Drop-Out Risk: Using Student, Faculty, Attendance, and Course Data with Regression to Alert Advising Teams

Nathan Dickmeyer, Jenny Zhu & Erez Lenchner
Institutional Research & Assessment
LaGuardia Community College
April 24, 2014

Problems with Academic Advising

- **CCSSE ratings significantly below all peer groups**
- **Noel-Levitz worst score in CUNY**
- **No progress in improving retention or graduation, despite reputation for innovation**
- **Students complain of being invisible**
- **Outcome of Achieving the Dream = Total Fragmentation of Services**

Initiative: Team Advising

- **Not enough resources for case load**
- **47 teams formed around each major**
- **Teams consist of**
  - **Faculty in major**
  - **Academic advising staff member**
  - **Admissions staff member**
  - **Career development staff member, and**
  - **Student life person**

Team Support

- **Provost heads both Academic Affairs and Student Affairs**
- **IT & IR&A heads sit on Advising Senior Leadership Team**
- **Provost request "Students at Risk" list for each team**
  - **Unable to specify how many on list**
  - **Unable to specify cut-off % for degree of risk**
  - **Unable to define risk:**
    - **Dropping out immediately?**
    - **Dropping out eventually?**
    - **Not graduating?**

Part I, Linear Regression Model

Modeling Data Specification

- Fall 2011 Degree Seeking Continuing Students
- Sample size: 10,962 students

Part I, Linear Regression Model

Dependent Variable (Outcome) Specification

- 1.00 – Dropped after one semester
- 0.85 – Dropped after two semesters
- 0.70 – Dropped after three semesters
- 0.55 – Dropped after four semesters
- 0.40 – Dropped after five semesters
- 0.25 – Still enrolled in the sixth semester
- 0.10 – Either graduated or transferred in three years

Part I, Linear Regression Model

Independent Variable Specification

- Age: Continuous
- Gender
- Full/Part Time
- Degree Type: AA/AS/AAS/Certificate
- Admission Type: Continuing/Readmit
- Late_Registration: Yes/No
- Exit From Remedial Math: Yes/No
- Cumulative Credits: Continuous
- GPA: Continuous
- Percentage of WU Grades Among all Grades Received: Continuous
- Ratio of Total Equated Credits Earned Vs. Equated Credits Attempted: Continuous

Part I, Linear Regression Model

Modeling Output

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	Intercept	1	0.57747	0.01674	34.50	<.0001	0
Age	Age	1	0.00251	0.00046881	5.36	<.0001	1.14483
Credits	Credits	1	-0.00532	0.00018665	-28.51	<.0001	1.21506
GPA	GPA	1	-0.09879	0.00469	-21.05	<.0001	1.63930
R_WU	R_WU	1	0.05293	0.03422	1.55	0.1219	1.26470
Gender (Male)		1	0.03012	0.00670	4.50	<.0001	1.02031
Exit from Remedial Math (No)		1	0.04005	0.00868	4.62	<.0001	1.22852
Full Time (No)		1	0.12055	0.00698	17.28	<.0001	1.11520
Late Registration (Yes)		1	0.05367	0.00719	7.47	<.0001	1.11108
Admission Type (Readmit)		1	0.08650	0.01364	6.34	<.0001	1.06396

Part I, Linear Regression Model

R-Squared

- Our Adjusted R-Squared: 22.3%
- R-Squared is the goodness of fit, the “percent of variance explained” by the model. That is, R-squared is the fraction by which the variance of the errors is less than the variance of the dependent variable.

Part I, Linear Regression Model

Independent Variables Proved to be Not Significant

- Degree Type
- Ratio of Total Equated Credits Earned Vs. Equated Credits Attempted

In Previous Modeling Attempts:

- GED
- Ethnicity
- Reading Remedial Needs
- Writing Remedial Needs

Part I, Linear Regression Model

Independent Variables Proved to be Significant for New Students

- Age: Continuous
- Gender
- Visa Type
- Full/Part Time
- Degree Type: AA/AS/AAS/Certificate
- Admission Type: Continuing/Readmit
- Late_Registration: Yes/No
- Math Remedial Needs
- Reading Remedial Needs

Part II, Logistic Regression Model

Modeling Data Specification

- Fall 2012, Spring 2013 and Fall 2013 Degree Seeking Students with Fewer than 45 Credits
- Sample size: 35,774 students

Dependent Variable (Outcome) Specification

- Whether Students Were Enrolled in the Third Semester Following the Base Semester

Part II, Logistic Regression Model

Independent Variable Specification

- Age: Continuous
- Gender
- Visa Type
- Financial Aid
- Full/Part Time
- Degree Type: AA/AS/AAS/Certificate
- Admission Type
- Late_Registration: Yes/No
- Exit From Remedial Math: Yes/No
- Cumulative Credits: Continuous
- GPA: Continuous
- Percentage of WU Grades Among all Grades Received: Continuous
- Ratio of Total Equated Credits Earned Vs. Equated Credits Attempted: Continuous

Part II, Logistic Regression Model

Modeling Output (Step 1, R-Squared = 0.1265)

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.0741	0.0634	1.3662	0.2425
Age		1	0.00535	0.00177	9.1843	0.0024
GPA		1	0.1988	0.0122	263.5139	<.0001
R_Credits		1	0.0821	0.0307	7.1295	0.0076
R_WU		1	-1.2303	0.1422	74.8537	<.0001
Gender	0	1	0.0900	0.0117	58.9153	<.0001
Math096	0	1	-0.1577	0.0131	145.3613	<.0001
FP	1	1	0.4666	0.0125	1384.2910	<.0001
Late_Reg	0	1	0.1175	0.0121	94.8296	<.0001
Admis	1	1	-0.0187	0.0279	0.4496	0.5025
Admis	2	1	0.2817	0.0319	77.7849	<.0001
Admis	3	1	0.2225	0.0449	24.6046	<.0001
Degree	11	1	0.0349	0.0256	1.8561	0.1731
Degree	12	1	0.1519	0.0243	38.9661	<.0001
Degree	13	1	-0.0527	0.0256	4.2472	0.0393
Fin_Aid	0	1	0.0288	0.0131	4.8483	0.0277
Visa	0	1	-0.2011	0.0343	34.3878	<.0001

Part II, Logistic Regression Model

Modeling Output (Step 2 with Course, R-Squared = 0.154)

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.5070	0.0656	59.7659	<.0001
Age		1	0.00656	0.00179	13.4560	0.0002
GPA		1	0.2153	0.0122	312.0951	<.0001
R_WU		1	-1.2954	0.1432	81.7874	<.0001
Gender	0	1	0.0965	0.0119	66.2800	<.0001
Math096	0	1	-0.1647	0.0132	155.1568	<.0001
FP	1	1	0.3032	0.0138	485.0706	<.0001
Late_Reg	0	1	0.1062	0.0122	75.5920	<.0001
Admis	1	1	-0.0288	0.0279	1.0652	0.3020
Admis	2	1	0.2929	0.0320	83.9279	<.0001
Admis	3	1	0.2233	0.0450	24.5825	<.0001
Degree	11	1	0.0140	0.0260	0.2923	0.5888
Degree	12	1	0.0955	0.0248	14.8540	0.0001
Degree	13	1	-0.0339	0.0259	1.7096	0.1910
Fin_Aid	0	1	0.0299	0.0132	5.1477	0.0233
Visa	0	1	-0.2080	0.0346	36.2336	<.0001
Course		1	0.2525	0.00898	789.9979	<.0001

Part II, Logistic Regression Model

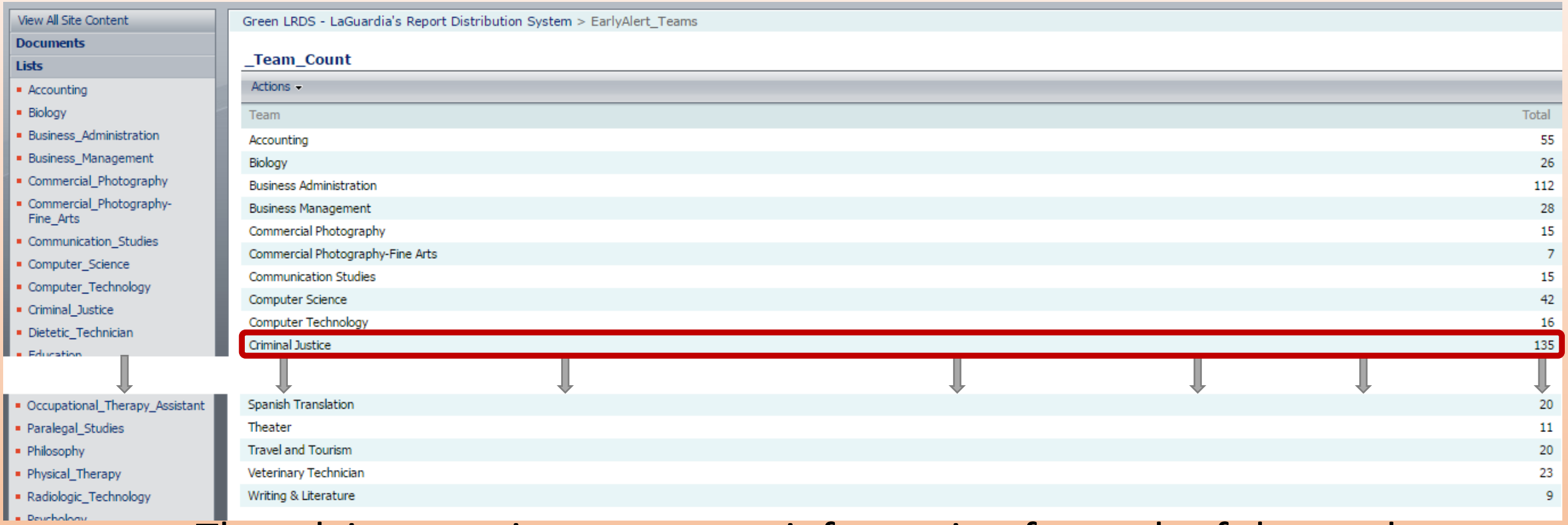
Modeling Output (Step 3 with Course & Faculty, R-Squared = 0.1734)

Parameter		DF	Standard Estimate	Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.7172	0.0680	111.3719	<.0001
Age		1	0.00678	0.00184	13.6078	0.0002
GPA		1	0.2151	0.0124	298.9364	<.0001
R_WU		1	-1.3210	0.1448	83.1706	<.0001
Gender	0	1	0.0970	0.0121	64.5870	<.0001
Math096	0	1	-0.1610	0.0134	143.4140	<.0001
FP	1	1	0.2391	0.0142	284.3193	<.0001
Late_Reg	0	1	0.0982	0.0125	62.2309	<.0001
Admis	1	1	-0.0349	0.0284	1.5090	0.2193
Admis	2	1	0.2789	0.0325	73.5178	<.0001
Admis	3	1	0.2382	0.0459	26.9178	<.0001
Degree	11	1	0.00854	0.0266	0.1031	0.7481
Degree	12	1	0.1025	0.0255	16.2065	<.0001
Degree	13	1	-0.0538	0.0267	4.0775	0.0435
Fin_Aid	0	1	0.0439	0.0135	10.5904	0.0011
Visa	0	1	-0.2037	0.0351	33.7158	<.0001
Course		1	0.2517	0.00958	690.7318	<.0001
Faculty		1	0.2163	0.00825	687.1277	<.0001

Applications for Students' Outreach (I)

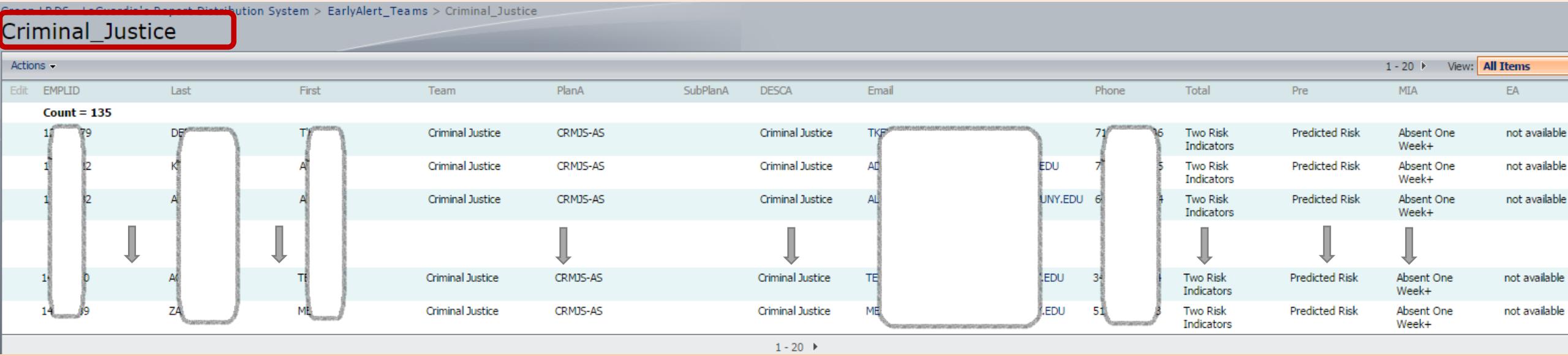
- Pre-Semester risk module provides one dimension to the student profile
- Student Real-Time semestrial performance provides on-time view of their risk profile
- Class attendance– May be a problem within a class, or across all the student classes.
  - Student attendance data results in student absent report
  - The report consolidates a list of all students who were *positively absent from all classes for at least seven days.*
- Pre-semester and Real-Time information are combined for a comprehensive view of student's profile
- Data Security- The information is distributed through Microsoft Sharepoint, in compliance with all college, state and federal data security requirements.

Advisor View (Updated Weekly)



- The advisor receives summary information for each of the student teams' s/he is responsible.
- By selecting each group, the advisor can find more information regarding the at risk students

Advisor View (Updated Weekly)



- Within each team, the advisor receives a detailed summary for each student.
- The summary specifies the reason(s) the student is listed: Pre-semester predicted risk, absence for classes for seven consecutive days (or more), and early alert (a college-based ‘bundle’ of programs) information.
- The list contains contact information, and the advisor can immediately create customized messages to a group (or sub-group) of interest
- The advisor can conduct basic filtering and segmentation within sharepoint, or further download the list for further sorting and segregation
- Given screen size and “human processing” limitations, certain information is available to the advisor only when the information is downloaded

Applications for Students' Outreach (II)

- Deployment Period-
  - Specific Targets
    - Presentation of Team's data on need-to-know basis
    - Simple, Single Logon Access
    - Automated or Semi-Automated update cycles
    - Usage Reports and Linkage to Student Outcomes
  - Creations of Advisors lists/teams and responsibility areas is crucial for setting of the correct authorizations
  - Distribution requires modifications in data consumption practices, in particular, direct access
  - Advisors' training and orientation workshop/module are crucial in the deployment of new practices
  - Handling Data Constraints

Applications for Students' Outreach (III)

- Constraints
  - Instructor's compliance with on-time attendance requires awareness to information usage
    - The information, usage plan and requirements presented to faculty in "opening session" by the Provost
  - Quality of the list is strongly dependent on Instructor's on-time reporting
    - On time reporting rate constantly increasing (Currently exceeding 75 percent)
  - Advisors require access to information that is not updated/accessible in the same data environment, resulted in modifications to data presentation and data retrieval changes
  - **Screen size and "human engineering"** limits the amount of information presented in a single view.
    - Advisors ranked the information needed most in a student's list.
    - Additional, detailed variables (e.g. Probation status) are presented only when an advisor downloads a list

Applications for Students' Outreach (IV)

- Future Developments
  - Single Login via "Data Store" (A single, secured, data distribution center)
  - Direct Linkage to students accounts for customized, automated messages
  - Creation of Standardized Reports for 'subgroups' as long as they can be defined using subset of variables (not to exceed three)
    - In testing phase: Athletics Program