WHEN BIG DATA BECOMES LITTLE DATA: PREDICTING FRESHMAN ATTRITION

Joseph B. O'Donnell and Paul L. Sauer

The BRC Academy Journal of Business 8, no. 1 (2019): 1–27.

<https://dx.doi.org/10.15239/j.brcacadj.2020.08.01.ja01>

WEB APPENDIX

<https://dx.doi.org/10.15239/j.brcacadj.2020.08.01.wa01>

Appendix 1 – Tables

Table 1

Impact of Environmental and Organizational Changes at Target College

Available data – big to little

- Discontinued use of self-report instrument and data
 - Faculty attrition
 - Staff attrition
 - Enrollment decline
 - Instrument decline
 - Psychographic – 79 items to 37 to none
 - Third Party Contracts
-

Table 2

Timeline – Organizational Periods – Forecasting Retention

-
- 1980 to 1995
 - Administrative database and survey instrument development
 - 1995 to 2005
 - Improvement and refinement of database
 - Use of varied estimation and modeling techniques
 - Departure of a key administrator champion
 - 2005 Departure of key faculty member responsible for database management and validation
 - 2006 to 2013
 - Administrator turnover – no consistent retention management policy
 - 2013 to 2018
 - Renewed interest of new key administrators and faculty
-

Table 3

The 12 Principle Component Self-Report Factors and Background Variables Years 2000-2003

<u>Variable Name</u>	<u>Variable Description</u>
Self-Report Factors	
F1	Academic Reasons for Choosing This College
F2	Unsure of Career
F3	Enjoy Politics
F4	Cost Important Factor in Choice of This College
F5	Moral/Religious Values
F6	Bad Academic Attitude in High School
F7	Good Study Habits in High School
F8	Good Relationships with High School Teachers
F9	Concern for the Disadvantaged
F10	Physical Fitness
F11	Expect Academic Problems in College
F12	Computer Skills
Background Variables – Survey Items	
concern1	Concern for financing education
offwork1	Off-campus hours worked per week
Background Variables Administrative Data	
Hsave	Eight-term high school academic average
Matrcage	Age at time of matriculation
Edparnts	Total of mother's and father's education
Marfut	Expect marriage 1 year from graduation?
Zipcode	Live in local MSA?
Parmar	Parents alive, married, and living together?
Gender	Male or female
Reside	Residence (1) Campus/with parents (2) Other

Table 4

Dropout Rates by Year

<u>Year</u>	<u>Dropouts</u>	<u>Dropout Rate</u>
2004	142	16%
2005	152	20%
2006	146	19%
2007	148	17%
2008	143	18%
2013	96	14%
2014	107	17%
2015	99	16%
2016	104	17%
2017	84	15%

Table 5

Background Variables - Years 2013-2017

<u>Variable Name</u>	<u>Variable Description</u>	<u>Category</u>
SATM	SAT Math	Academic
SATV	SAT Verbal	Academic
YR1_TERM_GPA	First semester GPA at the College - 4.0 scale basis	Academic
GENDER	Male or Female	Demographic
COMMUTER_ADMIT	Commuter Admit Indicator	Engagement
ATH_INDICATOR	Athlete Indicator	Engagement
WORK_AMOUNT	Total amount of work-study - federal (FWS) or other including	Engagement

institutional work study - need or
non-need based

TOTAL_INST_GRANT	Total institutional grant awarded - restricted and unrestricted, funded or unfunded	Financial
PELL_ELIG_IND	PELL Eligible Indicator	Financial
YR1_UNMET_NEED	Year 1 Unmet Need Amount	Financial
YR1_LOAN_AMT	Year 1 Loan Amount	Financial

Table 6

Classification Table Dropout vs Persist - Pooled 2013 to 2017 Data

<u>Observed</u>	<u>Predicted</u>		<u>Percentage Correct</u>
	<u>Dropout</u>	<u>Persist</u>	
Dropout	284	649	30.4%
Persist	112	1,938	94.5%
Overall Percentage			74.5%

Table 7.

Logistic Regression Dropout versus Persist Academic Years 2013 – 2017 Inclusive

<u>Variable Name</u>	<u>B</u>	<u>S.E.</u>	<u>Wald</u>	<u>df</u>	<u>Sig.</u>	<u>Exp(B)</u>
YR1_TERM_GPA	-0.93308	0.065272	204.3553	1	2.34E-46	2.542335
SATV	0.00657	0.000613	114.7926	1	8.74E-27	0.993454
ATHLETE_IND	-0.41161	0.161736	6.476745	1	0.010929	1.509246
COMMUTER_ADMIT	-0.31129	0.093359	11.11785	1	0.000855	1.365189
YR1_UNMET_NEED	0.05929	0.006575	81.32137	1	1.92E-19	0.942435
PELL_ELIG_IND	0.20922	0.102296	4.183157	1	0.040828	0.811213
YR1_LOANS	-0.02555	0.005924	18.60704	1	1.61E-05	1.025882
Constant	1.537368	0.359188	18.31947	1	1.87E-05	4.65233
Dependent Variable – Dropout: Yes (2) No (1)						