A GUI FOR ENHANCED INSIGHT INTO DATA IN A UNIVERSITY SETTING: SAS/EIS® AND SAS/AF® FRAME TO THE RESCUE

Cecil Hallum
Prof. of Mathematical Statistics
Sam Houston State University

Melinda Miller
Asst. Professor of Mathematical Statistics
Sam Houston State University

ABSTRACT

This paper highlights a "point-and-click" graphical user interface (GUI) capability for monitoring and assessing information from student, faculty and related university data bases that puts control in the hands of the user (faculty, administrative personnel, etc.). This capability utilizes the SAS/EIS and SAS/AF Frame to permit retrieval and reporting of strategic information in real-time. Specific enhancement of the ability for faculty and administrative personnel to have "ease of access" to key information in support of numerous tasks including that of student advising, diagnostic assessments, and standard information retrieval is achieved using this capability.

SYSTEM CHARACTERISTICS

Many academic institutions are moving toward data base management and information retrieval system software to process regularly scheduled reports of information from large, multifaceted, student and faculty data bases. While customized reporting is typically available, oftentimes, personnel that produce reports operate in an overload status and such requests take hours or days to process. Lacking on most campuses is the ability for faculty and administrative personnel to directly access such data, retrieve information in a user-friendly, highly informative manner, as well as generate their own customized reports, charts, plots, etc.

The described system is a GUI that maximizes reliance on a "point-and-click" environment with the overall system possessing the following key characteristics:

1) Emphasis is on "ease of access" for administration and faculty.
2) Summary diagnostic information is provided about students collectively at each level of the demographic variables.
3) Individual student information is available through profile charts and tables.
4) Campus-wide accessibility is achieved while maintaining data base security.
5) Icons and objects are configured and labeled in an attempt to be self-explanatory.
6) "Help" buttons are utilized through-out.

DATA BASE CHARACTERISTICS

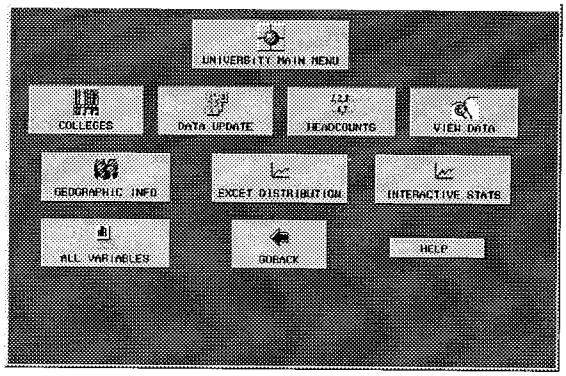
The system discussed herein manipulates the university student data base which is comprised of numerous variables including: student name, social security number, gender, ethnic background, birth date, SAT scores, ACT scores, TASP scores, transfer grade point averages, department, major, high school code, county residence, state residence, and numerous other key variables. The entire system is structured to minimize response time via preprocessed data sets that are of minimal size.

As is well-documented, (e.g., Polzin (1994) and Habich (1995)), special data preprocessing is warranted in regard to the handling of large data sets in SAS. If

users are to experience "real-time" access and manipulations of the various data, it is essential to preprocess data (e.g., using PROC SUMMARY or PROC SQL) to keep the size and number of variables to a minimum. The university data discussed in this paper is updated a couple of times per semester, consequently, the preprocessing program is re-run upon each update. The preprocessing not only includes the generation of summarized data from the use of PROC SUMMARY, but also utilizes the LENGTH statement and COMPRESS option to assist in keeping the data sets to a minimum. A decade of our data currently comprises less than 100 megabytes; running the preprocessing source code results in the overall raw data set being subdivided into 4 university-level summarized data sets (one for each college on campus), followed by a further subdivision down to the departmental level.

A SAMPLE OF SCREENS

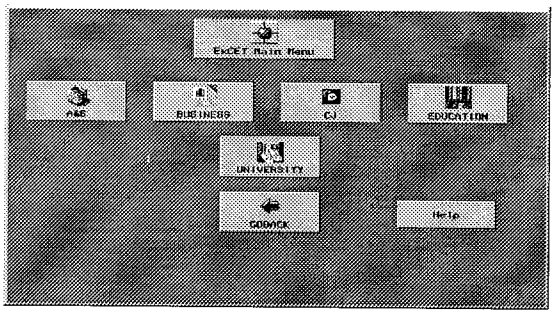
A sampling of screens is discussed for clarity. After a welcome screen the general screen below appears (Screen 1).



Screen 1: University Level Screen

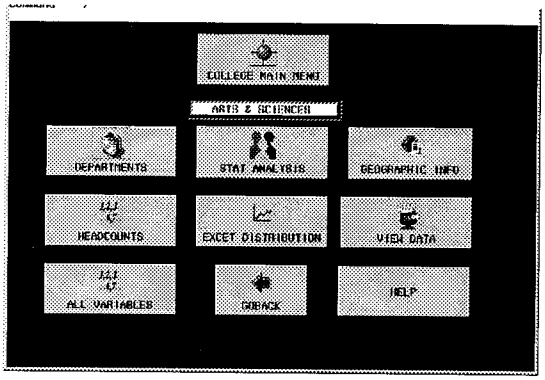
This screen allows the user to access information for a particular college (by indexing COLLEGES) or to remain at the university level by clicking on another icon. This screen is EIS-based and the icons are essentially self-explanatory.

Selecting the COLLEGES icon results in Screen 2. Clearly, the user begins the specialization of the search here with the college selection. Assuming the A&S (i.e., the College of Arts & Sciences) icon is



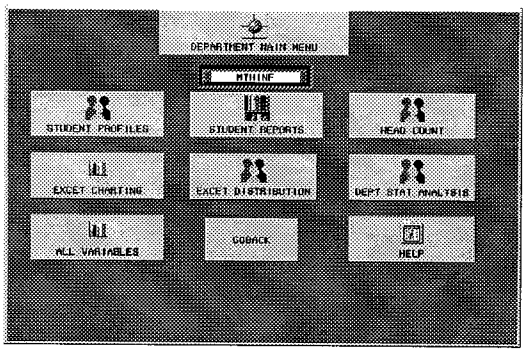
Screen 2: College Selection

indexed, the user is taken to screen 3, which provides a number of specialty looks at the college level information. Screen 3 is the standard screen that appears regardless of which college was chosen, except that the remainder of the search is specific to the selected college's students. If one now clicked on the "DEPT SPECIFICS" icon, a popup menu appears and displays all departments from the given college; after clicking on a particular department,



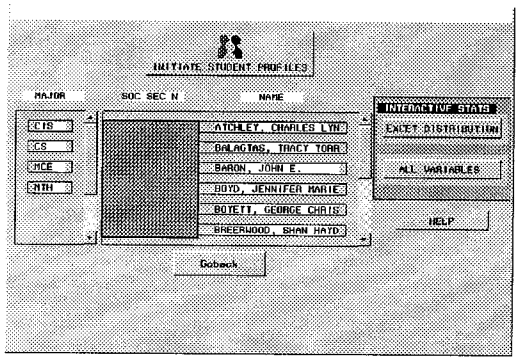
Screen 3: Arts & Sciences Screen

screen 4 appears and permits detailed departmental level student information to be accessed.



Screen 4: Departmental Screen

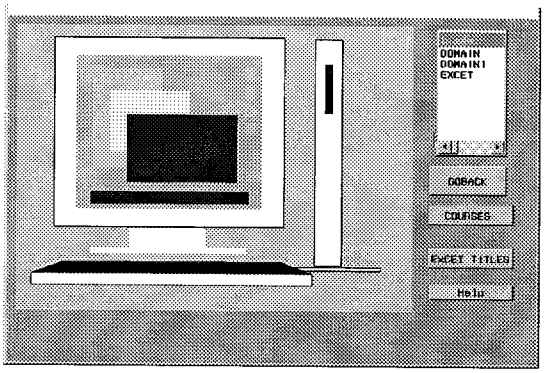
Continuing on this path, upon clicking on the "STUDENT PROFILES" icon, screen 6 appears. This screen makes use of two AF Frame extended tables; the user may click on the "major" for the student being sought in the first extended table. The second extended table adjusts immediately to all students of that "major" and shows the various students' names listed alphabetically as well as their social security numbers (both, of course, are excluded from the printout here).



Screen 5: Student Selection Screen

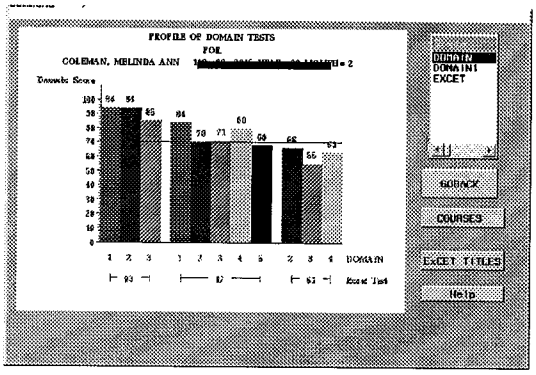
Clicking on a particular student's name leads to screen 6 which provides charting of this particular student's performance on a number of tests (note that this information is, again, excluded from the enclosed screen printout for security purposes). Numerous other outputs will be discussed in the conference paper.

Screen 6 makes use of several Frame objects including a listbox and graph output object.



Screen 6: Graph Output Object Screen

Note that the listbox is automatically populated with the names of generated PROC GCHART plots that provide information on the selected student. Upon clicking on one of the chart names (e.g., the name "DOMAIN"), the chart in screen 7 appears, allowing for visual insight into key information about this student.



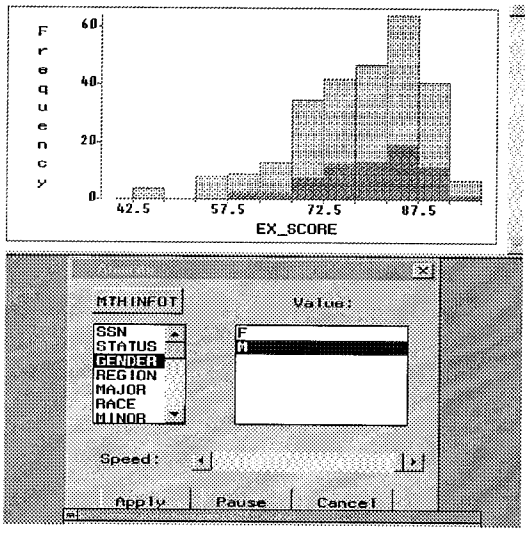
Screen 7: Student Profile Plot

Approximately 30 screens comprise the overall system, with only a select few included in this presentation. While the above screens provide something of a tour of the system several others are worth highlighting.

HIGHLIGHTED FEATURES

PROC INSIGHT:

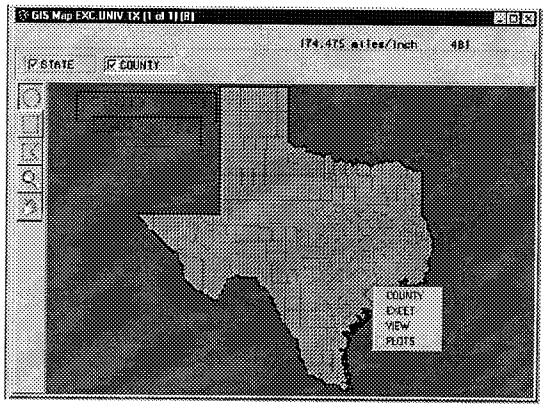
These include the highly useful GIS module in SAS for detailing geographic information as well as numerous SAS procedures including INSIGHT along with its ready-built "point and click" environment. In particular, screen 8 depicts a use of the "animate" part of INSIGHT along with the "BAR" distributional chart in seeing what part of a given distribution is attributable to differing levels of attribute variables (e.g., such as gender, ethnic background, etc.).



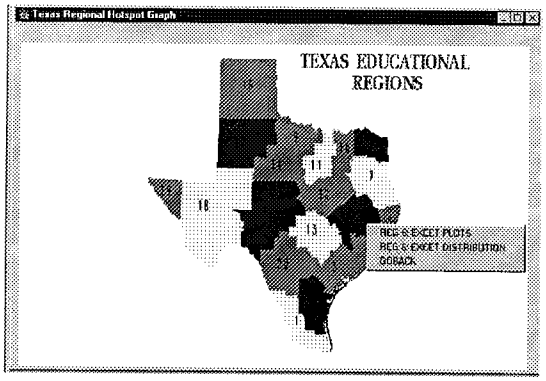
Screen 8: INSIGHT with Animate Option

GIS:

The use of the SAS GIS in conjunction with AF Frame and EIS provide key geographical information characterizing the university data. Screen 9 is the GIS screen of Texas counties with several options that are built in using the "actions" function of GIS and the hotspot and listmenu capability of EIS. Clicking on a particular county (or union of counties) and then one of several "actions" permits ease of attaining geographic information. A couple of selectable actions result in the running of underlying SAS programs while another simply uses PROC FSVIEW to view the selected data. Screen 10 shows the 20 educational regions in Texas that are hotspotted using the EIS Hotgraph to permit accessing regional information.



Screen 9: GIS County Level Data

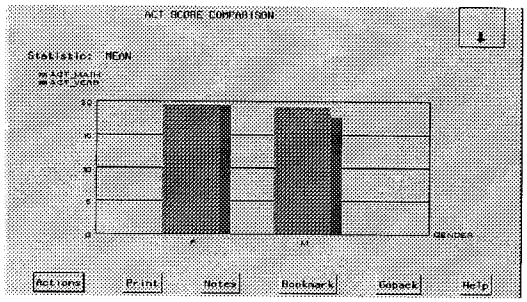


Screen 10: EIS Student Hotgraph

SAS 6.12 OBJECTS:

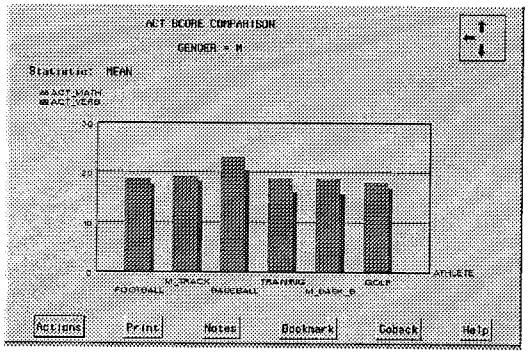
A part of this overall system was developed using SAS 6.11 on a UNIX-based machine; it was then transferred to the university PC DOS system using PROC CPORT and CIMPORT. At the initiation of this development, SAS 6.12 was not yet available for all UNIX environments. With 6.12 recently becoming available on our campus PC DOS machines running under Windows NT via the multiserver environment, we subsequently made use of a number of the

new objects in EIS such as the GROUPBAR object. An example of its application to the university athletic data is depicted in the follow screens. Screen 11 depicts the mean ACT Mathematics and ACT Verbal scores of all athletes by gender.



Screen 11: ACT Groupbar Screen

Clicking anywhere on the screen allows one to specify a gender for consideration. Choosing *M* leads to screen 12, which represents the mean scores of male athletes by sport.



Screen 12: Groupbar Hierarchy Screen

Using the arrows in the upper right corner, one may consider only females by clicking the left arrow. The down arrow leads to a screen containing mean scores of all male athletes, categorized by class. The up arrow leads to the previous screen. By clicking anywhere else in the screen, one may restrict the analysis to one sport. One may continue this process through a hierarchy of student variables.

KEY EIS AND AF PROGRAMMING SPECIFICS

Below are summary recommendations and insights gained from this effort that are of particular interest to pass along to future and current programmers that use EIS and AF Frame. The listed items are worth every minute one takes to become familiar. These are not listed in any particular order of priority:

1. The CALL DISPLAY function in combination with the ENTRY command saved considerable time and storage by permitting multiple usage of various screens, particularly with the passing forward of data sets and variables from one screen to another.
2. The use of a preprocessing program that makes all PROC SUMMARY runs each time the database is updated is mandatory to achieve efficiency (in speed and storage).
3. The discouraging task of becoming proficient with SAS AF Frame (in view of the "mountain" of available commands to master), is best accomplished by building from just the select few, initially, that gets the job done (i.e., minimize the "bells and whistles"). Enhancements will become obvious as you progress. Accelerating up the learning curve was accomplished very effectively by referencing users' papers from *SUGI Proceedings* (particularly those over the past 4 years – e.g., see Golby et. al. (1995), Timbers (1997), and Destiny Corp. (1997)).
4. Time deadlines will oftentimes require that tougher problems be addressed to the SAS technical consultants. Their assistance has proved to be invaluable, timely and most courteous.
5. Enough cannot be said about the use of INSIGHT in conjunction with the ANIMATE option as a most insightful

tool for visually depicting information related to attribute data. To do this, an icon is included on the majority of screens with underlying simple SAS source code that is as simple as:

```
PROC INSIGHT NOMENU
NOBUTTONS;
OPEN   STDNT.MATH\
NODISPLAY;
CLASS GENDER RACE MAJOR;
BAR GPA;RUN;
```

This approach was particularly useful to attain diagnostic information related to students' GPA, ACT, SAT, and TASP scores.

SUMMARY

It is hoped that the reader will benefit from these additional examples and experiences related to the use of SAS EIS and AF Frame in manipulating data bases in a university setting. The sample of screens and associated discussion, hopefully, effectively adds to the "snowballing" information about object oriented programming (OOP) and related capabilities available in SAS. The startup learning curve is steep and discouraging, but the payoffs appear to warrant the effort. A detailed laptop PC demonstration will be given at conference time.

REFERENCES

Destiny Corporation (1997), "Getting Started with SAS/EIS Software", *Proceedings of the Twenty-Second Annual SAS Users Group International Conference,* 96-103.

Golby, Loretta and Dupuis, Serge (1995), "Prevent Pane-Moving from MENU & PROGRAM Entries to FRAME Entries", *Proceedings of the Twentieth Annual SAS Users Group International Conference,* 1-5.

Habich, Candy R. (1995), "Tips, Techniques, and Good Habits for SAS/AF Application Developers", *Proceedings of the Twentieth Annual SAS Users Group International Conference,* 6-11.

Polzin, Jeffrey A. (1994), "DATA Step Efficiency and Performance", *Proceedings of the Nineteenth Annual SAS Users Group International Conference,* 1574-1580.

Timbers, Vincent L. (1997), "An Introduction to Developing Applications with SAS/AF Software FRAME Entries", *Proceedings of the Twenty-Second Annual SAS Users Group International Conference,* 368-374.

ACKNOWLEDGEMENTS

AUTHOR CONTACT

For further discussion or additional information, contact:

Cecil Hallum
Dept. of Math. & Info. Sciences
Sam Houston State Univ.
Huntsville, TX 77341-2206
Ph: (409)-294-3706
E-mail: hallum@fisher.shsu.edu
Fax: (409)-294-1882

Melinda Miller
Dept. of Math. & Info. Sciences
Sam Houston State Univ.
Huntsville, TX 77341-2206
Ph: (409)-294-1596
E-mail: mth_mam@shsu.edu
Fax: (409)-294-1882