

**Paper # P505
Florida Community College System
Putting Minds to Work**

**Jeanette C. Humphrey
Tallahassee Community College**

**Howard Campbell
And
Brian Walsh
Division of Community College
Department of Education
State of Florida**

In the Beginning:

In 1988 the legislators in the State of Florida mandated that the Florida Community College System, comprised of the 28 community colleges in Florida, begin reporting student type information to the state in a format to be defined by the Division of Community Colleges (DCC). The purpose of this directive was to give the state access to detail student level information, allowing the state to produce federal Integrated Postsecondary Education Data System (IPEDS) reports, state reports for funding, and conduct detail data research and analysis. Information Systems of Florida (ISF) was awarded the state contract to produce a long-range plan and database design specifications, including data definitions.

The Plan:

ISF's completed its study and suggested the following:

- The data should be reported by the community colleges annually to the state in the following manner:
 - a. Summer term data should be submitted once, in October;
 - b. Preliminary fall term data should be submitted once at the beginning of the term (October) and final fall data submitted after the term is completed (January.)
 - c. Preliminary spring term data should be submitted once at the beginning of the term (January) and final spring data after the term is completed (June.)
- For the Initial Phase of the project, the data should be used only for federal reporting, such as IPEDS. During this initial phase, the colleges should receive reports, which they could verify locally. After allowing the colleges a few years to verify their data, the state should then use Full-Time Equivalent (FTE) generated from the Student Data Base (SDB) for funding and research.
- The colleges should report the data in an ASCII format. The DCC could then use the software of its choice to read the ASCII files into a relational type database.
- The following data record definitions were suggested:
 - a. Demographic Record – (max occurs 1 time) with demographic type information, including accumulators for total student hours taken by term and

SAS SOLUTIONS AND VERTICAL PRODUCTS

- to date. Include only students enrolled, graduating, or receiving acceleration credit during the current term.
- b. Entry Level Test Record - (can occur multiple times) with most recent Entry Level Test (ACT SAT) information.
 - c. Acceleration Record – (can occur multiple times) with Credit by Exam taken during reporting term.
 - d. Program of Study Record – (can occur multiple times) with the Students declared program(s) of study.
 - e. Completion Record – (can occur multiple times) with the program completion information.
 - f. Course Record – (can occur multiple times) with information on course(s) taken during reporting term.
 - g. Financial Aid Record – (can occur multiple times) with information on financial aid received during reporting term.

ISF's plan was accepted, as proposed. Detail information on the current Student Data Base Dictionary is available at the following DCC web site:

<http://www.dcc.firm.edu/dccpubs.htm#deds>

From the college's perspective, the DCC realized the financial hardship created from expending resources to verify data that was produced primarily as a reporting tool. Producing the data was difficult enough; to expect the colleges to verify the data to the level desired was unrealistic. In a presentation to the *Florida Assn. for Institutional Research* in 1989, Howard Campbell shared his vision for the future: the state would use this data for reporting, funding, and research. In addition, the colleges could use this same data for research and local reporting purposes, utilizing the efforts expended for local benefit. The state would also benefit from college participation: the increased utilization by the local institution, would allow more data anomalies to be found, explained, and corrected.

After reviewing system requirements, the DCC chose the SAS programming language, primarily because of its efficiency, ease of use, and connectivity to IBM's Data Base DB2 relational database, which was mandated by the state legislature.

To assist the DCC in data related questions, the state created Management Information System Task Force (MISATFOR), an advisory group for the purpose of discussion of questions and problems related to the data. Currently, the President of each college appoints one member to MISATFOR. In addition, personnel from the DCC and the registrars group are included. MISATFOR meets six times a year.

As the years passed, the use of this data by the state has increased dramatically. To provide follow-up information on our students, this data is now matched against other state databases. For example, to determine success of our students after leaving the community college, the college data is matched to state unemployment files, the State University System, Private College Stipend files, Public High School files, Job Training Partnership Act (JTPA), WAGES, Public Assistance Files, and military files. Community colleges are now aware of how many of their students go on to get a high wage jobs, go into the military, or eventually graduate with a bachelors degree. In addition, the information fed back to the colleges identifies whether students were economically or academically disadvantaged.

Another use of the data is performance based funding and accountability. Much of the college's state revenue is now based on outcomes, not the number of seats filled in a classroom. As the funding became more and more tied to the data, the importance of accuracy increased.

Micro-computer Project – Sharing the Information:

In 1995, Carol Hawkins, Dean of Information Technology at Polk Community College (PCC) in Winter Haven, FL., wrote a proposal to the Division of Community College to fund the creation of a micro-computer system that would emulate the system and data stored at the DCC. Ms. Hawkins proposed that PCC provide the technical expertise for the project, in exchange the DCC would share costs for the project and distribute the resulting system to the other community colleges.

Phase I: The following year the DCC gave PCC the first of three contracts to produce a stand-alone point and click turnkey computer system to be shared with the other Florida Community Colleges. The programming language for this turnkey system was designated by the DCC to be SAS. The core of this system was to interact with a CDROM of all the data from all the community colleges for one academic year. The DCC would supply a CDROM to the colleges after the close of each reporting year.

Included in the initial Phase I were funds to purchase hardware, a Dec Alpha AXP; software, including the SAS Academic Computing Offer; and funds to provide SAS training. In addition, the contract included funds for personnel costs to design the system and develop a prototype. Jeanette C. Humphrey, Coordinator of Research and Reports at PCC, was designated as the project leader, analyst, and programmer. At this point in time, Ms. Humphrey had programming experience; however, had never used SAS.

Phase II: In Phase II, conducted the second year, Ms. Humphrey transported the prototype system to run on a microcomputer running Windows NT, which was now able to handle the large data files. In addition, she added additional reporting modules, on-line documentation, and a module to interact with 'Local' ASCII Student Data.

The system contained programs to load the ASCII data into SAS datasets with similar data names as those found on the CDROM. This ASCII data was the same ASCII student data sent by the college to the DCC each term and contained all the student information of enrollments at the local college, including student social security number and student name. This step was an important inclusion for a number of reasons. First, the data could be used in a timely manner. As soon as the data was pulled for submission to the state, the ASCII data could be read into SAS datasets and used by the college research personnel to verify the data submission, before the end of the submission window. Second, since this data contained the student social security number and name, the college mainframe system could be used to verify questionable data. Finally, for complex projects, like tracking systems, this was the best data, since it contained student identifiers that could be matched against additional data from the institution's mainframe and other local sources.

Phase III: The final Phase of the project provided funds to purchase a computer for each of the 28 community colleges, provide training, and user documentation to the colleges. In addition, enhancements included a Local side for the relatively new state mandated Personnel Data Base

SAS SOLUTIONS AND VERTICAL PRODUCTS

and Facilities Data Bases. College personnel were trained at PCC by Ms. Humphrey and returned to their home campus with their computer.

System Requirements:

The following is a synopsis of the important components of the Micro-computer Project.

Hardware/Software Requirements: The basic system provided to each of the 28 Florida Community Colleges was a COMPAQ 400 MHz Pentium with 12 gig hard drive, running Windows NT. Minimum SAS components necessary to run the system were: Base SAS, SAS/AF, SAS/FSP, SAS/Graph, SAS/Assist, and SAS/STAT. In addition, Word and Excel were also included. Application software required 2 gig of hard drive, leaving 10 gig for reports and SAS temporary data sets. In testing, we found that more memory the better, since SAS uses all the memory it can find. Also, a significant amount of free disk space was required when working with large datasets.

Data: The Micro-computer system used either SAS Datasets provided to the colleges on a CDROM or SAS datasets from the Local ASCII data. The CDROM contained all the data for all the community colleges for one year. The student's name was deleted from the dataset and an identifying number, generated by an algorithm, replaced the social security number. This same algorithm was also used in successive years, allowing the tracking of a specific student over the years, without identifying that student. The Local data created by the system was from the same ASCII student data provided term by term to the Division of Community Colleges.

Documentation/Training: Ms. Humphrey at PCC provided the initial training to the community colleges. In addition, a step-by-step documentation manual was included with each system. SAS Training by SAS Institute was highly recommended.

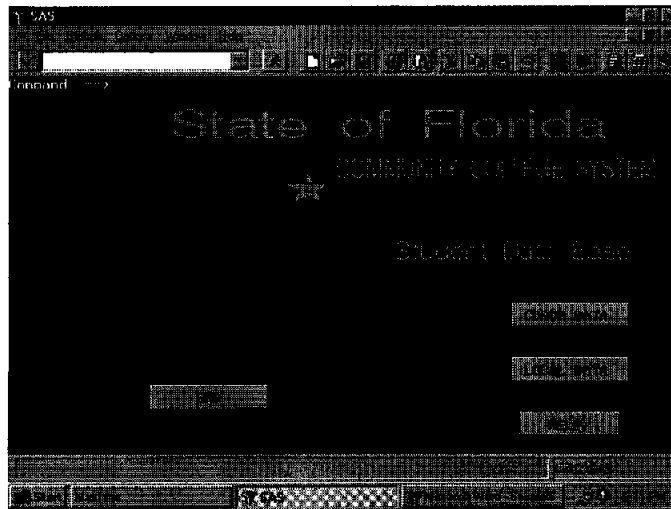
Updates/New Procedures: In the years following, the DCC provided new CDROMS of the current years data to the colleges at the close of each reporting year. These CDROMS have continued to work with little or no maintenance of the original system. Enhancements and updates to this system have been taken over by the DCC and are maintained at the state level. The state is currently evaluating the value of the new SAS/SCL, now SAS Component Language.

Knowledge: Without the excellent training by SAS Institute and support by SAS Technical Support, this project would not have been possible.

System Design:

Initial phase of the system included two parts: The CDROM section provided interaction with the Annual CDROM. This CDROM contained all the data for all the community colleges for one year. The Local Data section allowed interaction with the original ASCII student data sent term by term to the Division of Community College. This original data was read into SAS Datasets by the Micro-computer system.

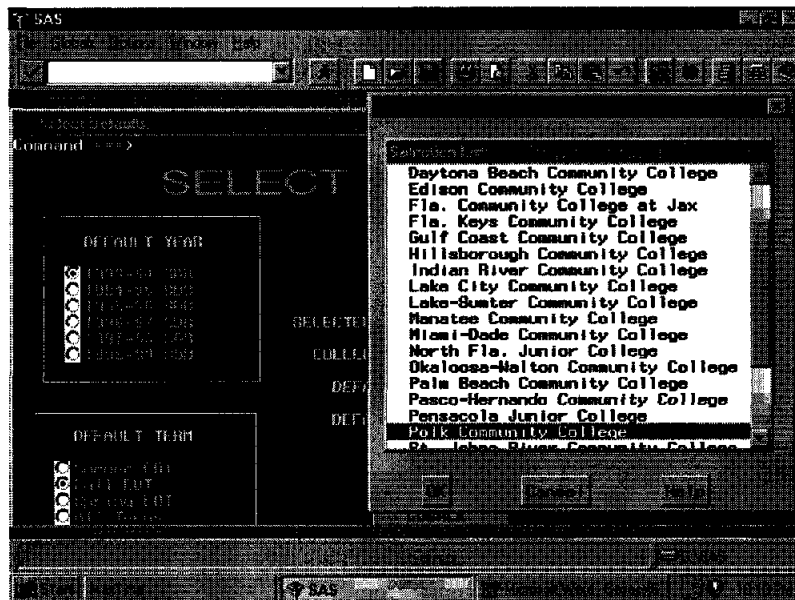
the Annual CDROM. This CDROM contained all the data for all the community colleges for one year. The Local Data section allowed interaction with the original ASCII student data sent term by term to the Division of Community College. This original data was read into SAS Datasets by the Micro-computer system.



System Data Structure: The location of the SAS data sets for use by the system was an integral part of the system design of the 'Local' side of the system. Each term of student data resided within a subdirectory named the actual term identifier. For example, the Fall Term, known as term 2, in the year 1999 would be identified as 299. Each of the subdirectories are located under C:\SAS_DATA, so Fall 1999 would be found under C:\SAS_DATA\299. Using this structure enabled the use of macros for processing with relative ease.

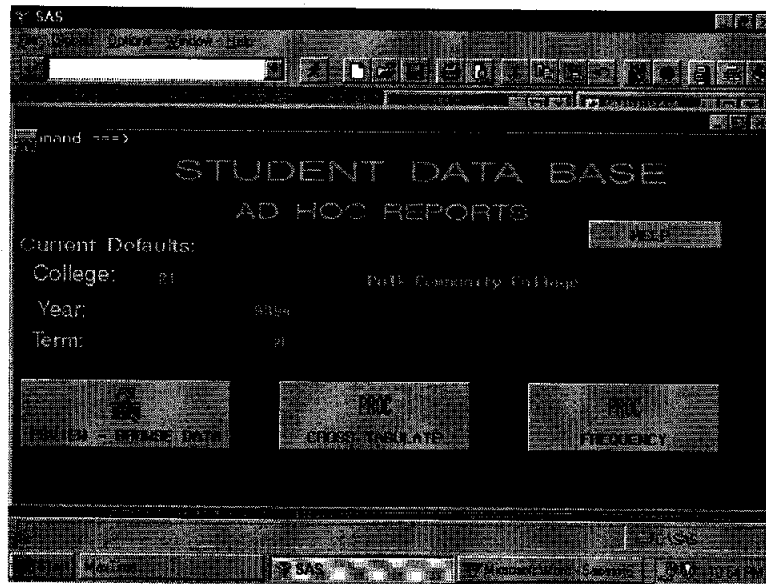
Similarly, in the CDROM side of the system, the datasets on the CDROM have the relevant year embedded as part of the name of the dataset. Using this method, verification that the correct CDROM resides in the drive can take place before reading a dataset.

The system took advantage of many of SAS/AF abilities to provide Pop-up selection lists and radio boxes.



SAS SOLUTIONS AND VERTICAL PRODUCTS

Another beneficial section of the system allowed the users to use FSVIEW to scroll through their data or produce cross tab or frequency tables. Again, the system utilized the point and click



features of SAS/AF.



Some Favorite SCL Code:

- VARLIST - to populate a pop-up list from a SAS Dataset, without hard coding variable names (see example above) ;
- PMENU – to provide interactive pull-down menus;
- PREVIEW WINDOWS – to view output;
- SUBMIT CONTINUE – to imbed SAS Code in SAS/SCL;
- REPLACE – to store string in a SAS/SCL Variable;
- Modular 'Link' Programming;
- MACRO Substitutions;

- Passing Variables between SAS/SCL Programs;
- Where statements applied to a SAS/SCL Dataset;
- On-line Help using the Help functions.

Conclusion:

As a result of this project, both the state and the community colleges have benefited. The most obvious benefit to the state has been more accurate reliable data. In addition to the state benefits, many colleges have experienced the following benefits:

- Establishing the Student Data Base as a research base has provided the Florida Community College System with a reliable source of data for research and improved the accuracy of the data used in state reporting and funding.
- Sharing of SAS programs between sister community colleges has increased good will among the sister colleges. Since all the colleges are using the same data structures, only minor modifications are necessary for the programs to run at the sister college. Many SAS programs have been shared between the reports/research personnel at the colleges, including some extensive SAS systems. For example, Tallahassee Community College recently shared a tracking system written in SAS using the Local SDB data. This system is macro driven and easily modified for use at any of the colleges. Seminole Community College also shared a comprehensive tracking system developed by consultant, W. M. Consulting in Tallahassee. They are hoping that future costs of enhancements to their tracking system may be shared by other community colleges. Of course, future benefits would also be shared.
- Using this data for research has increased the understanding of the databases, how they relate to the college, and how they relate to other state databases.
- Colleges are now able to replicate the state SAS programs, which are used to report data and produce funding. This ability allows colleges to further understand and verify the state's programs and identify reasons for data anomalies.
- Florida Community College SAS Users group was created to share resources and provide training. The Users Group meets in conjunction with the semi-monthly MISATFOR meetings and has successfully addressed many SAS data issues pertinent to the community college system.

Acknowledgements:

TCC for the time to prepare this paper and its support of the FCCS SAS Users Group.

Carol Hawkins, Vice-President of Institutional Effectiveness, Planning, and Information Systems of Seminole Community College, formerly employed by PCC, for sharing the vision, writing the grant proposal, and supporting the project.

Florida State Legislature for funding the projects.

SAS Technical Support for valuable support with quick response time.

Addresses:

Questions pertaining to this article should be addressed to:

Jeanette C. Humphrey
Tallahassee Community College
444 Appleyard Dr.
Tallahassee, Florida 32304-2895
(850)201-6083
Humphrej@tcc.cc.fl.us

Howard Campbell or Brian Walsh
Bureau of Research and Information Systems
Division of Community Colleges
Department of Education
1314 Turlington Building
325 W. Gaines Street
Tallahassee, Florida 32399-0400
(850)488-8597
Howard@flccs.org Brian@flccs.org