

Paper 013-2007

To Cube or Not to Cube: Is OLAP the Better Choice?

Evangeline Collado, University of Central Florida, Orlando, FL
M. Paige Borden, University of Central Florida, Orlando, FL

ABSTRACT

There is a continuing demand across the higher education sector for increased extraction and analysis of university data to facilitate strategic decision-making and management of university programs and activities. This paper will detail the migration of a historical Enrollment Profile detailing student headcount information designed in Base SAS® and SAS/IntrNet® to the creation of an Enrollment Cube using SAS® OLAP Cube Studio and SAS® Enterprise Guide software. A comparison will be made between the new Enrollment Cube and historical version of the Enrollment Profile, focusing on functionality, usability, and maintenance. We will present the pros and cons of each tool and allow the reader to decide which format provides better functionality for this type of reporting. The writers' experience presented in this paper is applicable to any audience interested in ways to provide aggregate data to its user community.

WHO WE ARE

MISSION

The mission of the Office of Institutional Research is to provide information of the highest quality, and which is both timely and easily accessible, and to facilitate and enhance decision-making, strategic planning, and assessment at the University of Central Florida.

PURPOSE

Institutional Research (IR) provides electronic and web-based dissemination of official information to the University community (including the Board of Trustees, the various colleges, departments and other academic and administrative units), external agencies, and the Florida Board of Governors (BOG). IR generates, supervises or develops all official University data reports and state-required reports, as well as providing end-user data solutions and training to facilitate decision-making processes. The director and staff serve on numerous university-wide committees and workgroups and assist with the collection and interpretation of institutional data, assist in planning academic programs, and participate in the implementation of evaluative procedures. The functions of the office support the entire university enterprise.

BACKGROUND

Institutional Research has, as one of its primary objectives, the task of developing and implementing a university data warehouse. This enterprise-wide data system will facilitate the creation, access, and dissemination, by internal and external stakeholders, of institutional knowledge pertinent to the university. In collaboration with the data warehouse project stakeholder group, which represents a cross-section of key data users, the IR staff has established a data source of ten years' worth of student reporting data. This warehouse of information serves as the foundation for the development of a wide variety of reporting applications using SAS Business Intelligence software.

INTRODUCTION

Historically, this university reported enrollment headcount information to its user community via a dynamic web-based application created five years ago that utilized Base SAS and SAS/IntrNet software technology. Various "views" of headcount numbers and percentages were available allowing the user to choose their preferred display format. After upgrading to the SAS Enterprise BI Server software suite, it was decided the ideal environment for a tool with this capability was an OLAP cube. OLAP, On Line Analytical Processing, is a mechanism to provide pre-summarized dimensional data that meets business reporting requirements in a user-friendly, drill-down environment with very fast response times. The information seekers are given the option to select how they wish to see the data, so they are free to create customized reports and charts, rather than relying on the Institutional Research office to provide pre-designed views of the output or require additional programming development every time a new "view" is requested.

HISTORICAL VERSION – ENROLLMENT PROFILE – SAS/INTRNET® VERSION

OVERVIEW

The initial Enrollment Profile website was originally conceived as a reporting solution that would allow the Institutional Research office to immediately respond to daily enrollment questions during key times of the year. The decision was made to utilize SAS/IntrNet technology to create a dynamic environment that provided information on student enrollment headcount by utilizing a series of unique "views." For example, this application shows enrollments by college, undergraduate vs. graduate, full-time vs. part-time status, gender, ethnicity, classification and major in a drill-

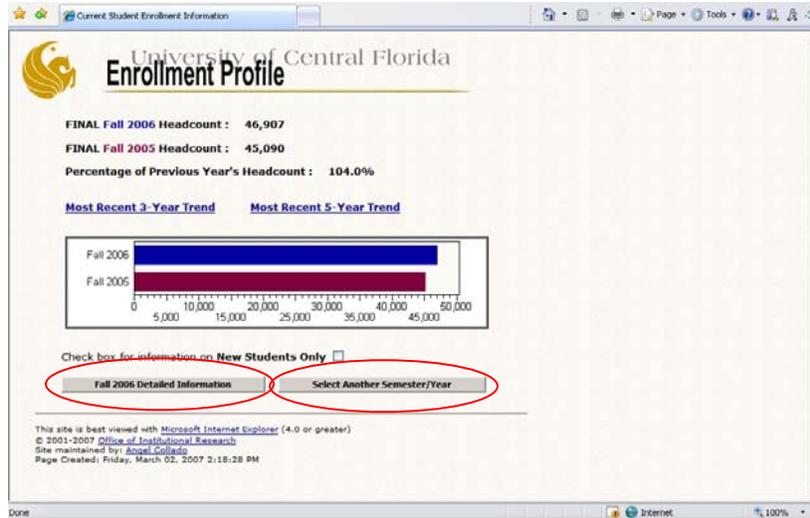
down fashion. This web application was designed to replace and enhance numerous hard copy reports and be accessible via the web in a user-friendly, dynamic and interactive environment.

At full implementation, after considerable feedback by the user community, the Enrollment Profile had more than 30 different tabular displays available by college or entire university. The users were provided with a mechanism to "slice-and-dice" enrollment headcount information in multiple ways. However, this functionality required the creation of numerous programs that demanded continuous maintenance, required complex programming upgrades and, in the end, was impacted by licensing issues.

DETAIL

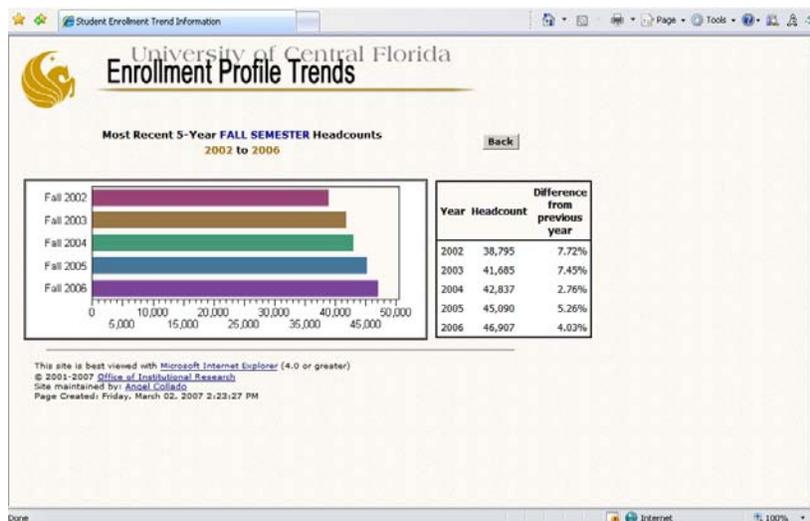
As a user navigates to the home page of the Enrollment Profile (Figure 1), he/she is provided with headcount statistics at the university level for the current term compared with the same term in the previous year. Options are available to select another term and year or continue to detailed, semester-based information.

Figure 1



Each "view", or tabular report, is a separate program that extracts the data needed using the SQL procedure, runs the TABULATE procedure to generate the table, and streams the output to the web browser in real time. Parameters based on user selections are passed to the program as macro variables to filter the data as required. In general, the rows in the resulting tabular display reflect college, level, or major and the columns represent gender, ethnicity, residency status, or academic load. The reports, as originally designed, are all term-based and, therefore, trends were not readily available for a particular college or major.

Figure 2



To incorporate trend functionality, additional programs were created to display the most recent five-year trends in multiple areas. Figure 2 above displays university trends, Figure 3 displays trends for the College of Business Administration, and Figure 4 displays trends for the undergraduate Finance program.

Figure 3

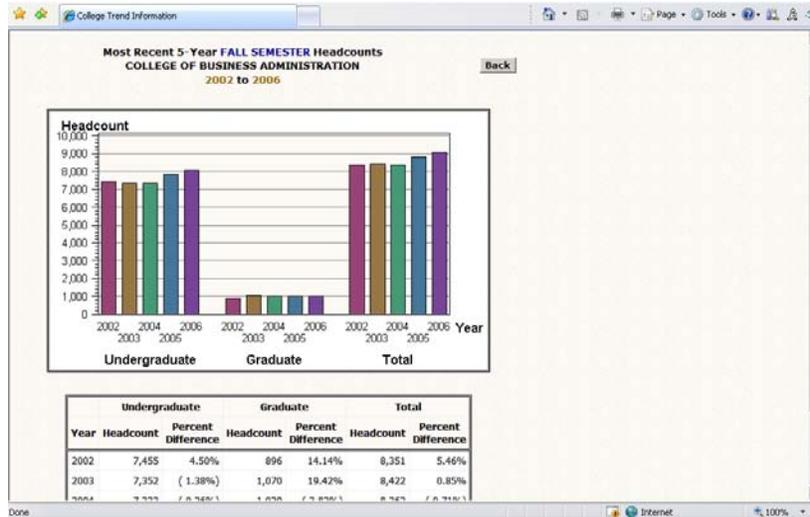


Figure 4

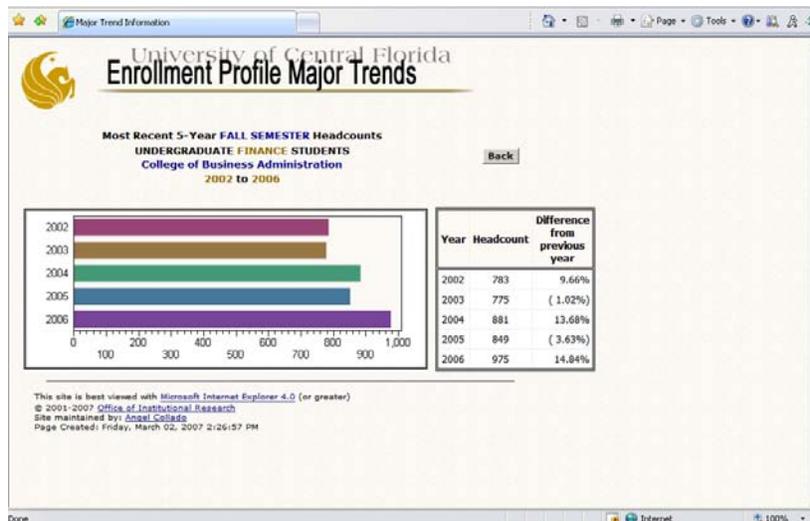
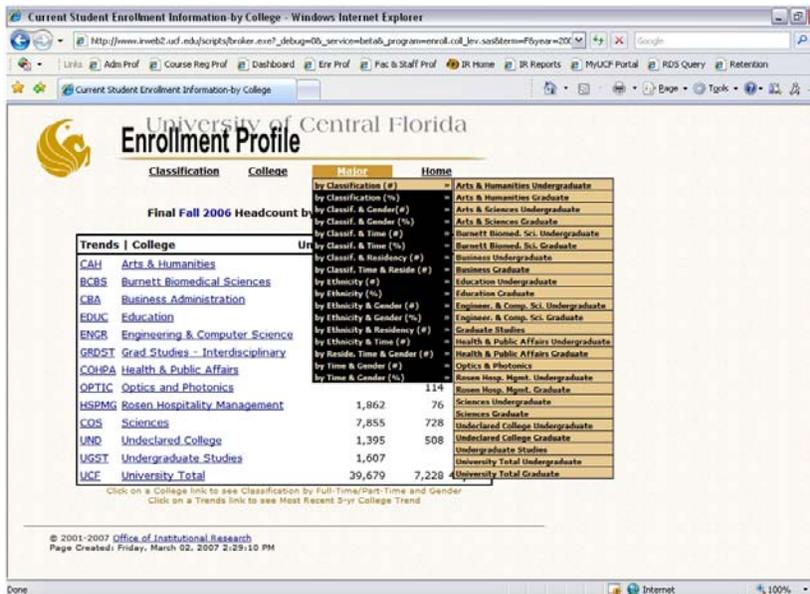


Figure 5



There are two ways to obtain information using this application. Figure 5 details the navigation menu that allows a user the added flexibility to go directly to a particular report. The alternative navigation method is from page to page in a drill-down fashion by clicking on the dynamically generated links in the report. The feedback from the user

community was very positive indicating that the application was easy to use and quickly provided information in a clear, concise format.

MAINTENANCE

Maintaining this application, however, is a long, multi-step process. First, the data needs to be updated by appending each new census file. Five files are submitted to the state each academic year and each file is often submitted more than once; therefore, this data maintenance process is performed subsequent to each file submission. Other reports are generated from this same data file. The previous process required the development of a Microsoft Access database table that was created with all the fields and then a subset was extracted containing the necessary fields for the Enrollment Profile application. The next step saw the data imported into Base SAS and put through various data quality checks, cleansed as needed and then finally stored as the official data set. In addition, the statistics for the college and major five-year trends were stored in separate SAS data sets as they were programmatically calculated each term.

The second part of maintaining this application concerned the overall program maintenance. Approximately ninety percent of the programs use macro variables that were passed to them requiring only limited maintenance as necessary. However, there were several programs that had default values and required a change from term to term. Other programs needed to be updated when new colleges and/or majors were added to the university inventory. Additionally, the navigation menu was programmed in JavaScript and required modification whenever a new view or report was added. On average, each cycle of maintenance and data update required two full days of programmer effort.

A special maintenance cycle occurred when the Enrollment Profile application was migrated from SAS version 8.2 to SAS version 9.1. The HTML standard in the latest version of SAS is HTML4; however, our applications were all coded using the HTML3 standard. We had to modify more than 30 programs to have the web display look as it did in v8.2. There is a possibility that this special step may continue to be necessary following each software upgrade. In addition to the added maintenance, there were financial considerations. To continue using the Enrollment Profile application requires an annual license renewal for SAS/IntrNet that is separate from, and in addition to, our license renewal fees for the Enterprise BI Server.

UPDATED VERSION – ENROLLMENT CUBE – SAS OLAP CUBE STUDIO

OVERVIEW

The upgraded Enrollment Cube design relied upon the SAS Enterprise BI Server suite. The SAS OLAP Cube Studio provided a tool to create and maintain OLAP cubes in a user-friendly, simplified interface. Following discussions with the data warehouse stakeholder group and other business analyst experts, dimensions and levels of detail required to support our application were designed into logical hierarchies and the data was summarized at these levels and stored.

The upgraded Enrollment Cube has multiple access options via SAS® BI tools; including SAS® Enterprise Guide, SAS® Web Report Studio, SAS® Information Delivery Portal, and the SAS® Web OLAP Viewer for Java and .NET. The cube can also be viewed as a pivot-table in Microsoft Excel via the SAS® Add-in to Microsoft Office.

DETAIL

For our initial attempt at creating an OLAP cube to replace the Enrollment Profile, it was decided to build an enrollment fact table that would be an extract from the warehouse dimension tables of the necessary data. SAS Enterprise Guide was the tool of choice to create a reusable project that would go through multiple steps to build the table (Figure 6). This method allows the freedom to schedule the job and have it run every term.

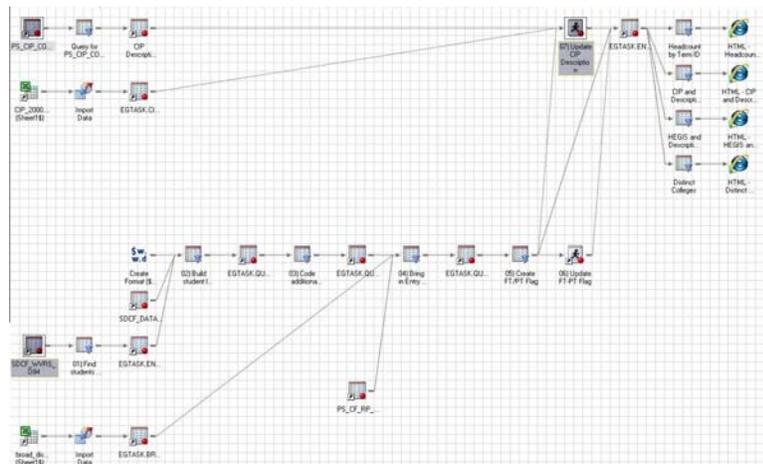
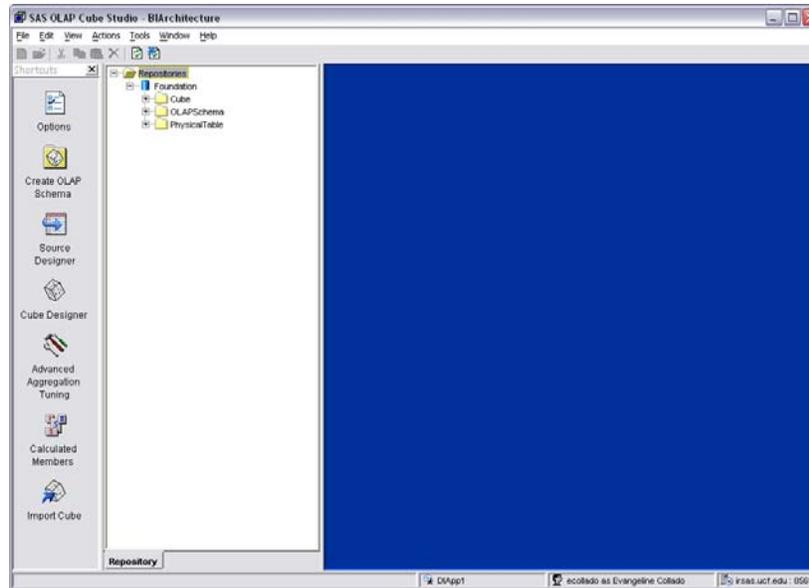


Figure 6

Since we were trying to mimic display options available in the previous Enrollment Profile, dimensions, levels, and hierarchies were defined that would provide the same robust output. SAS OLAP Cube Studio (Figure 7) is a drag-and-drop type tool that guides the user in the creation of an OLAP cube. The first step is to use the Source Designer to define the source data and the last step is to define the cube structure – dimensions, levels, and hierarchies – using the Cube Designer. Additional tools are available for calculating new measures not defined in the data or for tuning the steps needed to create aggregate information.

Figure 7



When using the Visual Data Explorer in the SAS® Information Delivery Portal, as shown below in Figure 8, the Enrollment Cube opens to a default view. Since the data is pre-summarized, the response time is very quick when alternate choices are made. Trend functionality is available immediately, as all years in the data set are shown when a particular semester is expanded or drilled through. Charts are added by the click of a button and automatically update as different user-selected options are altered in the data view. The user can customize the display to suit his/her preferences, create a bookmark to return to the exact view each time, and right-click on the display to create or append to Adobe PDF or export the table to Microsoft Excel.

Figure 8



There are two ways to alter the data view – drill down (⏏) or expand (⏏). Drill down isolates that particular value at a lower level whereas expanding keeps all other values in the display. Figure 9 shows the "College of Business Adm" expanded and Figure 10 shows the same group after clicking on the drill down icon. Note that the "Drill Path:" box now contains the academic group we have selected for a drill down. Navigating the cube is very easy; to change the display the user can add or remove available hierarchies, then click "Apply". Filtering can be done to limit the amount of data shown. Additional options are available by clicking on the row or column variable name. Figure 11 shows some of this detail.

Figure 9

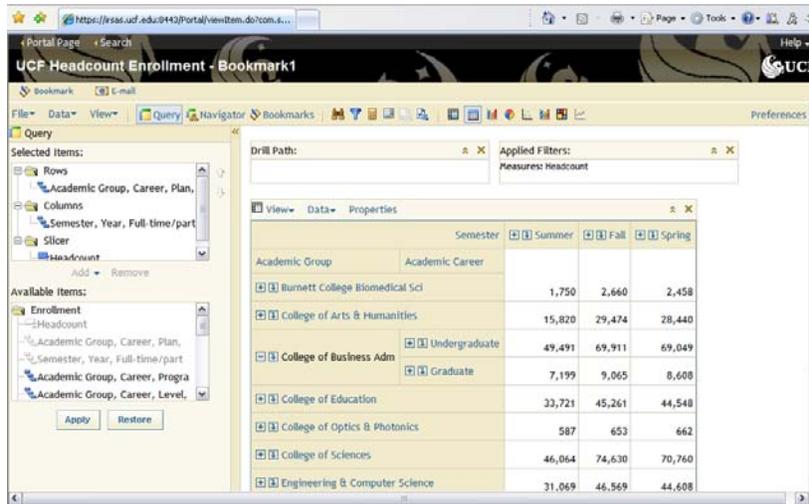


Figure 10

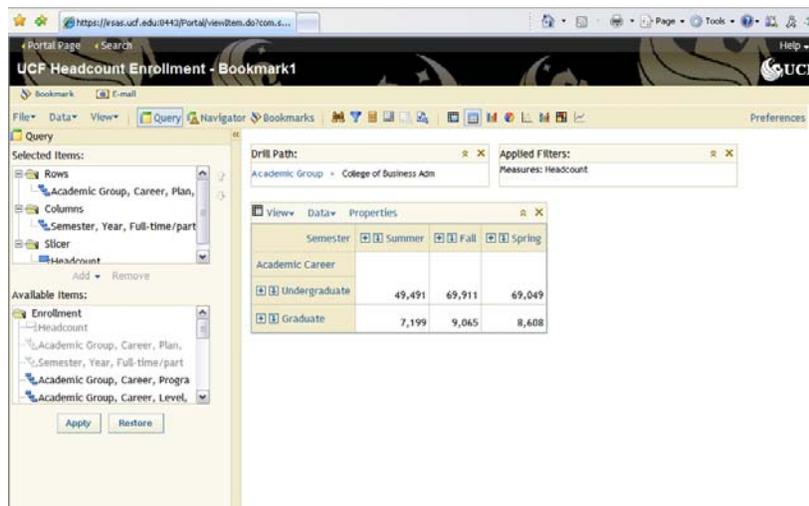
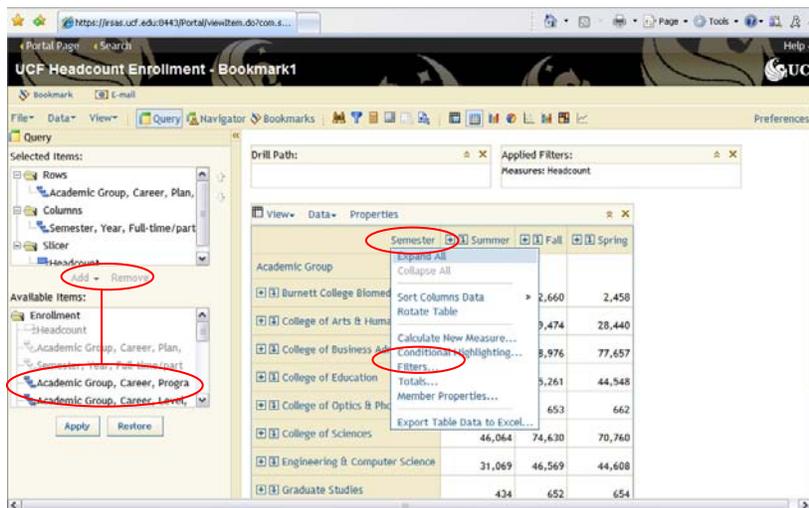


Figure 11



The Enrollment Cube development was guided by input from prominent Enrollment Profile users and the data warehouse stakeholder group. The testing feedback led to the development of additional data dimensions. Although the final application was recently delivered to the end-users, we anticipate continuing enhancements as part of the life-cycle of the product. Also, given the variety of tools available to end users to access the cube data, training will be an issue. The majority of the users will access the data via the Visual Data Explorer in the Information Delivery Portal. In the coming month, we plan to have a complete user guide developed, and although the drill down and

expand navigation is intuitive, training may be necessary for many of our users in order to fully utilize the available functionality.

MAINTENANCE

The maintenance process for the Enrollment Cube is very simple and is completed in two steps:

1. Run the job in Enterprise Guide to update the enrollment fact table
2. Rebuild the cube in OLAP Cube Studio

For more information on OLAP functionality provided by SAS, please visit <http://www.sas.com/technologies/bi/olap/index.html>.

COMPARISON OF APPLICATIONS

Although the original Enrollment Profile is very robust and provides ample information, the Enrollment Cube has more functionality, such as charting, user controlled custom views, trends at every dimension, and the ability to create PDF or export to Excel. The Enrollment Profile was simple to learn initially, whereas the Enrollment Cube will require some brief instruction to be navigated efficiently. The Enrollment Cube can also be viewed as a Pivot table in Excel using the SAS Add-in to Microsoft Office (Figure 12), as well as displayed and manipulated in SAS Enterprise Guide (Figure 13).

Figure 12

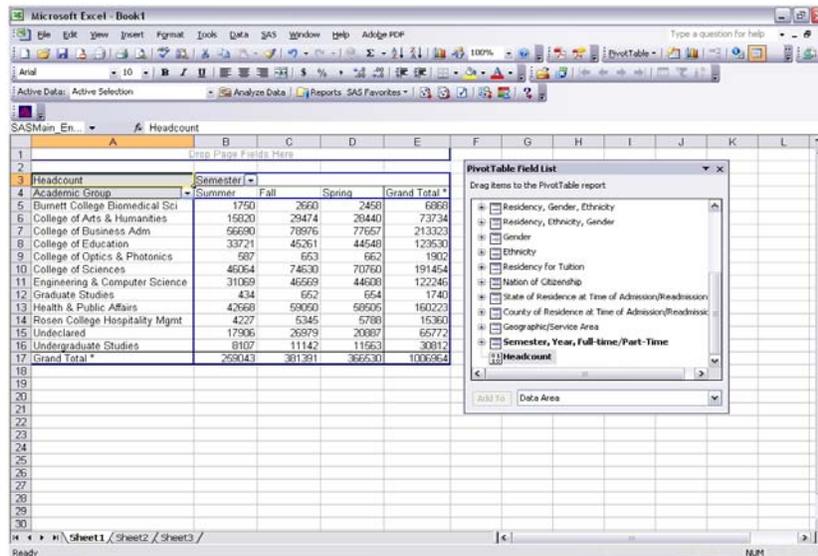
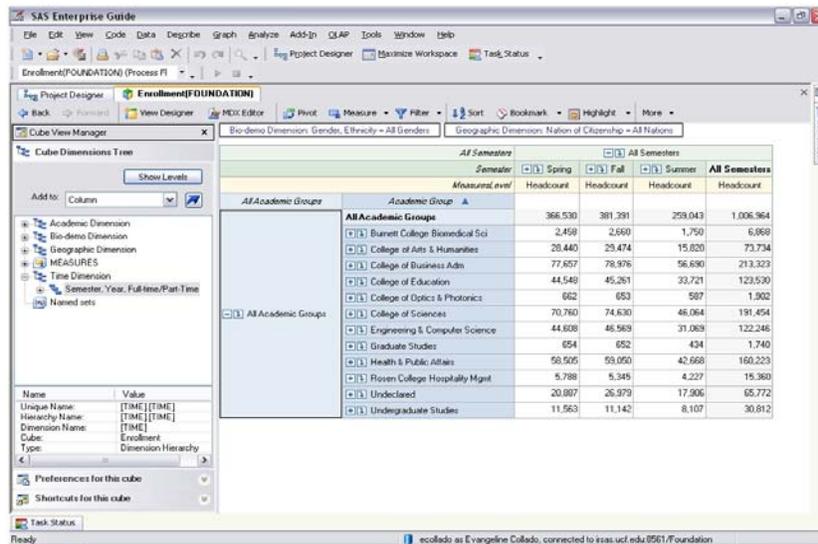


Figure 13



The most significant difference between the two applications is in the initial development and the maintenance. The

development programming and testing for the first set of views of the Enrollment Profile took approximately three months. The development of the Enterprise Guide project to build the enrollment fact table and the building of the cube required only two weeks of programming. The maintenance of the Enrollment Profile takes at least two days; whereas, the cube can be rebuilt in minutes. Of course, the maintenance estimates depend upon the amount of data – our data set has over one million records.

CONCLUSION

The Enrollment Cube has census data going back to the summer term of 1995. We have plans to add current data on a daily basis during critical times to enhance the decision-making processes at UCF. Other future plans include development of a cube to display degrees granted information in a similar fashion to the enrollment display. IR currently maintains several other applications driven by SAS/IntrNet technology that may benefit from OLAP functionality.

Institutional Research seeks to continuously improve the information provided, available functionality and end-user flexibility in all of our reporting. The university certainly benefits from a tool as robust as the SAS Enterprise BI Server software suite.

CONTACT INFORMATION

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

M. Paige Borden
Director and University Data Administrator
Office of Institutional Research
University of Central Florida
P.O. Box 160021
Orlando, FL 32816-0021
Work Phone: (407) 823-5061
Fax: (407) 823-4769
Email: mborden@mail.ucf.edu
Web: <http://www.iroffice.ucf.edu>

Evangeline Collado
Coordinator, Computer Applications – Data Warehouse Technical Lead
Office of Institutional Research
University of Central Florida
P.O. Box 160021
Orlando, FL 32816-0021
Work Phone: (407) 823-4968
Fax: (407) 823-4769
Email: ecollado@mail.ucf.edu
Web: <http://www.iroffice.ucf.edu>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.