Data Management

Texas Woman's University

Topics of Discussion

- **Software and IR**
- **SAS and IR**
 - **Accessing the Data**
 - **Cleaning the Data**
 - **Automating Reports and Surveys**
 - **Output to Word, Excel, or PDF**
 - **Graphing/Charting Data**

Software and IR

- **Most IR Offices Can't Function Without:**
 - Oracle SQL Developer, SQL Management Studio, or Colleague Uniquery (Colon Prompt)
 - for accessing and sometimes cleaning the data
 - Excel
 - for cleaning, reporting, or graphing the data
 - Access
 - for accessing, cleaning, or reporting the data
 - SPSS or SAS
 - for accessing, cleaning, reporting or graphing the data

SAS and IR

- **With SAS you can truly create a “One Stop Shop” for IR. You can:**
 - Access your data
 - Clean your data
 - Automate reports and surveys
 - Output to Word, Excel, or PDF
 - Graph/Chart your data
 - Create geographic maps of your data

Accessing Your Data with SAS

- **Direct Access to the Database**
 - PROC SQL (default)
 - PROC SQL (pass-through)
 - PROC PWENCODE
- **Importing from other sources**
 - Excel spreadsheet
 - Tab delimited, comma separated
 - Other raw data sources

Accessing Your Data with SAS

- PROC SQL (default)
 - Specific with SQL format
 - No functions
 - Can run slower (case statements, for example)
 - Only multiline comments (/* */)

Accessing Your Data with SAS

- PROC SQL (default)

```
3 libname coll oracle user="&myusername" password="&mypassword" path=colr schema=colleague;
4
5 PROC SQL;
6 create table nopass as
7 SELECT
8     X.XCBM_PERSON_ID AS ID LABEL=' ',
9     X.XCBM_TERM AS TERM2 LABEL=' ',
10    substr(X.XCBM_TERM,1,5) AS TERM LABEL=' ',
```

NOTE: Table WORK.NOPASS created, with 16697 rows and 28 columns.

```
79 QUIT;
```

NOTE: PROCEDURE SQL used (Total process time):

real time	3:46.27
cpu time	1:55.34

- query took 3m46s to complete

Accessing Your Data with SAS

- PROC SQL (pass-through)
 - Easier SQL structure
 - Full access to functions
 - Runs much faster than default
 - Still, only multiline comments

Accessing Your Data with SAS

- PROC SQL (pass-through)

```
3 PROC SQL;  
4 Connect to oracle(path=colr user="&myusername" password="&mypassword");  
5 Create table twelve as  
6     Select * from connection to oracle  
7     (  
8     SELECT  
9         X.XCBM_PERSON_ID AS "ID",  
10        X.XCBM_TERM AS "TERM2",  
11        substr(X.XCBM_TERM,1,5) AS "TERM",
```

NOTE: Table WORK.TWELVE created, with 16697 rows and 28 columns.

```
83 Disconnect from oracle;
```

```
84 Quit;
```

NOTE: PROCEDURE SQL used (Total process time):

real time	0.55 seconds
-----------	--------------

cpu time	0.12 seconds
----------	--------------

- Same query took 55s to run with pass-through

Accessing Your Data with SAS

- **PROC PWENCOD**
 - Allows sharing without compromise
 - Include a file with the encrypted password in each program

```
1 %include 'C:\Users\SAS_Credentials\mycreds.sas';  
2  
3 PROC SQL;
```

Accessing Your Data with SAS

- Importing from other files
 - Excel

```
4 proc import
5     out=majorcips
6     datafile='C:\Users\gengo\Documents\My SAS Files\9.3\SAS_Temp\Majors.xlsx'
7     dbms=xlsx replace;
8 run;
```

```
filename data 'X:\My Personal Folder\13.01_CSCI5663\Assignment\clinical.txt';

data clinical_data;
    infile data delimiter = ',' dsd;
```

Data Clean-up with SAS

- In IR clean data is not only crucial for coordinating board reports and IPEDS but also for surveys and internal reports.
- Examples of bad data:
 - Missing data (ex: no gender, ethnicity, or major)
 - Mismatched data (ex: Undergraduate student in a Graduate level major)
 - Old or obsolete data (ex: Student placed in a phased out major)
 - Just plain old wrong data (ex: Physical Therapy student at Denton campus)

Data Clean-up with SAS

In SAS you can programmatically find data errors and list them. The list can then be used to correct the database.

```
PROC REPORT DATA=twelve_nodup PS=60 LS=110 MISSING SPLIT='\ ' NOWINDOWS

TITLE1 'Wrong/Mismatched Levels';
COLUMN CLASS PROGRAM1 LEVEL PROG_LEVEL ID;
WHERE LEVEL NE PROG_LEVEL;

DEFINE CLASS      / GROUP ORDER=DATA WIDTH=12 'Class';
DEFINE PROGRAM1  / GROUP FORMAT=$MAJORFMT. ORDER=INTERNAL WIDTH=12 'Program';
DEFINE LEVEL      / GROUP ORDER=DATA WIDTH=12 'Stu Level';
DEFINE PROG_LEVEL / GROUP ORDER=DATA WIDTH=12 'Prog Level';
DEFINE ID         / GROUP ORDER=INTERNAL WIDTH=12 'Student ID';
```

```
RUN;
```

Wrong/Mismatched Levels				
Class	Program	Stu Level	Prog Level	Student ID
FR	COMMSCI.BSPB	UN	PB	999912
SO	COMMSCI.BSPB	UD	PB	999914
MM	ADLTHLTHNP.CER	GM	GC	999925
	ENGLISH.BA	GM	UD	999918
	NURSSCIENCE.PHD	GM	GD	999922

Data Clean-up with SAS

An important component of cleaning-up your data could be redefining variables but not making any actual changes to the database.

```
data sch2;
  set sch;

  TEMP = INPUT(SECTION_NO,2.)
  IF FB_SUBJECT = 'HSM' THEN FB_CAMPUS = 'DAL';
  ELSE IF CAMPUS IN ('AUX', 'DEN', 'CHC', 'OFF') THEN FB_CAMPUS = 'DEN';
  ELSE IF CAMPUS IN ('DPH', 'DED') THEN FB_CAMPUS = 'DAL';
  ELSE IF CAMPUS = '' AND (TEMP >= 30 AND TEMP < 50) THEN FB_CAMPUS = 'DAL';
  ELSE IF CAMPUS = 'HOU' THEN FB_CAMPUS = 'HOU';
  ELSE IF CAMPUS = '' AND TEMP >= 60 THEN FB_CAMPUS = 'HOU';
  ELSE FB_CAMPUS = 'DEN';

RUN;
```

SECTION_NO	CAMPUS	FB_SUBJECT	FB_CAMPUS
02	AUX	MATH	DEN
30	OFF	EDUC	DEN
62		NURS	HOU
60	HOU	OT	HOU
50	DED	HSM	DAL
10	DEN	KINS	DEN

Automating Reports/Surveys with SAS

A report that might take days to weeks to produce in another software can be reduced down to minutes once a program is written in SAS.

Academic Program Review											
Undergraduate Courses-Texas Woman's University											
Semester Credit Hours/ Full-time Equivalent		Fall Semesters									
		2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
UG Total	Student SCH	73,299	77,626	81,210	84,861	88,326	91,537	99,181	105,517	109,813	111,427
	Student FTE	6,108.3	6,468.8	6,767.5	7,071.8	7,360.5	7,628.1	8,265.1	8,793.1	9,151.1	9,285.6
	Faculty SCH	-	-	3,970	4,257	4,519	4,303	4,250	4,456	4,554	-
	Faculty FTE	-	-	291.8	311.7	333.0	321.8	307.7	309.4	327.6	-
	Student/Faculty Ratio	-	-	19/1	18/1	18/1	19/1	21/1	23/1	22/1	-
UGL	Student SCH	36,699	38,333	40,858	41,796	42,516	43,552	47,971	52,182	55,895	57,410
	Student FTE	3,058.3	3,194.4	3,404.8	3,483.0	3,543.0	3,629.3	3,997.6	4,348.5	4,657.9	4,784.2
	Faculty SCH	-	-	1,293	1,323	1,386	1,305	1,249	1,289	1,430	-
	Faculty FTE	-	-	106.1	105.1	110.3	107.4	100.7	98.9	113.1	-
	Student/Faculty Ratio	-	-	26/1	27/1	26/1	27/1	32/1	35/1	33/1	-
UGU	Student SCH	36,600	39,293	40,352	43,065	45,810	47,985	51,210	53,335	53,918	54,017
	Student FTE	3,050.0	3,274.4	3,362.7	3,588.8	3,817.5	3,998.8	4,267.5	4,444.6	4,493.2	4,501.4
	Faculty SCH	-	-	2,677	2,934	3,133	2,998	3,001	3,167	3,124	-
	Faculty FTE	-	-	185.7	206.6	222.7	214.4	207.0	210.5	214.5	-
	Student/Faculty Ratio	-	-	14/1	14/1	14/1	15/1	16/1	17/1	17/1	-

Automating Reports/Surveys with SAS

Academic Program Review Undergraduate-Texas Woman's University

Enrollment	Fall Semesters									
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Total Enrollment	5,826	6,266	6,675	7,014	7,420	7,837	8,481	9,010	9,443	9,515
Ethnicity										
White	3,314	3,452	3,626	3,651	3,770	3,914	4,046	4,148	4,118	4,046
African-American	1,253	1,299	1,370	1,470	1,551	1,621	1,777	1,903	2,020	2,015
Hispanic	754	876	1,005	1,201	1,330	1,462	1,707	1,914	2,083	2,226
Asian/Pacific Islander	275	346	409	475	562	616	683	726	811	852
Amer-Indian/Alaskan Native	50	57	64	62	65	67	86	103	127	123
International	177	212	181	145	131	140	126	94	96	99
Other	3	24	20	10	11	17	56	122	188	154
Gender										
Female	5,479	5,846	6,215	6,524	6,898	7,182	7,699	8,155	8,516	8,506
Male	347	420	460	490	522	655	782	855	927	1,009
Student Load										
Full-time	4,328	4,554	4,839	5,007	5,072	5,327	5,811	6,269	6,575	6,722
Part-time	1,498	1,712	1,836	2,007	2,348	2,510	2,670	2,741	2,868	2,793
Basis for Admission										
High School	7	4	0	1	0	102	128	186	173	211
FTIC Freshmen	625	672	783	723	797	740	984	1,074	1,217	1,201
UG Transfers	784	822	816	946	1,015	1,139	1,352	1,360	1,418	1,277
Continuing	4,410	4,768	5,076	5,344	5,608	5,856	6,017	6,390	6,635	6,826
Average Age (years)	25.7	25.7	25.6	25.6	25.7	25.7	25.5	25.4	25.3	25

Output to Word

```
OPTIONS NODATE orientation=landscape;
ODS rtf FILE='C:\Users\tstegmair\Documents\course_freq.rtf'
STARTPAGE=NO KEEP;
TITLE;
ODS ESCAPECHAR='^';
title "^S={font_size=11pt font_face='Times New Roman'
      font_style=italic font_weight=bold outputwidth=100%
      just=c} Accountability Report";

...PROC REPORT, PROC TABULATE or other code to output...

ods rtf startpage=now;
ODS RTF CLOSE;
```

Output to Excel

```
ods tagsets.excelxp
    file = 'c:\Accountability_Degrees.xls'
    STYLE=VwaExcel /*Personal PROC TEMPLATE, default is HTMLblue*/
options ( Orientation = 'landscape'
FitToPage = 'yes'
pages_fitheight = '4'
embedded_titles = 'yes'
frozen_headers = '5'
row_repeat = '1-5'
hidden_columns = '3'
row_heights = '25,16.1,0,9.5,0,0,19'
absolute_column_width =
'13,40,.5,5.7,5.7,5.7,5.7,5.7,5.7,5.7,5.7,5.7,5.7'
width_fudge = '.64');
ods escapechar='^';
    title "^S={font_size=11pt font_face='Times New Roman'
font_style=italic font_weight=bold outputwidth=100% just=c}
Accountability Report";
...PROC REPORT, PROC TABULATE or other code to output...
ods tagsets.excelxp close;
```

Output to PDF

```
options orientation=landscape leftmargin=".35in" rightmargin=".35in"
      topmargin=".25in" bottommargin=".25in" nodate;
ods escapechar='^';
ods pdf file = 'c:\Accountability_Enrollment.pdf' style=styles.margin
      startpage = YES;

ods proclabel " Accountability Report "; /*Bookmarks your PDF*/
title "^S={font_size=11pt font_face='Times New Roman' font_weight=bold
      outputwidth=100% just=c} Accountability Report";

...PROC REPORT, PROC TABULATE or other code to output...

ods pdf startpage = NOW /*Starts a new page*/

...PROC REPORT, PROC TABULATE or other code to output...

ods pdf close;
```

Graphing/Charting with SAS

```
ods rtf text="^S={font_size=12pt font_face='Times New Roman' font_weight=bold outputwidth=100% just=1}
           {Total New Undergraduate Enrollment for Falls - 2008 to 2012: FTIC vs Transfer }";

AXIS1 STYLE=1 WIDTH=1 MINOR=(NUMBER=1) LABEL=(a=90 'Enrollment'); /* Y-axis or vertical axis */
AXIS2 STYLE=1 WIDTH=1 LABEL=('Year'); /* X-axis or horizontal axis */
LEGEND1 LABEL=('New Type') FRAME POSITION = (BOTTOM CENTER OUTSIDE);

proc gchart data=VQ2;
vbar YEAR / GROUP=NEW
           CLIPREF
           FRAME
           TYPE=FREQ
           OUTSIDE=FREQ
           COUTLINE=BLACK
           RAXIS=AXIS1
           MAXIS=AXIS2
           LEGEND=LEGEND1
           PATTERNID=GROUP;

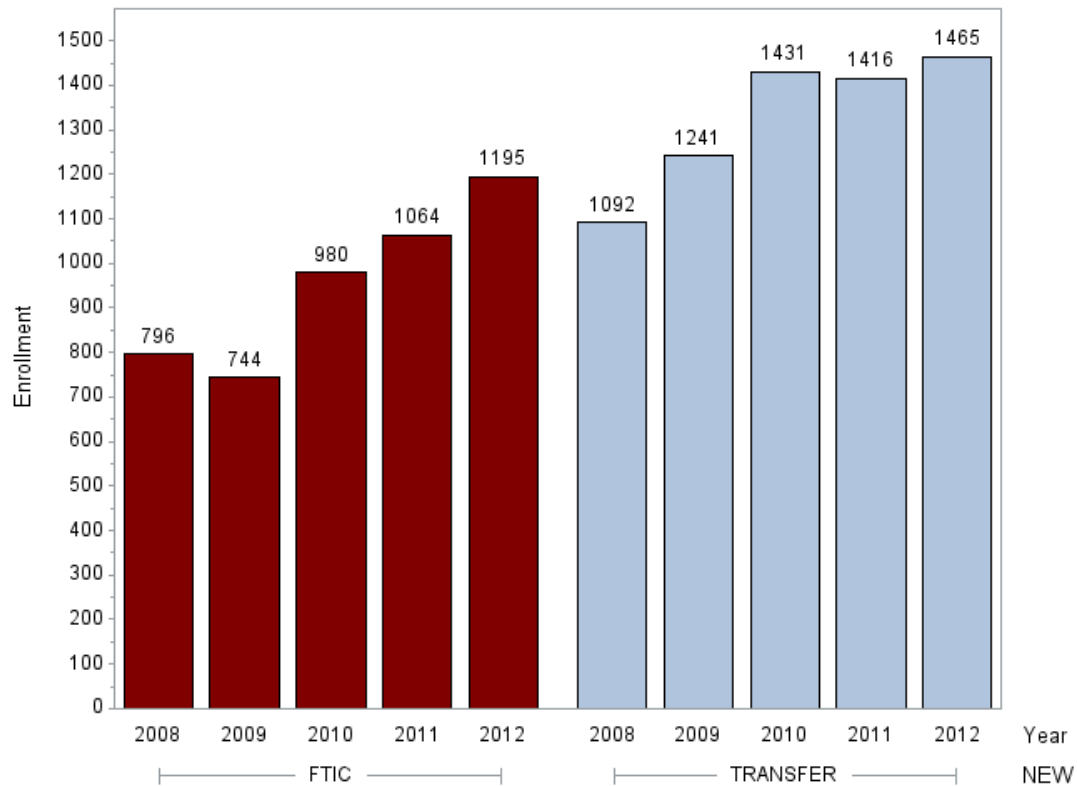
run;

PROC REPORT DATA=FINALGROWTH NOWINDOWS;
ods rtf text="^S={font_size=12pt font_face='Times New Roman' font_weight=bold outputwidth=100% just=1}
           {Percent change from previous year }";
COLUMN GROUP YEAR2009 YEAR2010 YEAR2011 YEAR2012 AVERAGE;

WHERE GROUP IN ('FTIC_CHNG', 'TRNS_CHNG');
FORMAT AVERAGE PERCENT7.1 GROUP $CHNG_GROUP. YEAR2009 PERCENT7.1
        YEAR2010 PERCENT7.1 YEAR2011 PERCENT7.1 YEAR2012 PERCENT7.1;
LABEL GROUP = "New Type" YEAR2009 = "2009" YEAR2010 = "2010" YEAR2011 = "2011" YEAR2012 = "2012";

RUN;
```

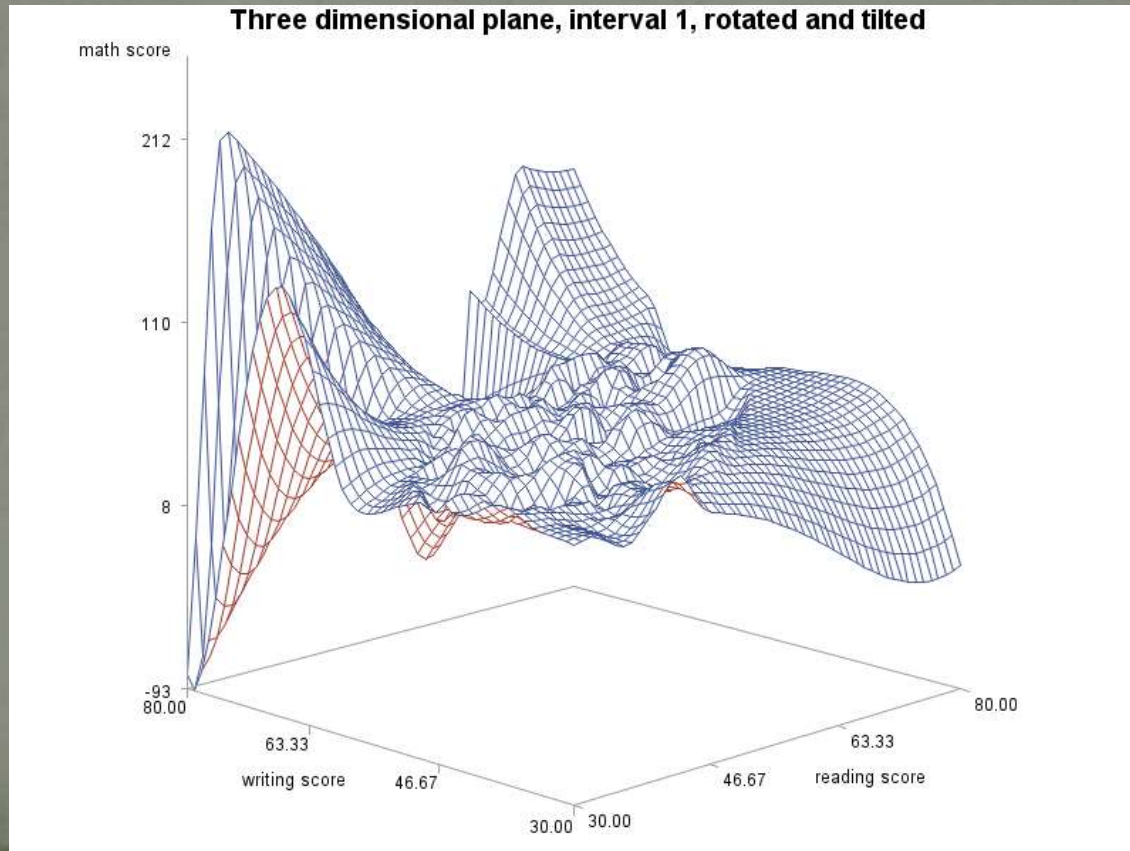
Graphing/Charting with SAS



New Type	2009	2010	2011	2012	AVERAGE
FTIC Change	(6.5%)	31.7%	8.6%	12.3%	11.5%
Transfer Change	13.6%	15.3%	(1.0%)	3.5%	7.8%

Graphing/Charting with SAS

```
proc g3d data=b;  
  title 'Three dimensional plane, interval 1, rotated and tilted';  
  plot write*read = math / rotate=45 tilt=80;  
run;
```



Thank You!

< Greg Gengo & Tracy Stegmair >

< ggengo@twu.edu & tstegmair@twu.edu >