Student Success Dashboard at California State University, Fullerton

Afshin Karimi  
Research Analyst  
Institutional Research & Analytical Studies  
California State University, Fullerton  
akarimi@fullerton.edu

Edward Sullivan  
Assistant Vice President  
Institutional Research & Analytical Studies  
California State University, Fullerton  
esullivan@fullerton.edu

Abstract - Improving student persistence, increasing graduation rates and lowering achievement gaps between underrepresented and non-underrepresented students are parts of a strategic goal at CSUF. The 'Student Success Dashboard', a business intelligence tool developed collaboratively by Institutional Research and Information Technology departments is a dynamic dashboard that helps faculty and administrators with these tasks. The key indicators displayed for given first-time full-time freshman as well as new transfer cohorts are numbers of ‘Graduated’, ‘Currently-enrolled’, and ‘Dropped-out’ students. These are live and up-to-date data and are shown both in a summary format, as well as a detailed student-level format and can further be sliced and viewed in different ways (by Ethnicity, Gender, Parent Education, Underrepresented Status, Prior Institution Type, Major at Entry, College at Entry, Latest Major, Latest College). The paper will focus on the collaborations between IR and IT in establishing the dashboard. Efforts to ensure data quality will be discussed. Immediate benefits of the dashboard (including identification of students for review that had completed their degree requirements but had not been cleared for graduation) will be described as will future enhancements to the dashboard.

Introduction

A performance dashboard is “a multilayer application built on a business intelligence and data integration infrastructure that enables organizations to measure, monitor, and manage business performance more effectively” (Eckerson 2006). At colleges and universities, institutional dashboards provide critical information about the current state of affairs of the institution in an easy to understand format. Terkla, Sharkness, Cohen, Roscoe and Wiseman (2012) draw analogies between monitoring and navigational aspects of institutional dashboards and automobile dashboards. They also refer to the selection of dashboard’s performance indicators as the first step in creating dashboards and also as the most critical step in the dashboard development process. These indicators, they argue, must be: a) easy to understand, b) relevant, c) strategic, d) quantitative, and e) up-to-date with latest information. Yigitbasioglu and Velcu (2012) refer to the problem of ‘information overload’, as increasing amount of data being generated from the organization’s Enterprise Resource Planning (ERP) system, performance scorecards and other reports generated from a variety of sources. They argue that administrators are often overwhelmed and distracted by these reports. Furthermore, authors believe performance dashboards offer a remedy to the information overload problem, since they provide relevant, easy to access information in one manageable place. Harel and Sitko (2003) also describe the difficulty of obtaining good institutional data, as data stored in institutional computer systems are usually hard to navigate and aggregate. Dashboards would allow users to focus on institutional priorities and track performance indicators by filtering out irrelevant information. The authors furthermore identify many of the key infrastructure elements necessary to build successful dashboards:

- A comprehensive data warehouse
- Security infrastructure, including a single sign-on process
- Campus-wide web-based transaction & reporting systems
- Business/Student portals
- Technical expertise for building dashboard components
- Excellent relationship and trust between functional and academic departments

Muntean, Sabau, Bologa, Traian & Alexandra (2010) define main characteristics of performance dashboards to be their ability to use visual components, to gather data from different sources, to have a drill-down capability to see underlying data source, to present dynamic view with timely data refreshes, to display key performance indicators in intuitive format, to help monitor indicators, and finally to be easy to use. Eckerson (2005) believes one of the most distinctive features of a dashboard is that it consists of three views or layers of information: (1) graphical representation to monitor the key performance indicators; (2) high-level dimension-model summary that can be “sliced/diced”; and (3) detailed low-level data that can be used to link a problem with actions that can be taken to resolve it. Arnold (2010) shows that applying business intelligence analytics at Purdue University improves student success, retention and graduation rates. Their dashboard displays actionable information that can be used by faculty and advisors to intervene at-risk students.

The Student Success dashboard at Cal State Fullerton tracks and compares performance of both first-time freshman and new transfer students cohorts broken down by gender, ethnic-race, parents’ education, underrepresented status, college at entry, latest college, and prior institution type. It furthermore allows the users to drill down on a particular subgroup of students in order to obtain detailed actionable student-level information that can be used for, among other things, intervention. Ability to drill down to this detailed student-level information is limited to certain users through the application’s security settings. The underlying data warehouse is refreshed daily and so all performance indicators are up-to-date as of the most recent refresh of the warehouse.

The dashboard was developed collaboratively by the Institutional Research (IR) and the Information Technology (IT) departments using the waterfall software development life cycle model. The dashboard was built using the Oracle Business Intelligence Enterprise Edition suite. The IR office originally defined the product requirements including the aggregate performance indicators, as well as the dashboard’s different tables and charts. The IT office then designed and implemented the underlying data warehouse and the ETL (Extract, Transform and Load) processes. The star schema was used for the data warehouse design. The IR performed the high-level query and dashboard design and implementation, as well as performing data validation and testing. The original product release as well as the follow-up maintenance releases went through the traditional development/staging/production environments. The product was deployed originally to a few select users. Dashboard is accessible via the employee’s portal where user authentication is done through a campus-wide single sign-on process.

**Dashboard Structure**

The Student Success Dashboard includes four summary tabs that are used to view the key indicators for different first-time full-time freshman cohorts and new transfer cohorts. These four tabs display overall cohort summary information, as well the same information broken by ‘College’, ‘Ethnicity’, and ‘Parent Education’. Each summary tab includes a bar graph and a table. Figure 1 below shows the bar graph displayed on the overall cohort summary tab.
In addition to the four cohort comparison summary tabs, there are two detailed tabs shown in figure 2 below. On each of the two detailed tabs, numbers can be displayed by gender, ethnicity, underrepresented status, prior institution type, parent education, latest college and college of entry.
Figure 3 below shows the table displayed on the first-time freshman cohorts detailed tab. This table shows the key indicators for different cohorts by major college. The key indicator values shown in this table can be clicked on in order to drill down to view student-level data.

<table>
<thead>
<tr>
<th>Cohort Description</th>
<th>College (major)</th>
<th>Size</th>
<th>Degree Count</th>
<th>% Graduated</th>
<th>Enrolled Count</th>
<th>% Enrolled</th>
<th>Not Graduated &amp; Not Enrolled Count</th>
<th>% Not Graduated &amp; Not Enrolled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arts</td>
<td>CBF</td>
<td>137</td>
<td>48.6%</td>
<td>23</td>
<td>6.9%</td>
<td>116</td>
<td>42.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COD</td>
<td>324</td>
<td>67.6%</td>
<td>22</td>
<td>4.6%</td>
<td>131</td>
<td>27.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ECS</td>
<td>59</td>
<td>31.1%</td>
<td>20</td>
<td>10.8%</td>
<td>107</td>
<td>57.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HHO</td>
<td>429</td>
<td>68.8%</td>
<td>48</td>
<td>7.7%</td>
<td>147</td>
<td>23.6%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HSS</td>
<td>487</td>
<td>57.1%</td>
<td>44</td>
<td>5.2%</td>
<td>322</td>
<td>37.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Misc</td>
<td>0</td>
<td>0.0%</td>
<td>0</td>
<td>0.4%</td>
<td>222</td>
<td>9.9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NSW</td>
<td>101</td>
<td>43.2%</td>
<td>24</td>
<td>10.3%</td>
<td>109</td>
<td>48.6%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arts</td>
<td>131</td>
<td>41.2%</td>
<td>62</td>
<td>19.5%</td>
<td>105</td>
<td>39.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBE</td>
<td>390</td>
<td>47.3%</td>
<td>131</td>
<td>15.6%</td>
<td>312</td>
<td>37.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COD</td>
<td>131</td>
<td>15.6%</td>
<td>62</td>
<td>27.6%</td>
<td>124</td>
<td>55.1%</td>
<td></td>
</tr>
</tbody>
</table>

Dashboard Development Process

The Student Success Dashboard was developed using the waterfall software development model. Waterfall is a model that involves stepwise progression of tasks that lead to the release of the software product. The development proceeds sequentially through a series of phases starting with system requirement analysis and definition and leading to product release and maintenance. There are feedback loops between each phase so that in case problems are encountered, it is possible to go back to a previous phase and make appropriate modifications. The development progress ‘flows’ from one phase to the next (like waterfall). Completing one phase before moving to the next helps with the long term stability and robustness of the product. Also, in general, the earlier bugs and defects are found, the cheaper they are to fix. McConnel (1996) estimates that a requirements defect that is left undetected until after the product is released will cost 50 to 200 times more to fix as it would have cost to fix early in the process. This includes the cost of investigating, fixing, testing and releasing the solution, as well as the cost associated with reputation and credibility damage. It is therefore important that every phase of the waterfall model (starting with requirements definition and analysis) is completed before moving on to the next phase. The sequential phases used for the dashboard development processes included requirements definition, design, implementation, testing, release, and maintenance.

Requirements Definition

The Student Success Dashboard’s primary purpose is to track performances of the cohorts (or cohort's subgroups) in real time. These include first-time freshman and new transfer cohorts. Three main key performance indicators (‘enrolled’ flag, ‘graduated’ flag and the ‘dropped out’ or ‘not graduated & not enrolled’ flag) are high-level indicators of student success. These three variables are mutually exclusive; each student exactly has one of these three flags set. If a student has graduated and has enrolled as a graduate or post-baccalaureate student in the current term, he/she would only be counted as having graduated. The percentages of the three aggregated indicators in a given cohort add up to 100% of the cohort and so this makes it easy to visually compare cohorts’ performances (see Figure 1 above -
green color on the horizontal bar represents percentage of students who have graduated, blue represents percentage of currently enrolled students and red represents percentage students who are neither currently enrolled nor have graduated).

It was decided to separate the comparison summary tabs from detailed tabs. Users interested in getting a quick idea about a cohort’s progress or in comparing cohorts’ performances could view information quickly on the summary tabs. On the other hand, users who would like to view detailed information about a particular subgroup of students would use the detailed tabs to drill down on that group. The drilled-down student-level tables would include students’ IDs, their names, their contact information, their demographic information, their initial and latest majors, their current GPAs, their earned units, their degree checkout terms and their academic standings. In addition, if a student is currently enrolled, his/her ‘current units attempted’, as well as ‘units taken for future semester’ (if any) would be displayed. For a student who has graduated, the additional ‘degree date’ column would be included in the table.

The requirements also called for two levels of user authorization. Users belonging to level one would only have access to the summary tabs (first four tabs). Over time, all active campus employees would belong to this group. Level two users (select administrators, deans, associate deans, program chairs, advisors) would have access to all tabs including the detailed data tabs. It was required for users to have been trained on the ‘Family Educational Rights and Privacy Act’ (FERPA) topic before being able to access the dashboard.

Oracle Business Intelligence Enterprise Edition (OBIEE) suite was chosen to be the platform for dashboard development. The requirements definition aspect of the project was handled by the IR office.

Design

Institutional Research office at Cal State Fullerton processes, cleanses and finalizes census student record files before transmitting them to the university chancellor’s office a few weeks after the start of each semester. These ‘Enrollment Reporting System’ files are also used to create data that is consumed by the dashboard’s ETL process and is loaded into the dashboard’s underlying data warehouse. This data includes static information (such as student’s ID, ethnicity, gender, parent’s education) of the students belonging to the new first-time full-time freshman cohort, as well as the new transfer students cohort. The complexity of transformations required to establish key cohort based categorical variables coupled with the potential for credibility risk associated with changes to these variables resulting from changes to the dynamic CMS (Common Management System) data made the use of static census level data preferable.

Once a semester, new students cohort data are fed to the data warehouse. A second ETL process is also run on a daily basis to extract student dynamic data from the Common Management System (CMS) and load into the data warehouse. This includes- among other variables- the student’s enrollment flag, degree flag, drop-out flag, GPA, units earned, units being attempted, units taken for next term, academic status, contact information, original and latest major, original and latest college, degree checkout term, and degree term. This dynamic data is extracted for students whose IDs were provided in the cohort’s static data file. In other words, the student IDs in the static file are used as keys to perform queries against different tables. Figure 4 shows the overall flow of data from CMS to the dashboard.
The underlying student data warehouse used the star schema architecture. Star schema is one of the most popular methods of implementing a multi-dimensional model in a relational database. In this model, data is organized into facts and dimensions (Sen, Sinha 2005). A fact table contains measures that are counted or aggregated. A dimension contains reference information about the fact. One or more fact tables are surrounded by a number of dimension tables. A fact table is related to a dimension table by using a foreign key relationship. The benefits of such a model include simpler and faster queries, as well as fast aggregations (such as count, sum, and average). Furthermore, these aggregations can be grouped and filtered easily (sliced & diced) by the dimensions. For this project, initially the three main key performance indicators (enroll flag, degree flag and drop-out flag) were placed in the fact table. Later, student GPA, units earned, units attempted (current) and units attempted (future) were added to the fact table. This fact table was linked to a number of dimension tables. Dimensions are used to look at facts by using the “by” conditions (for example, viewing currently enrolled students “by” ethnicity). Figure 5 shows the star schema used for the dashboard’s underlying data warehouse. The design phase of the project was handled by the IT department.
Implementation

The implementation phase included establishing the ‘Extract, Transform and Load’ (ETL) processes, creating the data warehouse, developing queries (built on top of data warehouse) and creating dashboard tabs (that included graphs and tables). The ETL process refers to the process of extracting data from data sources, transforming it to fit the operational needs (including joining data from multiple tables, aggregating data, translating coded values), and finally loading the data into the data warehouse. As shown in figure 4 above, there are two ETL processes implemented. The first one extracts and loads static and cleansed term cohort data and is run once a semester after that semester’s enrollment files are finalized. The second ETL process runs daily and pulls in dynamic data on a daily basis. This dynamic aspect of the dashboard is one of its important aspects. For example, student enrollments can be tracked on a daily basis before (or even after) start of a semester. In addition to implementing the ETL processes, the data warehouse fact and dimension tables that were identified during the design stage were implemented. Once the data warehouse was created and populated, the next step of the implementation process was to build queries that use the data in the warehouse to build tables and charts on different tabs. Also, the drill-down functionality of the detailed tabs was implemented. The ETL processes and the data warehouse were established by the IT department, while IR developed the queries and the dashboard tabs’ graphs and tables.
Testing & Data Validation

The major portion of the dashboard testing activity revolved around validating the ETL processes. The inputs to the ETL processes are data from CMS and the static cohort file and the ETLs’ output are loaded into the data warehouse. Having access to both ETLs’ input data and output data, their functionality was validated. The output data were reviewed by building OBIEE queries on top of the populated data warehouse. The input data were also checked by querying CMS directly. Stratified samples of students from both first-time freshman as well as new transfer cohorts were used for ETL validation. Students from different categories (graduated, enrolled, dropped-out) belonging to different colleges with different demographics were used for spot-checking. In addition to this approach, the aggregated numbers were also verified against the static cohort enrollment files. Another level of testing involved validating the ‘slicing & dicing’ aspects of the dashboard. This indirectly validated the star schema design of the data warehouse. Finally, the high level queries that were built on top of the data warehouse were tested. Few examples of common data quality issues that were encountered and resolved in this step are mentioned below:

- Adding a dimension or variable to a query resulted in reduced number of records in the result set. For example, a detailed dashboard table first showed 100 female students belonging to the Fall 06 cohort who are still enrolled. Adding the new ‘academic standing’ column to the table resulted in only 92 students listed, all having valid academic standings. The reason here is using an incorrect database ‘join’ operation (using inner instead of outer join or even using left outer join instead of right outer join or vice-versa). As a result, students with missing academic standings were dropped from the result set instead of simply showing blank academic standings.

- Some of the student records were not found in CMS for students listed in the static cohort file, or there were mismatches between the static cohort file and CMS. The reason was that some records were updated retroactively in CMS after the census enrollment files were finalized, causing mismatches between the two.

- While there were usually more than one method in which an action could be recorded in CMS (such as student major change process), the ETL logic would not always account for all those possible methods resulting occasionally in dashboard displaying incorrect data.

The testing and data validation phase of the process was performed by the IR team.

Release & Maintenance

All dashboard development and testing activities were carried on a development server using a backup version of production CMS. After different aspects of the dashboard were implemented and validated, it was moved to the staging environment (an environment that mirrored the production environment) for a final look before being moved to production. This process was followed when the dashboard was initially launched, as well as when it was re-released for follow up maintenance releases.

The dashboard was first launched to a few select users belonging to the two different user authorization groups (one having access to only the summary tabs and one having access to everything). The next step was to release the dashboard to one select college. The college’s dean and department chairs were first trained on dashboard usage and then were given access. The dashboard’s hyperlink was placed in the employee’s web portal (accessed through a campus-wide single sign-on process). The decision to provide a stepped release allowed IR and IT staff to learn from initial adopters how to better utilize from the dashboard tools. It also allowed a faculty voice to emerge that recognized the current and future value of the dashboard.
Usage Scenarios

Very early in the release it became apparent that the dashboard would allow greater scrutiny of existing campus efforts and procedures related to retention and graduation. In its initial release, institutional research identified multiple student records that had been overlooked during the degree clearing process. The dashboard allowed identification of students who had believed they would be graduating in May 2013 that had completed all degree requirements by the close of the fall 2012 semester to have degrees awarded in January 2013.

The dashboard allowed close scrutiny of freshman and transfer cohort members that had not enrolled in spring 2013 semester but showed in the dashboard with an anticipated degree date of May 2013. The detailed dashboard screens allowed the user to note the number of units completed, current GPA, current major, campus ID number, and contact information. Degree audit reports were run for students with units and GPAs beyond minimum degree thresholds. A list of students meeting all requirements was forwarded to the degree processing office for further review and posting of degrees.

The controlled release of dashboard to select department chairs resulted in increased scrutiny of the progress of majors in their discipline. A friendly competition also emerged among the department chairs as they worked to better understand and improve the success rates of students in their programs. Within a day the users had identified a data field error resulting from an incorrect pointer and requested multiple augmentations (active email links, last enrolled semester, and ability to direct run degree audits from dashboard). Within a week, the department chair identified and cleared the pathway for eight of the department students to successfully graduate in spring or summer 2013. Without the dashboard and the efforts of this department, it is likely that these students would not have met graduation requirements by the end of summer 2013.

The use of the dashboard changed a passive department graduation review process to an active process. The department began actively contacting its students, developed a faculty working group to focus on student success, and brought light to administrative hurdles facing students hoping to graduate. A review of the use suggests that with broader release and similar use, the dashboard could generate improved overall university graduation rates from greater data review and ownership at the department level.

An additional benefit of the dashboard was to provide users with a mechanism to identify students that fail to re-enroll during and after the priority registration period. Department chairs are able to identify and contact their students to see if there are things that could be done to facilitate re-enrollments. The dynamic nature of the dashboard provides a daily update of progress towards retaining majors. It also provides a mechanism through which to estimate potential remaining course demand by student level.

Prior to the development of the success dashboard, the department chairs relied on static lists drawn at the close of the registration period. The dashboard allows for a more timely intervention that can be enacted as priority windows for seniors and juniors close increasing the timeliness of the contacts by a couple weeks. This provides students who miss their initial registration priority window an opportunity to enroll in classes that may have been closed by the time the static post priority registration non-reenrolled lists were generated.

Future Work & Conclusion

The dashboard continues to evolve. It has been rolled out in stages to deans and department chairs as a tool in their efforts to ensure student success. With each additional group of users new elements and modification needs were noted. Updates to the existing process now are made in two ways, (1) immediate needs in the form of adding fields that currently exist in the data warehouse but not the dashboard and (2) developmental needs requiring new programming data extraction.
It is likely in a future period that the dashboard will be expanded to focus on graduate and post-baccalaureate success. A solution that allows a batch generation of degree audits from the dashboard remains in pre-development stages. It is expected that the dashboard will in the future use data from the student affairs co-curricular transcript to further augment the data available to users.

Initial successes in this project are due in large part to the collaborative nature of the project participants. An initial shared vision and recognition of the advantages gained by utilizing specialized talents allowed this project to progress from discussion to viable tool in a very short time span. Simple user friendly design and data quality were emphasized throughout the project development.

Shared vision was perhaps the greatest asset in the development phase of this project. It allowed an ongoing focus on the goal of delivering a core tool without the distractions of late arriving additions. The team was small (five persons). Decisions were made purposefully, and multiple converging workstreams were maintained so that progress was not slowed.

A beta version of the dashboard was shared with the leaders of the graduation initiative committee within a month of the onset of development to ensure project vision aligned with university needs. The review of the beta version provided an achieved checkpoint that allowed the emerging product to be shared with college deans for comment. Dean commentary resulted in recognition that there was a desire for a future iteration of the dashboard to address the need for graduate and post-baccalaureate student data.

The staged rollout to campus users allowed a controlled mechanism to learn more about how the dashboard worked and where there were opportunities for improvement. It also revealed the various levels of understanding of the existing processes related to student success and the general use of technology among department chairs and deans. The adoption, use, and recommendations related to the dashboard by the department chairs have exceeded expectations.

In the end, the student success dashboard has been well received and provided immediate benefits. The awareness of students that have stopped out or left is a valuable resource to the departmental advisor. Although unlikely to occur in all departments, the augmented attention to student success at the academic department level that evolved from use of the dashboard suggests that the right tool in the right hands at the right time can improve the likelihood that students get timely advice that will result in more students graduating on-time.
References


