University of Illinois at Urbana-Champaign

Women In Engineering

STAT427 Project

Yaxian Xie, Anqi Li, Yilin Zhao and Jing Gu
5-11-2015

Table of Contents

1. Introduction

College of Engineering at University of Illinois at Urbana-Champaign is one of the top ranked colleges in US. Currently College of Engineering offers 14 engineering degrees. In 2014, there are 1375 (18.1) women undergraduates and 571 (7.5%) Under Represented Minority (URM)[1]. In recent years, College of Engineering provides Women in Engineering programs to attract and retain female students. Specifically, there are two women programs including Women in Engineering (WIE) Orientation and Girls Adventures in Mathematics, Engineering, and Science (GAMES) Camp.

The objectives of this Women in Engineering project include two aspects. First, we aims to compare the *performance/outcomes of* female vs. male students in *College of Engineering* regarding Retention Rate, GPA, etc. Second, we will explore factors related to the differences between female and male students including major, enrollment credentials, etc.

In the following sections, we proposed specific research questions based on different datasets. In general, we covered four topics including Retention Rate in College of Engineering, Effectiveness of Women in Engineering programs, Retention Rate across Colleges as well as Performance and Enrollment. For each topic, we presented data description, analysis methods and results. Besides, we briefly summarized the key points for research questions in each section. Finally, we recaptured our key findings and suggested directions for future study,

2. Retention Rate in College of Engineering

2.1 Research questions

This section focuses on exploring the relationship between gender and retention rate as well as the relationship between race and retention rate. R*etention Rate* is defined as Number of students who graduate in the same major divided by Number of students who graduate from Engineering.

Here we seek to answer the following questions:

1) Are there significant differences in Retention Rate between female and male students within College of Engineering?
2) Does the proportion of female students in each major have significant effect on their *Retention Rate*?
3) Does GPA have significant effect on *Retention Rate* of female students?
4) Are there significant differences in Retention Rate between URM and Non-URM students within College of Engineering?
5) Does GPA significantly affect the *Retention Rate* of URM students?

To answer these questions, first we focused on the relationship between gender and retention rate. And then we explored the differences between URM and Non-URM students.

[1] Under Represented Minority (URM) - Black/African American, Native American, Hispanic

2.2 Data on Gender

We created a new data set called *subdata*, which contains 148 observations and 7 variables (see in Table 1).

Table 1 Data description

Variable	Type	Meaning
CohortTerm	Factor	There are 5 cohort terms from 2003-2007
CohortFirstMajor	Factor	There are 15 majors in college of engineering. These are the majors students enroll in.
Enroll GPA	Numeric	It's the 4 year GPA of students who enroll in college of engineering
Remain GPA	Numeric	It's the 4 year GPA of students who graduate in the original major
Female Ratio	Numeric	This is the proportion of female students in the specific major. Number of female students/Number of all students
Retention Rate	Numeric	Number of students who graduate in the same major/Number of students who graduate from Engineering
Gender	Factor	Students gender

2.3 Gender - Descriptive Analysis

We used Pearson r's Correlation to check the correlation between each pair of the predictors. Here is the ellipse correlation plot. The shade of the color represents the strength of the correlation. The deeper the color is, the stronger the correlation is.

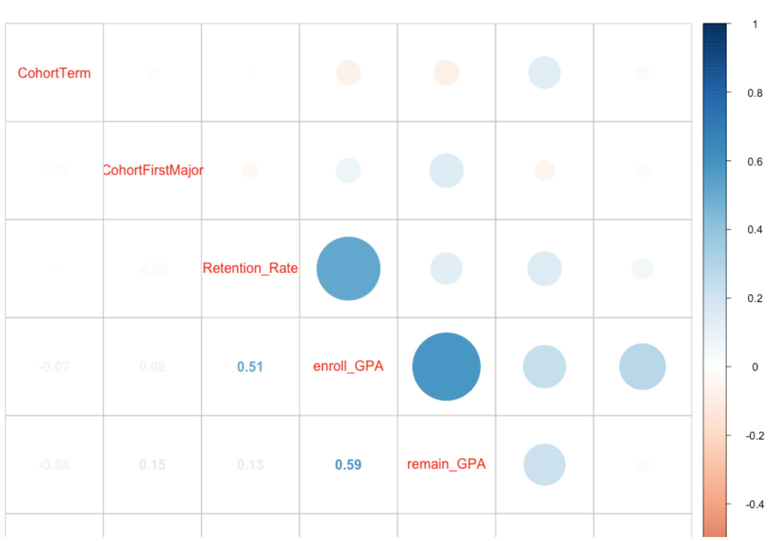

Figure 1 Ellipse Correlation Plot

Figure 1 indicated that there is strong correlation between *enroll_GPA* and *remain_GPA* as well as *enroll_GPA* and *Retention Rate*. Next, we took a look at the contingency table. Based on Chi-square test (p-value=0.23) and likelihood test (p-value=0.24), we concluded that Gender and Retention are independent (see R results in Appendix 7.1).

Table 2 Contingency Table Gender and Retention

Retention Or Not	Male	Female	Sum
Yes	3240	2318	5558
NO	638	494	1132
Sum	3878	2812	6690

2.4 Gender Effects: ANOVA and ANCOVA

1) Gender and Retention Rate

First we conducted ANOVA analysis for Gender and Retention rate in all cohort majors and terms. Based on the ANOVA results, we can see that there is no significant difference in Retention Rate between Female and Male students (p-value=0.203). Based on the boxplot in Figure 2, we can see that the mean values of male and female students are close. But Female students have larger variance.

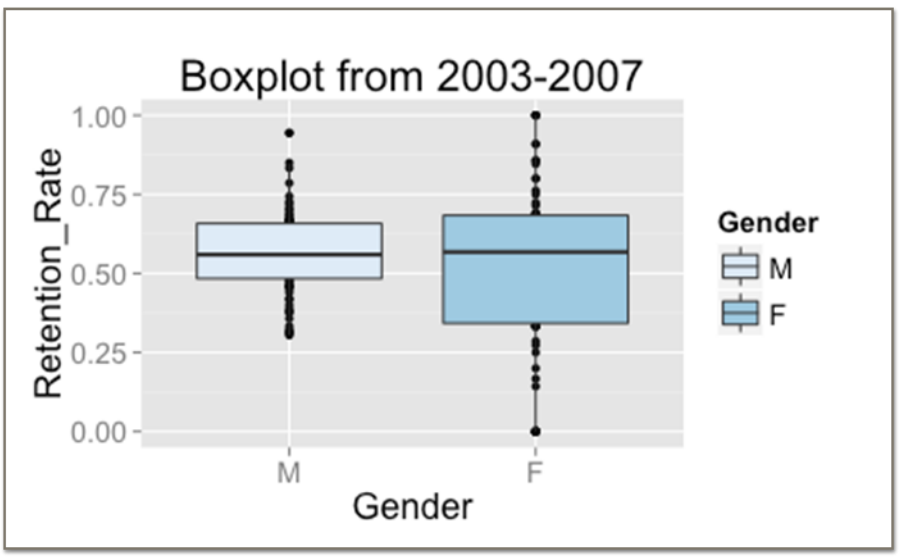

Figure 2 Boxplot for Gender and Retention Rate

Next we explored the relationship between gender and retention rate in each major (see in Table 3). As a result, we found that there is significant difference in Retention Rate between female and male students in two majors, that is, Engineering Physics (p-value=0.01) and General Engineering (p-value=0.002). Specifically, male students have higher Retention Rate than female students in these two majors. And the results for the two major are shown in Figure 3. In terms of other majors, there is no significant difference between each gender.

Table 3 ANOVA analysis for each major

Major	Estimate	Std Error	T.value	P-value
ABE	-0.2074	0.1272	-1.6302	0.1471
Aero	-0.0639	0.0596	-1.0726	0.3190
CS	-0.0177	0.0604	-0.2928	0.7782
Chem E	0.0520	0.0627	0.8289	0.4345
Civil	-0.0090	0.0459	-0.1970	0.8494
Comp E	0.0024	0.1037	0.0232	0.9821
E Mech	0.0288	0.2375	0.1211	0.9070
E Phys	-0.2636	0.0814	-3.2368	**0.0143**
EE	0.0283	0.0354	0.7996	0.4502
Gen Eng	-0.1212	0.0259	-4.6763	**0.0023**
Ind Eng	0.0583	0.1523	0.3827	0.7133
MatSE	0.0797	0.1204	0.6619	0.5292
Mech E	-0.0934	0.0739	-1.2632	0.2470
NPRE	-0.1150	0.1417	-0.8113	0.4439
Bioen	-0.0154	0.0443	-0.3474	0.7424

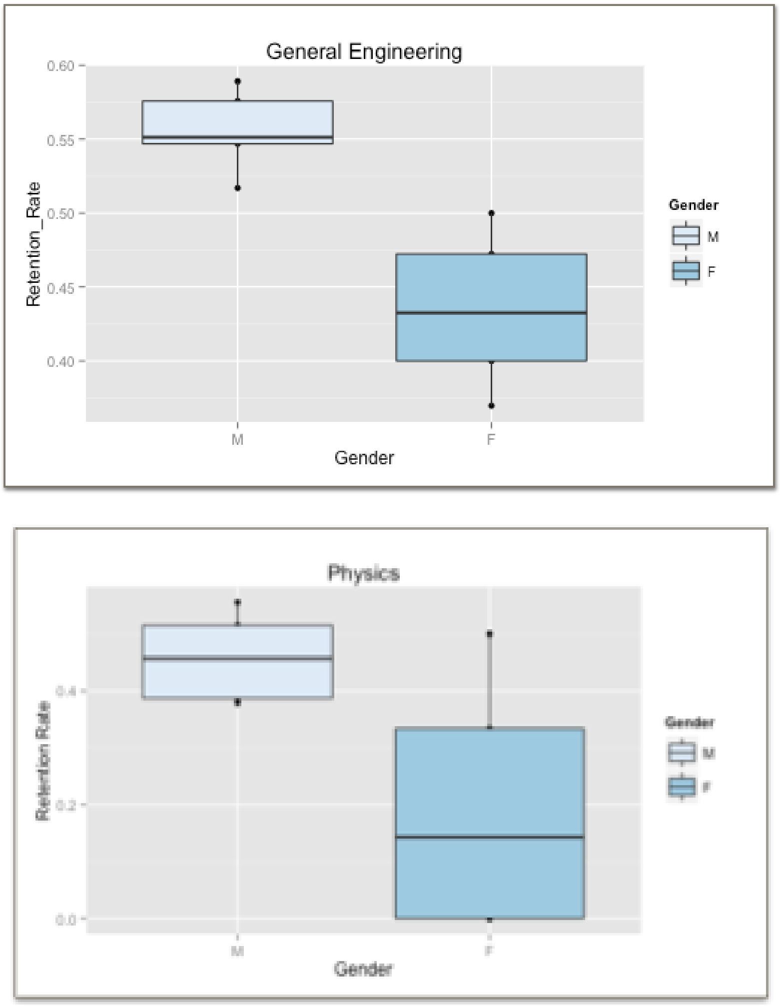

Figure 3 Box plot for Physics and General Engineering

2) Proportion of Female Student and Female Retention Rate

In this part, we further explored retention rate of female students. Here we defined female ratio as the proportion of female students. Specifically, we fit an ANCOVA model for retention rate with female ratio, enroll_GPA, remain_GPA, CohortFirstMajor and CohortTerm as predictors. We need to check the correlation between female ratio and other 3 predictors since this question focus on the effect of female ratio (see in Table 4).

Table 4 Association with Female Ratio in Subfemale

variables	**Method**	**Coefficients**	**Interpretation**
enroll_GPA	correlation analysis & linear regression	correlation coefficient r=0.263 p-value of linear regression is 0.0184 < 0.05	correlation analysis shows a weak positive correlation while linear regression shows association between enroll_GPA and female ratio.
remain_GPA	correlation analysis & linear regression	correlation coefficient r=0.269 p-value of linear regression is 0.029 < 0.05	correlation analysis shows a weak positive correlation while linear regression shows association between remain_GPA and female ratio.
CohortFirstMajor	ANOVA	p-value = 2.27E-09	CohortFirstMajor is associated with female ratio.
CohortTerm	ANOVA	p-value = 0.309	CohortTerm has no correlation with female ratio.

Based on results above, we dropped enroll_GPA, remain_GPA and CohortFirstMajor. And our final model is: $Retention\_Rate \sim Female\_Ratio + CohortTerm$ (see R results in Appendix 7.1). The effect of Female Ratio is significant (p-value=0.01). But the size of effect appears marginal.

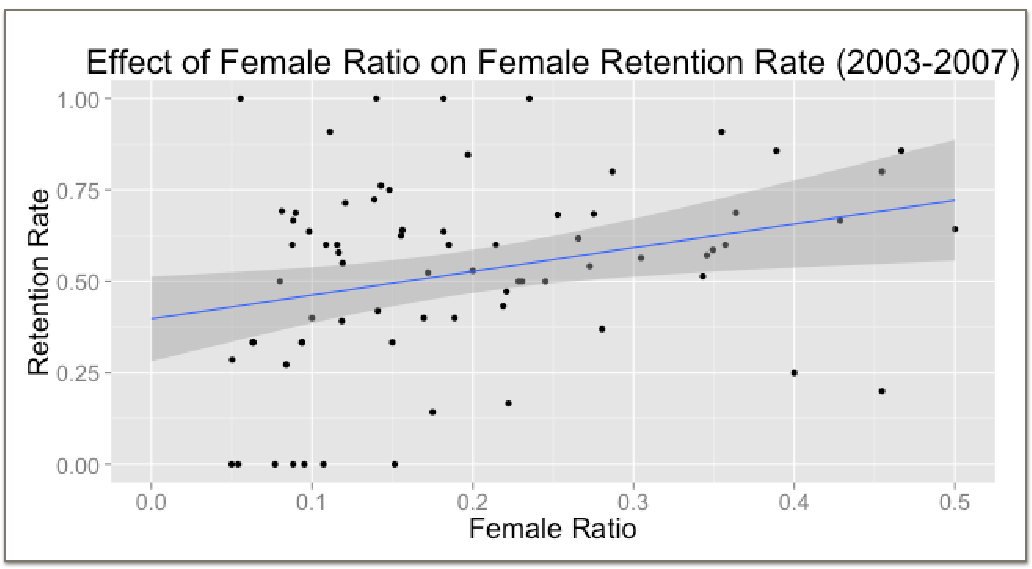

Figure 4 Linear Regression Line for Female Ratio and Retention Rate

3) GPA and Female Retention Rate

First we tested the effect of enroll_GPA on retention rate for female students. Next, we checked the correlation among predictors, and then we built model with enroll_GPA, CohortFirstMajor, CohortTerm (see R results for each major in Appendix 7.1).

$$Retention\ Rate \sim enroll\_GPA\ +\ CohortFirstMajor\ +\ CohortTerm$$

The estimate of enroll_GPA is significant (p-value < 0.01). When there is 0.1 increase in enroll_GPA, we expect female's retention rate will increase 4.47% (see regression line in Figure 5).

Figure 5 Linear Regression Plot for enroll_GPA and Retention Rate

Similarly, we built a model with remain_GPA, CohortFirstMajor, CohortTerm (see R results in Appendix 7.1).

$$Retention\ Rate \sim remain\_GPA\ +\ CohortFirstMajor\ +\ CohortTerm$$

The estimate of remain_GPA is insignificant (p-value = 0.38). That is, remain_GPA has no effect on retention rate. It seems reasonable since remain_GPA refers to GPA of students who stay in the same major.

2.5 Data on Race

To explore the effects of race (i.e. URM vs. Non-URM), we created a new data set ("URMdata") including 148 observations and 6 variables (see in Table 5).

Table 5 Data Description for URMdata

	Type	Meaning
CohortTerm	Factor	There are 5 cohort terms from 2003-2007
CohortFirstMajor	Factor	There are 15 majors in college of engineering. These are the majors students enroll in.
Enroll GPA	Numeric	It's the 4 year GPA of students who enroll in college of engineering
Remain GPA	Numeric	It's the 4 year GPA of students who graduate in the original major
Retention Rate	Numeric	Number of students who graduate in the same major/Number of students who graduate from Engineering
Race	Factor	URM or Non URM students

2.6 Race – Descriptive Analysis

We used Pearson r's Correlation to check the correlation among predictors (see results in Figure 6). The shade of the color represents the strength of the correlation. The deeper the color is, the stronger the correlation. This plot indicates that there is strong correlation between *enroll_GPA* and *remain_GPA* as well as *enroll_GPA* and *Retention Rate*. There is moderate correlation between *enroll_GPA* and *Race* as well as *Remain_GPA* and *Race*.

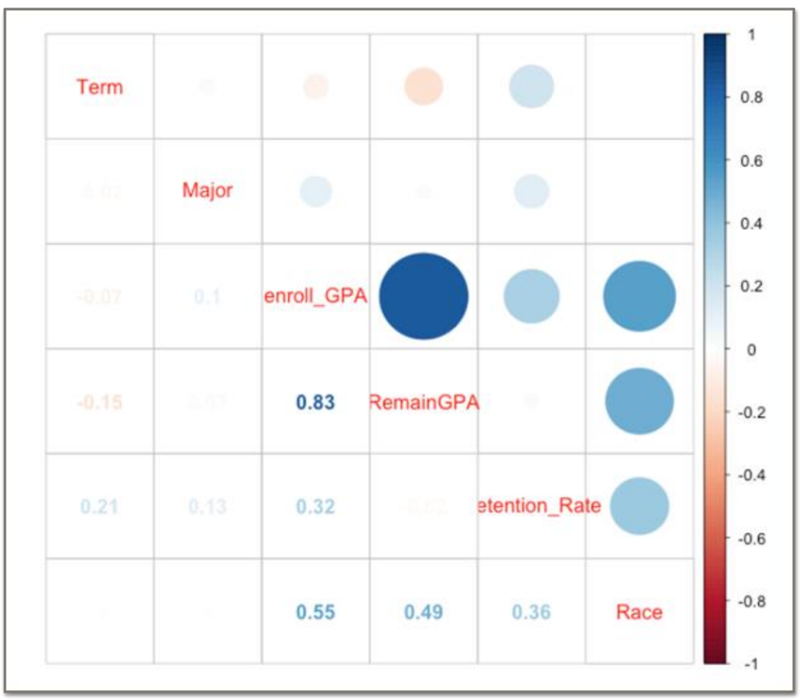

Figure 6 Ellipse Correlation Plot

2.7 Race Effects: ANOVA and ANCOVA

1) Race and Retention Rate

First we ran ANOVA analysis for Race and Retention rate in all cohort majors and terms.

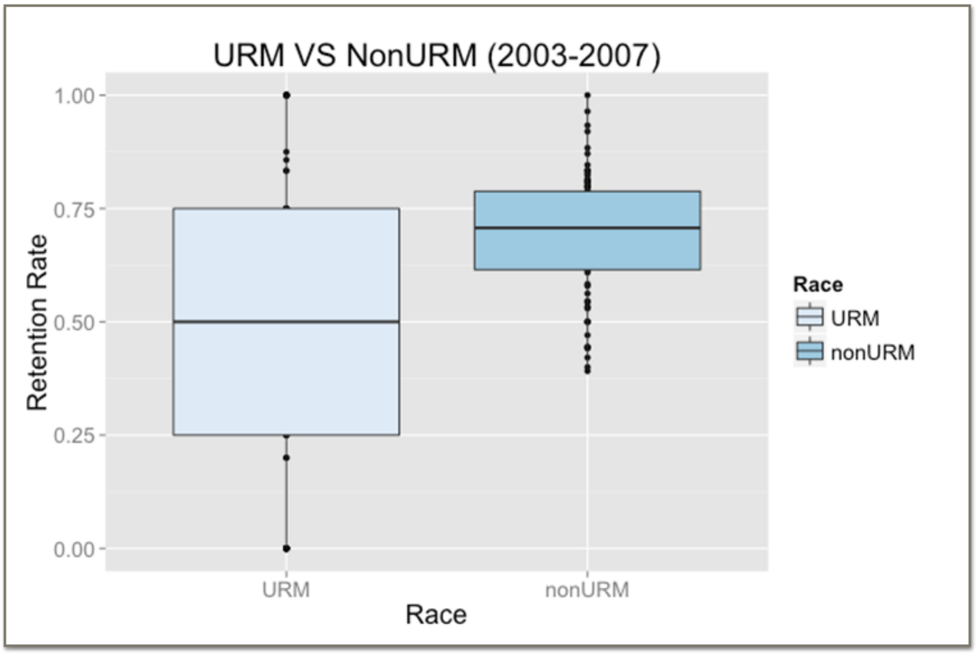

Figure 7 Boxplot for Race and Retention Rate

We can see that there is significant difference in Retention Rate between URM and Non-URM students (p-value<0.05). Specifically, URM students have lower retention rate than Non-URM students. In the boxplot (Figure 7) we can see that the mean value of Retention Rate for URM students is smaller than that of Non-URM students. Also, Retention rate of URM students have larger variance.

Next we explored the relationship between Race and Retention Rate for each major. However, there is no URM students for certain majors. Accordingly, we ruled out these majors for further analysis. In this part, we built ANOVA models for CS, Aero, Civil, Electronic Engineering, General Engineering, and Mechanical Engineering (see results in Table 6). Based on the results below, we found that the effects of Race on Retention Rate are significant for three majors including CS (p-value=0.026), EE (p-value=0.011) and Mech E (p-value=0.027). For these majors, URM students have significantly lower retention rate than Non-URM students. Regarding other majors, there is no significant difference between URM and Non URM students.

Table 6 ANOVA: Race and Retention Rate for Each Major

Major	Estimate	Std.Error	t-value	P-value
Aero	0.1031	0.0829	1.2445	0.2485
CS	0.3013	0.1107	2.7218	**0.0262**
Civil	0.0918	0.1138	0.8071	0.4430
EE	0.3266	0.0989	3.3025	**0.0108**
Gen Eng	-0.0819	0.0960	-0.8528	0.4186
Mech E	0.1457	0.0539	2.7019	**0.0270**

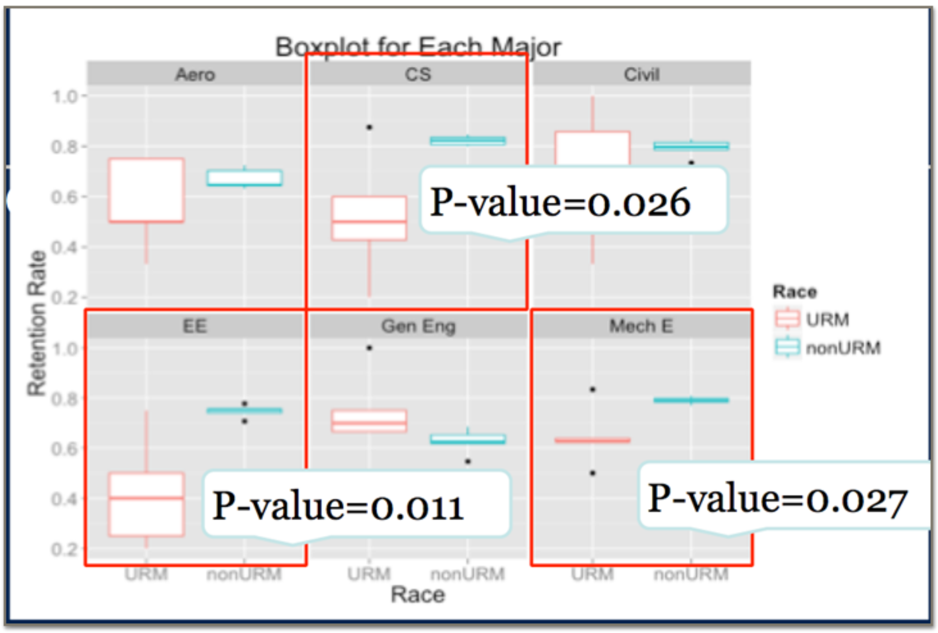

Figure 8 Boxplot for each major

2) Race and GPA

In this part, we are interested in URM students. First we explored effects of enroll_GPA for URM students. And then we checked association among predictors and built model with enroll_GPA, CohortFirstMajor, CohortTerm as follows:

$$Retention\ Rate \sim enroll\_GPA\ +\ CohortFirstMajor\ +\ CohortTerm$$

And the estimate of enroll_GPA is significant (with p-value <0.01). When there is 0.1 unit of increase in enroll_GPA, we expect female's retention rate will increase 3.40%.

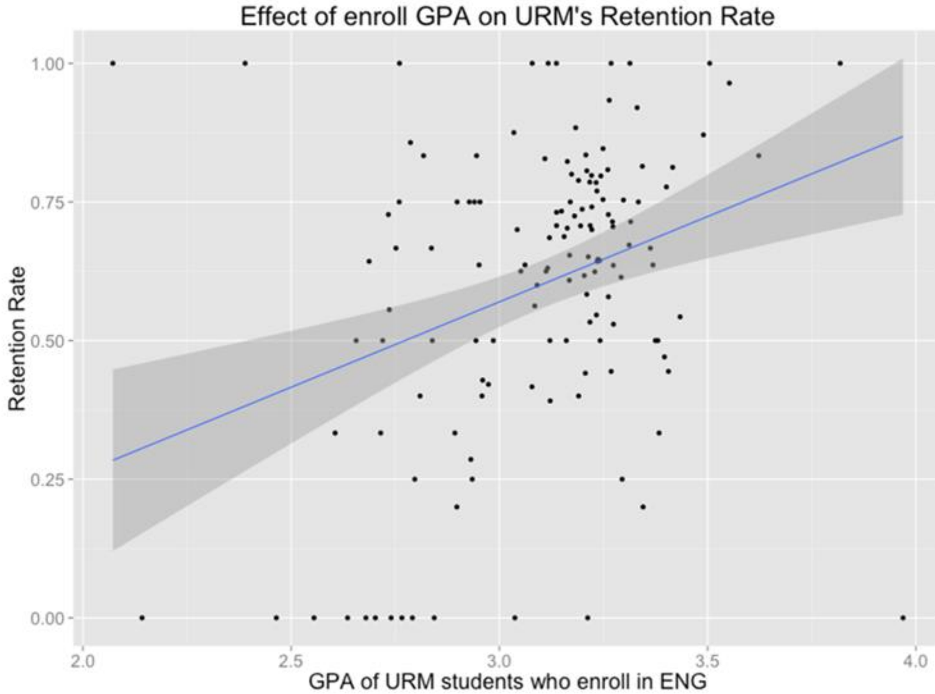

Figure 9 Linear Regression Plot for enroll_GPA and Retention Rate

Next we took a look at the effect of remain_GPA. After checking the correlation between remain_GPA and other predictors, we built model with remain_GPA, CohortFirstMajor, CohortTerm. The estimate of remain_GPA is insignificant(p-value = 0.6982). So remain_GPA has no effect on retention rate for URM students (see R results in Appendix 7.1).

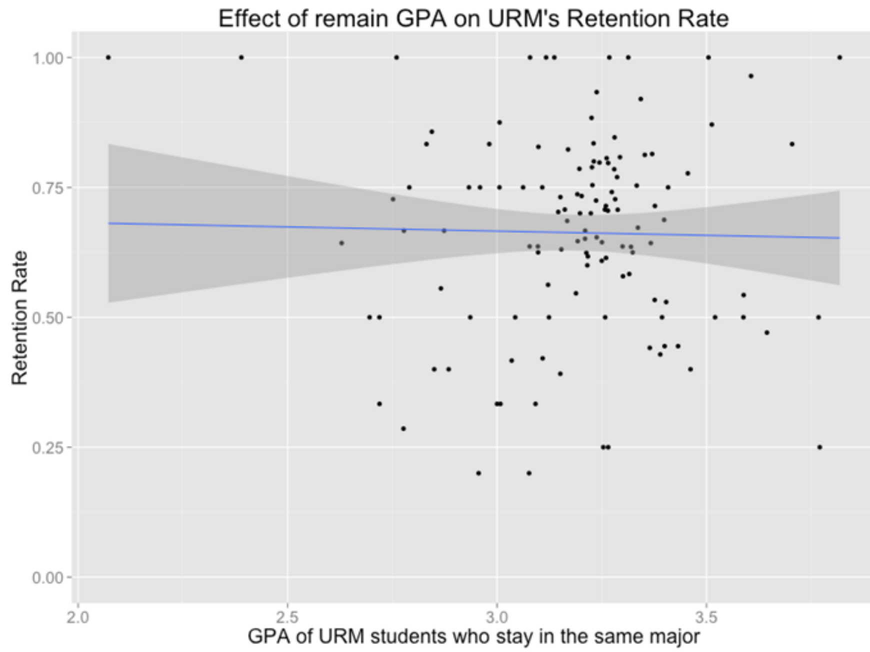

Figure 10 Linear Regression Plot for remain_GPA and Retention Rate

2.8 Brief Summary

Overall, there is no significant difference in Retention Rate between female students and male students. But in E Physics and General Engineering, female students have significantly lower Retention Rate. Also, the proportion of female students has significantly positive effect on female's Retention Rate. Besides, GPA(especially Enroll GPA) has significantly positive effect on female's Retention Rate.

In general, Non-URM students have significantly higher Retention Rate than URM students. And we can conclude that in CS, EE, Mechanical Engineering, Non-URM students have significantly higher Retention Rate than URM students. GPA (especially Enroll GPA) has significantly positive effect on URM's Retention Rate.

3. Effectiveness of Women in Engineering Programs

3.1 Research questions

Women in Engineering program aims to provide a welcoming and supportive environment for female students in the College[2]. For prospective students, Girls Adventures in Mathematics, Engineering, and Science (GAMES) Camp is designed to give high school girls an opportunity to explore exciting engineering and scientific fields. For admitted students, Women in Engineering - Freshman Orientation (WIE Camp) is an orientation for incoming women undergraduates in the College of Engineering. The

[2] Please see more details from http://engineering.illinois.edu/wie/index.html

program helps the female students adjust to college and gain confidence in their engineering abilities.

In this section, we attempt to answer the following research questions: Is the Women in Engineering program including WIE orientation and GAMES camp effective based on different indicators including GPA, Retention Intention and Time to Graduation?

Specifically, we define the related terms as follows: 1) GPA is measured by Last Engineering Cumulative GPA and Last UIUC Cumulative GPA; 2) Retention in Engineering refers to a case that a student enters and graduates from College of Engineering; 3) Retention in major means that a student graduates from initial major in College of Engineering; 4) Time to Graduation measures how long a student takes to graduate in semesters.

3.2 Data

In order to answer the questions above, we obtained raw data including 3243 female student in College of Engineering from 2003 to 2014. Within this dataset, there are 1698 graduates and 510 students who transfer to other colleges. Here is the list of variables in this raw data (see in Table below). Next, we created several new variables for further analysis as follows:

- *WIE*: 1 if participate, 0 otherwise;
- *Games*: 1 if participate, 0 otherwise;
- *MajorRe*: 1 if Initial Major= Last Major for graduates, 0 otherwise;
- *CollegeRe*: 1 if LastUIUCCollege= Engineering or LastMajor=Agricultural & Biological Engr/Chemical Engineering for graduates; 0 otherwise;
- *GradTerms* (/terms): DegreeTerm-InitialTerm if graduates, ‘Not Graduate’ otherwise;
- Graduation: ‘On Time’ if graduation time is within 4 years; ‘Postpone’ otherwise.

Table 7 Variables in Dataset 2

Variables	**Meaning**	**Values**
InitialTerm	Initial term for each student	Categorical values; For example, 120138 refers to Fall 2013 at Urbana-Champaign campus
InitialMajorCode	Major code for initial term	Categorical values
InitialMajor	Initial major for each student	Categorical values
LastEngrTerm	Last Engineering term for each student	Categorical values
LastUIUCTerm	Last UIUC term for each student	Categorical values
DegreeTerm	Degree term for each student	Categorical values
LastUIUCMajorcode	Major code for last UIUC major	Categorical values
LastUIUCMajor	Last UIUC major	Categorical values
LastUIUCCollCode	College code for Last UIUC college	Categorical values
LastUIUCCollege	Last UIUC college	Categorical values
LastEngrCumGPA	Last engineering cumulative GPA	Numeric values
LastUIUCCumGPA	Last UIUC cumulative GPA	Numeric values
GamesCamp	Games Camp participation	Categorical values: Y – participated; Null – Not participate
WIECamp	WIE Orientation participation	Categorical values: Y – participated; Null – Not participate

3.3 Descriptive Analysis

In this part, we took a look at the data and checked basic descriptive analysis. Given that GAMES Summer Camp is relatively new, there is very few female students who have participated GAMES and graduated from College of Engineering. Accordingly, we focused on WIE Orientation as an example of Women in Engineering program. Here we are interested in WIE Orientation & Retention, GPA and Graduation.

1) WIE and Retention in Major/College

Based on results below (Table 8 and Table 9), among 1698 female students who graduated, 796 students (47%) have participated WIE Camp and 1096 (65%) students stayed in the original Engineering major. Moreover, the Chi-Square test (p-value=0.01) indicated WIE participation and Retention in Major are not statistically independent.

Table 8 Contingency Table: WIE and Retention in Major

Table of WIE by MajorRe			
WIE(WIE Camp Participation)	MajorRe(Retention in Major)		
Frequency Percent	0	1	Total
0	345 20.32	557 32.80	902 53.12
1	257 15.14	539 31.74	796 46.88
Total	602 35.45	1096 64.55	1698 100.00

Table 9 WIE and Retention in Major: Chi-Square Test

Statistic	DF	Value	Prob
Chi-Square	1	6.5678	0.0104
Likelihood Ratio Chi-Square	1	6.5841	0.0103
Continuity Adj. Chi-Square	1	6.3099	0.0120
Mantel-Haenszel Chi-Square	1	6.5639	0.0104
Phi Coefficient		0.0622	
Contingency Coefficient		0.0621	
Cramer's V		0.0622	

Also, we checked the relationship between WIE participation and Retention in College (see results below). Similarly, we found that WIE participation and Retention in College are associated.

Table 10 Contingency Table: WIE and Retention in College

Table of WIE by CollegeRe			
WIE(WIE Camp Participation)	**CollegeRe(Retention in College)**		
Frequency Percent	**0**	**1**	**Total**
0	220 12.96	682 40.16	902 53.12
1	156 9.19	640 37.69	796 46.88
Total	376 22.14	1322 77.86	1698 100.00

Table 11 WIE and Retention in Major: Chi-Square Test

Statistic	DF	Value	Prob
Chi-Square	1	5.6327	**0.0176**
Likelihood Ratio Chi-Square	1	5.6599	0.0174
Continuity Adj. Chi-Square	1	5.3582	0.0206
Mantel-Haenszel Chi-Square	1	5.6294	0.0177
Phi Coefficient		0.0576	
Contingency Coefficient		0.0575	
Cramer's V		0.0576	

2) Retention and Time to Graduate

Based on the contingency tables below, we can see that students who have stayed in the same major/college are more likely to graduate on time. And the Chi-Square tests (p-value<0.01) supported such statements.

Table 12 Contingency Table: Retention in Major and Time to Graduate

Table of MajorRe by Graduation			
MajorRe(Retention in Major)	**Graduation**		
Frequency Percent	**On Time**	**Postpone**	**Total**
0	474 27.92	128 7.54	602 35.45
1	982 57.83	114 6.71	1096 64.55
Total	1456 85.75	242 14.25	1698 100.00

Table 13 Contingency Table: Retention in College and Time to Graduate

Table of CollegeRe by Graduation			
CollegeRe(Retention in College)	**Graduation**		
Frequency Percent	**On Time**	**Postpone**	**Total**
0	297 17.49	79 4.65	376 22.14
1	1159 68.26	163 9.60	1322 77.86
Total	1456 85.75	242 14.25	1698 100.00

3) WIE and GPA

Based on results below, we can see that the mean and variance of Last Engineering GPA seem close among students who have participated WIE and those who have not participated WIE. In other words, GPA performance is not related to WIE participation. We think the result is reasonable because the objective of WIE orientation is not to help female students improve performance directly.

Table 14 Basic Descriptive Statistics

	LastEngrCumGPA		
	Mean	**Std**	**N**
WIE Camp Participation			
0	3.11	0.62	1494
1	3.11	0.61	1740

3.4 ANOVA

In this part, we examined the effectiveness of WIE Orientation in terms of Retention in Major/College, GPA and Time to Graduate.

1) WIE and GPA

First, the results below show that WIE participation had no significant effects on both Last Engineering GPA (p-value=0.72) and Last UIUC GPA (p-value=0.60). That is, a female student' GPA performance is not related to her participation in WIE orientation.

Table 15 ANOVA: WIE and Last Engineering GPA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.048188	0.048188	0.13	0.7208
Error	3232	1219.413277	0.377294		
Corrected Total	3233	1219.461465			

Table 16 ANOVA: WIE and Last UIUC GPA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.0507023	0.0507023	0.27	0.6023
Error	1696	316.5371246	0.1866375		
Corrected Total	1697	316.5878269			

2) WIE and Retention in Major/College

Regarding retention in major/college, we found that there are significant effects of WIE participation on Retention in Major (p-value=0.01) and Retention in College (p=0.02) at 0.05 level. That is, female students are more likely to stay in the same major and college when they participate WIE orientation vs. not participate.

Table 17 ANOVA: WIE and Retention in Major

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1.5029759	1.5029759	6.59	0.0104
Error	1696	387.0671066	0.2282235		
Corrected Total	1697	388.5700824			

Table 18 ANOVA: WIE and Retention in College

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.9710947	0.9710947	5.64	0.0176
Error	1696	291.7685991	0.1720334		
Corrected Total	1697	292.7396938			

3) WIE and Time to Graduate

Besides, we tested the effects of WIE participation on Time to Graduate. And the results below indicated that there is no significant relationship between WIE participations and Time to Graduate (p-value=0.76).

Table 19 ANOVA: WIE and Time to Graduate (/terms)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.111976	0.111976	0.10	0.7570
Error	1696	1982.548212	1.168955		
Corrected Total	1697	1982.660188			

4) Retention in Major/College and Time to Graduate

Moreover, we explored the effects of Retention in Major/College on Time to Graduate. The results below supported that both Retention in Major (p<0.0001) and Retention in College (p=0.0055) have significant effects on Time to Graduate. In other words, female students are more likely to graduate on time when they stay in the same major/college.

Table 20 ANOVA: Retention in Major and Time to Graduate (/terms)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	22.809363	22.809363	19.74	<.0001
Error	1696	1959.850826	1.155572		
Corrected Total	1697	1982.660188			

Table 21 ANOVA: Retention in College and Time to Graduate (/terms)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	8.995355	8.995355	7.73	0.0055
Error	1696	1973.664833	1.163717		
Corrected Total	1697	1982.660188			

3.5 Logistic Regression Model

Since we have data at individual student level, we are able to analyze how different factors affect Retention in Major/College and Time to Graduate. Because these response variables are binary, it's appropriate to build logistic regression models. Besides, we used stepwise method to select explanatory variables.

1) Retention in Major

Based on results below, we found that both WIE (p-value<0.01) and Last Engineering GPA (p-value<0.01) have significant effects on Retention in Major. Particularly, female students who have participated WIE orientation are more likely to stay in the same major. Specifically, the odds ratio of Retention in Major increases by 32.3% when female students participate in WIE Orientation vs not participate.

Table 22 Logistic Regression on Retention in Major

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-3.1413	0.3524	79.4596	<.0001
WIE	**1**	1	0.2802	0.1059	7.0004	0.0081
LastEngrCumGPA		1	1.1503	0.1104	108.6139	<.0001

Table 23 Retention in Major: Odds Ratio Estimates

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
WIE 1 vs 0	1.323	1.075	1.629
LastEngrCumGPA	3.159	2.545	3.922

2) Retention in College

Next, we found that there are similar results regarding Retention in College. Both WIE (p-value=0.02) and Last Engineering GPA (p-value<0.01) have significant effects on Retention in College at 0.05 level. Particularly, female students who have participated WIE orientation are more likely to stay in the same college. Specifically, the odds ratio of Retention in College increases by 36% when female students participate in WIE Orientation vs not participate.

Table 24 Logistic Regression on Retention in College

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-4.4150	0.4122	114.7210	<.0001
WIE	**1**	1	0.3072	0.1266	5.8844	0.0153
LastEngrCumGPA		1	1.8050	0.1349	179.0658	<.0001

Table 25 Retention in College: Odds Ratio Estimates

Odds Ratio Estimates			
Effect	**Point Estimate**	**95% Wald Confidence Limits**	
WIE 1 vs 0	1.360	1.061	1.743
LastEngrCumGPA	6.080	4.668	7.920

3) Time to Graduate

Here we measured Time to Graduate by Graduation on time or not. And results showed that both Retention in Major (p-value<0.01) and Last Engineering GPA (p-value<0.01) have significant effects on Graduation. Particularly, female students who have stayed in the same major are more likely to graduate on time. Specifically, the odds ratio of postpone graduation (more than 4 years) decreases by 39.2% when female students stay in the same major.

Table 26 Logistic Regression on Time to Graduate

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	2.2113	0.4312	26.2930	<.0001
MajorRe	**1**	1	-0.4982	0.1506	10.9412	0.0009
LastEngrCumGPA		1	-1.2149	0.1492	66.2744	<.0001

Table 27 Time to Graduate: Odds Ratio Estimates

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
MajorRe 1 vs 0	0.608	0.452	0.816
LastEngrCumGPA	0.297	0.221	0.398

3.6 Brief Summary

Based on the analysis above, we summarized that WIE Orientation is generally effective to increase female students' retention in both major and college. But WIE Orientation is not related to performance indicators (GPA and Time to Graduate). On the other hand, the odds ratio of female students who postpone their graduation (more than 4 years) decreases 40% when they stay in the same majors. That is to say, WIE Orientation is likely to motivate female students to stay in the same major/college, which further decrease their likelihood to postpone graduation.

4. Retention Rate across Colleges

4.1 Research questions

This part delves into how College of Engineering compares to other colleges regarding retention rate of men, women, URM and non-URM students. Specifically, we aims to answer the following questions: Is College of Engineering different from other colleges regarding Retention Rate of men, women, URM and non-URM students? And how?

4.2 Data

The data sets include number of male, female, URM and non-URM students graduated from different colleges from 2002 to 2008. There are consistently eight colleges in total, which include Agricultural, Consumer & Environmental Sciences (ACES), Applied Life Sciences (ALS, now as Applied Health Sciences), Fine & Applied Arts (FAA), Liberal Arts & Science (LAS), Aviation, Business, Education and Engineering.

This part consists of seven variables: Male Retention Rate, Female Retention Rate, Under-Represented Minority (URM) Retention Rate, non-URM Retention Rate, Ratio of

Female Retention Rate to Male Retention Rate (Female Retention Rate/Male Retention Rate), Ratio of URM Retention Rate to non-URM Retention Rate (URM Retention Rate/non-URM Retention Rate) and College. While in calculating the URM Retention Rate and non-URM Retention Rate, we skip the unknown part for those who are unwilling to identify their races/ethnicity. For colleges, as Institute of Aviation has much fewer students, we exclude it from our analysis with seven colleges finally included.

4.3 Descriptive Analysis

1) Male Retention Rates of Different Colleges

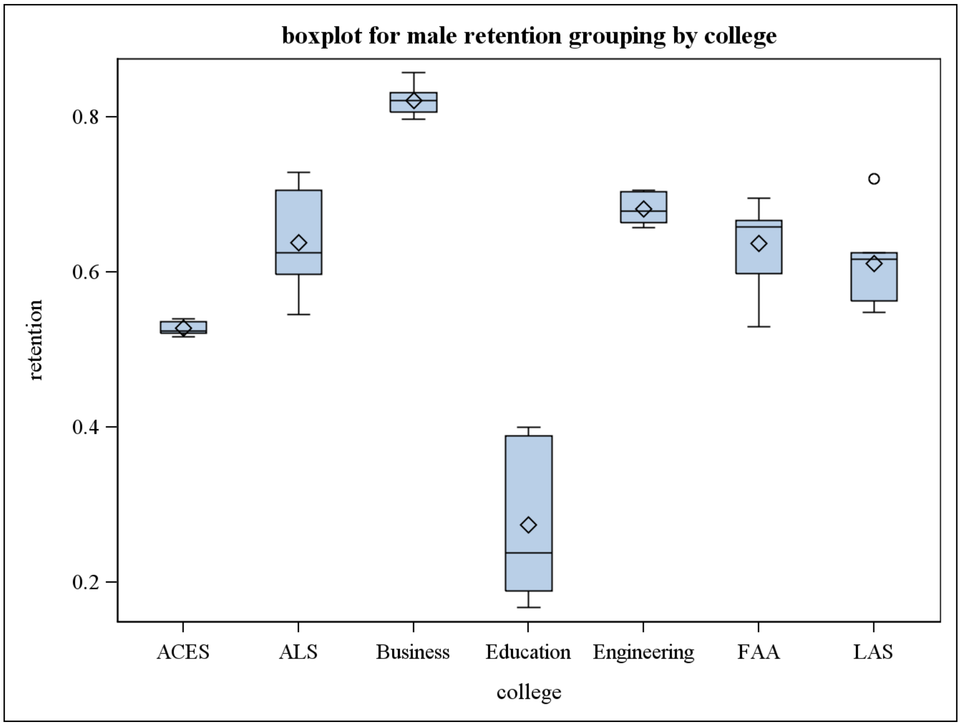

Figure 11 Boxplot: Male Retention Rate across Colleges

Figure 11 displays the boxplot of male retention rate for different colleges. We can see that College of Business has the highest male retention rate while ALS and College of Engineering have relatively higher male retention rate. Compared to ALS, College of Engineering has relatively smaller variance. And among all colleges, College of Education has smallest male retention rate.

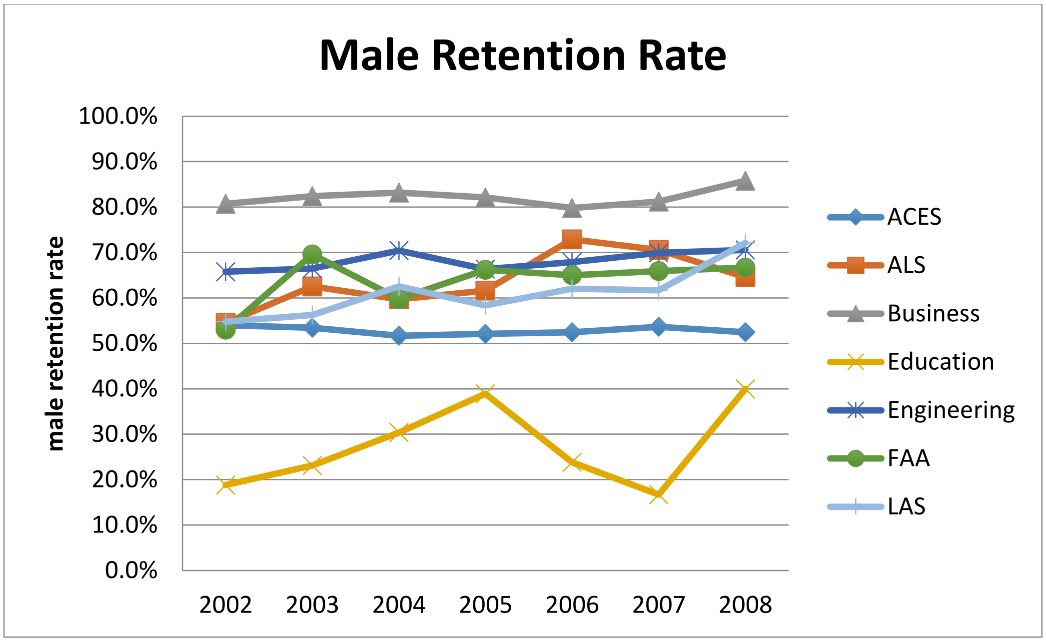

Figure 12 Trend of Male Retention Rate across Colleges

Figure 12 shows the trend of male retention rate for different colleges from 2002 to 2008. Obviously College of Business stays highest through all these years. College of Engineering performs better than other colleges in most years while ALS and FAA are also competitive. Through these years, male retention rate of College of Engineering stays stable while having small increase.

2) Female Retention Rates of Different Colleges

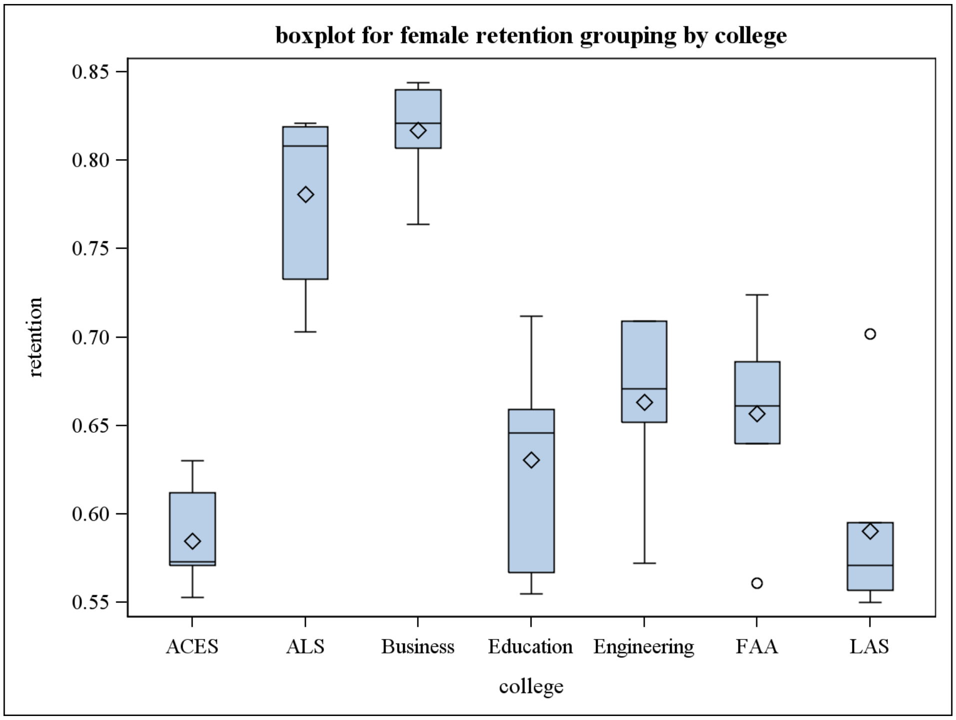

Figure 13 Boxplot: Female Retention Rate across Colleges

Figure 13 displays the boxplot for female retention rate of different colleges. College of Engineering performs better than most colleges except for College of Business and ALS. College of Business and ALS are far in advance than other colleges.

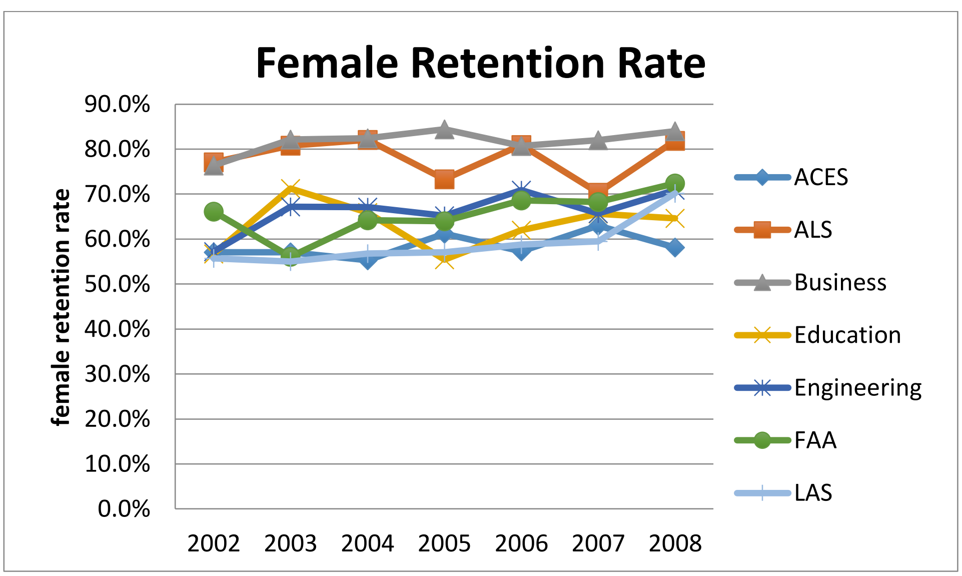

Figure 14 Trend of Female Retention Rate across Colleges

Figure 14 shows how female retention rate of different colleges changes over time. We can find similar results. College of Business takes the lead in female retention rate in all these years while ALS is almost the equal as College of Business except in year 2005 and 2007. For College of Engineering, it is relatively high among other colleges and has a slightly increasing trend.

3) Ratios of Female Retention Rate to Male Retention Rate of Different Colleges

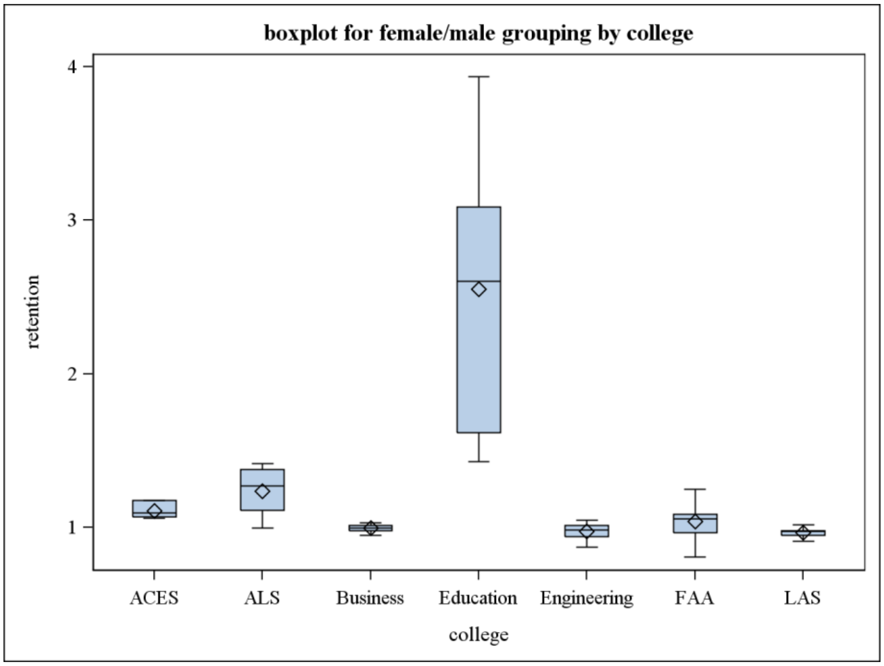

Figure 15 Boxplot: Female Retention Rate/Male Retention Rate

Figure 15 presents the boxplot for ratio of female retention rate to male retention rate. When the ratio is larger than one, it means female retention rate is larger than male retention rate. College of Education has much higher ratio than other colleges, which is far above one. The ratios of ACES and ALS are a little bit above one while ratios of other colleges stay around one. College of Engineering is not quite competitive in performance of ratio.

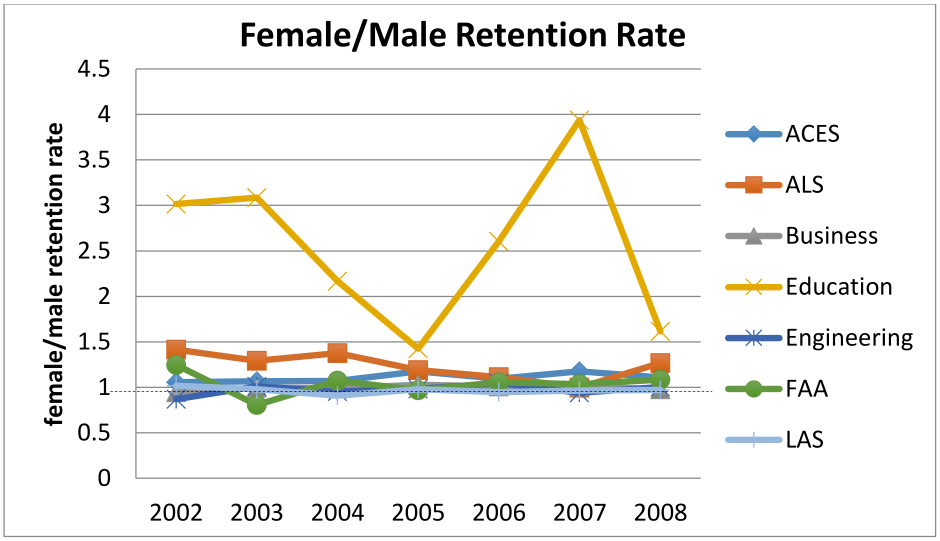

Figure 16 Trends of Female Retention Rate/Male Retention Rate

From Figure 16, we can see how ratio of female to male retention rate changes from 2002 to 2008. It is obvious that College of Education is highest in all these years. Also ALS is relatively higher compared to other colleges. Colleges except for Education and ALS are around one. College of Engineering slightly increases since 2003.

4) URM Retention Rates of Different Colleges

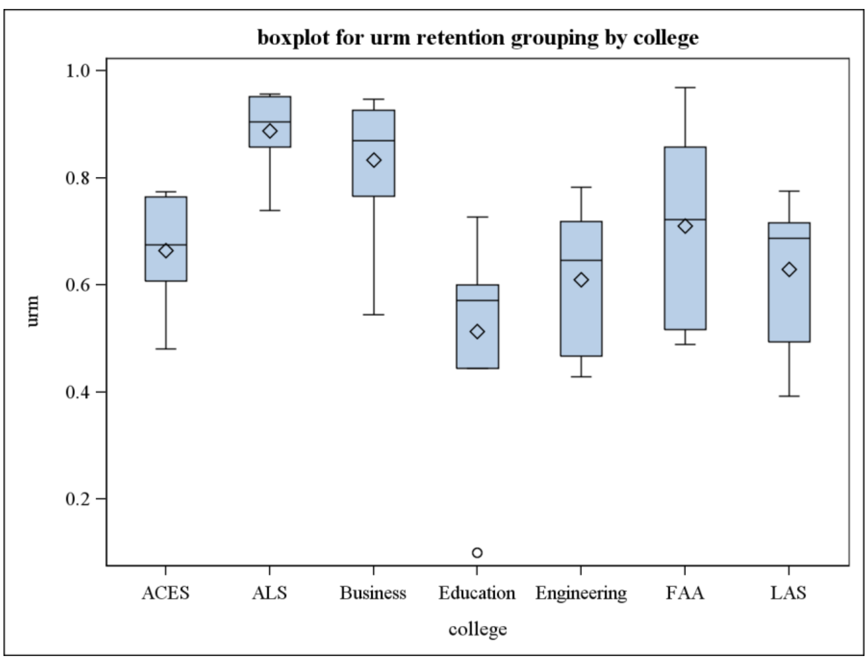

Figure 17 Boxplot: URM Retention Rate across Colleges

Figure 17 performs boxplot of URM retention rate for all colleges. ALS and College of Business are relatively higher than other colleges while College of Engineering is relatively lower among all colleges.

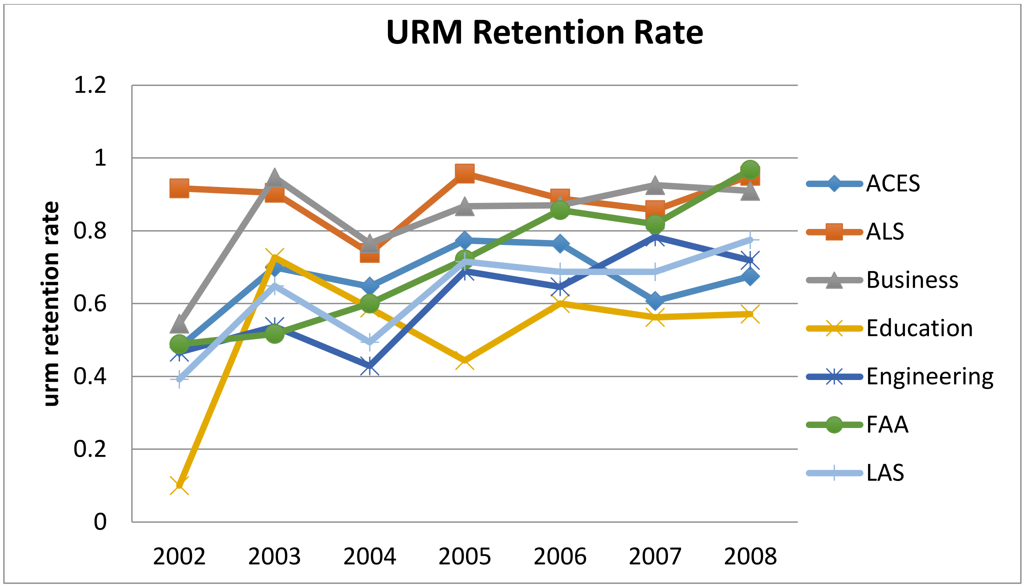

Figure 18 Trends of URM Retention Rate across Colleges

From Figure 18, we can get similar results. ALS and College of Business are relatively higher from 2002 and 2008. It needs to be pointed out that College of Engineering keeps increasing from 2002 to 2008 and in 2008 it is the highest among all these colleges.

5) Non-URM Retention Rates of Different Colleges

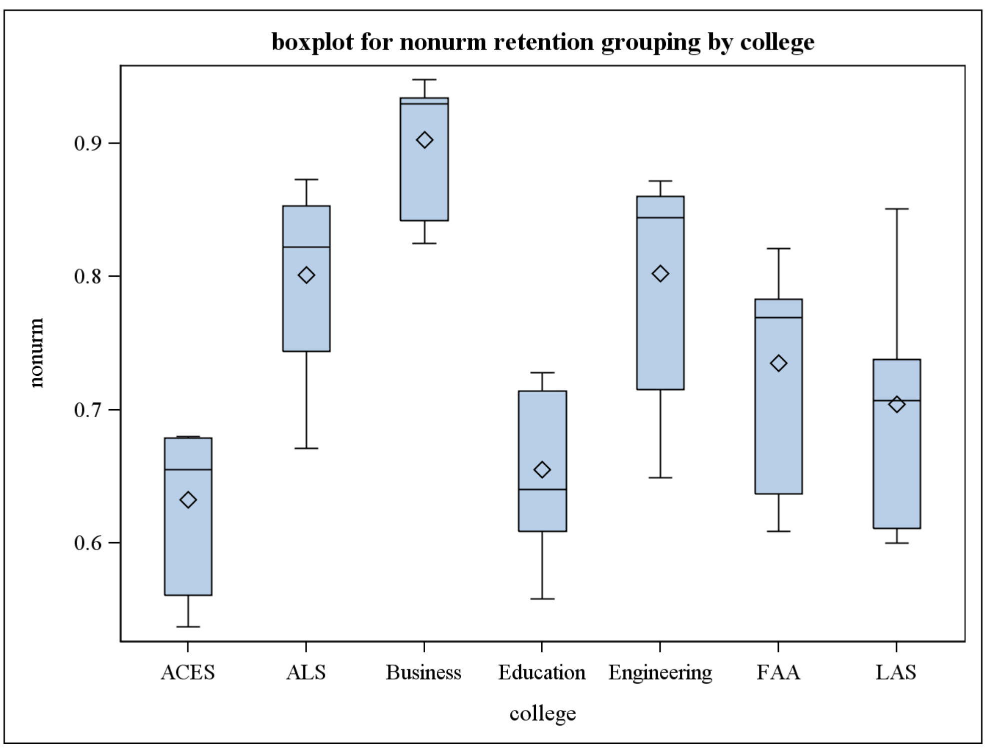

Figure 19 Boxplot: Non-URM Retention Rate across Colleges

Figure 19 shows boxplot for non-URM retention rate grouped by colleges. College of Business is the highest among all colleges. Both ALS and College of Engineering perform higher among the rest colleges.

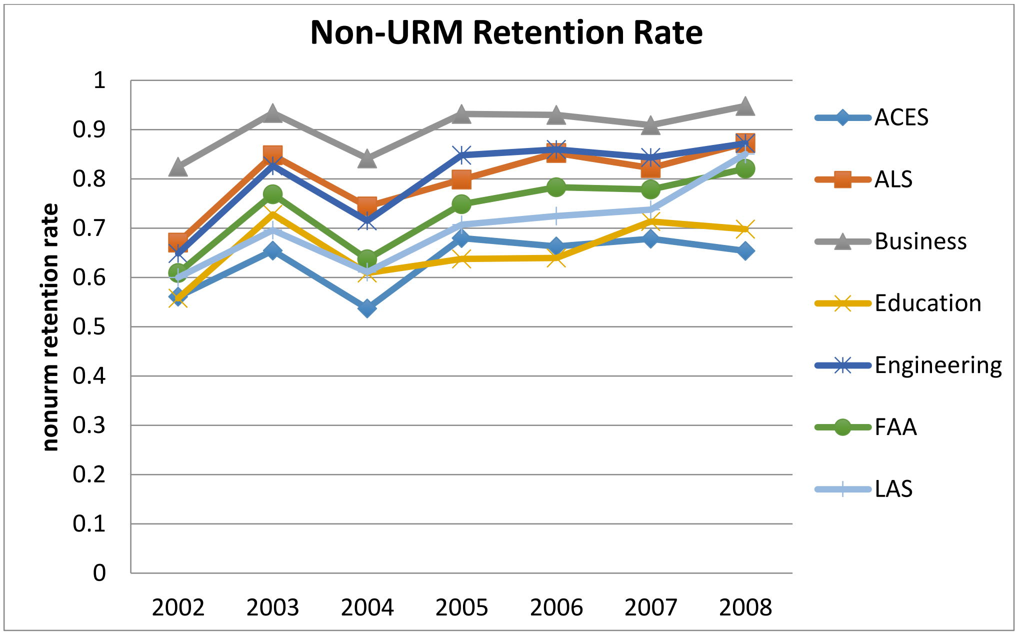

Figure 20 Trends of Non-URM Retention Rate across Colleges

Figure 20 presents the trend of non-URM retention rate for different colleges from 2002 to 2008. It seems that the trends of most colleges are parallel and increasing. College of Business appears to be the highest from 2002 to 2008. ALS and College of Engineering are above other colleges while keeping the trend of increasing.

6) Ratios of URM Retention Rate to non-URM Retention Rate of Different Colleges

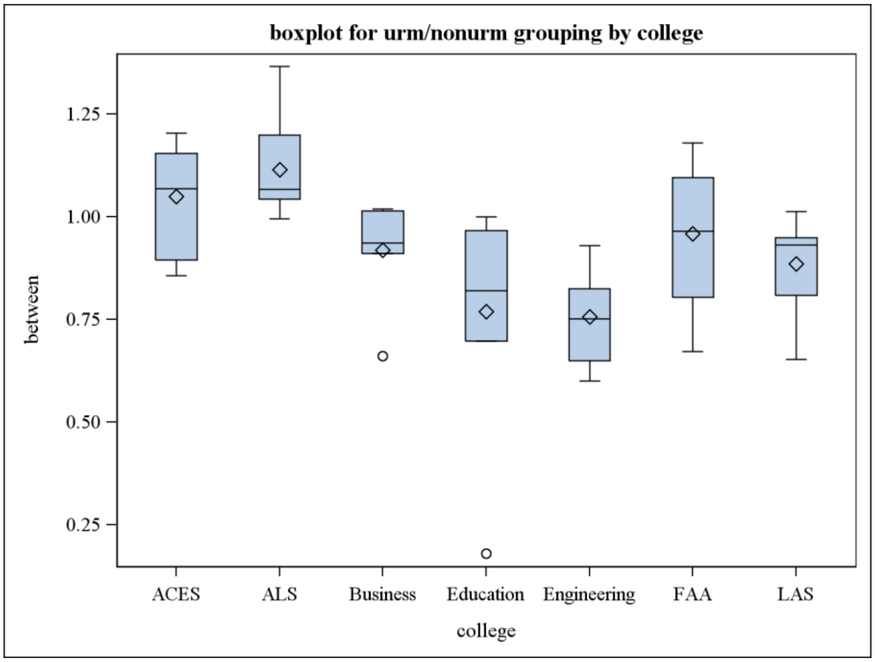

Figure 21 Boxplot: URM Retention Rate/ Non-URM Retention Rate across Colleges

From Figure 21, which shows boxplot for ratio of URM to non-URM retention rate of all colleges, we can see that the most colleges have this ratio around one. Ratio lower than one means that URM retention rate is less than non-URM retention rate. For ALS, URM retention rate is higher than non-URM retention rate with all values above one. For ACES, College of Business and FAA, their ratios are around one, which means in most cases URM retention rate and non-URM retention rate are quite similar. For College of Engineering, College of Education and LAS, their ratios are below one, which means that their URM retention rate is less than their non-URM retention rate. College of Engineering is lower in this ratio compared to other colleges.

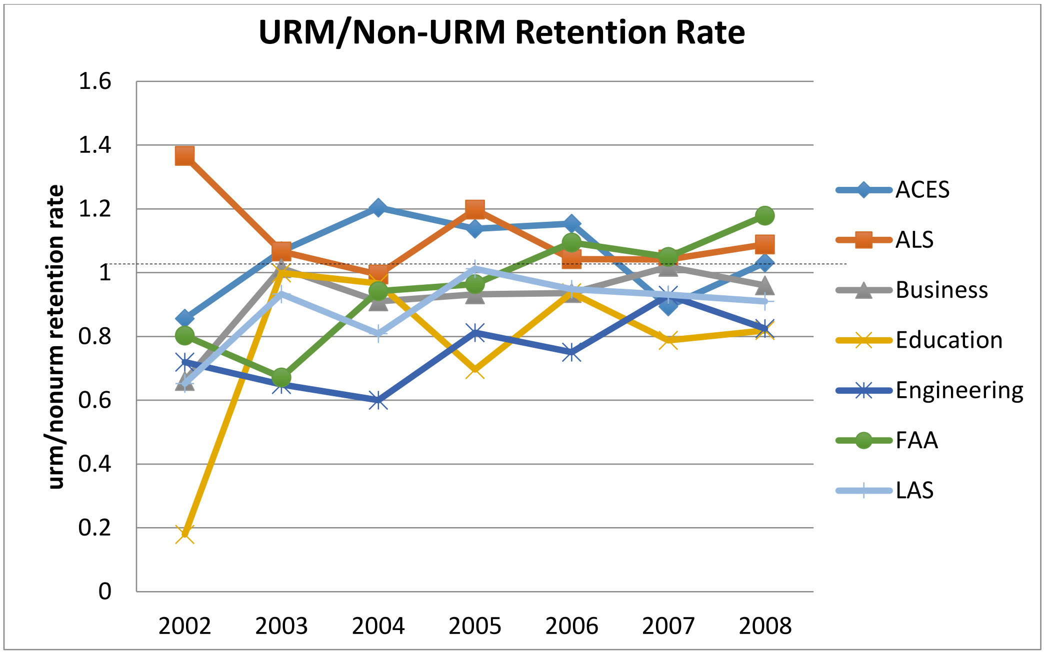

Figure 22 Trends of URM Retention Rate/ Non-URM Retention Rate across Colleges

Figure 22 presents how ratio of URM to non-URM retention rate changes with year. We can conclude similarly with what we get from Figure 3.11. The trends for different colleges vary a lot. ACES and ALS has ratio more than one in most years. College of Engineering is lower than other colleges, however, it keeps increasing these years and approaches one.

4.4 Analysis of Variance (ANOVA)

1) Effect of College on Male Retention Rate

Table 28 shows the result of testing whether there is significant difference in male retention rate among different colleges. As p-value is less than 0.0001, we should reject the null hypothesis that male retention rate is the same. That is, male retention rate is significantly different among colleges. Thus we march along to ask which pairs are different using Tukey's HSD test.

Table 28 ANOVA: Male Retention Rate vs. Colleges

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	1.19324334	0.19887389	70.02	<.0001
Error	42	0.11929714	0.00284041		
Corrected Total	48	1.31254049			

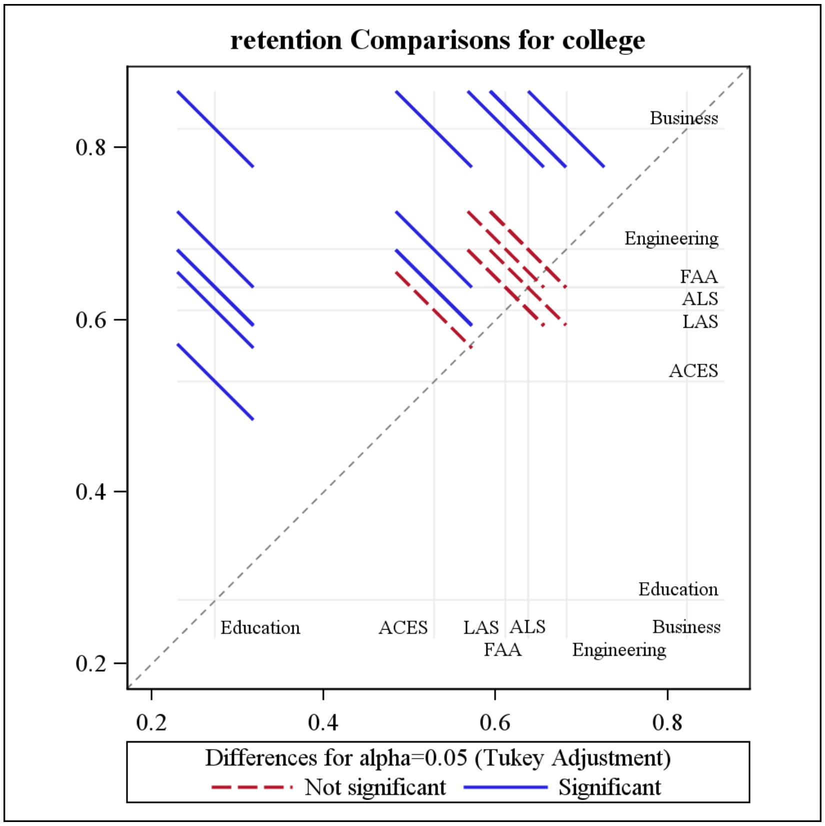

Figure 23 Pairwise Comparison: Male Retention Rate vs. Colleges

Figure 23 presents which pairs of colleges are significantly different from each other in male retention rate while Table 29 below shows the mean of male retention rate for different colleges. Focusing on College of Engineering, we figure out that it differs from ACES, College of Education and College of Business significantly in male retention rate. College of Engineering is significantly higher than ACES and College of Education while significantly lower than College of Business.

Table 29 Male Retention Rate: Mean Comparison across Colleges

college	retention LSMEAN	LSMEAN Number
ACES	0.52800000	1
ALS	0.63771428	2
Business	0.82185714	3
Education	0.27385714	4
Engineering	0.68214286	5
FAA	0.63742857	6
LAS	0.61128571	7

2) Effect of College on Female Retention Rate

From the ANOVA Table 30, we conclude that there is significant difference in female retention rate among all colleges (p-value < 0.0001). Then we turn to test which pairs are different through Tukey's HSD test.

Table 30 ANOVA: Female Retention Rate vs. Colleges

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	0.34435698	0.05739283	28.52	<.0001
Error	42	0.08452457	0.00201249		
Corrected Total	48	0.42888155			

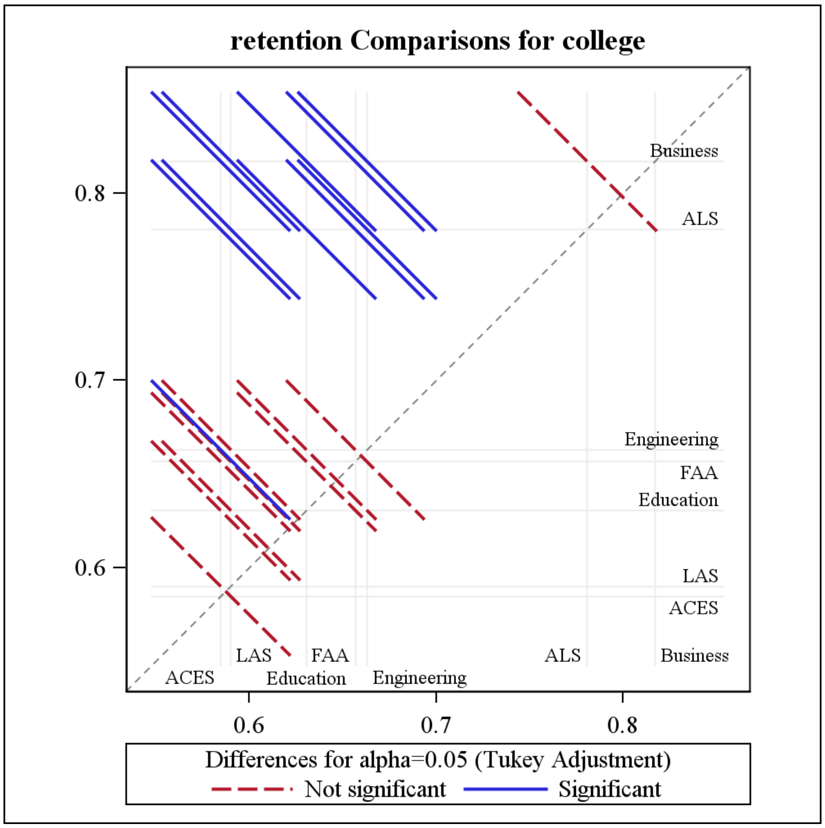

Figure 24 Pairwise Comparison: Female Retention Rate vs. Colleges

Figure 24 shows which pairs of colleges are different in female retention rate while Table 31 shows the means of different colleges. Combining the results from Figure 24 and Table 31, we find that College of Engineering is significantly higher than ACES and lower than ALS and College of Business regarding female retention rate.

Table 31 Female Retention Rate: Mean Comparison across Colleges

college	retention LSMEAN	LSMEAN Number
ACES	0.58457143	1
ALS	0.78071428	2
Business	0.81714286	3
Education	0.63071429	4
Engineering	0.66314286	5
FAA	0.65671428	6
LAS	0.59014286	7

3) Effect of College on Retention Rate Ratio

Table 32 presents the result testing whether colleges are different in ratio of female to male retention rate. As p-value is less than 0.0001, we conclude that significant difference exists for ratio of female to male retention rate across colleges.

Table 32 ANOVA: Retention Rate Ratio vs. Colleges

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	13.84567593	2.30761265	19.39	<.0001
Error	42	4.99778719	0.11899493		
Corrected Total	48	18.84346311			

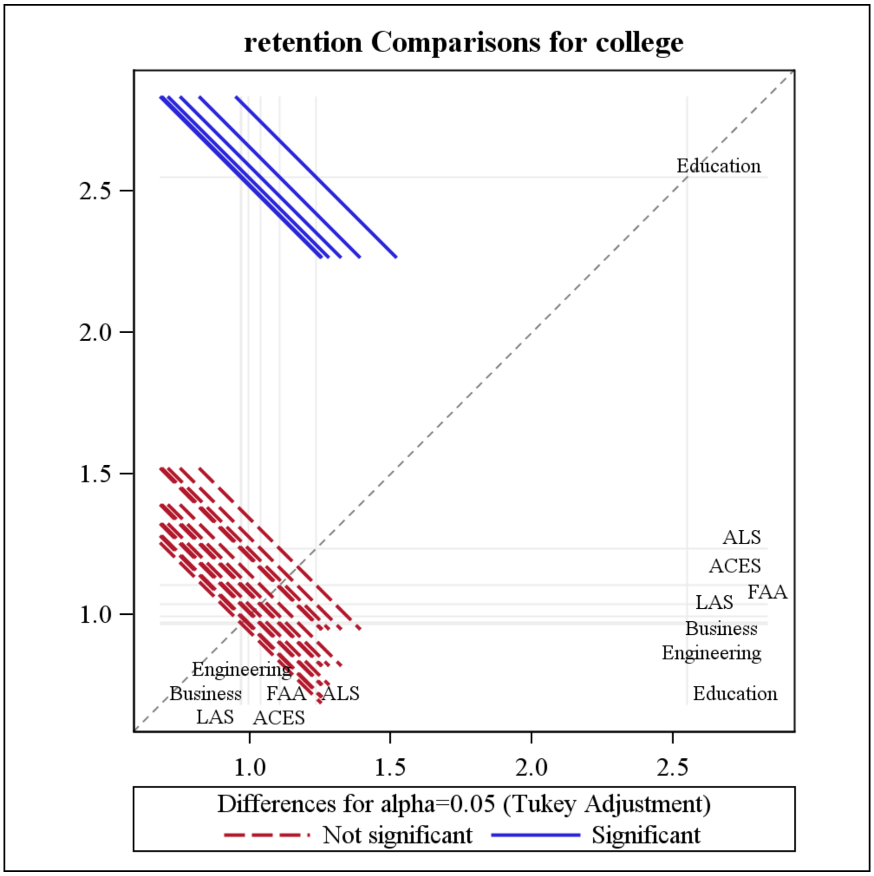

Figure 25 Pairwise Comparison: Retention Rate Ratio vs. Colleges

Applying Tukey's HSD method, we get Figure 25, which reveals which pairs of colleges are different in female to male retention ratio, and Table 33, which shows the mean of this ratio for all colleges. From Figure 25 and Table 33, we conclude that College of Engineering is significantly lower than College of Education and have no significant difference from other colleges.

Table 33 Retention Rate Ratio: Mean Comparison across Colleges

college	retention LSMEAN	LSMEAN Number
ACES	1.10670240	1
ALS	1.23527552	2
Business	0.99450644	3
Education	2.55056630	4
Engineering	0.97200260	5
FAA	1.03862720	6
LAS	0.96652966	7

4) Effect of College on URM Retention Rate

Table 34 presents that URM retention rates are significantly different among all colleges with p-value 0.0002, which is less than 0.05. Thus we keep to check which pairs are different and how.

Table 34 ANOVA: URM Retention Rate vs. Colleges

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	0.71455428	0.11909238	5.76	0.0002
Error	42	0.86838171	0.02067576		
Corrected Total	48	1.58293600			

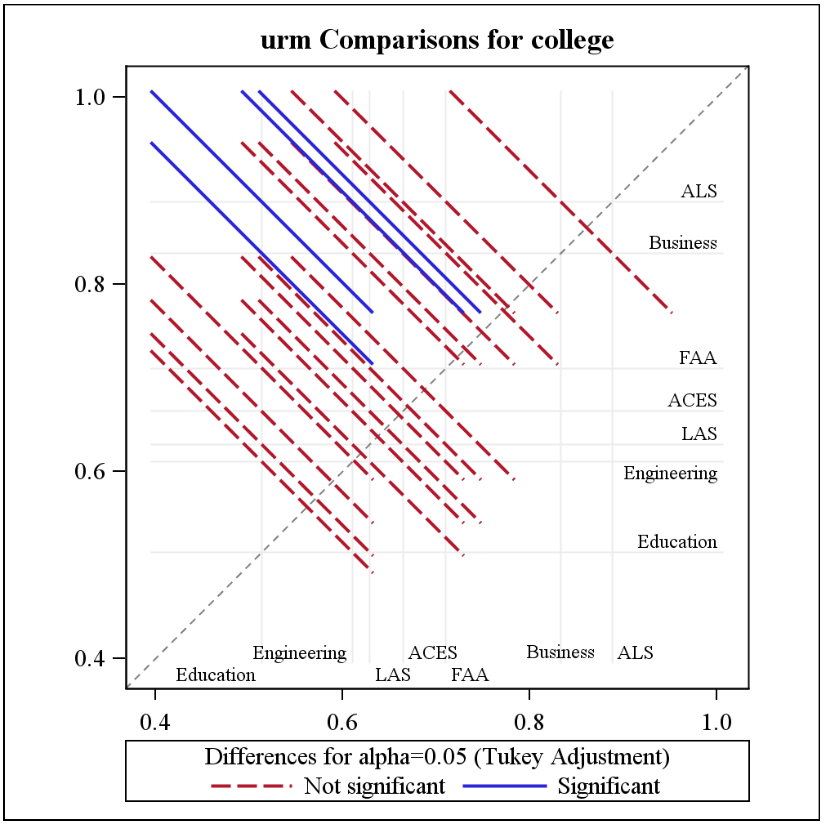

Figure 26 Pairwise Comparison: URM Retention Rate vs. Colleges

Figure 26 presents significantly different pairs of colleges on URM retention rate, showing that College of Engineering is significantly different from College of Business and ALS while have no significant difference with others. We can see that College of Engineering is significantly lower than ALS and College of Business. Though College of Engineering is not significantly different from most colleges, its mean of URM retention rate is relatively lower than most colleges.

Table 35 URM Retention Rate: Mean Comparison across Colleges

college	URM LSMEAN	LSMEAN Number
ACES	0.66400000	1
ALS	0.88785714	2
Business	0.83314286	3
Education	0.51328571	4
Engineering	0.61000000	5
FAA	0.71014286	6
LAS	0.62857143	7

5) Effect of College on Non-URM Retention Rate

Table 36 shows the result of testing whether colleges are different in non-URM retention rate. With p-value less than 0.0001, we conclude that non-URM retention rates are different across colleges. Then we take a look at which pairs are different and how they compare to each other.

Table 36 ANOVA: Non-URM Retention Rate vs. Colleges

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	0.37662310	0.06277052	12.36	<.0001
Error	42	0.21329229	0.00507839		
Corrected Total	48	0.58991539			

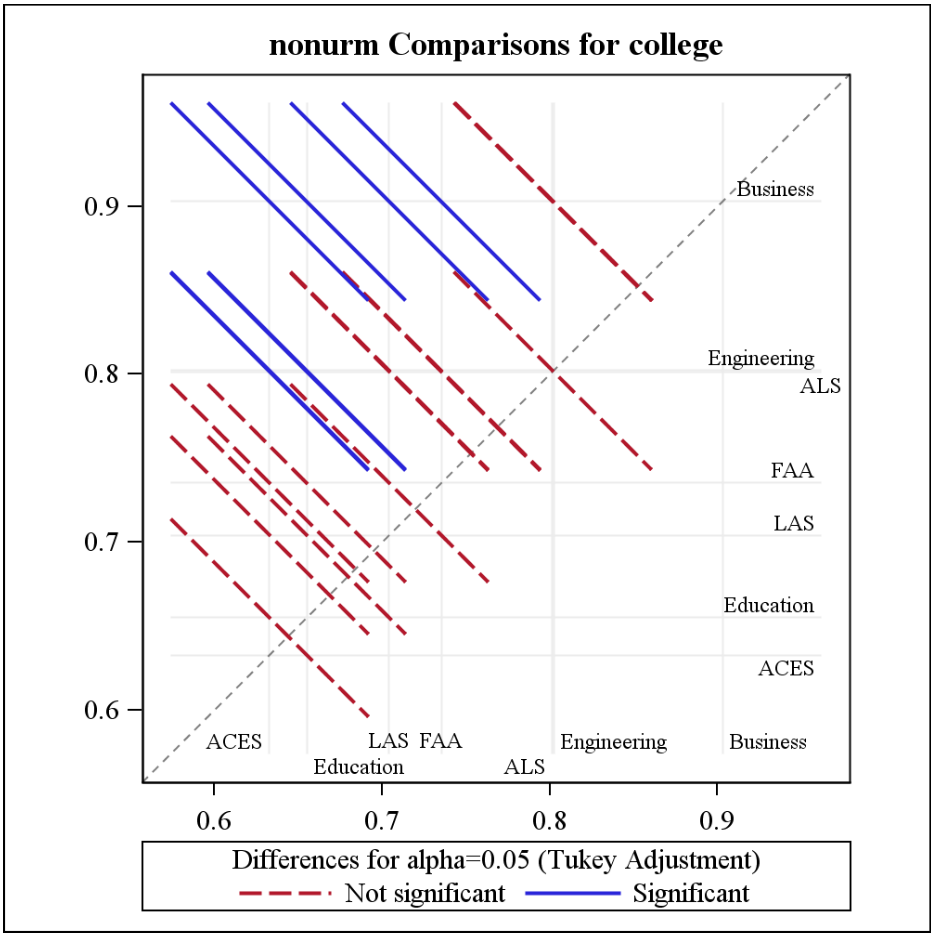

Figure 27 Pairwise Comparison: Non-URM Retention Rate vs. Colleges

Figure 27 presents significantly different pairs of colleges regarding non-URM retention rate while Table 37 presents their means of non-URM retention rate. As for College of Engineering, it is significantly higher than ACES and College of Education while not significantly different from other colleges.

Table 37 Non-URM Retention Rate: Mean Comparison across Colleges

college	nonurm LSMEAN	LSMEAN Number
ACES	0.63271428	1
ALS	0.80142857	2
Business	0.90285714	3
Education	0.65500000	4
Engineering	0.80214286	5
FAA	0.73528571	6
LAS	0.70400000	7

6) Effect of College on URM Retention Rate Ratio

Table 38 shows the result of testing whether ratios of URM to non-URM retention rate are significantly different across colleges. Given p-value is 0.0009 which is less than 0.05, we find that the ratio of URM to non-URM retention rate is significantly different among colleges. Next we are going to check how colleges differ from each other in this ratio.

Table 38 ANOVA: URM Retention Rate Ratio vs. Colleges

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	0.74880563	0.12480094	4.73	0.0009
Error	42	1.10845600	0.02639181		
Corrected Total	48	1.85726163			

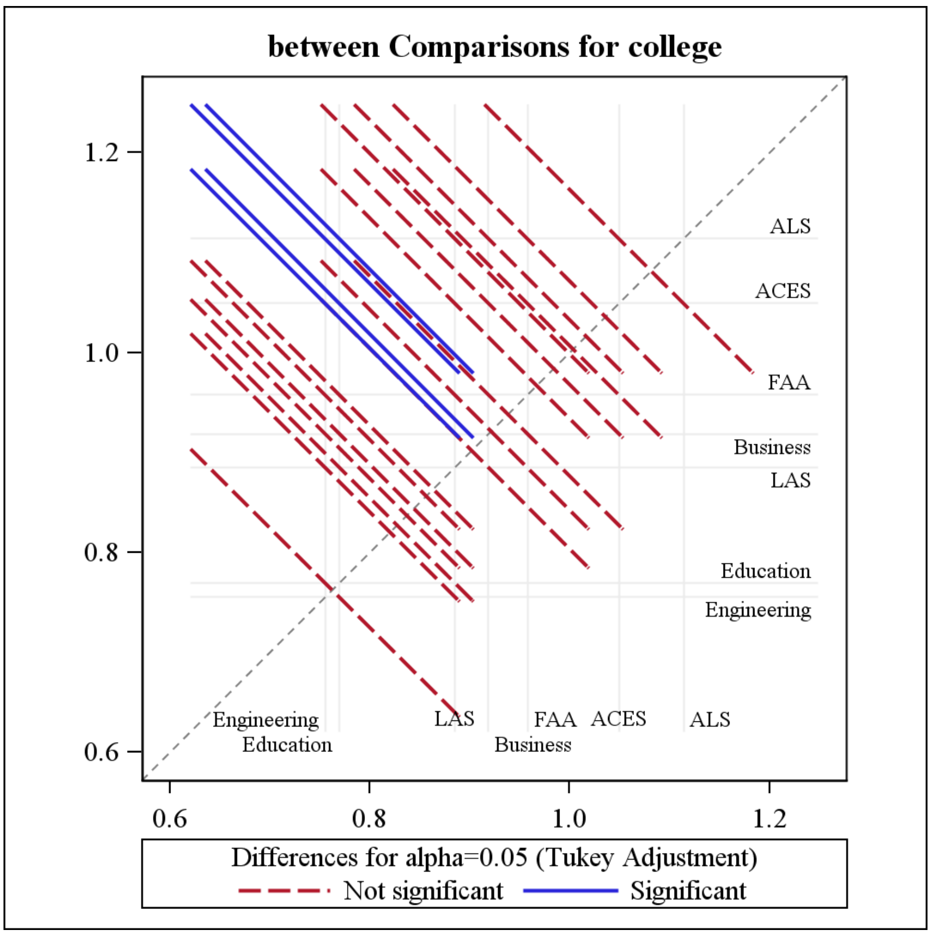

Figure 28 Pairwise Comparison: Non-URM Retention Rate Ratio vs. Colleges

Figure 28 presents which pairs are different from each other in ratio of URM to non-URM retention rate while Table 39 presents means of this ratio for all colleges. We can figure out that College of Engineering is significantly lower than ALS and ACES while having no significantly difference from other colleges. However, it should be pointed out that College of Engineering has lowest mean in this ratio.

Table 39 Non-URM Retention Rate Ratio: Mean Comparison across Colleges

college	between LSMEAN	LSMEAN Number
ACES	1.04942857	1
ALS	1.11414286	2
Business	0.91871428	3
Education	0.76928571	4
Engineering	0.75514286	5
FAA	0.95785714	6
LAS	0.88514286	7

4.5 Brief Summary

To sum up, we find that College of Engineering performs well in all these retention rates and two ratios.

For retention rates and ratio regarding sex, College of Engineering is better than most colleges and keeps the trend of increasing from 2002 to 2008. However, College of Business performs better than College of Engineering in both male retention rate and female retention rate. Besides, ALS is better than College of Engineering in female retention rate while College of Education is better than College of Engineering in ratio of female to male retention rate.

For retention rates and ratio regarding under-represented minority, College of Engineering have different performances but have increasing trends in all cases. It is quite competitive in non-URM retention rate. However, for URM retention rate, College of Engineering is significantly lower than ALS and College of Business but has relatively lower mean than other colleges. Similar situation happens in ratio of URM to non-URM retention rate: significantly lower than ACES and ALS, College of Engineering has lower means compared to other colleges.

5. Performance and Enrollment

5.1 Research Questions

This part will focus on exploring two questions about the performance and enrollment of students in college of engineering. We have two datasets for analysis. The first dataset will be used to check whether there are significant differences regarding GPA when comparing female and male students as well as URM and Non-URM students for each major in college of engineering. The second dataset will be analyzed to see if GPA is related to the admission criteria. And if the relation exists, we want to check how the relationships vary between female and male students.

5.2 Data

The first dataset consists of the GPA statistics for current and graduated students. We have 6457 current observations and 5946 graduated observations. For each observation, we will use the information of major, year of entry or graduate, gender, race and GPA to conduct analysis.

To answer the first question, we analyze the difference between female and males students and between URM and Non-URM students for current students and graduated students. For current students, we conduct further analysis to see the differences for different education levels. The method we use here is mainly ANOVA analysis to check if GPA correlates with gender or URM.

5.3 ANOVA: Gender and Race in Each Major

Table 40 shows the ANOVA results for graduated students categorized by each major. If the difference is 0, it means the performance difference between female and male students or between URM and Non-URM students is not significant. If the difference is positive for difference, it means female students have better performance than male students on average and the mean of the difference is the absolute value of difference. If the difference is negative for difference, it means male students have better performance than female students on average and the mean of the difference is the absolute value of difference. If the difference is positive for URM, it means URM students have better performance than Non-URM students on average and the mean of the difference is the absolute value of difference. If the difference is negative for URM, it means Non-URM students have better performance than URM students on average and the mean of the difference is the absolute value of difference. The interpretations for Table 41-46 are the same as Table 40.

Table 40 shows that for graduated students, only two majors, civil engineering and electrical engineering have performance differences between female and male students. And female students have better performance than male students for both majors. About half the majors have differences between URM and Non-URM students. And among all those majors which have differences regarding URM, Non-URM students have better performance than URM students. Overall, for graduated students, there is no performance difference between female and male students but Non-URM students have better performance than URM students.

Table 40 ANOVA: Gender and Race for Graduates in Each Major

	Female	Difference	URM	Difference
Total	17.69%	0	6.27%	-0.272
Agri Eng	16.67%	0	4.17%	0
Civil Eng	27.52%	0.071	7.39%	-0.221
Comp Eng	7.76%	0	6.39%	-0.419
Comp Sci	10.84%	0	4.02%	-0.271
Eletr Eng	13.48%	0.092	6.69%	-0.275
Engr Mecha	12.07%	0	6.03%	0
Engr Phy	9.49%	0	7.30%	-0.324
General Eng	21.60%	0	6.17%	0
Industr Eng	25.77%	0	7.36%	-0.35
M. Sci Eng	24.75%	0	3.96%	0
Mecha Eng	9.77%	0	5.83%	-0.238
Nuclear Eng	10.71%	0	7.14%	0
Chem Eng	30.88%	0	7.14%	-0.316
Bio Eng	41.91%	0	6.62%	0
Aero Eng	14.49%	0	8.48%	-0.229
Ag Bio Eng	25.20%	0	6.30%	0
N.P.R Eng	14.58%	0	7.64%	-0.531

Table 41 shows that for current students, only three majors, civil engineering, mechanical engineering and nuclear, plasma and radiological engineering have performance differences between female and male students. And female students have better performance than male students for civil engineering. For the other two majors, male students have better performance than female students. About half the majors have differences between URM and Non-URM students. And among all those majors which have differences regarding URM, Non-URM students have better performance than URM students.

Table 41 ANOVA: Gender and Race for Current Students in Each Major

	Female	Difference	URM	Difference
Civil Eng	25.82%	0.16373	11.88%	0
Comp Eng	9.74%	0	6.67%	-0.291
Comp Sci	16.59%	0	4.09%	-0.378
Eletr Eng	12.96%	0	6.97%	0
Engr Mecha	9.78%	0	11.96%	-0.314
Engr Phy	8.16%	0	7.65%	-0.310
General Eng	24.76%	0	6.80%	0
Industr Eng	22.37%	0	4.11%	0
M. Sci Eng	29.18%	0	5.78%	-0.266
Mecha Eng	16.42%	-0.101	9.69%	-0.261
Chem Eng	28.79%	0	10.05%	-0.266
Bio Eng	45.69%	0	6.47%	0
Aero Eng	8.59%	0	12.37%	0
Ag Bio Eng	28.46%	0	12.20%	0
N.P.R Eng	18.02%	-0.304	11.71%	0
Undeclared	37.93%	0	10.34%	0

Table 42 shows that, for current students, there is no performance difference between female and male students but Non-URM students have better performance than URM students. For the four education levels, only junior level has performance difference between female and male students. Non-URM students have better performance than URM students for the four education levels.

Table 42 ANOVA for Current Students by Education Level

	Female	Difference	URM	Difference
Freshman	22.39%	0	8.42%	-0.250
Sophomore	18.21%	0	8.94%	-0.251
Junior	17.39%	0.067	7.93%	-0.129
Senior	18.27%	0	6.51%	-0.166
Total	19.16%	0	8.04%	-0.204

Table 4.4 shows that for current freshmen, there is no major which has performance between female and students. The results may come from the reason that the students are only in their first year and differences are not so significant. Only three majors have differences between URM and Non-URM students. They are computer engineering, computer science and mechanical engineering. For these three majors, Non-URM students have better performance than URM students. It should be noted that for freshman, the engineering mechanics major has no female students and only students in total.

Table 43 ANOVA for Freshman by Each Major

	Female	Difference	URM	Difference
Civil Eng	27.43%	0	11.50%	0
Comp Eng	11.60%	0	6.08%	-0.513
Comp Sci	27.43%	0	4.87%	-0.780
Eletr Eng	14.85%	0	5.68%	0
Engr Mecha	0%	0	29.41%(17ttl)	0
Engr Phy	9.68%	0	6.45%	0
General Eng	15.00%	0	10.00%	0
Industr Eng	33.33%	0	6.35%	0
M. Sci Eng	30.00%	0	6.25%	0
Mecha Eng	20.10%	0	10.55%	-0.352
Chem Eng	26.35%	0	9.46%	0
Bio Eng	43.84%	0	4.11%	0
Aero Eng	11.28%	0	15.04%	0
Ag Bio Eng	24.39%	0	14.63%	0
N.P.R Eng	24.00%	0	4.00%	0
Undeclared	37.93%	0	10.34%	0

Table 44 shows that for current sophomores, only three majors, civil engineering, computer science and mechanical engineering have performance differences between female and male students. And female students have better performance than male students for civil engineering. For the other two majors, male students have better performance than female students. About half the majors have differences between URM and Non-URM students. And among all those majors which have differences regarding URM, Non-URM students have better performance than URM students.

Table 44 ANOVA for Sophomore by Each Major

	Female	Difference	URM	Difference
Civil Eng	25.86%	0.200	13.22%	0
Comp Eng	9.06%	0	7.48%	-0.325
Comp Sci	19.28%	-0.216	3.61%	-0.447
Eletr Eng	12.55%	0	9.52%	0
Engr Mecha	13.51%	0	10.81%	0
Engr Phy	4.35%	0	6.52%	-0.785
General Eng	15.09%	0	7.55%	-0.590
Industr Eng	26.53%	0	0%	0
M. Sci Eng	28.92%	0	6.02%	0
Mecha Eng	16.84%	-0.191	12.24%	-0.287
Chem Eng	28.02%	0	12.64%	-0.346
Bio Eng	33.96%	0	3.77%	0
Aero Eng	7.02%	0	12.28%	0
Ag Bio Eng	44.00%	0	22.00%	0
N.P.R Eng	18.18%	0	12.12%	0

Table 45 shows that for current juniors, only one major, civil engineering, has performance differences between female and male students. Female students have better performance than male students. And only two majors have differences between URM and Non-URM students. The two majors are chemical engineering and nuclear, plasma and radiological engineering. Non-URM students have better performance than URM students for these two majors.

Table 45 ANOVA for Junior by Each Major

	Female	Difference	URM	Difference
Civil Eng	25.00%	0.193	11.49%	0
Comp Eng	11.71%	0	5.85%	0
Comp Sci	8.53%	0	3.88%	0
Eletr Eng	11.79%	0	6.67%	0
Engr Mecha	9.52%	0	9.52%	0
Engr Phy	10.64%	0	10.64%	0
General Eng	37.50%	0	5.36%	0
Industr Eng	17.86%	0	3.57%	0
M. Sci Eng	30.00%	0	7.50%	0
Mecha Eng	12.30%	0	7.49%	0
Chem Eng	29.41%	0	10.08%	-0.264
Bio Eng	53.70%	0	12.96%	0
Aero Eng	8.00%	0	12.00%	0
Ag Bio Eng	21.88%	0	15.63%	0
N.P.R Eng	12.90%	0	22.58%	-0.433

Table 4.7 shows that for current seniors, only one major, material sciences engineering, has performance differences between URM and Non-URM students. Non-URM students have better performance than URM students for material sciences engineering seniors.

Table 46 ANOVA for Senior by Each Major

	Female	Difference	URM	Difference
Civil Eng	25.34%	0	10.96%	0
Comp Eng	5.71%	0	7.14%	0
Comp Sci	11.28%	0	4.10%	0
Eletr Eng	12.27%	0	5.52%	0
Engr Mecha	11.76%	0	0%	0
Engr Phy	7.32%	0	7.32%	0
General Eng	28.07%	0	5.26%	0
Industr Eng	9.80%	0	5.88%	0
M. Sci Eng	27.91%	0	3.49%	-0.674
Mecha Eng	16.15%	0	8.07%	0
Chem Eng	31.88%	0	7.25%	0
Bio Eng	51.92%	0	5.77%	0
Aero Eng	6.76%	0	8.11%	0
Ag Bio Eng	28.00%	0	4.00%	0
N.P.R Eng	18.18%	0	4.55%	0

5.4 Enrollment Data

The second dataset consists of the GPA statistics and enrollment statistics for current and graduated students. It has 18298 observations in total. For each observation, we will use the information of GPA, gender, enrollment index and entry term to conduct analysis. We have three kinds of enrollment index. They are campussi, which is used from June 2003 to January 2004, campuspgpa, which is used from June 2004 to January 2012, and Academic Index, which is used from May 2012 till now. It should be noted that for each term of enrollment, the calculation equation for enrollment index is different from term to term. Thus the analysis can be only conducted within each enrollment term.

5.5 Enrollment Data: Descriptive Analysis

Table 47 shows the enrollment terms and number of students enrolled in that particular term. We can see that for each enrollment year, the August enrollment terms will have most of the students. Thus we mainly focus on these August enrollment terms and the following analysis will be conducted based on these terms.

Table 47 Enrollment Terms and Number of Students

Term	Number of Students	Term	Number of Students	Term	Number of Students
06/2003	14	05/2007	28	05/2011	1
08/2003	1232	08/2007	1382	08/2011	1662
01/2004	3	01/2008	1	01/2012	6
06/2004	9	05/2008	49	05/2012	24
08/2004	1301	08/2008	1430	08/2012	1648
05/2005	9	01/2009	3	01/2013	1
08/2005	1318	05/2009	31	05/2013	59
01/2006	3	08/2009	1445	08/2013	1793
05/2006	23	01/2010	2	01/2014	1
08/2006	1284	05/2010	12	05/2014	59
01/2007	3	08/2010	1666	08/2014	1702

5.6 Regression models: Enrollment and GPA

We will use three regression models to check if the enrollment index correlates with GPA performance. The first model is the simple linear model treating enrollment index as predictor. The second model treats the gender as a predictor to check if there is performance difference between female and male students. The third model combines the previous two predictors together and add an interaction term of enrollment index and gender, to test whether the index effects on GPA performance are difference between female and male students. Index is the three kinds of enrollment index. Gender is an indicator of male and female. EngineeringGPA is the GPA performance of the students.

- Model 1 $\boldsymbol{EngineeringGPA} \sim \boldsymbol{Index}$
- Model 2 $\boldsymbol{EngineeringGPA} \sim \boldsymbol{Gender}$
- Model 3 $\boldsymbol{EngineeringGPA} \sim \boldsymbol{Index}+\boldsymbol{Gender}+\boldsymbol{Index}*\boldsymbol{Gender}$

Table 48 shows the regression results of the three models above. If a cell has no value, it means the effect of that variable is not significant. For the column named Female, if the value in the cell is positive, it means female students have better GPA performance than male students. If the value in the cell is negative, it means male students have better GPA performance than female students.

From the index column of model 1 and model 3, we can say that enrollment index has a positive effect on the GPA performances, which means if a student has a higher enrollment index, the GPA performance of the student will be higher. For detailed interpretation of index effect in model 3, we can say that one unit increase in enrollment index will result in 1.56 unit increase in GPA for students entered in August 2003 when holding other factors constant.

For the gender effect, we can see the female column of model 2 and model 3, some of the terms have significant differences between female and male students. There is no consistent trend of the difference between female and male students. In model 3, the female students have poorer GPA performance than male students in term August 2011 and better GPA performance than male students in term August 2012.

For the interaction effect between enrollment index and gender in model 3, only two terms have significant results. For August 2011, the interaction effect in positive, which means female students have a higher enrollment index effect on GPA performance than male students when holding other factors constant. For August 2012, the interaction effect is negative, which means female students have a lower enrollment index effect on GPA performance than male students when holding other factors constant.

Thus based on the results from Table 48, we can say the enrollment index is positively related with the later GPA performance, which means if a student has a higher enrollment index, the GPA performance of the student tends to be higher. Only in some years there exist significant GPA differences between female and male students. And there is no difference of the enrollment index effect on GPA performance between female and male students.

Table 48 Regression Results for Enrollment

Term	**Model 1**	**Model 2**	**Model 3**		
	Index	Female	Index	Female	Index*Female
08/2003	1.617	0.1435	1.56	/	/
08/2004	0.667	/	0.71	/	/
08/2005	1.352	0.156	1.306	/	/
08/2006	1.52	/	1.49	/	/
08/2007	1.21	/	1.20	/	/
08/2008	1.44	0.117	1.36	/	/
08/2009	1.24	0.136	1.25	/	/
08/2010	1.24	0.137	1.23	/	/
08/2011	1.3	0.099	1.197	-2.22	0.69
08/2012	0.026	0.119	0.029	1.00	-0.01
08/2013	0.049	/	0.048	/	/
08/2014	0.047	/	0.049	/	/

5.7 Further Exploration on Enrollment

For the academic index the college uses, we want to conduct further analysis to check if the college can lower the index for enrollment or not. We will use Academic Index as the enrollment index, which is used from May 2012 till now. It should be noted that for each term of enrollment, the calculation equation for enrollment index is different from term to term. Thus the analysis can be only conducted within each enrollment term. We will analyze the term August 2012, August 2013 and August 2014. In order to see whether the college can lower the academic index to admit students, we will divide the data into several subsets using 5%, 15%, 25%, 50%, 75% and 100% percentiles cutoff. Table 49 shows the thresholds for each term.

Table 49 Percentile Threshold

Term	mean	Min	max	5%	15%	25%	50%	75%	100%
08/2012	87.25	63	107	77	81	83	88	91	107
08/2013	88.04	62	107	78	82	84	88	92	107
08/2014	90.45	69	108	82	85	87	91	94	108

Figure 29 and 30 show the summary statistics for academic index and engineering GPA for term August 2012 divided into four subsets using 25%, 50% and 75% cutoff values. We have the similar summary statistics plot for the term August 2013 and term August 2014.

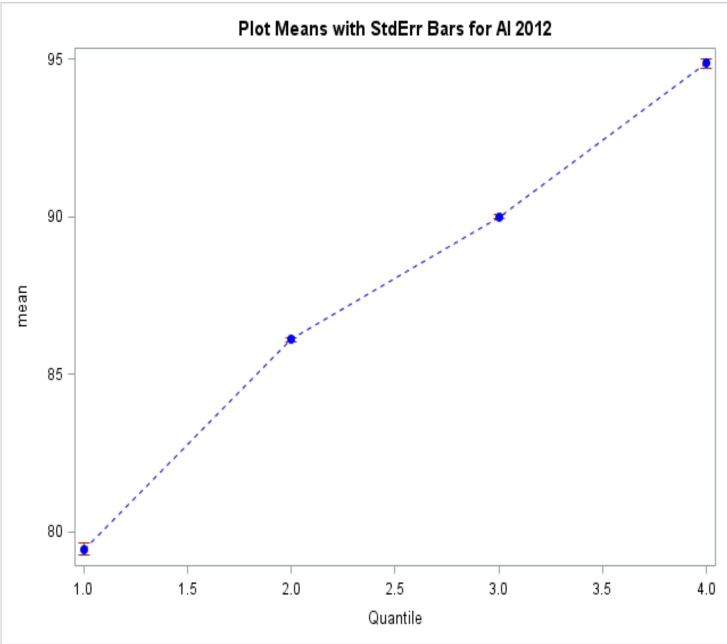

Figure 29 Academic Index for 2012

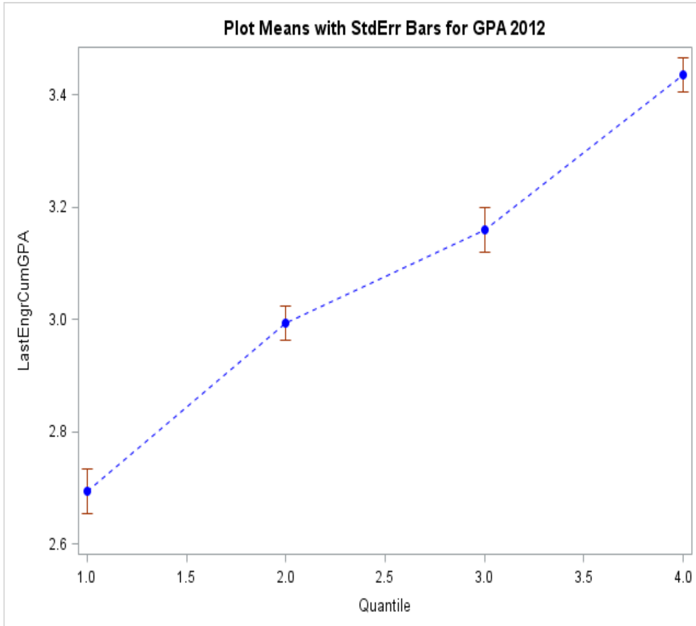

Figure 30 Engineering GPA Performance for 2012

Figure 31 shows the four regression results and related fitted lines for the four groups using 25%, 50% and 75% cutoff values. Only the regression for third subset (50%-75%), which is in green color, is not significant with a p value of 0.8644. For the other three groups, academic index have significant positive effects on the GPA performance. And we can see that the effects of academic index for these three groups are different since the slope of the fitted lines are different.

Figure 32 shows the four regression results and related fitted lines for the four groups using 5%, 15% and 25% cutoff values. Only the regression for fourth subset (25%-100%), which is in grey color, is significant. The other three regressions are not significant. Thus we can say the academic index is not a great predictor for students in the group of 0-5%, 5%-15% and 15%-25% quantiles in academic index.

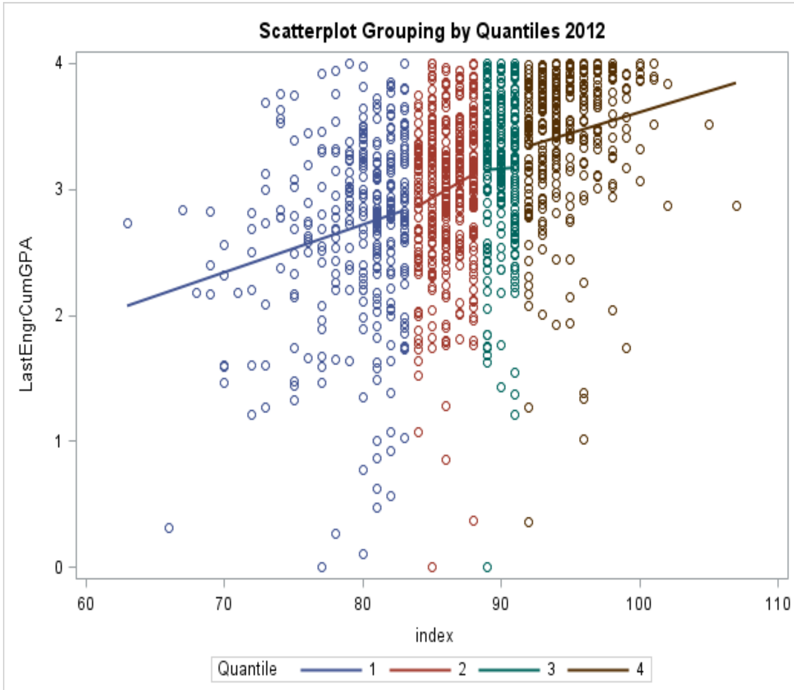

Figure 31 Academic Index vs. GPA Performance for 2012

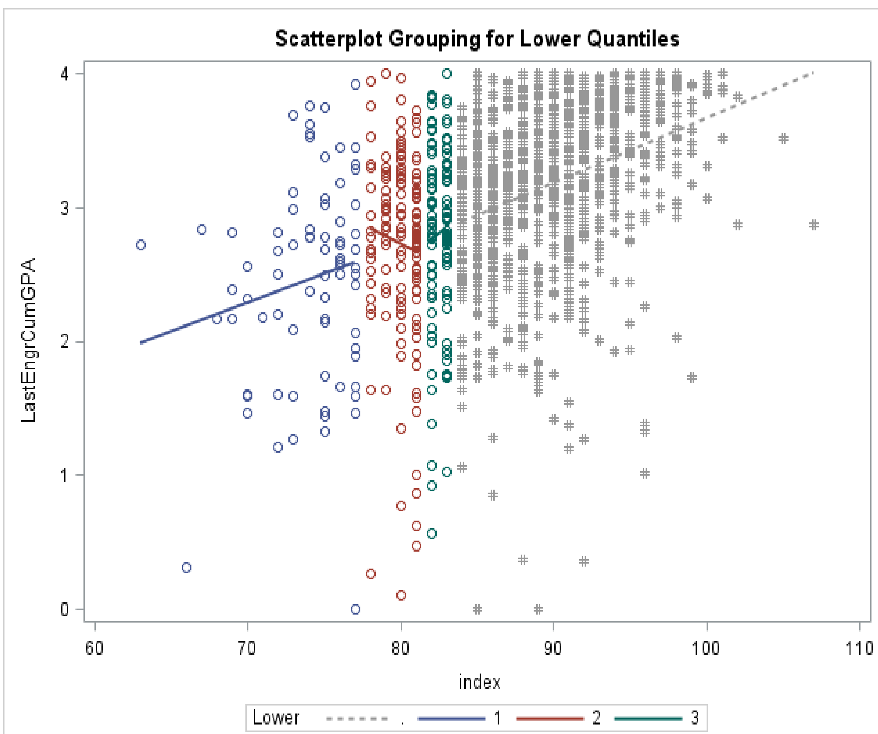

Figure 32 Academic Index vs. GPA Performance for 2012

Figure 33 shows the four regression results and related fitted lines for the four groups using 25%, 50% and 75% cutoff values. Only the regression for second subset (25%-50%), which is in red color, is not significant with a p value of 0.8282. For the other three groups, academic index have significant positive effects on the GPA performance. And we can see that the effects of academic index for these three groups are different since the slope of the fitted lines are different.

Figure 34 shows the four regression results and related fitted lines for the four groups using 5%, 15% and 25% cutoff values. Only the regression for fourth subset (25%-100%), which is in grey color, is significant. The other three regressions are not significant. Thus we can say the academic index is not a great predictor for students in the group of 0-5%, 5%-15% and 15%-25% quantiles in academic index.

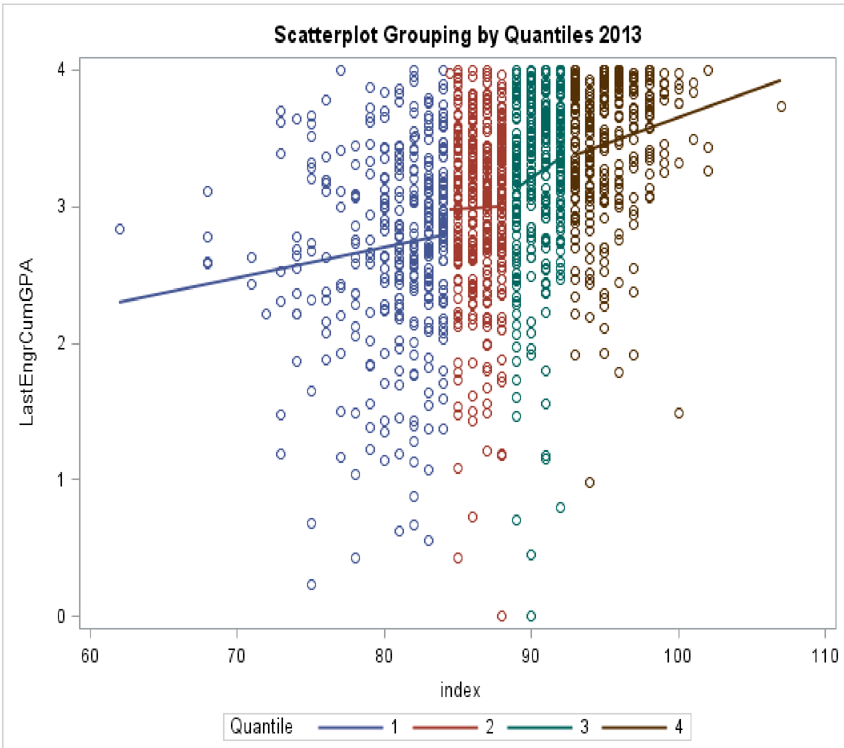

Figure 33 Academic Index vs. GPA Performance for 2013

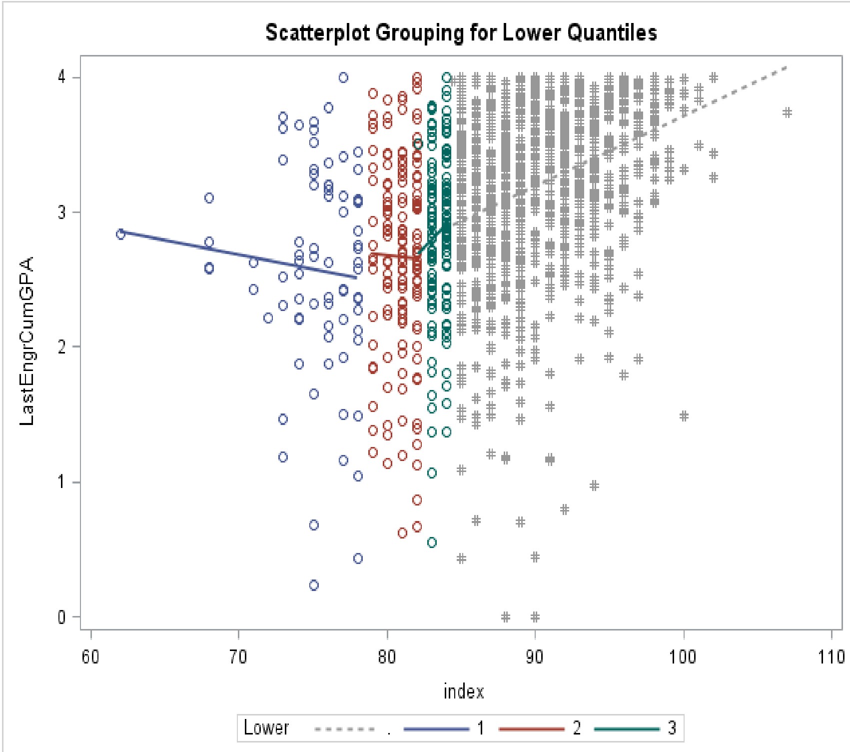

Figure 34 Academic Index vs. GPA Performance for 2013

Figure 35 shows the four regression results and related fitted lines for the four groups using 25%, 50% and 75% cutoff values. Only the regression for second subset (25%-50%) and third subset (50%-75%), which are in red and green, are not significant with p values of 0.0764 and 0.6114 respectively. For the other two groups, academic index have significant positive effects on the GPA performance. And we can see that the effects of academic index for these two groups are different since the slope of the fitted lines are different.

Figure 36 shows the four regression results and related fitted lines for the four groups using 5%, 15% and 25% cutoff values. Only the regression for fourth subset (25%-100%), which is in grey color, is significant. The other three regressions are not significant. Thus we can say the academic index is not a great predictor for students in the group of 0-5%, 5%-15% and 15%-25% quantiles in academic index.

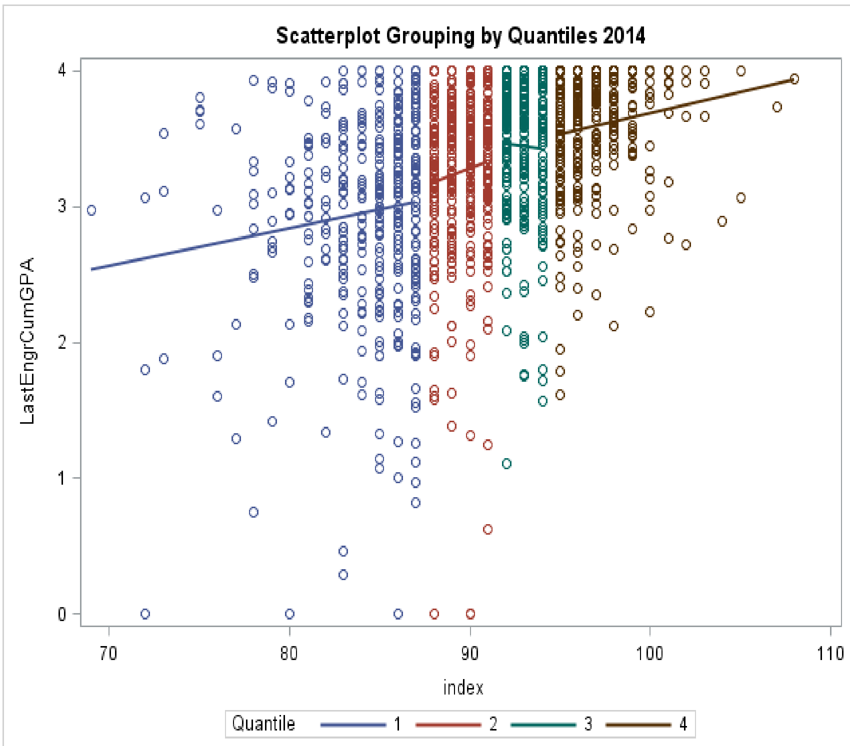

Figure 35 Academic Index vs. GPA Performance for 2014

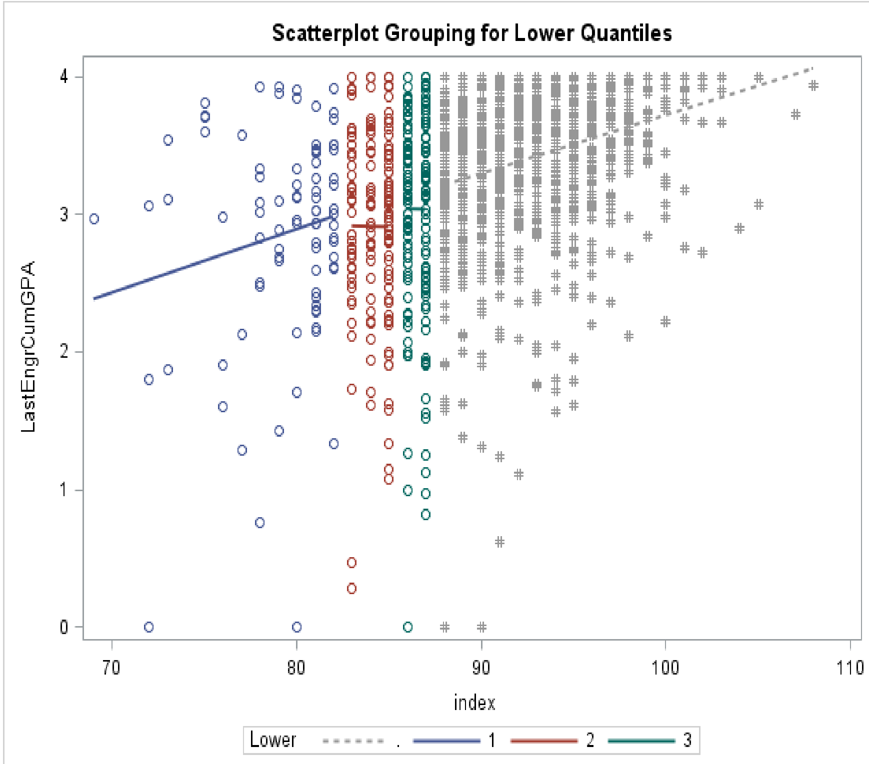

Figure 36 Academic Index vs. GPA Performance for 2014

5.8 Brief Summary

In performance analysis, we find that there exists no significant difference of the GPA performance between women and men, but significant differences between URM and non-URM students. For both graduated and current students in College of Engineering, the percentage of women does not correlate with the performance differences.

In enrollment analysis I, we find that enrollment index is positively related with the later GPA performance, which means if a student has a higher enrollment index, the GPA performance of the student tends to be higher. Only in some years there exist significant GPA differences between female and male students. And there is no difference of the enrollment index effect on GPA performance between female and male students.

In enrollment analysis II, we find that academic index is not a good predictor of future performance for students with lower academic index. On average their performances are poorer than other students. If want to lower the academic index for admission, possibly need other criteria besides academic index.

6. Discussions

The Women in Engineering project aims to examine whether female students in College of Engineering are significantly different from male students in terms of different indicators, such as GPA and Retention. Furthermore, we attempted to explore related factors that affect those differences. Specifically, we covered retention rate in College of

Engineering, effectiveness of women programs, retention rate across colleges and enrollment. And here are our key findings:

1) Female retention rate and URM retention rate are different in some engineering majors;
2) WIE Camp is effective to retain female students, who are more likely to graduate on time;
3) Retention rates of female students are significantly different across colleges;
4) Academic index is positively related to GPA performance.

Besides, there are some limitations of this study. First, there are limited data points for certain questions, for example, retention rate across colleges. The conclusions will be more robust and convincing when there are more data available. Second, most of data are aggregated at college level or year level. Accordingly, we suggest that longitudinal analysis at individual level would establish stronger causal relationships. For example, we could collect data on each student in each semester.

7. Appendix

7.1 Section 2: Results in R

1) Gender and Retention Rate: Likelihood test

```
> chisq.test(ConTable)

   Pearson's Chi-squared test with Yates' continuity
correction

data:  ConTable
X-squared = 1.3653, df = 1, p-value = 0.2426

> likelihood.test(ConTable)

   Log likelihood ratio (G-test) test of independence without
correction

data:  ConTable
Log likelihood ratio statistic (G) = 1.4398, X-squared df = 1,
p-value = 0.2302
```

2) Gender and Retention Rate: ANOVA

```
> summary(mod.1)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.56207    0.02445  22.986   <2e-16 ***
GenderF     -0.04419    0.03458  -1.278    0.203
Residual standard error: 0.2104 on 146 degrees of freedom
Multiple R-squared:  0.01106, Adjusted R-squared:  0.004287
F-statistic: 1.633 on 1 and 146 DF,  p-value: 0.2033
```

3) Female ratio and Female Retention Rate: ACONOVA

```
> ratio.4 <- lm(Retention_Rate ~ FemaleRatio+CohortTerm,
data=subfemale)
> summary(ratio.4)

Call:
lm(formula = Retention_Rate ~ FemaleRatio + CohortTerm, data =
subfemale)

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        0.45987    0.08320   5.527 5.77e-07 ***
FemaleRatio        0.66694    0.26524   2.514   0.0143 *
CohortTerm120048 -0.04697    0.09525  -0.493   0.6236
CohortTerm120058 -0.07904    0.09684  -0.816   0.4173
CohortTerm120068 -0.05782    0.09586  -0.603   0.5485
```

```
CohortTerm120078 -0.14254    0.09535  -1.495   0.1396
```

4) GPA and Female Retention Rate: ANCOVA

- Enroll GPA

```
> summary(GPA.1)
Call:
lm(formula = Retention_Rate ~ enroll_GPA + CohortFirstMajor +
CohortTerm, data = subfemale)
Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)            -1.0486234  0.3876801  -2.705 0.009168 **
enroll_GPA              0.4466223  0.1260362   3.544 0.000833 ***
CohortFirstMajorAero    0.2829956  0.1306156   2.167 0.034779 *
CohortFirstMajorBioen   0.4509026  0.1506478   2.993 0.004188 **
CohortFirstMajorCS      0.3173808  0.1326964   2.392 0.020347 *
CohortFirstMajorChem E  0.2850136  0.1335523   2.134 0.037480 *
CohortFirstMajorCivil   0.3754735  0.1324944   2.834 0.006492 **
CohortFirstMajorComp E  0.2339726  0.1312050   1.783 0.080271 .
CohortFirstMajorE Mech  0.1431928  0.1370937   1.044 0.301000
CohortFirstMajorE Phys -0.2247351  0.1445622  -1.555 0.125995
CohortFirstMajorEE      0.2780951  0.1381671   2.013 0.049234 *
CohortFirstMajorGen Eng 0.1600182  0.1322522   1.210 0.231668
CohortFirstMajorInd Eng 0.2448503  0.1372387   1.784 0.080129 .
CohortFirstMajorMatSE   0.3673870  0.1362955   2.696 0.009397 **
CohortFirstMajorMech E  0.3142728  0.1315924   2.388 0.020525 *
CohortFirstMajorNPRE    0.1425880  0.1422692   1.002 0.320781
CohortTerm120048       -0.0112644  0.0778883  -0.145 0.885557
CohortTerm120058       -0.0636480  0.0806466  -0.789 0.433499
CohortTerm120068        0.0009919  0.0782878   0.013 0.989938
CohortTerm120078       -0.0750672  0.0795006  -0.944 0.349335
```

- Remain GPA

```
GPA.1 <- lm(Retention_Rate ~ remain_GPA+CohortFirstMajor+CohortTerm,
data=subfemale)
> summary(GPA.1)

Call:
lm(formula = Retention_Rate ~ remain_GPA + CohortFirstMajor +
    CohortTerm, data = subfemale)

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)             0.893457   0.456465   1.957   0.0564 .
remain_GPA             -0.126668   0.143756  -0.881   0.3828
CohortFirstMajorAero   -0.005385   0.165134  -0.033   0.9741
CohortFirstMajorBioen   0.405846   0.182073   2.229   0.0307 *
CohortFirstMajorCS      0.144311   0.168531   0.856   0.3963
CohortFirstMajorChem E  0.124527   0.167783   0.742   0.4617
CohortFirstMajorCivil   0.182434   0.165313   1.104   0.2755
CohortFirstMajorComp E  0.011685   0.166047   0.070   0.9442
CohortFirstMajorE Mech  0.374086   0.187516   1.995   0.0520 .
CohortFirstMajorE Phys -0.098897   0.193407  -0.511   0.6116
CohortFirstMajorEE      0.191248   0.171529   1.115   0.2707
CohortFirstMajorGen Eng -0.030859   0.166571  -0.185   0.8538
CohortFirstMajorInd Eng 0.155965   0.174527   0.894   0.3762
```

```
CohortFirstMajorMatSE     0.264128   0.173337   1.524   0.1344
CohortFirstMajorMech E    0.106681   0.166510   0.641   0.5249
CohortFirstMajorNPRE      0.021007   0.176866   0.119   0.9060
CohortTerm120048         -0.023359   0.074594  -0.313   0.7556
CohortTerm120058          0.007196   0.081072   0.089   0.9297
CohortTerm120068         -0.032640   0.074623  -0.437   0.6639
CohortTerm120078         -0.066899   0.076390  -0.876   0.3857
```

5) Race and Retention Rate

```
> summary(URMmod_1)

Call:
lm(formula = Retention_Rate ~ Race, data = URMdata)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.49484    0.03161  15.656  < 2e-16 ***
RacenonURM   0.19457    0.04301   4.524 1.32e-05 ***
```

6) Race and GPA

- Enroll GPA

```
> summary(URMmod_3)

Call:
lm(formula = Retention_Rate ~ enroll_GPA + Term + Major,
data = subURM)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.62833    0.24974  -2.516   0.0132 *
enroll_GPA    0.33983    0.08160   4.164    6e-05 ***
Term120048   -0.06377    0.06852  -0.931   0.3540
Term120058    0.04756    0.06724   0.707   0.4808
Term120068    0.06677    0.06664   1.002   0.3184
Term120078    0.16587    0.06694   2.478   0.0146 *
MajorAero     0.13591    0.11509   1.181   0.2400
MajorBioen    0.05294    0.13551   0.391   0.6968
MajorCS       0.18872    0.11510   1.640   0.1038
MajorChem E   0.09897    0.11338   0.873   0.3845
MajorCivil    0.30144    0.11347   2.656   0.0090 **
MajorComp E   0.09776    0.11314   0.864   0.3893
MajorE Mech   0.06403    0.13265   0.483   0.6302
MajorE Phys  -0.03777    0.11976  -0.315   0.7530
MajorEE       0.11679    0.11433   1.022   0.3091
MajorGen Eng  0.19768    0.11440   1.728   0.0866 .
MajorInd Eng  0.04257    0.12495   0.341   0.7339
MajorMatSE    0.18044    0.11452   1.576   0.1178
```

```
MajorMech E    0.25043    0.11432   2.191   0.0305 *
MajorNPRE      0.14483    0.11618   1.247   0.2151
```

- Remain GPA

```
> summary(URMmod_4)

Call:
lm(formula = Retention_Rate ~ RemainGPA + Term + Major,
data = subURM)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.63185    0.25219   2.505  0.01378 *
RemainGPA    -0.02976    0.07656  -0.389  0.69828
Term120048   -0.04905    0.05567  -0.881  0.38031
Term120058    0.04563    0.05504   0.829  0.40899
Term120068    0.02194    0.05369   0.409  0.68362
Term120078    0.09617    0.05415   1.776  0.07865 .
MajorAero     0.05964    0.09761   0.611  0.54251
MajorBioen    0.36180    0.12626   2.865  0.00504 **
MajorCS       0.11116    0.09721   1.143  0.25546
MajorChem E   0.03584    0.09897   0.362  0.71800
MajorCivil    0.18191    0.09695   1.876  0.06343 .
MajorComp E   0.01745    0.09902   0.176  0.86044
MajorE Mech   0.11370    0.11503   0.988  0.32522
MajorE Phys   0.05898    0.10917   0.540  0.59017
MajorEE       0.02312    0.09721   0.238  0.81250
MajorGen Eng  0.10407    0.09699   1.073  0.28579
MajorInd Eng  0.05538    0.11005   0.503  0.61586
MajorMatSE    0.16602    0.09931   1.672  0.09759 .
MajorMech E   0.15584    0.09708   1.605  0.11146
MajorNPRE     0.15030    0.09928   1.514  0.13308
```

7.2 SAS Codes

Please see the STAT427 Code_Women in Engineering _ 20150511.txt file in the attachment.