The Perfect Marriage: Using SAS® Enterprise Guide, the SAS® Add-In for Microsoft Office, and Excel to support enrollment forecasting at a large university

Andre Watts and Lisa Sklar, University of Central Florida

Abstract

This paper will explore the steps taken by a large public research institution to develop a five year enrollment forecasting model to support the critical enrollment management process at the institution. A key component of the process is providing university stakeholders with a self-service, secure, and flexible tool that enables them to quickly generate different enrollment projections using the most up-to-date information as possible in Microsoft Excel. This paper will show how we integrated both SAS® Enterprise Guide and the SAS® Add-In for Microsoft Office to support this critical process which had very specific stakeholder requirements and expectations.

Introduction and Rationale

The Office of Institutional Research at the University of Central Florida is tasked with supporting the enrollment management process of the institution by providing five year enrollment forecasting of various enrollment measures. University stakeholders involved in enrollment planning were in need of a streamlined enrollment forecasting model that was more understandable and easier to interface with then what was currently available. The institutional research office analyzed the then current enrollment model along with supporting business processes, and developed a new model that was more accurate and flexible than previous.

One of the primary tasks of the projects was to integrate the model into an interface where stakeholders could manipulated certain factors, during their discussions, in order to see how certain external trends or potential university initiatives would affect future enrollments and credit hour output. The data preparation, integration and used within Microsoft excel will be the main focus, with some discussion on the model, its history, and genesis.

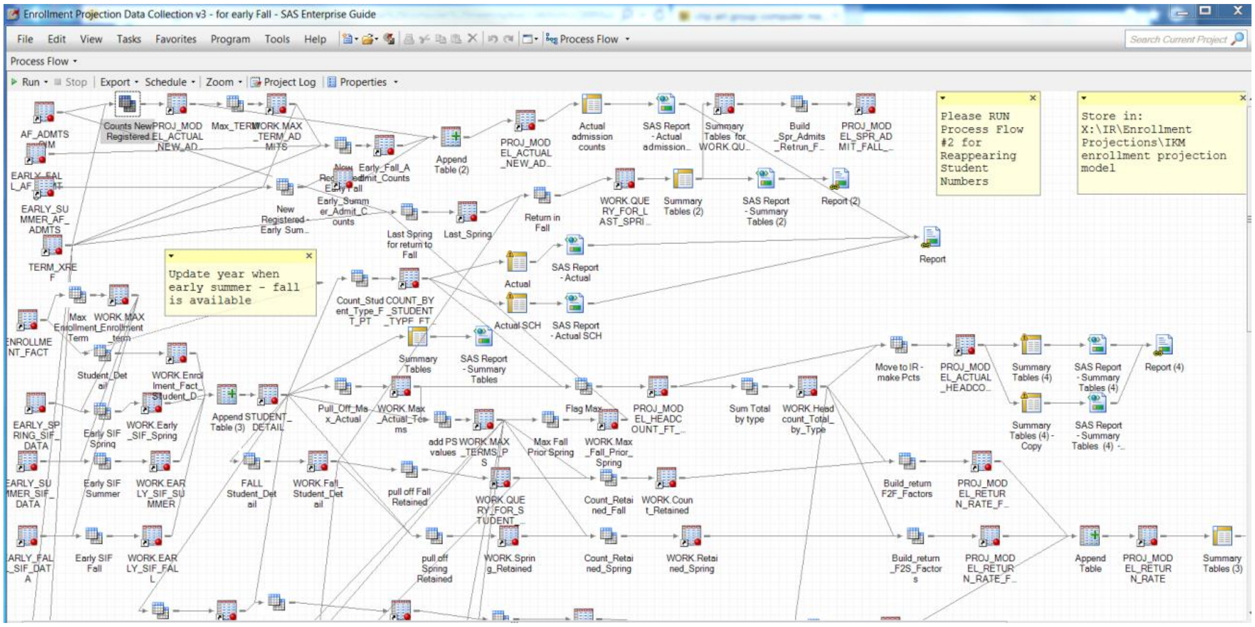
Why Excel?

Microsoft has long been used for analytical analysis and in more recent years has been used more and more in interactive and dashboard type applications. While the web provides a more natural environment for interactivity, the development time required for this medium would exceed the time needed for this project. The main reason for the use of Excel in this project was it gave the stakeholders a familiar interface to work with as they manipulated the factors and other parts of the projection model. In addition, The SAS® Add-In for excel provided an out of the box solution to easily import data from an existing SAS® library without the need for an additional data conversion process such as exporting to a csv, xls file, or even another database system for use with a true web solution.

Data Preparation

Preliminary data modeling was done on some known potential variables that would possibly be included in the final model routine. After this analysis, the variables chosen were retrieved from the internal data warehouse that stores information from the university's student information system (SIS) which is updated every semester. Since the data warehouse maintains a quality standard for all of its contained information, there were no concerns regarding the accuracy of the data. However, having accurate data does not ensure that all the variables needed for the model were retrieved, which is an issue that will be addressed during the discussion of the modeling analysis.

Once all the variables were identified, SAS® Enterprise Guide, our tool of choice, was used to retrieve, calculate, and store the variables in an SAS® library so that it could be updated and retrieved via the SAS® plug-in for Excel as needed. The following graphic shows a screen shot of the SAS® Enterprise Guide project.



SAS® Enterprise guide provided the easiest method of data retrieval and manipulation of the data. The necessary data transformations and retrieval could be accomplished by other methods, however these methods may be more complex and lengthen this part of the process to a large degree.

Modeling

The enrollment model is comprised of two parts:

1. Headcount projections for each semester
2. Student Credit Hour projections for each semester

The focus for this paper will be on the headcount portion and specifically the fall headcount projection. The variables used for the fall headcount projections are as follows:

1. Previous fall semester headcount
2. Current fall-to-fall return rates
3. Current spring admits to fall return rates
4. Estimated spring, summer, and fall new admits

After analyzing the accuracy of the initial model results using the above variables against previous years data, another subset of students that didn't present themselves in the preliminary analysis surfaced – reappearing students. These are students that do not fit into neither the returning nor new admission students and take time off such as transients or readmitting students. Upon learning this, the previous data preparation process was adjusted to retrieve those students and subsequently add them to the set of inputs. Once these reappearing students were retrieved and input into the model, the accuracy of the projections improved and the final model algorithm was put into Microsoft Excel.

Integrating Data into MS Excel

The SAS® add-in for Excel is used to bring in the needed SAS® datasets into different tabs of the worksheet. Flags in the dataset for items such as current year, model year, and most recent term eliminate the need to manually update filter criteria when bringing the data into Excel. This constant filter criteria ensures that the numbers of rows and column in tables do not change, which allows adjacent cells to be used.

After the datasets are updated, all of the SAS® data in Excel can be refreshed using a macro, which then updates the model with the most current available data. When the datasets are updated for a new model year, the refresh in Excel automatically rolls the model forward.

Several background worksheets are used to perform the calculations required by the model to create the projections, primarily using lookup, math, and logic functions. Because refreshing the SAS® data does not change the layout in Excel, formulas to concatenate fields for lookup helper columns do not need to be adjusted or monitored. Initially, array formulas were considered instead of lookup functions, but were ultimately determined to be impractical given the propensity of array formulas to dramatically decrease performance speed.

Two tabs on the Excel workbook are visible to the stakeholders. One shows the summary of the projections and allows users to manipulate factors. Users input the estimated number of incoming students for each term and student type for the years being projected, or the growth factor expected. Growth factors entered are applied to the previous year's projection.

THIS TABLE REQUIRES MANUAL INPUTS															
	Actual Entering Heads		Estimated Future Entering Students												
	2014-15	2015-16	2015-2016		2016-2017		2017-2018		2018-2019		2019-2020		2020-2021	
	Max Full Yr	Current Yr	Heads	Growth Factor	Heads	Growth Factor	Heads	Growth Factor	Heads	Growth Factor	Heads	Growth Factor	Heads	Growth Factor
1- FTIC															
Summer	2,767	2,902	2,768		2,810	1.015	2,852	1.015	2,909	1.02	2,967	1.02	3,011	1.015
Fall	3,745	2,713	3,745		3,782	1.01	3,820	1.01	3,878	1.015	3,955	1.02	4,014	1.015
Spring	412	-	413		419	1.015	425	1.015	434	1.02	443	1.02	449	1.015
2- FCS Transfer															
Summer	1,093	1,108	1,093		1,104	1.01	1,126	1.02	1,143	1.015	1,160	1.015	1,172	1.01
Fall	5,789	2,633	5,789		5,847	1.01	5,935	1.015	6,024	1.015	6,114	1.015	6,175	1.01
Spring	3,193	-	3,193		3,225	1.01	3,289	1.02	3,339	1.015	3,389	1.015	3,423	1.01
3- Other Transfer															
Summer	84	97	73		74	1.01	75	1.02	77	1.02	78	1.015	79	1.01
Fall	510	341	475		480	1.01	487	1.015	497	1.02	504	1.015	509	1.01
Spring	206	-	160		162	1.01	165	1.02	168	1.02	171	1.015	172	1.01
4- 2nd Deg Undergrad															
Summer	163	161	125		125	1	126	1.01	128	1.015	130	1.015	131	1.01
Fall	325	188	250		250	1	253	1.01	256	1.015	260	1.015	263	1.01
Spring	218	-	150		150	1	152	1.01	154	1.015	156	1.015	158	1.01
5- Unclassified															
Summer	9	4	9		9	1	9	1	9	1	9	1	9	1
Fall	21	24	21		21	1	21	1	21	1	21	1	21	1
Spring	1	-	1		1	1	1	1	1	1	1	1	1	1
6- Master's															
Summer	325	311	312		309	0.99	309	1	312	1.01	318	1.02	328	1.03
Fall	1,642	968	1,576		1,561	0.99	1,561	1	1,576	1.01	1,608	1.02	1,656	1.03
Spring	598	-	577		571	0.99	571	1	577	1.01	588	1.02	606	1.03
7- Doctoral															
Summer	43	49	44		46	1.03	47	1.04	49	1.03	51	1.04	52	1.03
Fall	390	135	402		414	1.03	430	1.04	443	1.03	461	1.04	475	1.03
Spring	58	-	60		62	1.03	64	1.04	66	1.03	69	1.04	71	1.03
8- Other Graduate															
Summer	253	271	253		257	1.015	261	1.015	265	1.015	269	1.015	273	1.015
Fall	587	213	547		555	1.015	564	1.015	572	1.015	581	1.015	589	1.015
Spring	462	-	400		406	1.015	412	1.015	418	1.015	425	1.015	431	1.015

This ability to plan to this degree of specificity was one of the key requests made by the stakeholders. When the model is being rolled over to the next model year, a macro is used to shift all of the data so that the stakeholders do not lose the estimates that they had previously made for each given year.

When the model was first demonstrated for the stakeholders, they requested the ability to adjust return rates for each term, year, and student type. A table displays the most recent return rates and allows for manual inputs. If no inputs are entered, the return rate defaults to the most recent actual rate. These inputs are also rolled over for the newest model year.

THIS TABLE REQUIRES MANUAL INPUTS							
	Most Recent Actual	Estimated Return Rate Enter a value to override the most recent actual return rate					
		2015-2016	2016-2017	2017-2018	2018-2019	2019-2020	2020-2021
1- FTIC							
Summer	0.5282						
Fall	0.6859	0.7300	0.7400	0.7400	0.7400	0.7400	0.7400
Spring	0.9132	0.9100	0.9200	0.9200	0.9200	0.9200	0.9200
2- FCS Transfer							
Summer	0.5445						
Fall	0.5091	0.5700	0.5900	0.5900	0.5900	0.6000	0.6000
Spring	0.8211	0.8195	0.8250	0.8300	0.8300	0.8300	0.8300
3- Other Transfer							
Summer	0.5807						
Fall	0.5772		0.6000	0.6050	0.6100	0.6150	0.6200
Spring	0.8474		0.8500	0.8550	0.8550	0.8600	0.8600
4- 2nd Deg Undergrad							
Summer	0.5725						
Fall	0.3913						
Spring	0.7293						
5- Unclassified							
Summer	0.22						
Fall	0.1927						
Spring	0.4982						
6- Master's							
Summer	0.5724		0.5700	0.5750	0.5800	0.5850	0.5850
Fall	0.4457		0.5080	0.5100	0.5100	0.5150	0.5200
Spring	0.8175		0.8200	0.8250	0.8250	0.8300	0.8300
7- Doctoral							
Summer	0.7141		0.7400	0.7000	0.7400	0.7400	0.7400
Fall	0.2897		0.7650	0.7700	0.7700	0.7700	0.7700
Spring	0.9188		0.9200	0.9200	0.9200	0.9200	0.9200
8- Other Graduate							
Summer	0.2597						
Fall	0.2143						
Spring	0.5014						

When these manual inputs are adjusted, the summary tables update dynamically to project the long-term effect on headcount and credit hour production. The other tab available to the user contains graphs of historical and projected headcounts and credit hours, which also dynamically update with stakeholder inputs.

Conclusions

Our process is one of many ways that an enrollment forecasting activity could be managed and displayed to stakeholders. The use of SAS® Enterprise Guide, the SAS® Add-In for Excel, and Excel streamlined the process eliminating complex data retrieval and manipulation, copying and pasting data into excel, and most importantly, maintaining a complex web application environment.

The only part of the process outside the scope of these tasks was the actual modeling exercise that was done with an outside statistical application such as SAS® Enterprise Miner or R. The results from the statistical analysis have to be translated into Excel with the necessary data elements correctly mapped from the data brought in by the SAS® add-in. However, this task is not normally performed every time the projection is run and only executed after a long period of time or when the model is showing signs of abnormal decay.

Contact Information

Andre Watts
University of Central Florida
Office of Institutional Research
Orlando, Florida 32826
407-823-2307
andre.watts@ucf.edu
http://ikm.ucf.edu

Lisa Sklar
University of Central Florida
Office of Institutional Research
Orlando, Florida 32826
407-823-1352
lisa.sklar@ucf.edu
http://ikm.ucf.edu