

Creating Student Academic Profiles

Janice K. McBee, Virginia Tech, Blacksburg, VA

ABSTRACT

University academic advisors rely on data from various sources ranging from a student's experiences prior to entering the university and a student's performance at the university. High school rank, grade point average, and SAT, ACT or other such scores typically make up some of the data elements. Transfer of credits from advanced placement high school courses or from courses taken at other institutions, e.g., community colleges adds to the student profile. In addition, performance at the university should be included in a student profile. The purpose of this paper is to join/merge data from various sources (MS Excel files, Banner Oracle tables and SAS datasets) using SAS on data from various sources ranging from a student's experiences prior to entering the university and a student's profile to be used in academic advising. SAS Enterprise Guide tasks and SAS procedures were used to develop a decision model to aid the advisors of freshman students in a particular department.

INTRODUCTION

This paper is intended for the beginning or intermediate SAS or SAS Enterprise Guide user. It is an example of using SAS Enterprise Guide to join SAS datasets with a client's (end-user's) provided data. The resultant SAS dataset becomes the data for a new SAS Enterprise Guide project and can be exported for additional use by the client.

The data that we use comes in various forms. We have student census extracts (SAS datasets and flat files) that are created each term. Clients will have their own data that they want merged with university data. It is important to know your data. When one joins datasets the variable/column descriptions should match to avoid erroneous results. To develop a student profile first, edit and verify your data, join datasets, and verify the resulting dataset.

STUDENT CENSUS DATA

Prior to Summer 2000 the student census data was legacy data and the flat files were converted to SAS datasets. As of Summer 2000, student census extracts (SAS datasets) are generated. A SAS version 8 program with PROC SQL was used by members of another department to join Banner Oracle tables and create student census SAS datasets. Our Decision Support Services team modifies the student census data by creating variables and adding labels in SAS version 6.12. This accommodates our Mac users/clients and creates SAS datasets that are compatible with Strategic Enrollment Management, a software product from the

SCT Corporation. The student census extract is created every term and is used for official university reporting.

CLIENT DATA: MS EXCEL FILES

MS Excel files import directly into SAS Enterprise Guide (Figure 1). After importing, one can use the Tasks to analyze the data (Figure 2). Guide accesses data from various forms, from SAS datasets to HTML files (Figure 3). The data then can be saved as SAS datasets and merged with existing data.

Figure 1.

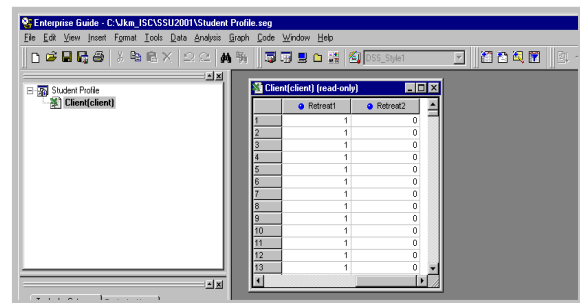


Figure 2.

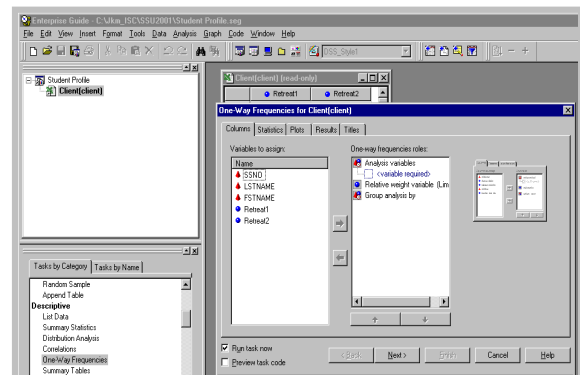
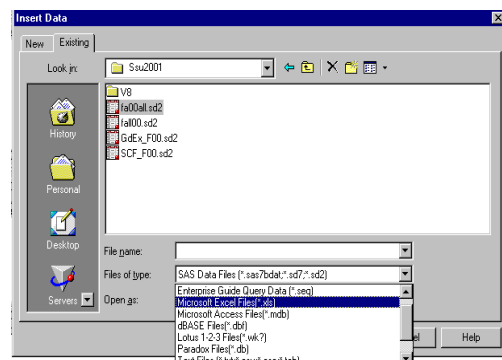


Figure 3.



The MS Excel file from the client originally had student identification (SSNO) defined as General. I changed the column attribute to text; however, when inserted into the Guide project, SSNO was interpreted as a numeric data type. Our student census files have student identification (SSNO) defined as character. Guide does not allow joining of data on mixed typed variables. Therefore, I wrote a program to modify the imported file to change the format of the SSNO to character and saved it at a SAS dataset [Client] .

JOIN/MERGE DATA

When data is in SAS datasets SAS Enterprise Guide is a versatile tool to use to merge data and then to analyze the resulting dataset. With Enterprise Guide one can merge SAS version 6.12 data with version 8 data. Note that it is better to keep SAS version 6 datasets in a different folder than SAS version 8 or one can convert datasets to a common release of SAS. In this study the census (version 6.12) data and term grade (version 8) were merged with the client's data (MS Excel converted to SAS version 8). In addition, term course grade data (version 6.12) was used to provide a current term profile (grade and description) of courses.

To merge the SAS datasets to be joined were placed in the project. I did the following:

- Placed data in project (optional)
- In Query Builder (Filter under Data) clicked on the Tables Tab (a in Figure 4)
- Clicked on Add Data button (b in Figure 4)
- In Add Data clicked on SCF_F00
- In Add Data clicked on grades_f00
- Renamed Query (a in Figure 5)
- Clicked OK (b in Figure 5) to create the Query
- Save joined data (optional)

Figure 4.

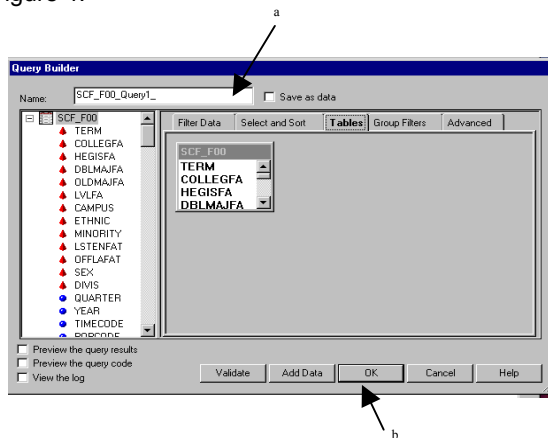
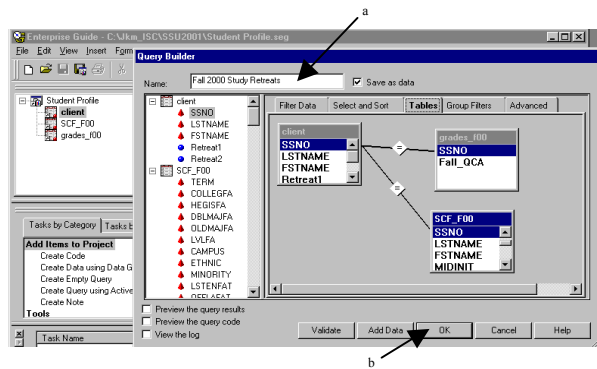


Figure 5.



ANALYSES

SAS Enterprise Guide was used two ways to assist academic advisors. One provided a profile of courses taken for the term and overall grade average. The second was used for comparison and prediction.

Profile

Faculty viewed a student's profile in Guide using the Student Term Profile project. The project uses grade extract data where each row of data represents a course that a student took during the current term. To create the profile the following was done:

- Selected Filter from the Data pull down menu (Figure 6)
- Filtered on a student's identification number (PIDM) or on name (Figure 7). I filtered on PIDM and renamed the query with the student's name. Please note that the identification numbers and names have been modified in the figures.
- Selected List Data under Descriptive in the Tasks box and chose variables (Figure 8)

Figure 6.

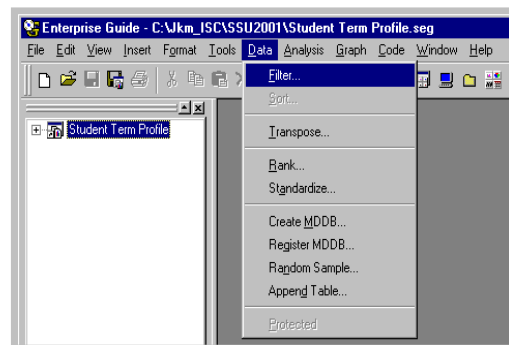


Figure 7.

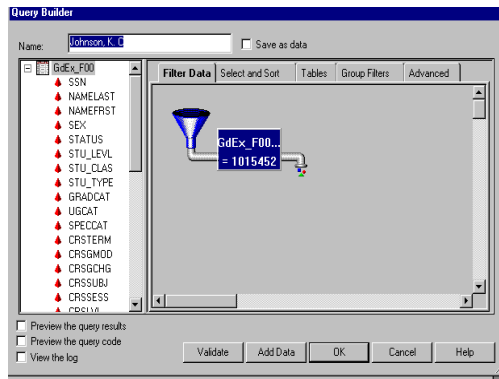
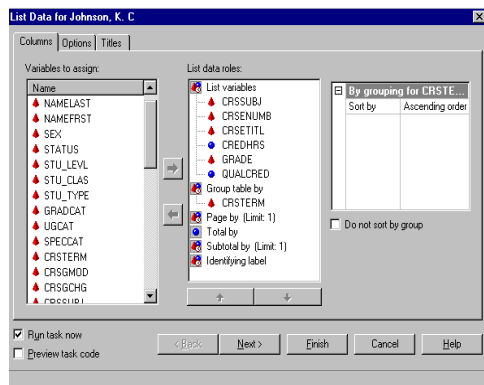


Figure 8.



An example of a profile is shown in Figure 9. An advisor can set up a filter for each advisee and resubmit the List Data or Summary during a consultation to view the current data. If the faculty member wants all of the advisees in one report, then instead of filtering the data and creating a task for each advisee, he/she may group by last name of student. It becomes a matter of personal preference. Since the advisor accesses the saved project and the data is updated, the advisor gets the most recent information for a student by re-running the task from the project window. In addition advisors can join Grade Extracts over several terms to maintain a complete course activity profile for their students.

Figure 9.

Obs	Course_Subj	Course Number	Course Title	Credit Hours	Course Grade	Quality Credits
1	MSE	3134	X-Ray Diffraction	4	A	4.0
2	MSE	4414	Physical Ceramics	3	B-	2.7
3	MSE	4424	Physical Ceramics Lab	1	A-	3.7
4	MSE	4984	SS:Kinetic Processes	3	C	2.0

Summary of Grades

Quality credits are the numeric equivalent of a letter grade. To calculate quality credit average (QCA = $\frac{\sum \text{quality credits} \times \text{credit hours}}{\text{Total credit hours}}$), I selected Summary under Descriptive in the Tasks box, chose quality credits for analysis variables, and chose credit hours as the relative weight variable (Figure 10). See Figure 11 for sample output.

Figure 10.

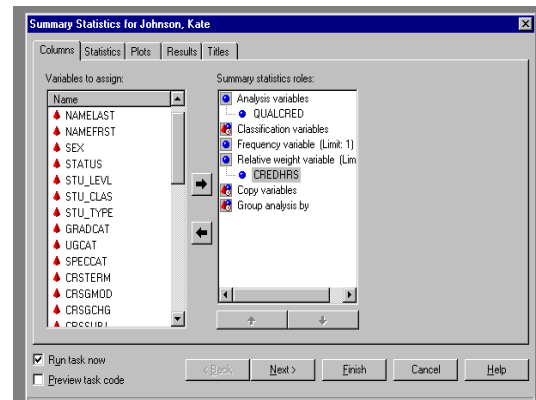


Figure 11.

Summary Statistics Results				
The MEANS Procedure				
Analysis Variable : QUALCRED Quality Credits				
Maximum	Mean	Minimum	N	Std Dev
4.0000000	3.0727273	2.0000000	4	1.6022711

Prediction

Not only is it valuable to be able to view a student's profile, but it is possible to evaluate the effectiveness of a new initiative with Enterprise Guide. In this study, the advisor wanted to know if attendance in the two offered study retreats had a beneficial effect on quality credit hours (QCA). Fall 2000 was the first time that the study retreats were offered. Students could attend both study retreats; most students chose to attend one of the sessions. The merged dataset [Fall 2000 Study Retreats] was the merge of Fall 2000 student census data with client study retreat data and Fall term grade and is listed in the Student Profile Project. The data was filtered to capture only those freshmen having major 'AZ' and having actually registered for classes in the fall. Retreats 1 and 2 were coded 1 for attending and 0 for not attending. Correlations with scatterplots were run in Guide to check for outliers. The independent variables, Retreat1, Retreat2, high school grade point average, SAT verbal score, SAT math score, percentile rank in high school, sex (dummy coded: 0=Female, 1=Male), and honors student (dummy coded: 0=No, 1=Yes) were used to predict freshman fall QCA. Students missing data were excluded from the analysis.

The correlations of Retreat 1 and Retreat 2 with Fall QCA were not significant at the .05 level of significance. Results from the initial correlations indicated the correlation of those who attended Retreat 2 were slightly higher with Fall QCAs than the correlation with attending Retreat 1. As expected, high school grade point average had the highest correlation (.34) with freshmen Fall QCA. Although the regression model was significant ($p < .0001$), neither Retreat 1 nor Retreat 2 contributed significantly ($p < .05$) to the prediction of the Fall QCA of students in 'AZ'. The model explained only 21 percent of the variance of Fall QCA. Results from the regression may be viewed in Figures 13 and 14.

Figure 12.

Effect of Study Retreats--Fall 2000					
Program = AZ					
The REG Procedure					
Model: Linear_Regression_Model					
Dependent Variable: Fall_QCA QCA					
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	124.61771	13.84641	32.32	<.0001
Error	1045	447.75446	0.42847		
Corrected Total	1054	572.37217			

Root MSE	0.65458	R-Square	0.2177
Dependent Mean	2.58498	Adj R-Sq	0.2110
Coeff Var	25.32242		

Figure 13.

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	-0.95571	0.29717	-3.22	0.0013
Retreat1		1	-0.00413	0.05161	-0.08	0.9362
Retreat2		1	-0.00540	0.05303	-0.10	0.9189
SATV	Verbal SAT Score	1	-0.00003309	0.00033075	-0.10	0.9203
SATM	Mathematical SAT Score	1	0.00225	0.00037206	6.05	<.0001
HS GPA	High School GPA	1	0.47555	0.06909	6.88	<.0001
PERCENT	Percentile High School Rank	1	0.38442	0.21904	1.76	0.0795
TRANHRS	Transfer Credit Hours	1	0.00446	0.00243	1.84	0.0668
SEX_NUM		1	0.05651	0.05386	1.05	0.2943
HONNORNO		1	0.35200	0.07706	4.57	<.0001

CONCLUSION

The SAS version 8.2, coupled with Enterprise Guide 1.2, is a versatile tool. Clients can custom design their projects by renaming datasets and joining data from various sources. Someone very familiar with the structure of the Oracle tables that one is using should either create or assist in the development of tables/views. When data changes or is updated the project links to the updated data. Faculty advisors can merge departmental data to existing university data and create their own prediction models and make their own comparisons. One can save the source code, make modifications and run additional analyses.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Janice K. McBee
 Decision Support Services
 Virginia Tech
 1700 Kraft Drive, Suite 2000
 Blacksburg Virginia 24060
 Work Phone: 540/231-3601
 Fax: 540/231-2429
 Email: jmcbee@vt.edu

SAS is a registered trademark or trademark of SAS Institute Inc. in the USA and other countries.