The Midwest SAS Users Group (MWSUG) Conference 2013
Columbus, Ohio – September 22-24, 2013

Paper CI-01-2013
Few SAS PROCs that Help Improving College Recruitment and Enrollment Strategies

By:

Harjanto Djunaidi and Monica Djunaidi[1]

Abstract

This paper discusses selected SAS PROCS which can be used by US colleges to improve their student recruitment and enrollment management strategies. Data were pulled from different sources such as NCES/IPEDS to show and illustrate the potential benefits of such applications. While the canned program might be able to accomplish the reporting needs and to some extend to produce certain basic statistical output by a matter of point-and-click, SAS will be the ultimate way to go in the future. The ability to translate strategies and decision makers' objectives into different SAS codes are imperative in a more volatile education industry which has been experiencing phenomenal structural changes recently such as the College Affordability Rating (CAR), a newly proposed government regulation. These changes occurred in fast rates with little or no time for adjustments. Therefore, it requires greater flexibility and ability to cope with them. Though the IRI (Institutional Research Intelligence) paradigms and education analytics concepts were just recently introduced to the public by the Association of American Education Analytics, the needs of people with such skills, expertise and experience have surged enormously as it can be seen from recent job postings. Increasing in demand for IRI or education analytics experts is derived partly from colleges' increasing needs to cope with phenomenal changes which have occurred in the competitive environments where they are operating. The higher learning institutions as well as the public reacted positively on these newly introduced mindsets, paradigms, approaches and analytical tools which can be applied to increase their operational efficiency, to outsmart their competitors and to retain, or increase their CAR such that federal aid money such as Pell grant can be secured. This paper and several others lay out examples how the IRI paradigms can be applied to help colleges to survive the on-going fundamental changes.

[1] Founder and co-founder, the Association of American Education Analytics (http://www.AAEA.us). Authors of the first education analytics book in the market (Institutional Research Intelligence: Go beyond Reporting) and College Affordability Rating (CAR). Innovator of Institutional Research Intelligence paradigms and Chief Editor, American Journal of Education Analytics. Contact harry.djunaidi@IRIntelligence.org

Introduction

Increasing the rate of college student enrollment growth is the biggest dream that most US college administrators have. This objective is more essential in recent years than it was in the past few years. There are several reasons why US colleges are working so hard to increase their enrollment numbers. First, colleges need to support and keep up with their continuous increasing operation cost. Second reason, declining revenue or income generated from alumni or third parties contributions has put college expansion as well as other planned investment on hold. Third, enrollment growth is used as an important yardstick by the school Board to judge the administrators' success. Fourth, increasing competition from for-profit universities has made the market share of traditional higher education institutions to shrink. Fifth, apparent shift of freshmen year students' preference. They prefer to take general education classes at a much cheaper place such as Community Colleges than at four-year colleges. Sixth, greater uncertainty on regulator's policies on student loans interest rates and tougher requirements to take the loans.

All these factors and increasing uncertainty on cost of future education have significant effects on the elasticity of demand for education services. These phenomenal changes have never occurred before. Uncertainly on student loans interest rate and increasing college cost have caused students and their family to be more careful in making their final decisions. This new development and shift in preference may have negatively impacted certain demand for education services. For example, should or should not the students attend a four-year college or will it be cheaper to take on-line classes?

Because of these dynamic changes in the industry, among others, the US colleges have been struggling in recent years to get their classes filled. When the enrollment revenue is down, it did not follow automatically by reduction in cost. Increasing cost of operation due to inflation or life-time labor contracts (the tenure system) or matching retirement fund among other factors may potentially create liquidity problems if colleges' source of income got disrupted. The Association of American Education Analytics has recently completed a study on this issue. The results suggest that 519 US higher education institutions are facing a liquidity and solvency problem as measured by their debt-to-equity ratio (www.aaea.us). The total liabilities over the total net assets of these institutions are $137 billion. The challenges in managing colleges are even more intense, challenging and difficult when the administrator introduces the College Affordability Rating (CAR). Colleges with lower CAR receive lower federal money compared to those with higher rating. Pressures are coming from all directions to urge colleges' administrators to increase and improve their operational efficiency so that tuition and student debts can be minimized. Perhaps, this is the sign that the heyday for those who hold the position as college administrators may soon be over.

This paper demonstrates how IRI (Institutional Research Intelligence) paradigms and education analytics help US colleges to identify and solve their enrollment challenges while at the same time increase their operational efficiency. Selected SAS PROCs combined with the IRI concepts can potentially get the job done nicely, efficiently and painless. This procedure may help identifying group of potential students that are highly probable to enroll. Therefore, recruitment efforts can be focused toward ensuring higher admissions yield. Applying the IRI mindsets are even more urgent nowadays than before. The CAR regulation in addition to the new law on student loans interest rate may add greater volatilities on college student enrollment. This is especially true when the rate is tied with commercial or market interest rate. When the economy gets better and private businesses are expanding their activities which will affect the competition to get the loans. It is just demand and supply. Increasing demand on commercial funding will certainly jack up the price for money i.e., interest rate. Therefore, colleges have to be ready to face higher demand volatility for their services than what it was in the past. Things are more uncertain because there are no financial instruments in the market where colleges can hedge and transfer some of their potential financial risks or exposures to the third parties, other markets or institutions who are willing to put their bet on it[2]. Unlike in the agricultural sector, colleges have to absorb totally their financial shortfalls. While corn farmers in Iowa or Illinois, for example can hedge their crop against price volatilities or weather through the futures market, it is not possible for colleges to shift the risk to the other parties if its enrollment decreases. The student loans perhaps can be bundled into financial product which then can be sold in the market. However, can one make a bet on enrollment risks and transform them into financial products which can be sold to the public?

The potential of financial failures facing by the colleges are pretty obvious. The risk is even more apparent for those who are borrowing from the commercial market and have outstanding loans to finance campus expansions. Therefore, the problem of enrollment failure is real and it needs to be addressed vividly.

The following SAS PROCs may be useful to identify the recruitment strategies. First, the college may need to know their past performance and who their competitors are. This paper will use publicly available data to illustrate how they can get the benefits by utilizing the approach. This study uses example on what has happened on colleges' enrollment located in Suffolk County, Massachusetts. This place is chosen where the city of Boston is located. It is one of the centers of learning where college students from different parts of the world come to study in the US. Theoretically, it implies that each college in this area should not have much of a challenge to recruit students for demand is growing regardless of how the US economy is. Therefore, one will expect that every college, private or public, profit or not-for profit higher learning institutions located in

[2]New profession will be created because of this need. Applying statistical and risk analyses in the decision making process are becoming more relevant in the near future. Therefore, it justifies the applications of IRI paradigms.

Boston should have experienced a positive enrollment growth.

First Step:

To focus the study in this area, one needs to create a small data file such that the effort can be done more efficiently. From a larger data set (ALL_Data), one can use the following SAS code to make a smaller one. Let us name this data set as Boston_0. ALL_Data are created from pulling the information from NCES/IPEDS data base. The following SAS statements will do the job nicely:

```
DATA MWSUG.Boston; SET All_Data;
IF (STABBR='MA' and CITY='Boston');
RUN;
DATA Boston_0; SET MWSUG.Boston;
AVE_MA=(P_GROWTH07+P_GROWTH08+P_GRO
WTH09+P_GROWTH10+P_GROWTH11)/(5);
RUN;
Proc Sort data=Boston_0
OUT=Boston_OUT; BY UNITID;
RUN;
PROC Print DATA=Boston_OUT;
VAR INSTNM COUNTYNM P_GROWTH07
P_GROWTH08 P_GROWTH09 P_GROWTH10
P_GROWTH11 AVE_MA;
RUN;
```

The codes above are written because the original data set has all enrollment information from the whole US colleges. Therefore, in order to only study enrollment growth in the Boston area, one needs to tell SAS to look and create a smaller data file which is a subset of the bigger data set. This can be done by specifying the state and the city or the county name as additional restrictions. SAS will create smaller file and name it Boston which is stored in MWSUG directory. As one can see, even with this simple task, there are three different SAS PROCs needed to create the small data file named Boston. In this first step, we sort (PROC SORT) the created data set based on the institution ID (UNITID), such that they can be used later for statistical analyses without worrying to sort them again. The PROC PRINT is embedded into the codes. This will give the researcher information of which institutions are listed in the newly created file. Including the PROC PRINT into the codes depends on the preference of the IRI expert who writes the codes. She or he can see the smaller data set by looking and examining them in the data set itself, without the need to print them as what we have shown above. The `P_GROWTH07` variable refers to the enrollment growth rate in 2007 academic year which is created using the enrollment information found in the ALL_Data file. The `AVE_MA` variable indicates the average of the enrollment growth. This average growth information will be used later in the analyses to other strategic information so that right decisions can be made. The results of running the above SAS codes are shown in Table 1. Some colleges in the Boston area may not show up in this study for not enough information found in the NCES/IPED data base.

Second Step:

The first step gives invaluable enrollment growth information to the respective school administrators. The average of enrollment growth information gives pretty good idea to the school management on their report card. For example, the administrators at Massachusetts College of Pharmacy and Health and Bay State College should be very happy for consistently maintaining their positive enrollment growth. On the other hand, things may not be so bright for those with negative numbers. The next question that one may ask is what next? The answer depends on the research agenda and the research questions. Supposed that Boston University's administrators wanted to know if their best strategies are to increase their student enrollment? What will be the odd to achieve that objective? Student enrollment growth number for this institution suggests that it may have reached the peak point and therefore it is operating on the plateau (growth rate line with slope close to zero or the limit of the line is approaching to zero). Therefore, instead of increasing its enrollment and build more classrooms, perhaps they should pay more attention on the quality of the enrolled students as well as improving its class room instructional activities. Rather than to employ cheap labor such as graduate students (teaching assistant/TA), the school needs to focus more on quality instructor and research.

Let us turn our attention to Emmanuel College that has an average enrollment growth equals to -0.80%. Suppose that the decision makers are trying to study their internal data. Just for the illustration purposes, this study is going to make up some "cooked" data so that the example can be completed. Only the College has the real numbers. Suppose the school administrators are asked by the Board to recruit a better student starting in the fall semester 2014 as measured by both ACT composite assessment score and HSGPA (high school GPA) for its newly created Honors College. How the Enrollment and Recruitment Office (ERO) needs to focus its efforts to meet the administrators' request? Since the real data are not available, this study has made up a data set which did not represent the institution under studied. Let us call this cooked data set as Zip. This data set contains four variables namely City, Zip Code, ACT composite scores and High School GPA.

If the administrators plan to increase its admissions standard for their newly Honors program, where can they find such high caliber applicants? Potential students who are admitted into the program will be given full financial supports. They will only be admitted with ACT composite score at least 28 and a high school GPA greater than 4[3]. The following codes can be written to create such a data set which only contains applicants' information that satisfies the admissions policies.

```
DATA High_ACT; SET Address;
If ACT >=28 and HSGPA > 4;
RUN;
Proc Freq data=High_ACT;
Table (ACT*Zip)/NOCOL NOROW
NOPERCENT;
TITLE 'This Information on ACT
by Zip';
Run;
```

Based on these results, the ERO can move on. The next research step consists of two actions to uncover the mystery. The first action would be to tell SAS to create yet another smaller data file as shown above. Let us call this data set High_ACT. The second action is to understand where the recruitment efforts need to be directed to

[3] Only those who took AP classes and AP test would be able to make a GPA higher than 4 out of 4.

achieve such goals. This market penetration effort is generated by the PROC FREQ. This statement will give information of the applicants' school Zip code.

There are many other analyses that can be done using the same data set. This is the advantage of having the skill to write codes in SAS. Unlike any other canned software, where researcher only needs to point and click to run a standard analysis, with the ability to write her or his SAS codes, a researcher can do other interesting studies as well. This example shows how much more ERO can do to analyze the data so that they can best serve the college and decision makers' interest in efforts to outsmart other competitors and to improve its report card and the College Affordability Rating (CAR).

The second step in the analyses aims to find the location of those students who satisfy the admissions standard for the Honors program.

The data set has to fulfill three criteria: (1). High School GPA > 4.00 and (2). ACT composite score is at least 28 and (3). The number of applicants who came from the same city has to be at least 3 people. To achieve this objective, then one can write the conditions as shown below. The important trick to learn in this step is to ask SAS to sum the number of applicants that came from the same high schools located in the same city. The information can be generated by asking SAS to compute the total number (NN) of applicants using the PROC MEANS and save the output (NEW1_OUT) so that it can be used later in the analyses.

Once the information is available, then merge the total number of the students in each city (NN) which has been saved in (NEW1_OUT) data set with (PL_OUT). In the process, the researcher needs to make sure that NN showed up in each city will be unique (no duplicate). This can be done before or after the PROC SQL statement. This study accomplishes this task after the two data sets were merged. In will be more efficient, had it been done prior of merging the data sets.

```
Data PL; Set Address;
If ACT >=28 and HSGPA > 4.00;
Run;
Proc Sort Data=PL Out=PL_OUT; By City;
Run;
Proc means n noprint Data=PL;
Class City;
Var ACT;
Output Out=NEW1 N=NN ACT;
Run;
Proc Sort Data=NEW1 Out=NEW1_OUT; By City;
RUN;
Proc sql;
        create table NEW_DATA as
             select *
        from NEW1_OUT as KUL1 left join
             PL_OUT as KAL1
        on KUL1.City = KAL1.City
        order by City;
quit;

Libname reflib 'SAS-data-library';
goptions reset=global gunit=pct border
cback=white
         colors=(black blue green red)
         ftitle=swissb ftext=swiss htitle=6
htext=4;

DATA PP; SET NEW_DATA;
DUPKEY=City;
If NN > 2;
RUN;
PROC SORT DATA=PP OUT=PP_OUT NODUPKEY; BY
City;
Run;
PROC Print DATA=PP_Out;
Var City NN;
RUN;
```

The PROC PRINT above will produce results and list the top city where the students might come from (an institution can use their past applicants' data to generate this important information). The following are the results of the PROC PRINT statement:

Obs	City	NN
1	Boston	9
2	Scituate	3
3	Springfield	4

The codes have asked SAS to produce results where NN is greater than 2 **(If NN > 2)**. That is the reason why only three cities make the list as shown above.

Given this important information, the ERO might be able to concentrate its recruitment efforts in these three cities. Most Students who are able to fulfill the new admissions standard are coming from high schools in Boston with total applicants equals to nine. So, the office might be able to direct its best recruitment efforts more effectively in the area. There is another way, how one can disclose the same strategic information. The following codes will give the same information, but in another form i.e., a graph as shown in Figure 1.

```
title1 h=4 'Applicants with ACT >= 28 and
HSGPA > 4';
title2 h=4 'Strong Applicants Cohort';

axis1 ORDER=(1 TO 10 BY 1)
      label=None
      offset=(5,5)
      width=3
      value=(height=5);
axis2 order=(0 to 10 by 1)
      label=None
      major=(height=1.5)
      minor=(height=1)
      width=3
      value=(height=5);

Proc Gplot data=PP_Out;
   format ACT NN;
   bubble NN*CITY=NN/ haxis=axis1
                             vaxis=axis2
                             vminor=1
                             bcolor=red
                             blabel
                             bfont=swissi
                             bsize=8
                             caxis=blue;
Run;
```

The size of the bubble represents the number of students who meet the requirement of ACT at least 28 and high school GPA greater than 4. Based on this cooked data set, high schools in Boston, Scituate and Springfield show up on the list. This information can be useful for the ERO to direct, target and max out their recruitment efforts.

Closing Notes

While to increase student enrollment might be the ultimate goals that most US college administrators are trying to accomplish in recent years, this objective can only be justified if the college is still operating at the decreasing part of its average cost (AC_0). For detailed discussions, please refer to the following paper (Harjanto Djunaidi and Monica Djunaidi: Increasing College Cost and Its Impacts on Student Loans which will be presented at *2013 Southeast SAS Users Group Conference* in St. Pete Beach, Florida, October 20-23, 2013). The messages convey by this graph is that a college may need to increase their student enrollment if it has not reached point A_0, where student enrollment equals to OA_0. Any effort or expansion beyond this point will certainly increase the college operational cost (perhaps, Boston University case). Any policy to increase student enrollment needs to be backed by intensive and solid research, and therefore, not just a try-and-error thing that college administrators can play with. It is important, therefore, for any higher institution to generate this graph based on its real data. This effort never been done and may not be necessary in the past, where resources are available more abundantly and competition is not as stiff as now. However, as the phenomenal structural competitive environments where the colleges are operating have changed tremendously in the past few months, US higher institutions have no option, but to use more IRI paradigms or education analytics that might help them to run their shops more efficiently and to deliver best education value to their clientele and the society or tax payers where they are serving.

Figure 2 - College Cost and Tuition

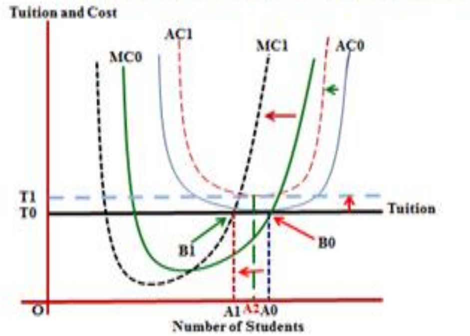
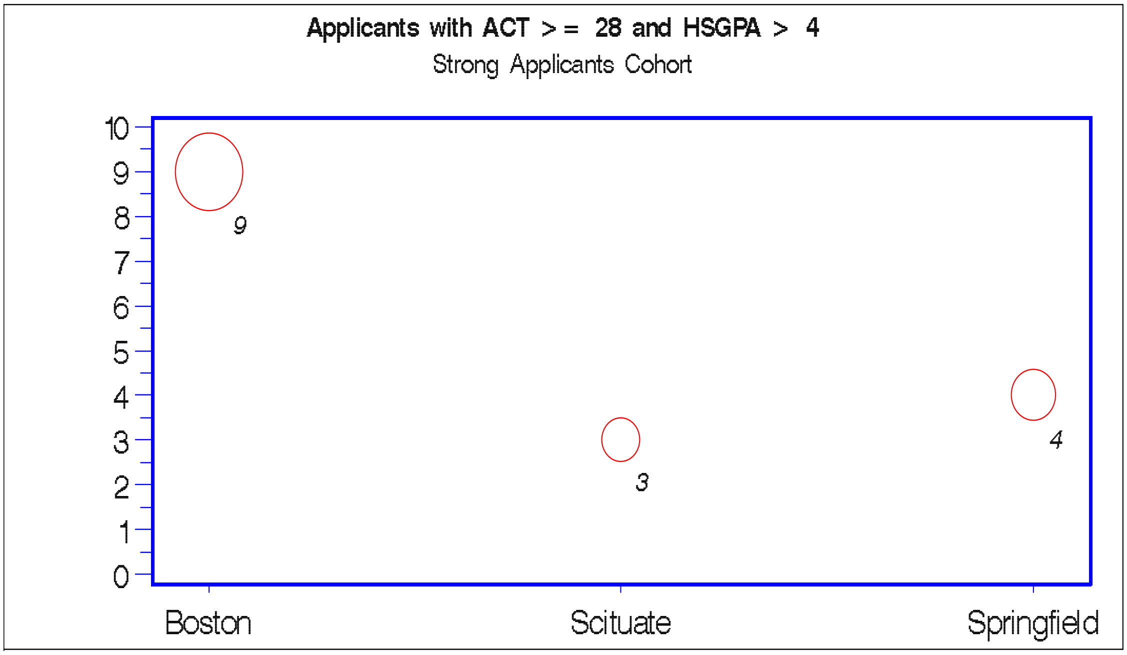


Table 1 – Student Enrollment Growth of Selected Colleges in the Boston Area

INSTNM	COUNTYNM	P_GR0WTH07	P_GR0WTH08	P_GR0WTH09	P_GR0WTH10	P_GR0WTH11	AVE_MA
Boston Baptist College	Suffolk County	0.72%	15.71%	-16.67%	17.78%	-11.32%	1.25%
Bay State College	Suffolk County	54.95%	-19.80%	20.00%	82.94%	5.41%	28.70%
Berklee College of Music	Suffolk County	-2.12%	1.22%	-0.04%	3.47%	-4.54%	-0.40%
Empire Beauty School-Boston	Suffolk County	19.54%	-1.68%	25.18%	-3.32%	-8.69%	6.21%
Boston Architectural College	Suffolk County	43.68%	-5.36%	2.54%	3.87%	-7.46%	7.45%
The Boston Conservatory	Suffolk County	50.51%	9.46%	3.29%	2.19%	0.58%	13.21%
Boston University	Suffolk County	-1.10%	0.46%	-0.24%	-1.44%	2.57%	0.05%
Bunker Hill Community College	Suffolk County	5.07%	5.10%	6.02%	14.51%	7.91%	7.72%
Butera School of Art at Fisher College	Suffolk County	-9.68%	1.79%	-50.88%	-14.29%	0.00%	-14.61%
Laboure College	Suffolk County	-13.56%	1.40%	4.97%	0.13%	8.54%	0.30%
Emerson College	Suffolk County	0.32%	-1.98%	1.45%	-0.61%	2.88%	0.41%
Emmanuel College	Suffolk County	7.54%	-6.91%	-0.83%	-4.28%	0.48%	-0.80%
Fisher College	Suffolk County	2.70%	3.12%	-4.46%	34.80%	12.00%	9.63%
Benjamin Franklin Institute of Technolog	Suffolk County	19.61%	13.73%	16.58%	4.79%	14.01%	13.74%
Gibbs College-Boston	Suffolk County	-18.06%	-4.55%	-53.07%	13.37%	33.73%	-5.72%
University of Massachusetts-Boston	Suffolk County	1.77%	5.90%	3.21%	4.67%	2.55%	3.62%
Massachusetts College of Pharmacy and He	Suffolk County	40.81%	24.78%	7.71%	16.30%	28.24%	23.57%
Massachusetts College of Art and Design	Suffolk County	3.90%	-2.30%	2.22%	0.00%	-1.59%	0.45%
School of the Museum of Fine Arts-Boston	Suffolk County	-11.43%	2.38%	-5.62%	0.87%	-7.91%	-4.34%
The New England Conservatory of Music	Suffolk County	0.26%	1.55%	-6.60%	11.96%	0.49%	1.53%
New England School of Photography	Suffolk County	17.39%	20.00%	-3.70%	0.00%	-22.44%	2.25%
North Bennet Street School	Suffolk County	-5.52%	-1.17%	-2.37%	-4.24%	4.43%	-1.78%
Northeastern University	Suffolk County	0.72%	0.42%	1.73%	5.31%	1.33%	1.90%
Simmons College	Suffolk County	9.29%	2.46%	1.94%	-9.54%	3.85%	1.60%
Suffolk University	Suffolk County	8.44%	2.53%	6.83%	0.87%	-1.57%	3.42%
Wentworth Institute of Technology	Suffolk County	-0.94%	3.43%	3.63%	-0.89%	-0.59%	0.93%
Wheelock College	Suffolk County	-11.28%	8.69%	3.22%	-6.59%	25.87%	3.98%
Urban College of Boston	Suffolk County	-0.44%	-9.88%	-2.81%	-3.39%	-0.93%	-3.49%
Average		**7.61%**	**2.52%**	**-1.31%**	**6.04%**	**3.14%**	**3.60%**

Figure 1 – Most Prepared Applicants' City for the Honors Program



References

Harjanto Djunaidi and Monica Djunaidi., College Affordability Rating: Strategy to Increase Federal Financial Aid. Academy Data Analytics Publisher. First Edition, 2013. (http://www.aaea.us/recent-education-policy-changes/).

Harjanto Djunaidi., *Various Latest Research Results*. Association of American Education Analytics. (http://www.aaea.us and http://www.IRIntelligence.org).

Harjanto Djunaidi and Monica Djunaidi., SAS PROC REPORT: How It helps Improving Student Retention Rate. *2013 MidWest SAS Users Group Conference*. Columbus, Ohio, September 22-24, 2013.

Harjanto Djunaidi and Monica Djunaidi., Increasing College Cost and Its Impacts on Student Loans. *2013 Southeast SAS Users Group Conference*. St. Pete Beach, Florida, October 20-23, 2013.

Harjanto Djunaidi., Predicting Students Enrollment Using SAS. *2012 South Central SAS Users Group Conference*. Houston, Texas, November 5-6, 2012.

Harjanto Djunaidi and Monica Djunaidi., *Institutional Research Intelligence: Go Beyond Reporting*. Academy Data Analytics Publisher. First Edition, 2013.

Harjanto Djunaidi and Monica Djunaidi., *Education Analytics and its Applications*. Academy Data Analytics Publisher. Upcoming book, 2014.

Harjanto Djunaidi and Monica Djunaidi., *Student Loans: Blessings or Courses*. Academy Data Analytics Publisher. Upcoming book, 2014.

Harjanto Djunaidi., Predicting Students Enrollment Using SAS. *2012 South Central SAS Users Group Conference*. Houston, Texas, November 5-6, 2012.

Harjanto Djunaidi., Data-driven Decision: Managing Higher Education Institutions more Effectively. *2012 North Carolina Community College System Conference*. Raleigh, October 7-9, 2012.

Harjanto Djunaidi., Institutional Research Intelligence: Go Beyond Reporting. *2012 South Central SAS Users Group Conference*. Houston, Texas, November 5-6, 2012.

SAS/STAT® 9.3