Paper 3510-2015

SAS® Visual Analytics: Emerging Trend in Institutional Research

Sivakumar Jaganathan, Thulasi Kumar
University of Connecticut

ABSTRACT

Institutional research and effectiveness offices at most institutions are often the primary beneficiaries of the data warehouse (DW) technologies. However, at many institutions, building the data warehouse for growing accountability, decision support, and the institutional effectiveness needs are still unfulfilled, in part due to the growing data volumes as well as the prohibitively expensive data warehousing costs built by UIT departments. In recent years, many institutional research offices in the country are often asked to take a leadership role in building the DW or partner with the campus IT department to improve the efficiency and effectiveness of the DW development. Within this context, the Office of Institutional Research and Effectiveness at a large public research university in the north east was entrusted with the responsibility to build the new campus data warehouse for growing needs such as resource allocation, competitive positioning, new program development in emerging STEM disciplines, and accountability reporting. These requirements necessitated the deployment of state-of-the-art analytical decision support applications, such as SAS® Visual Analytics (reporting and analysis), SAS® Visual Statistics (predictive), in a disparate data environment, including PeopleSoft (student), Kuali (finance), Genesys (human resources), and homegrown sponsored funding database. This presentation focuses on the efforts of institutional research and effectiveness office in developing the decision support applications using the SAS® Enterprise Data Warehhouse and analytical solutions. With users ranging from nontechnical to advanced analysts, greater efficiency lies in the ability to get faster and more elegant reporting from those huge stores of data and being able to share the resulting discoveries across departments. Most of the reporting applications were developed based on the needs of IPEDS, CUPA, Common Data Set, US News and World Report, graduation and retention, and faculty activity, and deployed through an online web-based portal. The participants will learn how the University quickly analyzes institutional data through an easy-to-use, drag-and-drop, web-based application. This presentation demonstrates how to use SAS® Visual Analytics to quickly design reports that are attractive, interactive, and meaningful and then distribute those reports via the web, or through SAS® ® Mobile BI on an iPad® or tablet.

INTRODUCTION

Making strategic decisions using solid business analytics using various data sources in an organization has proven to be a highly successful strategy for any business. This is especially true when it comes to higher education industry where the operating metrics are constantly changing and evolving. Several higher education institutions have adopted various data warehouse and business intelligence strategies to achieve this strategic decision making process. Most of the universities have tasked their institutional research offices to build this infrastructure with some level of success, but not fully satisfied with the outcome. This paper addresses the gap in business intelligence strategies and how the current in-memory analytical solutions are helping to fill those gaps. The Office of Institutional Research and Effectiveness (OIRE) at the University of Connecticut has implemented SAS® Visual Analytics along with the Data Warehouse solution that has proven to provide highly desirable results. This paper explains why this approach is now adopted by several higher education institutions to obtain the desired results with a reasonable budget.

TRADITIONAL BUSINESS INTELLIGENCE APPROACH

The traditional data warehouse methodologies for the last decade has been to utilize a data integration as the primary tool for getting the business intelligence reports that drive strategic decision.

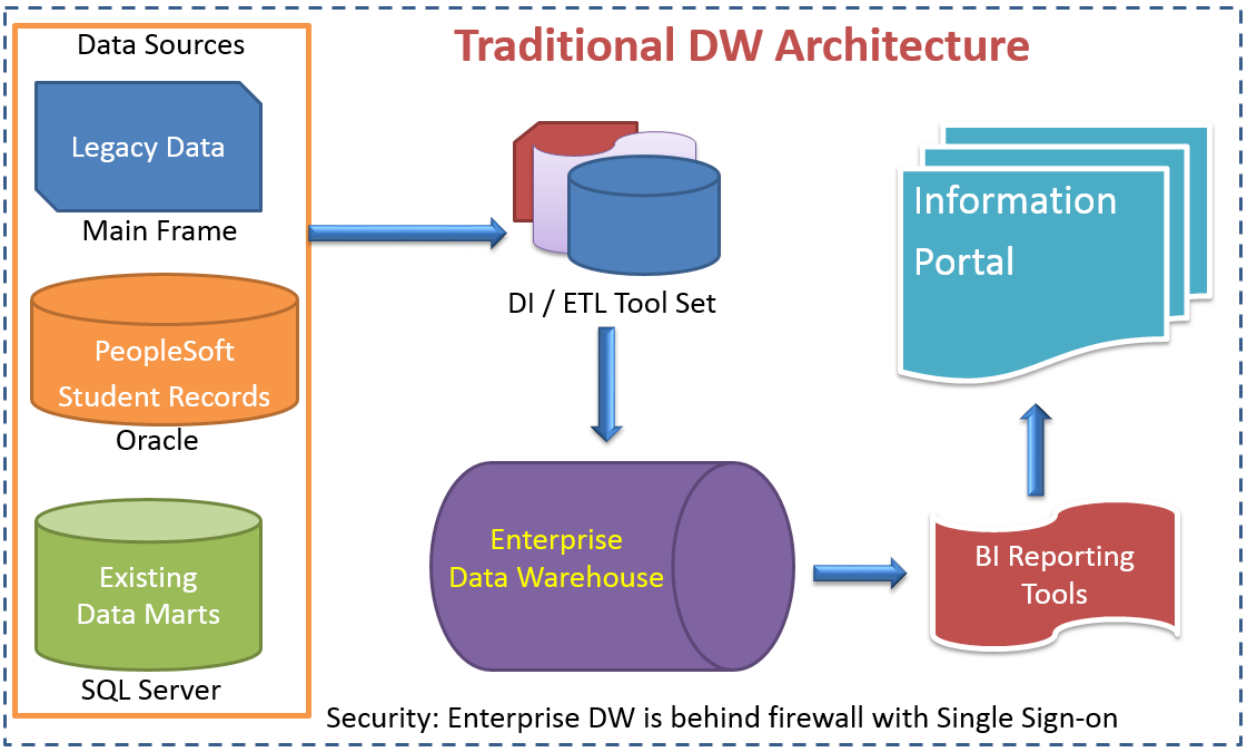

Figure 1. Traditional Higher Education Data Warehouse (DW) Architecture

Figure 1 shows the traditional architecture in a higher education institution. Here the primary data sources are used in creating a data model (typically a star schema) for the enterprise data warehouse. A data integration (DI/ETL) tool is employed in transforming, de-normalizing and loading the data tables to the target Data Warehouse tables. Upon successfully creating data warehouse tables, the specs and information maps are developed for Business intelligence (BI) reports and the development of such reports starts. Also, the individual developers should be trained on a specific development toolsets such as Web Report Studio, Information Maps and Dash Boards. These reports are then published on to the BI information portal. This is a time consuming approach and also the reports are very narrow in scope and can be utilized for a specific types of requests. In an every changing higher education environment this may not sufficient in obtaining strategic business metrics and analytics for a quick and accurate decision making due to the following reasons:

- The approach is tedious and time consuming
- Development toolset deployment and maintenance
- Patching and troubleshooting of the development tools
- Very little reporting flexibility
- No possibility of changing reporting flexibility on-the-fly

At the University of Connecticut, we were looking for a solution that is robust, accurate and broader in scope for getting better business insight. Our new approach explained in the following section proves to be a trend that more universities are adopting in successfully implementing a Data Warehouse (DW) and Business Analytics environment.

BUSINESS INTELLIGENCE APPROACH WITH VISUAL ANALYTICS

In this new approach, the burden of using Data Integration (DI/ETL) as the primary tools for developing Business Intelligence (BI) report is lifted. The new UCONN DW/VA architecture is shown in Figure 2. The data sources are directly visible to the SAS® Visual Analytics toolset as well as the Data Warehouse tables through the traditional DI/ETL methods. This approach provides immense flexibility in generating Business Analytics directly from the source data and the metrics can be modified during the data exploration phase. Also, Visual Analytics provides rich data visualization capabilities and examples of few such reports are provided in the following sections.

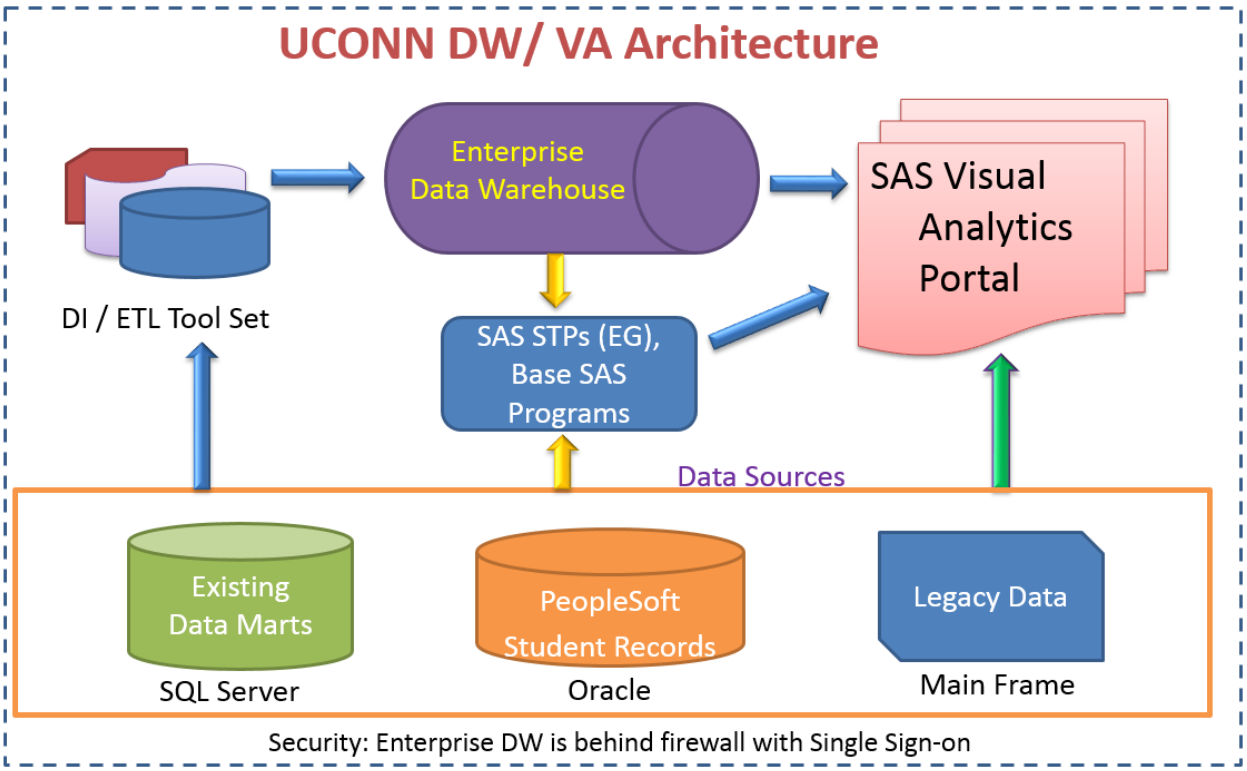

Figure 2. UCONN Data Warehouse and Business Analytics (DW/VA) Architecture

This new approach also has room for traditional Data Warehouse DI/ETL option for long-term standard data warehouse reporting purposes as shown in Figure 2. The Data Warehouse tables that is resulting from the Data Integration jobs are also available for consumption from the SAS® Visual Analytics environment. Apart from that, the SAS® Stored Processes (STPs) developed using SAS® Enterprise Guide (EG) and any Base SAS® programs can be distributed to consumption via the SAS® Visual Analytics portal. The SAS® Visual Analytics administrator provides easy method for setting up security for the reports and data explorations.

UCONN DATA WAREHOUSE AND VISUAL ANALYTICS (DW/VA) TIMELINE

At the OIRE we started the discussions on our Data Warehouse implementation strategies in the summer of 2014. Figure 3 shows the timeline with specific milestones and outcomes defined. The project metrics were developed, Data Warehouse needs and Business Analytics outcomes were defined by the end of summer. The project sponsors (Provost at UCONN) approved the project and funds necessary by early

fall 2014 and DW team started to build. The software vendor and the products were finalized based on our needs and the defined outcomes. The server sizing was carried out in collaboration with SAS® and server specifications were developed.

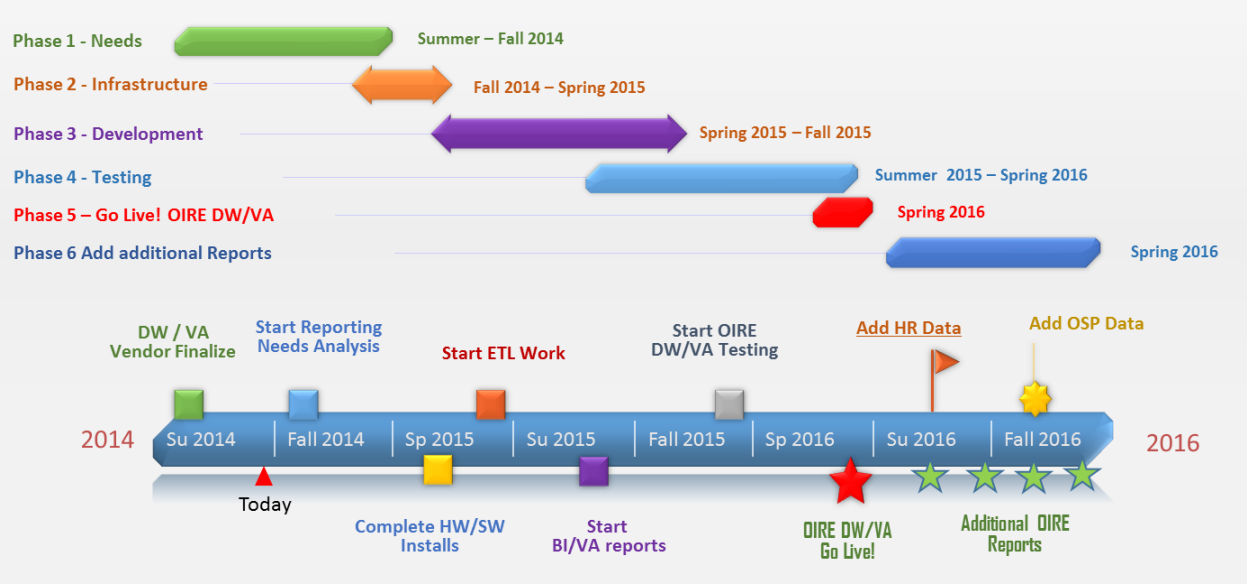

Figure 3. UCONN Data Warehouse and Business Analytics (DW/VA) Timeline

SAS® Visual Analytics relies heavily on the server specification as it is an in-memory solution. Which means the Server RAM memory will be relatively large as shown in Table 1.

SAS® Visual Analytics Server Specifications	
Operating System	Redhat Linux Ver 6.0
Server Hardware Manufacturer	IBM-Lenovo
Server Model	IBM System x3550 M5
Server Processors	Intel Xeon E5-2697v2
Server RAM Memory	512 GB
Total Hard Disk Space	2.4 TB
Server Software installed	SAS® Visual Analytics Ver 7.1

Table 1. SAS® Visual Analytics Server specifications

We also have the traditional SAS® Office Analytics products such as SAS® Enterprise Guide, SAS® Metadata server and Data Management Standard (DI/ETL). The specifications for the

SAS® Office Analytics Server Specification	
Operating System:	Redhat Linux Ver 6.0
Server Hardware Manufacturer	IBM-Lenovo
Server Model	IBM System x3650 M5
Server Processors	Intel Xeon E5-2623
Server RAM Memory	128 GB
Total Hard Disk Space	7 TB
Server Software installed:	SAS® Office Analytics Version 9.4
	SAS® Enterprise Guide Ver 7.1
	SAS® Data Integration Studio ver. 4.9
	SAS® Management Console 9.4

Table 2. SAS® Office Analytics Server specifications

By the end of fall 2014, the servers were ordered and the reporting needs analysis was carried out by the DW team. Currently (Spring 2015), we are working with SAS® to complete the Hardware and Software installs and starting to develop the SAS® Visual Analytics reports based on our needs analysis. The traditional ETL/DI work is also carried out for standard reporting purposes.

SAMPLE SAS® VISUAL ANALYTICS REPORT

In this section we will explore a sample report that was created using the SAS® Visual Analytics using existing data source in SQL Server. This report provides the fall enrollment census data using the Student Administration (SADM) database in SQL Server environment. Previous to the introduction of SAS® VA/OA products, this was done using Hyperion BI suite. The first step in reporting this data using VA was to replicate the Hyperion census query in SAS® EG as a SAS® stored process, then using the output of the SAS® stored process as the source, data was loaded into the LASR server via the VA Data Builder. The three VA Designer visualizations help the audience at UConn to see fall enrollment trends in a visual, easy to grasp, multi-dimensional and user configurable view, unlike the raw numbers previously reported in Microsoft Excel.

The following screen shots show the steps involved in creating this report and the whole report and the data visualization task took minutes to develop and distribute via Internet Explorer or Chrome web browsers. This report can also be consumed in the SAS® Add-in for Microsoft Excel.

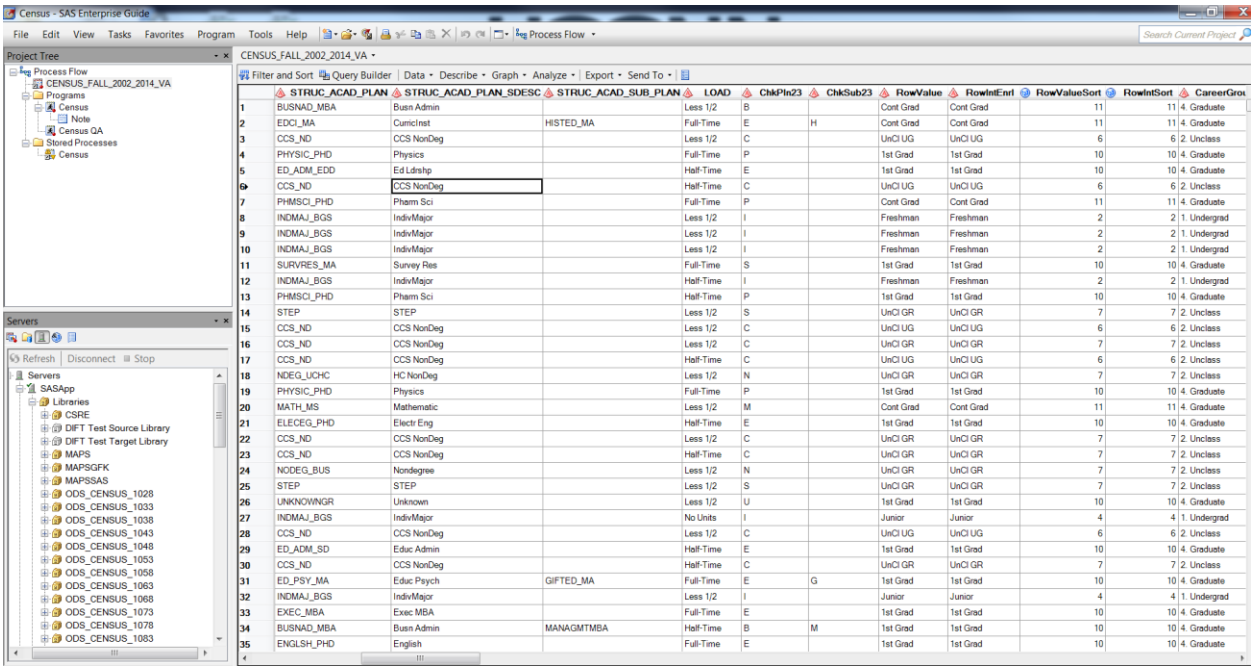

Figure 4. Census Dataset Created in SAS® Enterprise Guide

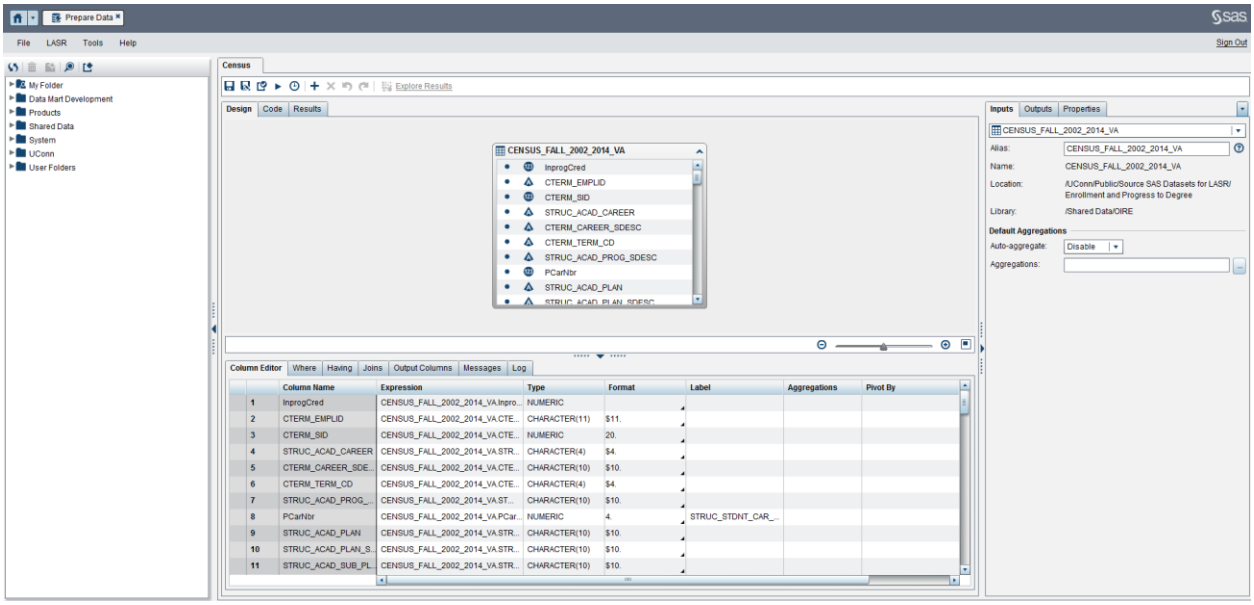

Figure 5. SAS® Enterprise Guide Output Loaded in LASR Server in SAS® Visual Analytics

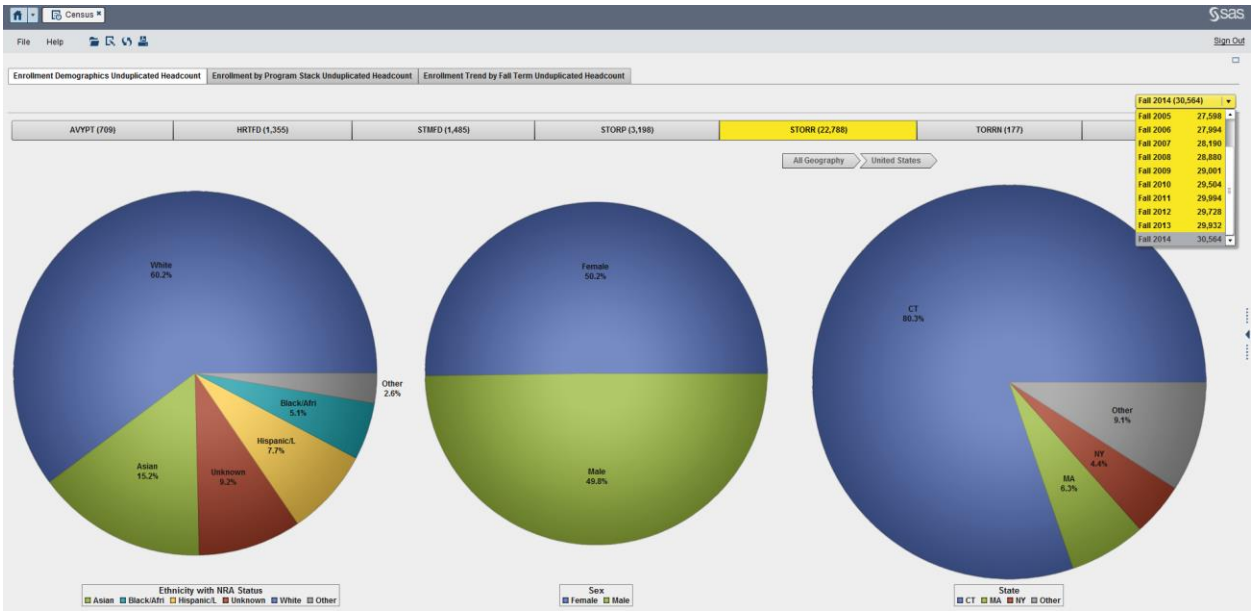

Figure 6. Enrollment Demographics with Unduplicated Headcount of Students

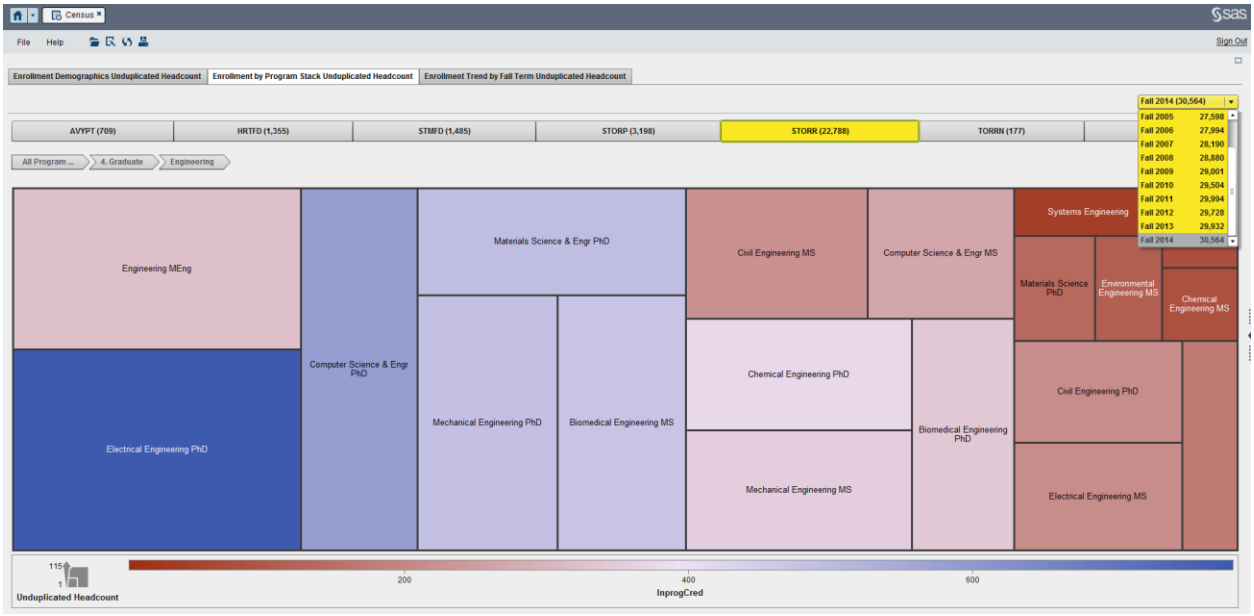

Figure 7. Enrollment by Program Stack with Unduplicated Headcount of Students

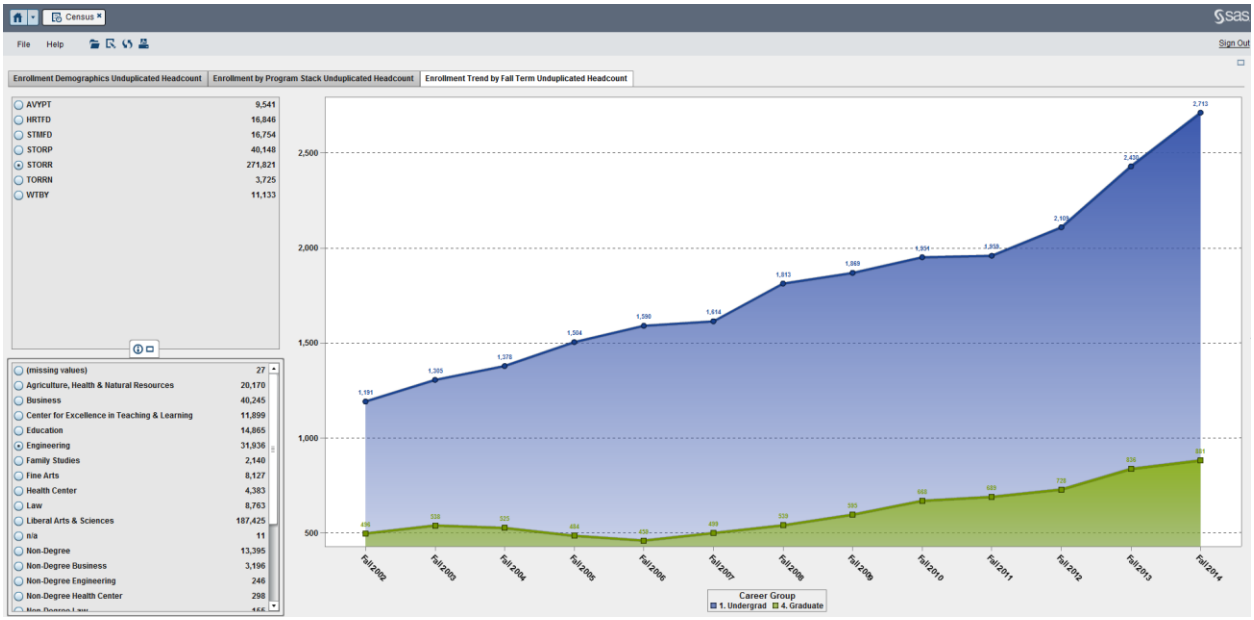

Figure 8. Enrollment Trend by Fall Term with Unduplicated Headcount of Students

This data exploration and visualization shows the power of SAS® Visual Analytics and the efficiency of using the in-memory analytics.

CONCLUSION

Developing a comprehensive Data Warehouse and Business Analytics solutions at any organization is a challenging project and a successful implementation is a key to future growth. When these systems are implemented flawlessly, they provide accurate information quickly and efficiently. The outcome of such critical and key information at the right time could be very beneficial to the organization. Based on the SAS® Visual Analytics implementation at UCONN here are the advantages of using this new trend.

- Ease of executive decisions using report flexibility
- Raw data and qualified Data Warehouse data are readily available for consumption
- Traditional DI/ETL can be accommodated to add value
- No special development tools need to be installed and managed
- Tremendous resource and cost savings along with quick implementation time-frame

We are diligently following our timeline and try to be ahead of schedule meeting the expectations of the project sponsors.

REFERENCES

Ariyachandra, T. and J.H. Watson (2008). "Which Data Warehouse Architecture is the Best?" Communications of the ACM 15(10), pp 146-147.

Breslin, M. (2004). "Data Warehousing Batlle of the Giants: Comparing the Basics of the Kimblaa and Inmon Models" Business Intelligence Journal 9(1), pp 6-20.

Brown, V. and J. Carmichael (2002). "Overcoming Organizational Challenges in Data Warehousing" [WWW document] http://www.tdwi.org/research/display.aspx?ID=6514 (accessed April 10, 2009).

Drucker, P. F. (1999). Management challenges for the 21st century. Butterworth Heinemann, Oxford.

Finnegan, J. D. and P. L. Willcocks (2007). Implementing CRM: From Technology to Knowledge. John Wiley & Sons Ltd, West Sussex.

Hakansson, H. and Snehota, I. (1995). Developing Relationships in Business Networks. Routledge: London.

Holland, P. C. and Light, B. (1999). "A Critical Success Factors Model for ERP Implementation" IEEE Software 16(3), pp 30-36.

Inmon, W.H. (2005). Building the Data Warehouse, 4th ed. Wiley: Indianapolis.

Kimball, R. and Ross, M. (2002). The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. Wiley, New York.

Murtaza, H.A (1999). "A Framework for Developing Enterprise Data Warehouses" Information Systems Management 15(4), pp 21-26.

Parzinger, M.J. and M.N. Frolick (2001). "Creating Competitive Advantage Through Data Warehousing" Information Strategy 17(4), pp 10-15.

ACKNOWLEDGEMENT

Authors acknowledge Anshul Puri, Dan Sokol and Lauren Jorgensen at the University of Connecticut for their support in providing SAS® Visual Analytics Reports.

RECOMMENDED READING

- *Base SAS® ® Procedures Guide*
- *SAS® ® For Dummies®*

CONTACT INFORMATION

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

Sivakumar Jaganathan, Thulasi Kumar
University of Connecticut
860-486-4240
Sivakumar.Jaganathan@uconn.edu