W5: A Primer on Analytics and Visualization in Higher Education

John Stanley, University of Hawai'i – West Oahu
Ken Nelson, Loma Linda University

WSCUC ARC Pre-Conference Workshop, San Diego, CA, April 19, 2017

Session objectives

1. Define "analytics" and discuss challenges.
2. Examine several case studies that have used analytics to improve institutional effectiveness.
3. Examine the use of analytics in the following key areas:
   a) Program review and assessment
   b) Enrollment management
   c) Surveys
   d) 'What-if' decision-making

2

Two institutions, one mission

University of Hawai'i West O'ahu

Loma Linda University

PacAIR Pacific Association for Institutional Research

AIR Association for Institutional Research

WASC Senior College and University Commission

California Association CAIR for Institutional Research

3

Activity

Think – Pair – Share Activity

3 minute questionnaire + 5 minute share

1. Area(s) of notable strength?
2. Area(s) for improvement?
3. If IR analytical capacity were **enhanced**, **what campus/program issue could be better addressed or potentially solved?**

4

Challenges for Institutional Research

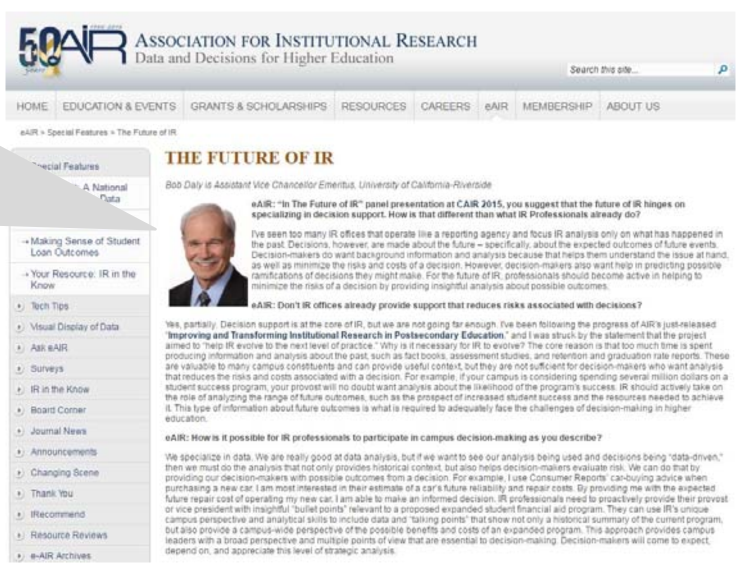
- Compliance vs. Self-Improvement
- Developing a culture of evidence
- From reporting to analysis
- Converting data into 'actionable' information
- Follow highest standards, best practices
- Know your customers, mission
- Leverage technology, stay abreast of tech
- Empower staff, continuous honing of skills
- Effective senior-management support working with IR (and IT)

5

AIR Newsletter March 2016

"I've seen too many IR offices that operate like a reporting agency and focus IR analysis only on what has happened in the past. Decisions, however, are made about the future – specifically, about the expected outcomes of future events… For the future of IR, professionals should become active in helping to minimize the risks of a decision by providing insightful analysis about possible outcomes."

– Bob Daly, eAIR Newsletter, March 2016



SOURCE: https://www.airweb.org/eAIR/specialfeatures/Pages/The-Future-Of-IR.aspx

6

What is analytics?

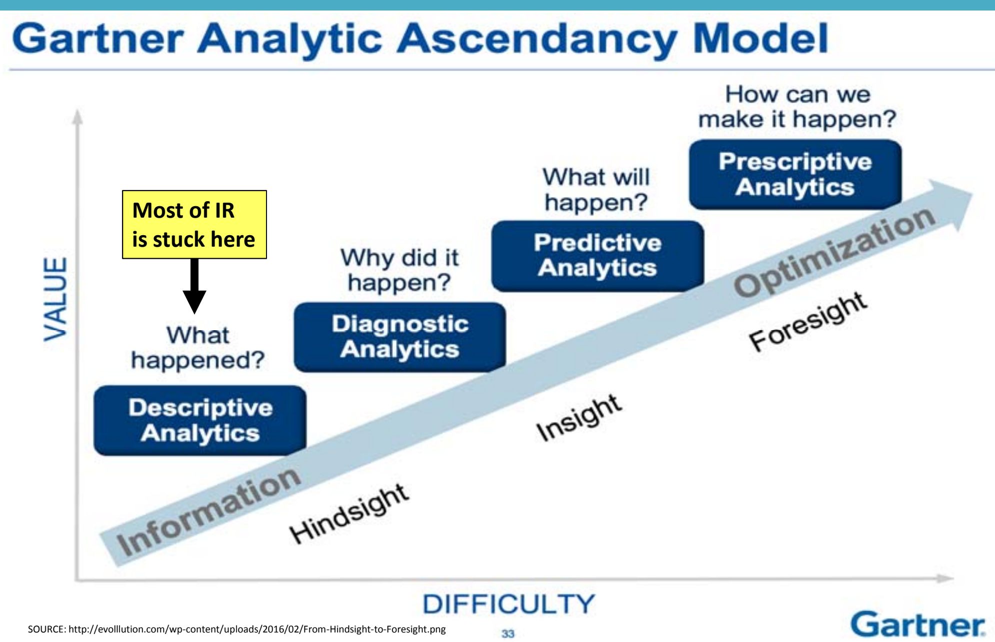
"Analytics is the use of data, statistical analysis, and explanatory and predictive models to gain insights and act on complex issues."

-EDUCAUSE Center for Applied Research

EDUCAUSE Center for Applied Research Video: "What is Analytics?"
Downloaded from: http://www.educause.edu/ero/article/video-what-analytics

7

From hindsight to foresight



SOURCE: http://evolllution.com/wp-content/uploads/2016/02/From-Hindsight-to-Foresight.png

8

IR tasks applied to Gartner Ascendency Model

DESCRIPTIVE	DIAGNOSTIC	PREDICTIVE	PRESCRIPTIVE
1. Disaggregating student retention rates by gender, ethnicity, Pell, first generation status, etc.	1. Using inferential statistics to determine if there are statistically significant differences between groups and identify important drivers of behavior.	1. Building a prediction model to identify which students are 'at-risk' of dropping out.	1. Delivering dropout risk assessment lists to student support services in order to provide actionable information.
2. Descriptive data in tables, charts, and graphs.	2. Interactive dashboards with slice and dice capability, drop downs.	2. Interactive dashboards with 'what-if' capability for key decision-makers.	2. Interactive dashboards used to push dropout risk data to academic advisors.
3. A cross-tabulation showing retention rates for students in learning communities versus non-learning community students.	3. A counter-factual analysis that controls for self-selection bias using student matching techniques.	3. Using learning community as a variable in a prediction model for retention or time-to-degree.	3. A more precise gauge of the impact of learning communities given to LC coordinators.

9

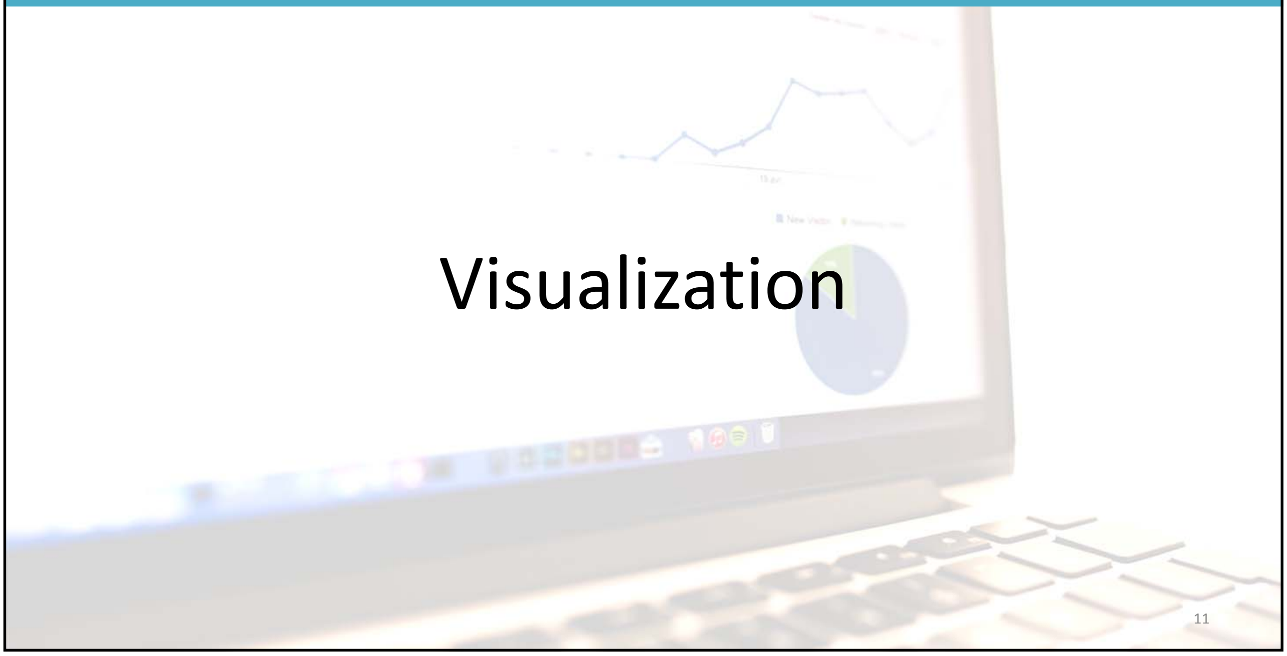
Analytics typology

Covered Today

Descriptive	Optimization	Explanatory	Action
Diagnostic	*Predictive*	Business Intelligence	A.I.
Academic	Prescriptive	*Visualization*	Big Data
Exploratory	Action	Business	*Learning*

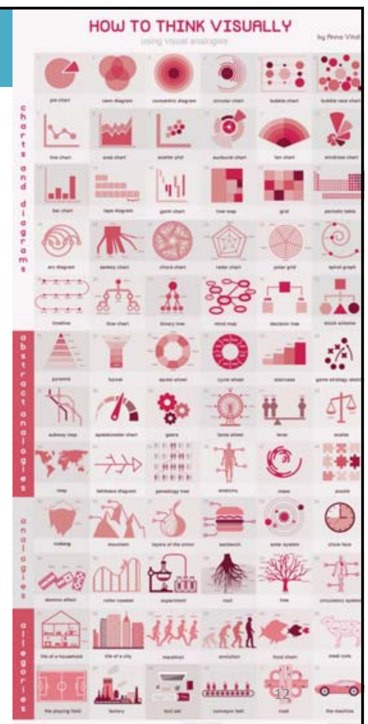
10

Source: https://library.educause.edu/~/media/files/library/2012/1/eli3026-pdf.pdf



Visualization

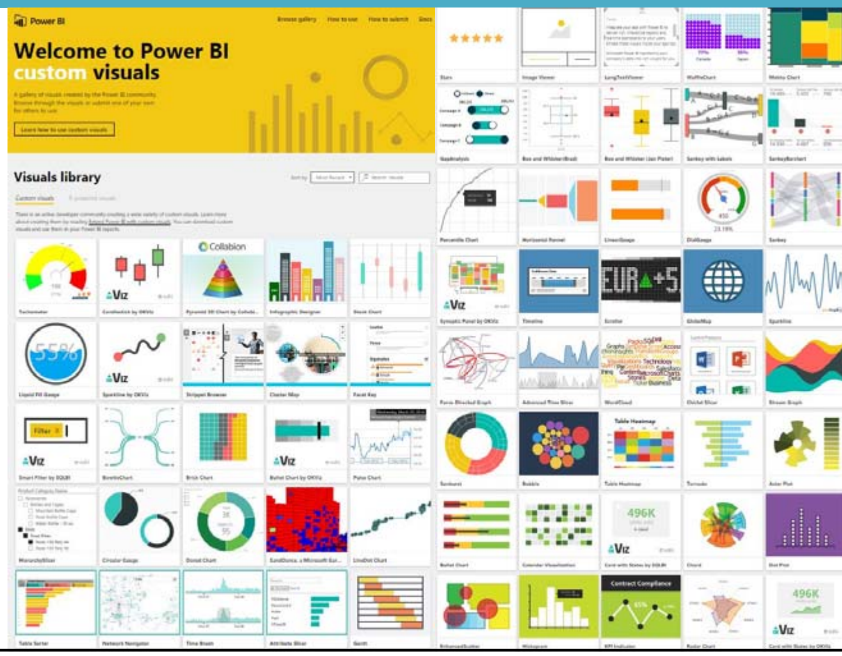
Visualization



- Data visualization is used to communicate data or information by representing it as visual objects (e.g., heatmaps, chords, sankeys).
- The goal is to communicate information clearly and efficiently to help users make:
  - faster insights
  - clearer choices
  - faster decision making

Source: http://designtaxi.com/news/376549/Infographic-72-Ways-To-Think-Present-Your-Ideas/

Visualization with Microsoft Power BI (free)



Source:
https://app.powerbi.com/visuals/

13

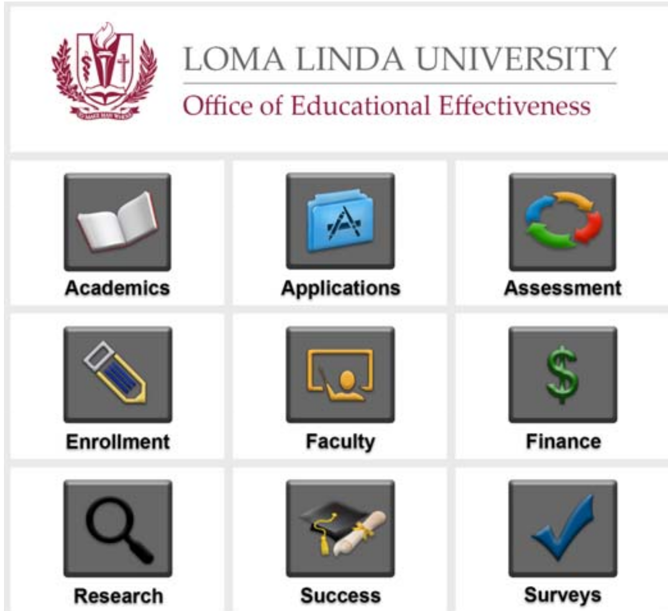
Gartner Magic Quadrant for BI (2017)

SOURCE: http://optimalbi.com/blog/2017/02/17/gartner-magic-quadrant-for-business-intelligence-2017-cloud-is-coming-slowly/

14

Visualization with Microsoft Power BI

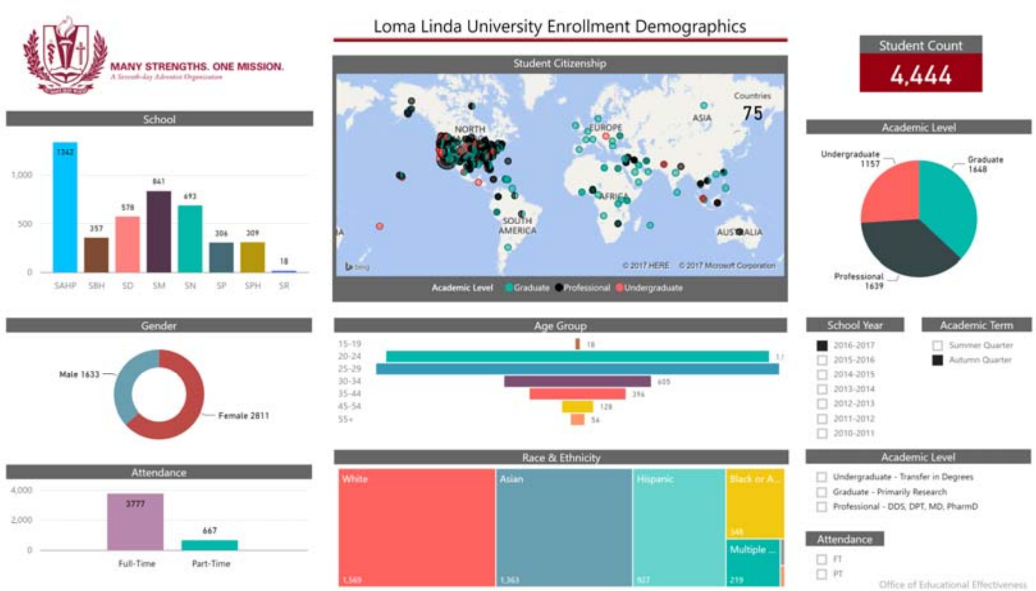
Dashboards as a Collection of Tiles Providing Links to Underlying Reports



15

Visualization with Microsoft Power BI

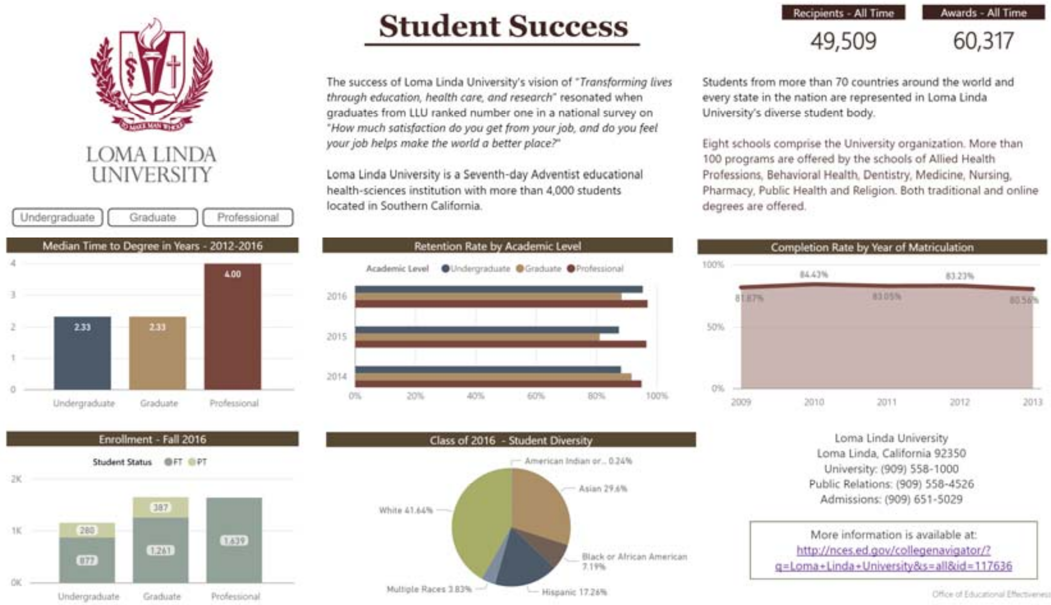
Public informative dashboards. In this example Loma Linda University's enrollment demographics are displayed on the University's website.



Source: http://home.llu.edu/academics/academic-resources/educational-effectiveness/institutional-research

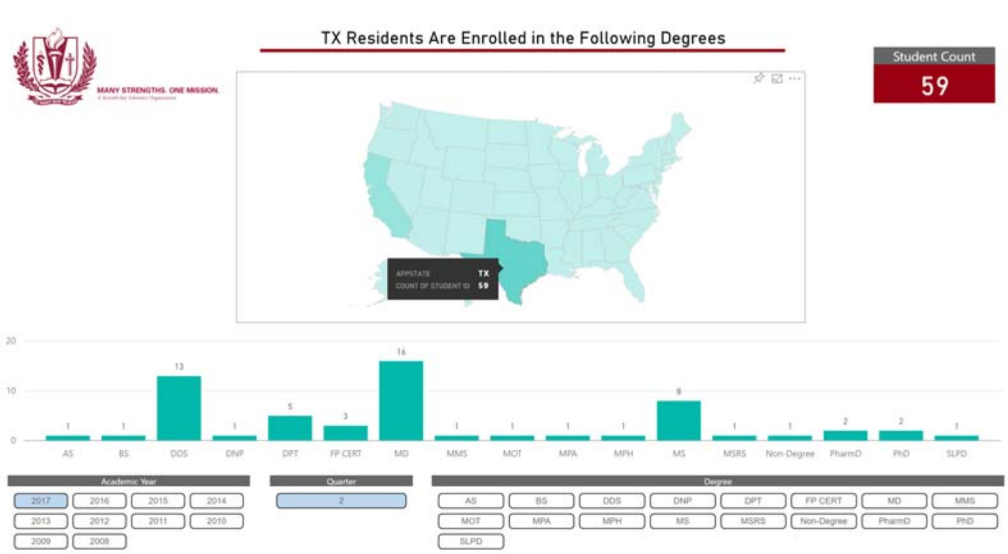
Visualization with Microsoft Power BI

Interactive charts and graphs with explanatory text. This example provides student success information as required by the regional accreditor.



Visualization with Microsoft Power BI

Interactive maps that allow disaggregating data by selecting areas of interest. In this example program enrollment is displayed by state selected.

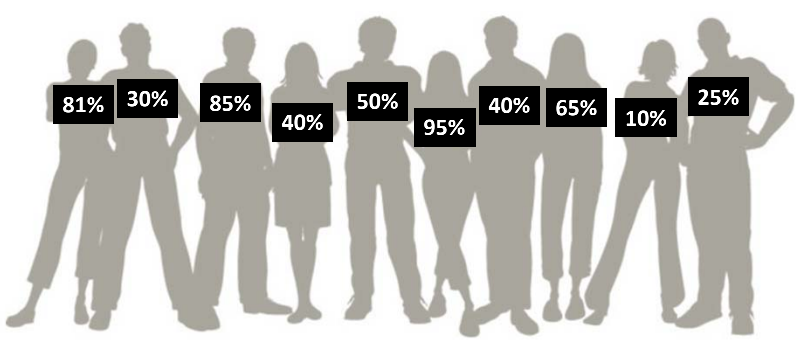


Predictive Analytics

19

Predictive Analytics

- Uses historical data to predict or forecast future behaviors, trends, or outcomes
  (i.e. enrollment likelihood, retention, course pass/fail, degree completion, gainful employment, etc.)



20

Benefits of predictive analytics

- Can generate "actionable" data (i.e., data used by academic support services to effectively assist students).
- Powerful and accurate predictive models can be constructed using matriculation data from your Student Information System (SIS).

21

Possible uses of predictive analytics

- Admissions recruitment
  - Predict which students are likely to enroll at your institution (Goenner & Pauls, 2006)
- Identifying at-risk students
  - Predict which students are likely to drop out or fall behind (Herzog, 2006 ; Sujitparapitaya, 2006)
- Students' price responsiveness to tuition increases or financial aid incentives (Des Jardins, 2001; Herzog & Stanley, 2017)
- Other uses?
  - Student Learning
  - Strategic Planning
  - Finance

22

A few examples of colleges using predictive analytics…

University of Nevada, Reno

Georgia State University

Florida State University 1851

Missouri State

UC San Diego

University of Hawai'i West O'ahu

SJSU San José State University

ASU Arizona State University

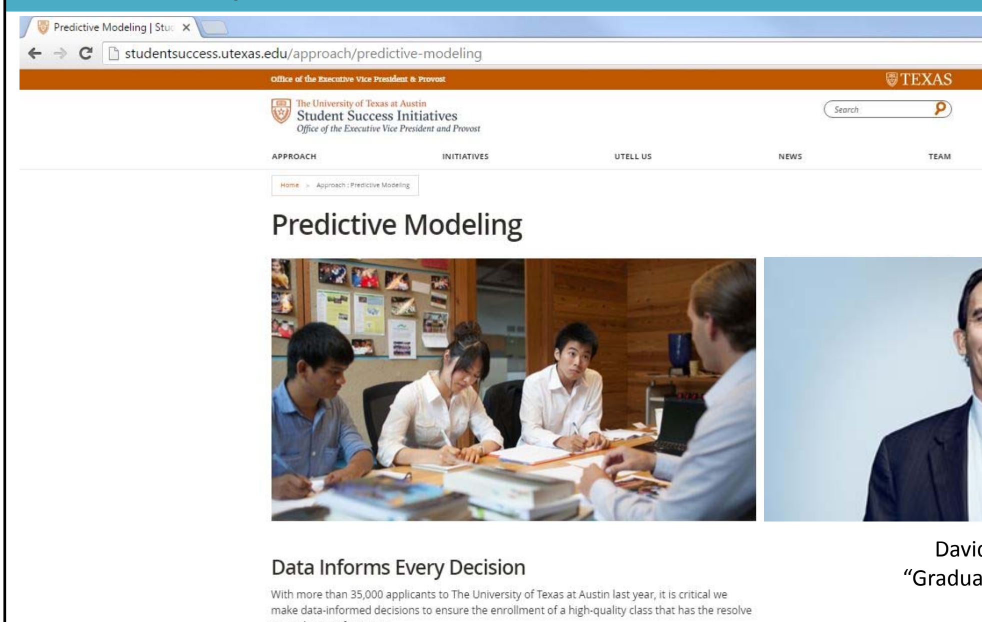
TEXAS The University of Texas at Austin

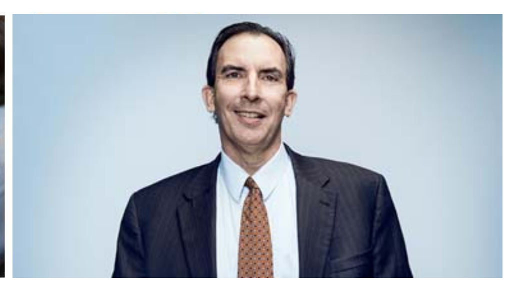
UC Davis University of California

Virginia Tech

23

University of Texas at Austin





David Laude, PhD
"Graduation Champion"

24

University of Texas at Austin

For the Class of 2017:

- 94.6 percent retention, up from 93.6 percent prior year, resulting in the highest one-year retention rate in the university's history for returning freshmen.
- Average GPA of 3.28, up from 3.22 for the previous class.
- Students enrolled in and passed more SSH (average 13.32 hours passed) than any entering class in the past five years. Taking more credit hours each semester will help these students stay on track to graduate in four years.


The New York Times

Who Gets to Graduate?

By PAUL TOUGH MAY 15, 2014

For as long as she could remember, Vanessa Brewer had her mind set on going to college. The image of herself as college student appealed to her — independent, intelligent, a young woma full of potential — but it was more than that; it was a chance to rewrite the endi to a family story that went off track 18 years earlier, when Vanessa's mother, then a high-achieving high-school senio in a small town in Arkansas, became pregnant with Vanessa.

Vanessa's mom did better than most teenage mothers. She married her high-

Vanessa Brewer Bill McCullough for The New York Times

Source: https://www.nytimes.com/2014/05/18/magazine/who-gets-to-graduate.html?_r=0

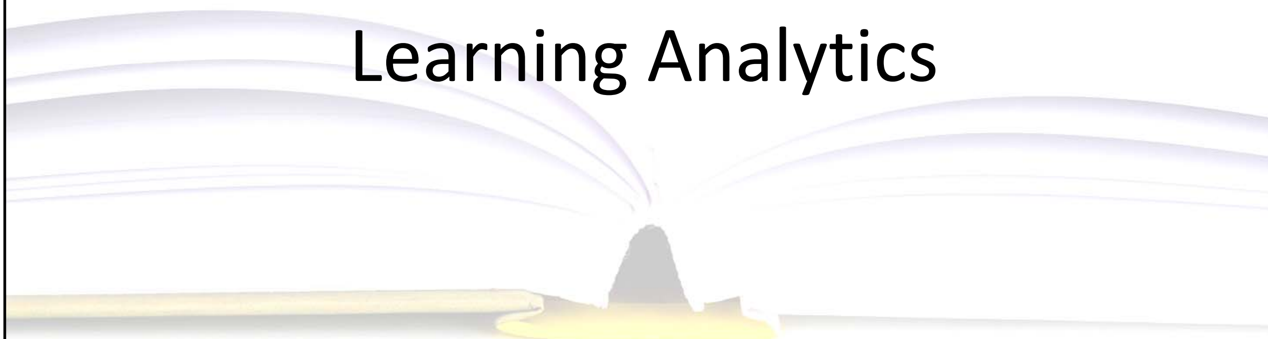
25

Other Noteworthy Examples

1. Georgia State University (reduced achievement gaps, featured nationally)
2. University of Nevada, Reno (early pioneer, raised retention 4% pts, featured nationally)

…See syllabus for more references

Learning Analytics

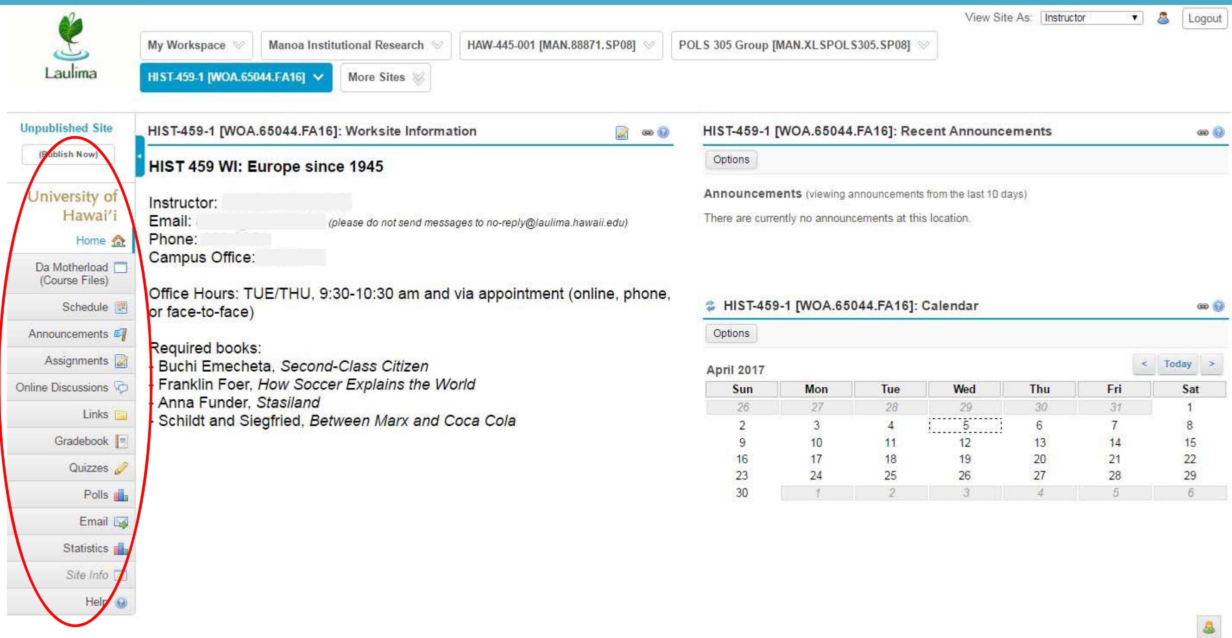


27

What is Learning Analytics?

- Defined as “the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs”
  - (Long & Siemens, 2011, p. 32).
- Data may be available in your institution’s Learning Management System (LMS).

Learning Analytics via Sakai LMS



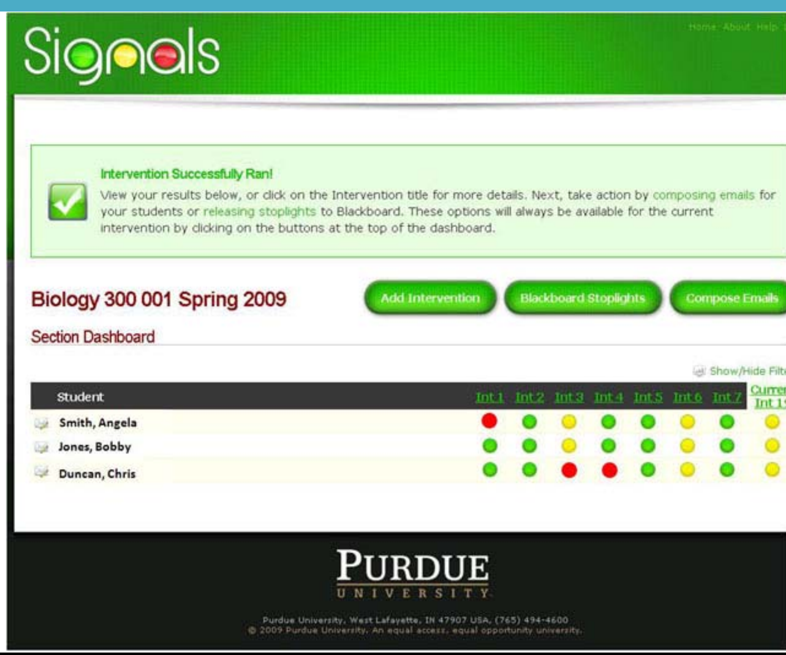
LA predictors commonly used

- Performance/Activity in Class (LMS)
  - Number and frequency of LMS logins
  - Amount of time spend on course website
  - Number of discussion posts
  - Responses to class polls
  - Grades and formative quiz scores
  - Percentage of points earned in course to date
  - Change between past and current test/quiz scores
- Student In-Class Assignments (LMS)
  - Blogs, discussion forum posts,
  - Essays, written assignments
- Student Learning Outcomes (LMS)
  - Measurements of student achievements in core competencies in class.
- Matriculation Predictors (SIS)
  - Demographics (age, gender & ethnicity), GPA, pre-collegiate HS GPA, standardized tests scores, first-generation, socio-economic & financial need

LA can help answer questions like…

- What is the likelihood that a particular student will pass the class?... or master a certain learning outcome?
- Are there dispositional characteristics that predict or explain performance in certain classes (i.e. Do males outperform females in STEM classes or vice versa)
- Can LMS data be used with SIS data to predict student persistence and degree completion?

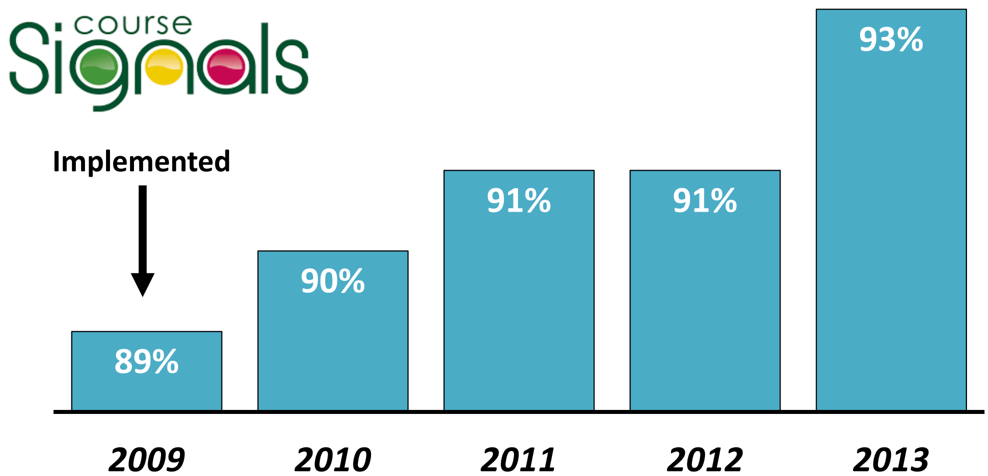
LA at Purdue University



http://www.nbcnews.com/id/3032619/vp/32634348#32634348

Student Retention at Purdue

>3 % point rise since 2009



Other Noteworthy Examples

1. Rio Salado College - “RioPACE”
2. University of Michigan - “Ecoach”
3. University of Maryland, Baltimore – “Blackboard Learn”

…See syllabus for more references

Challenges to using analytics

Affordability

- Infrastructure
- Technology
- People/Expertise
- Opportunity Costs

Data availability

- Student Information System
- Learning Management System
- Budget/ Human Resource Silos

Predictive data

- Culture change
- Wary of misuse of data
- Questions about data used to generate scores
- Students' access to risk scores
- Self-fulfilling prophecy

35

Combined Activity + Break

Think – Pair – Share Activity

5 minute discussion over break

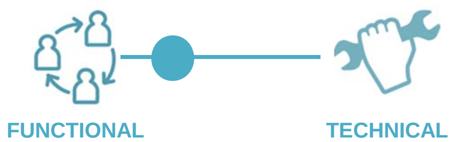
1. What are some challenges at your institution to supporting a culture for analytics **(i.e., affordability, data availability, expertise, etc)?**

2. What strategies may be helpful for overcoming these challenges?

36

Analytics Application #1

Program Review and Assessment

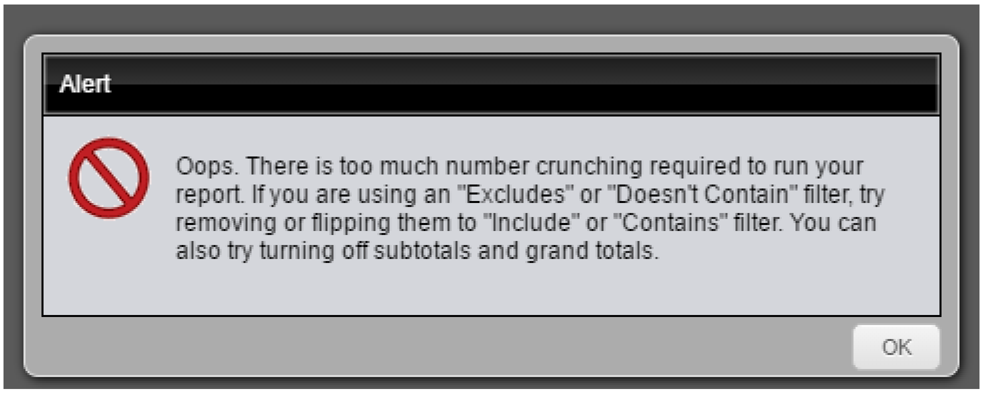


Process and Purpose

- Process of Program Review on campus
- Variable skillset of faculty and staff
- Curriculum mapping and rubric assessment
- Collecting assessment data
- Purpose
  - **Bringing it all together**
  - **Using the data to make changes where needed - “closing the loop”**
- Power BI examples

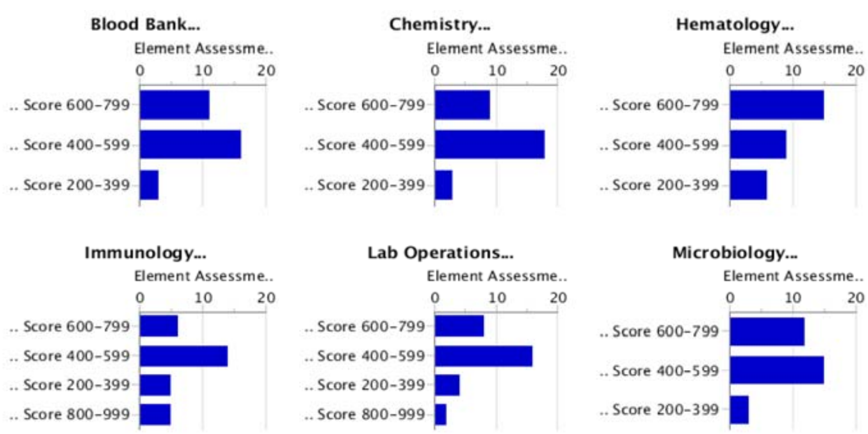
Rubric Element Scoring – Institution Example

- Challenge –
  - Assessment software couldn't provide average rubric value
  - Assessment software didn't allow complete data extract
  - Attempting to use the assessment software built-in analytics resulted in:



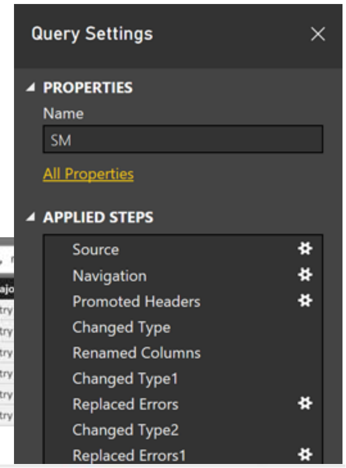
Rubric Element Scoring – Institution Example

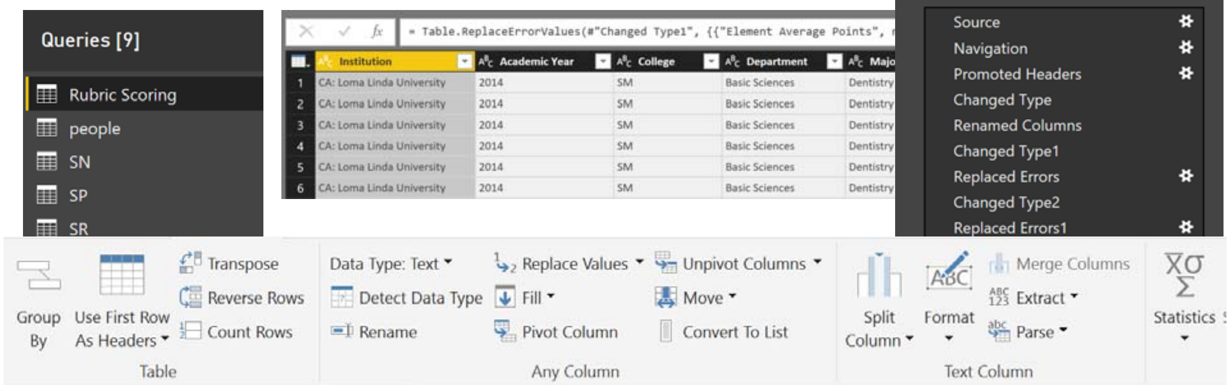
- Challenge –
  - Extract needed student IDs
  - Rudimentary analytics provided by assessment software



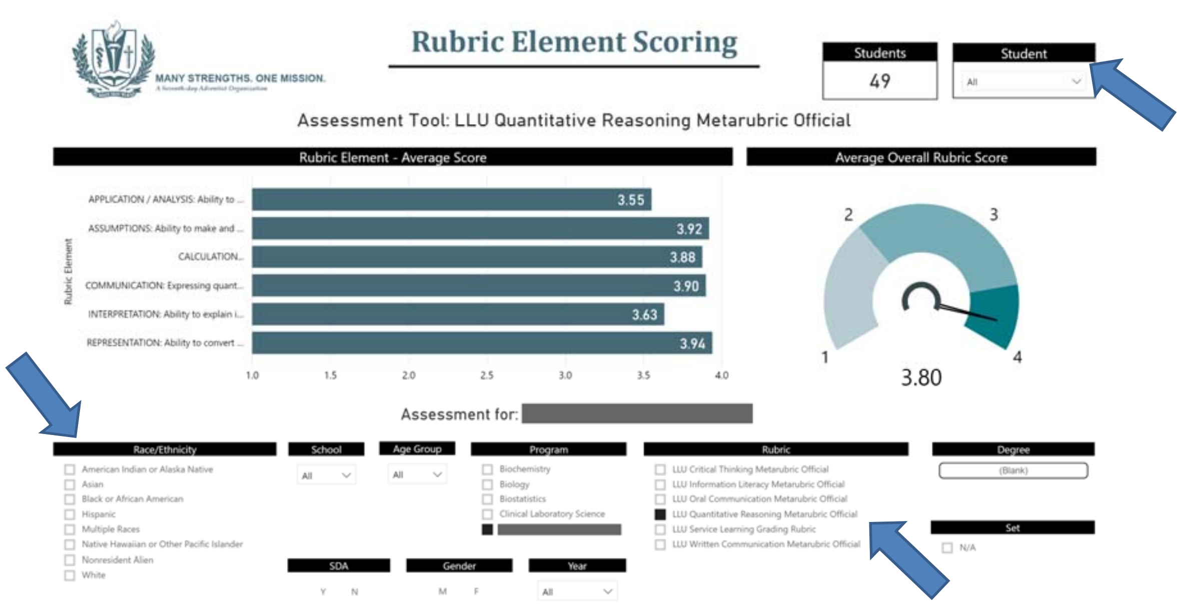
Rubric Element Scoring – Institution Example

- Solution –
  - Create small reports in the assessment software
  - Download csv of created report
  - Recombine using Power BI query editor



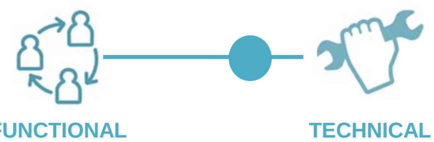


Rubric Element Scoring – Institution Example



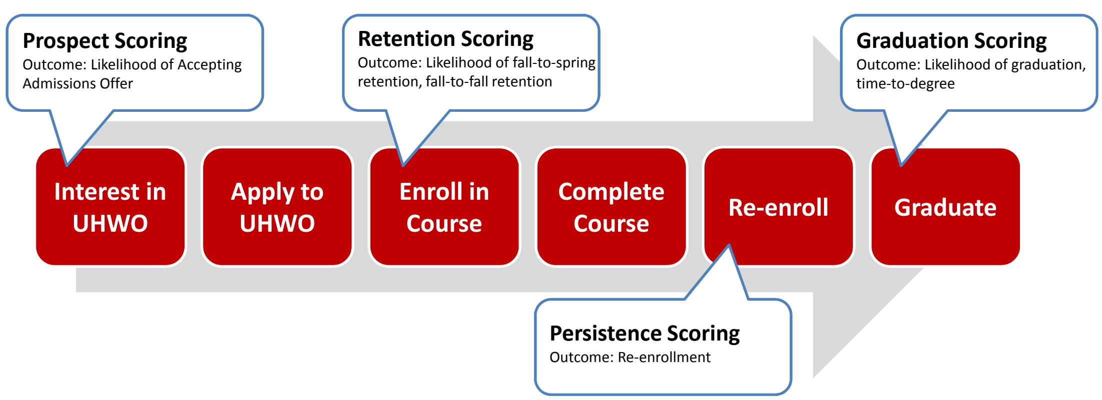
Analytics Application #2

Enrollment Management



43

Predictive analytics for the enrollment funnel



University of Hawai'i West O'ahu

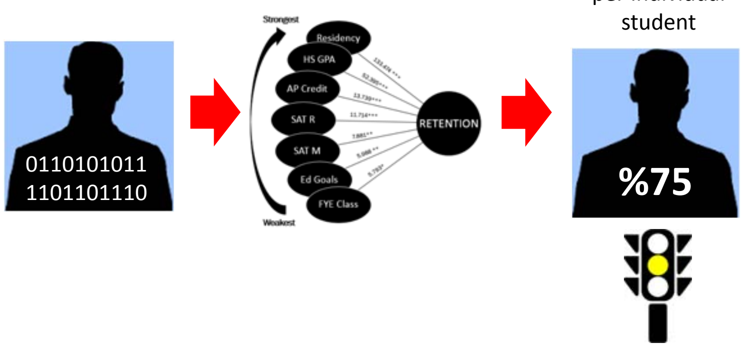
How does predictive modeling work?

Student Dropout Risk Example

1. Student data collected from SIS

2. Build prediction model using standard regression techniques

3. Prediction estimates are calculated per individual student



Student John Doe at the beginning of the semester:

- First Generation Student
- < Average high school GPA (3.00)
- Attempting 12 credits (12)
- Low % of financial need met (65%)
- Undeclared major
- Not enrolled in a campus learning community
- No educational goals in survey
- Not working on campus

Probability of Dropping Out:
75%

45

A taxonomy of SIS data available for prediction

RETENTION / GRADUATION

DEMOGRAPHICS
- Gender
- Age
- Ethnicity
- Residency
- Geographic Origin

CAMPUS EXPERIENCE
- On Campus Employment
- Housing
- Student Life Activities
- Athletics
- STAR Usage
- Average Class Size

PRE-COLLEGE
- High School GPA & Rank
- SAT
- AP CLEP
- Educational Goals
- Transfer GPA
- # Transfer Credits

FINANCIAL NEED
- Need Based Aid
- Non-need Based Aid
- Pell Grant
- Work Study
- % of Aid Met

ACADEMIC
- Major
- Credit Load
- Credits Earned
- First Term GPA
- Distance Education
- Dual Enrollment
- High Failure Rate Courses
- Courses Taken (including Math & English)

MILE-STONES
- Credits earned
- Credits attempted
- Credit Completion Ratio
- Math/English Enrollment/Completion
- Continuous Enrollment
- Milestone metrics

INTERACTIONS
- Ethnicity by Geographic Origin
- Employment by Housing
- High School GPA by First Term GPA
- Residency by Need Based Aid
- Ratio of Successful Adds to Drops

46

Sample data for advisors/success coaches

ID	LAST NAME	FIRST NAME	EMAIL	CURRENT CREDITS	RESIDENT	AP/ CLEP	HS GPA	WORK ON CAMP	1st YR EXP CLASS	% FIN NEED MET	STAR LOGINS	ADVISOR PREVIOUS CONTACT
001				15	HI	6	3.80	Y	Y	77%	5	Y
002				14	HI	0	3.33	N	Y	63%	3	N
003				12	CA	6	3.00	N	N	45%	0	N

ID	AGE	GENDER	ETHNICITY	COLLEGE	MAJOR	DEGREE	Ed Goal Specified	Relative Risk Value	Risk Level
001	18	F	CH	CA&H	ART	BA	Yes	14.92	LOW
002	18	F	HW	CSS	SOC	BA	Yes	36.88	MEDIUM
003	18	M	UNDEC	UNDEC	UNDEC	UNDEC	No	89.18	HIGH

Academic advising intervention example

John

- First Generation
- SAT-M/SAT-R = 900
- H.S. GPA = 3.00
- 85% Unmet Financial Need
- Undeclared
- 12 credits registration

- Dropout risk probability: **60%**
- Risk group: **6 of 10**

Intervention strategy:

- Proactive advising
- Meta major pathway mapping
- Revisit financial aid support and other support
- Check for possible ill-advised registration choices

Ken

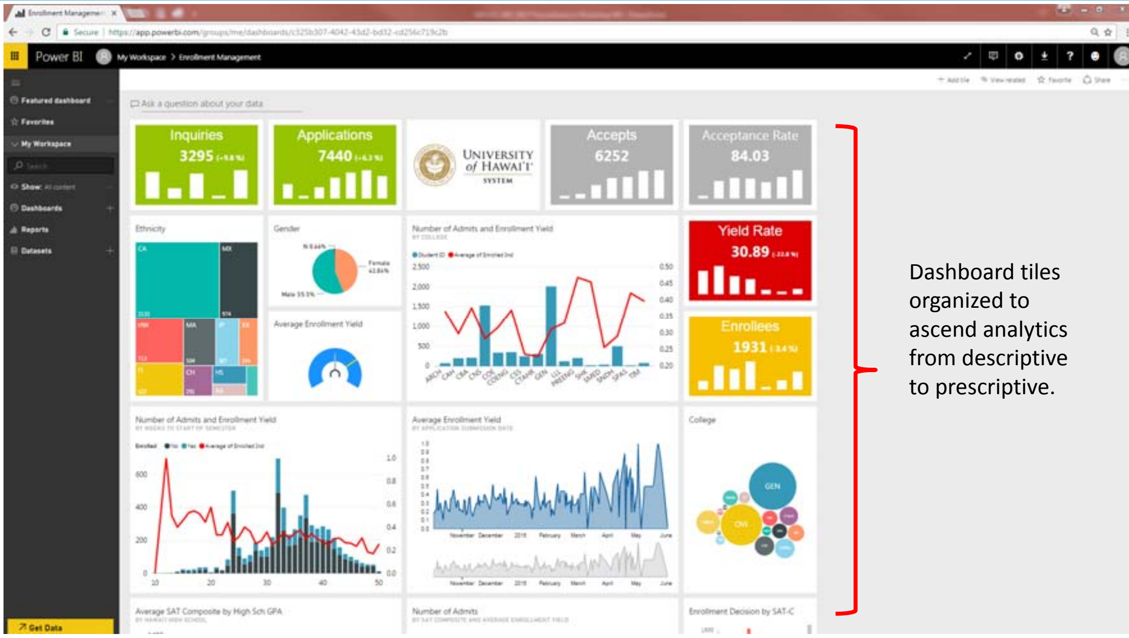
- SAT-M/SAT-R = 1100
- H.S. GPA = 3.50
- Declared major (Accounting)
- Local address within 5 miles
- 15 credits registration
- Educational Goals = "Earn B.A."

- Dropout risk probability: **15%**
- Risk group: **1 of 10**

Intervention strategy:

- Monitor *Starfish* reporting
- Mid-semester check-in
- Re-assess dropout risk at end-of-semester

Analytics system for enrollment mgmt.

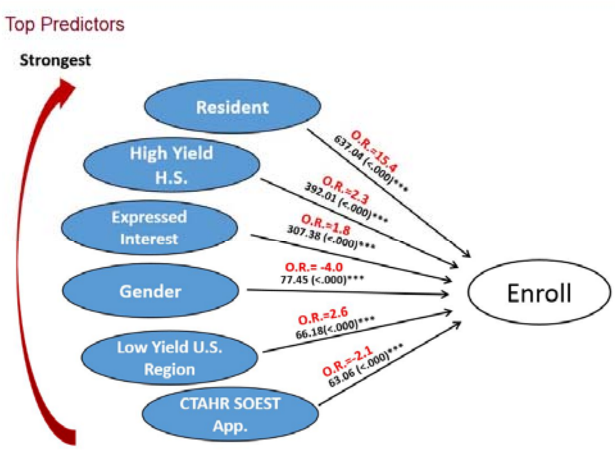


Dashboard tiles organized to ascend analytics from descriptive to prescriptive.

49

Predictive analytics

Fall 2016 Holdout Results
University of Hawai'i



These variables account for approximately 26% of the variance in a student's likelihood of completing a degree (Pseudo R Square = .263).

Model Summary

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
23658.222a	0.19	0.26

Classification Table

Outcome	Did Not Enroll	Enrolled	Percentage Correct
Did Not Enroll	3517	784	81.80
Enrolled	908	921	50.40
Overall			72.40

Parameter Estimates for Fall 2016 Enrollment Yield

Variables in Equation	B	S.E.	Wald	Sig.	Exp(B)
*Constant	-3.78	0.12	923.36	0.00	0.02
Resident	2.73	0.11	637.04	0.00	15.39
Historically high yield Hawaii HS	0.82	0.04	392.01	0.00	2.27
Expressed interest	0.57	0.03	307.38	0.00	1.77
Parent gross income less than 85k	0.48	0.03	210.28	0.00	1.61
Female	-0.29	0.03	77.45	0.00	0.75
Low yield US region	0.96	0.12	66.18	0.00	2.60
CTAHR SOEST applicant	-0.66	0.08	63.06	0.00	0.52
Nor Cal	-0.28	0.04	37.72	0.00	0.76
High yield Isle region	0.30	0.05	31.07	0.00	1.35
Application date 270 days or older	-0.19	0.04	25.00	0.00	0.83
Number of applications	0.15	0.04	13.98	0.00	1.16
Miles from campus over 20	-0.14	0.04	10.44	0.00	0.87

University of Hawai'i System

50

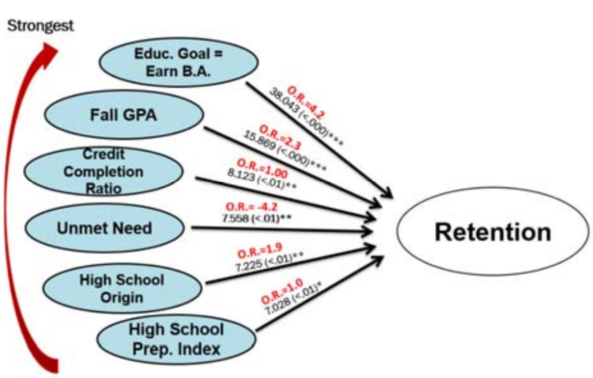
A prediction system for retention

Dashboard tiles organized to ascend analytics from descriptive to prescriptive.

Predictive analytics

UNIVERSITY *of* HAWAI'I
WEST O'AHU

Retention Prediction Model
Fall 2015 Cohort Holdout Results



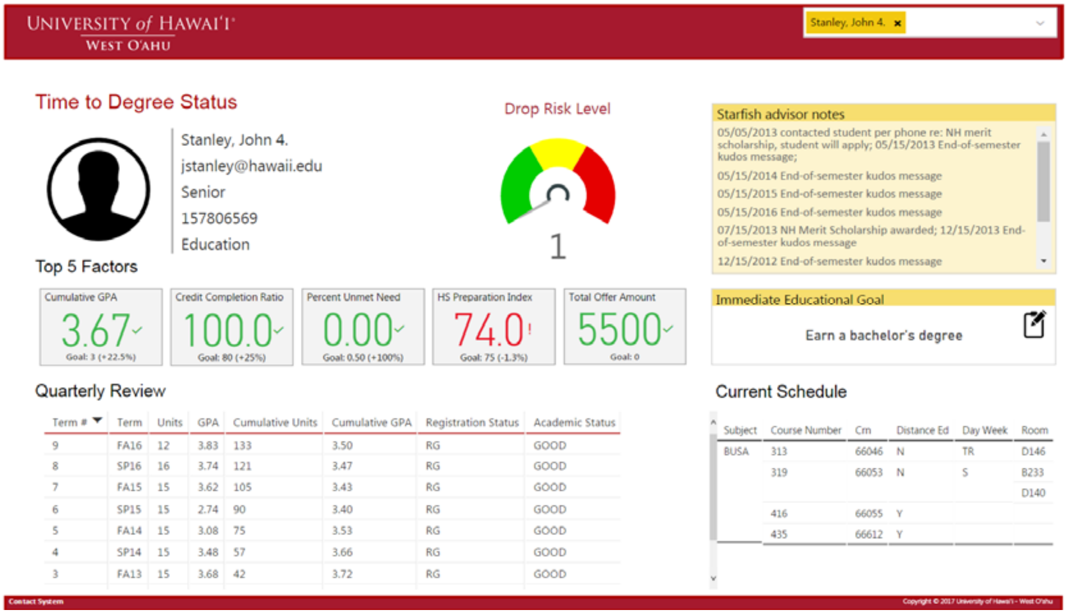
These variables account for approximately 47% of the variance in a student's likelihood of completing a degree (Pseudo R Square = .469).

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
491.806a	0.34	0.47

Outcome	Did Not Enroll	Enrolled	Percentage Correct
Did Not Enroll	44	18	71.00
Enrolled	33	136	80.50
Overall			77.90

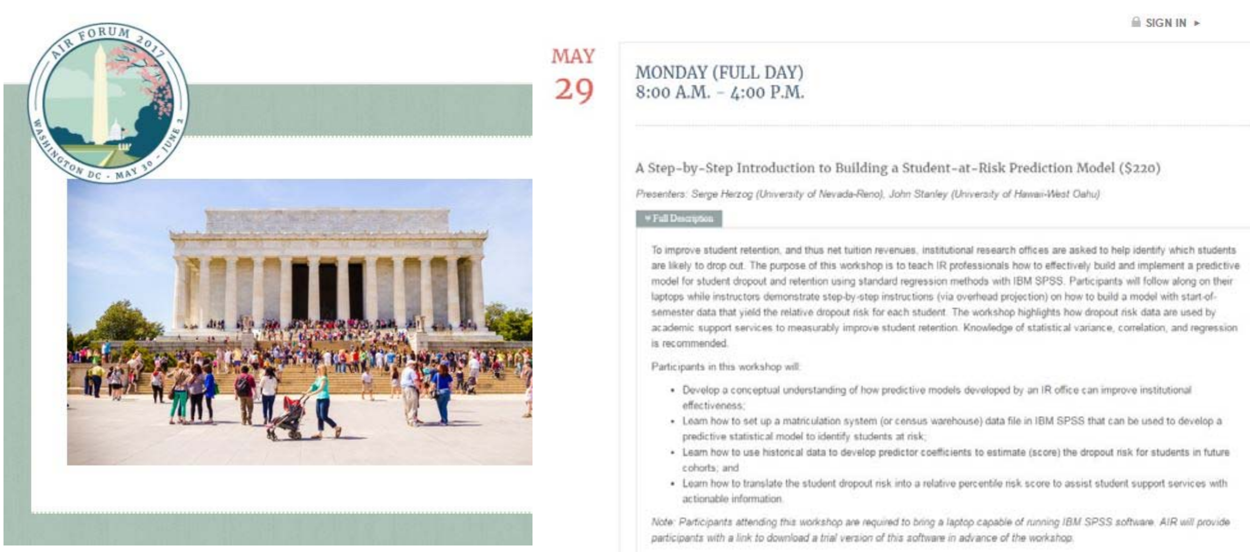
Variables in Equation	B	S.E.	Wald	Sig.	Exp(B)
*Constant	-2.73	1.30	4.40	0.04	0.07
Age less 18 Flag	0.29	0.27	1.16	0.28	1.34
Dual Enrolled UH Campus Flag	0.63	0.32	3.92	0.05	1.89
End-of-semester Fall CCR	0.03	0.01	8.12	0.00	1.03
End-of-semester Fall GPA	0.85	0.21	15.87	0.00	2.34
Historically High Retained HS	0.66	0.25	7.22	0.01	1.94
HS Prep Index	-0.05	0.02	7.03	0.01	0.95
Immed Ed Goal Earn BA End-of-Semester Flag	1.43	0.23	38.04	0.00	4.20
Math First Term Flag	0.30	0.24	1.54	0.21	1.34
STAR Usage End-of-semester	0.03	0.01	6.46	0.01	1.03
Total Offer Amount	0.00	0.00	6.59	0.01	1.00
Undeclared Major Flag	-0.42	0.24	2.99	0.08	0.66
Unmet Need Percent	-1.44	0.52	7.56	0.01	0.24

Prescriptive analytics



53

Technical training available for predictive analytics



AIR Workshop: <u>A Step-by-Step Introduction to Building a Student-at-Risk Prediction Model</u>

Instructors: Serge Herzog (University of Nevada-Reno) and John Stanley (University of Hawai'i – West Oahu)

54

Prescriptive analytics activity

Post hoc Activity

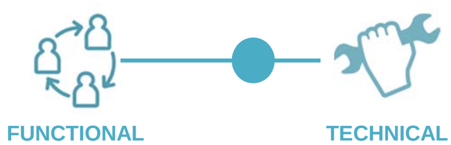
Read *Retention data use . . .*, see spreadsheet, and then

A. Generate list of reasons from multiple perspectives.
B. Draft techniques/strategies to address reasons.
C. Share with a neighboring table.

55

Analytics Application #3

Surveys



56

Power BI Solution for Qualtrics Survey Data

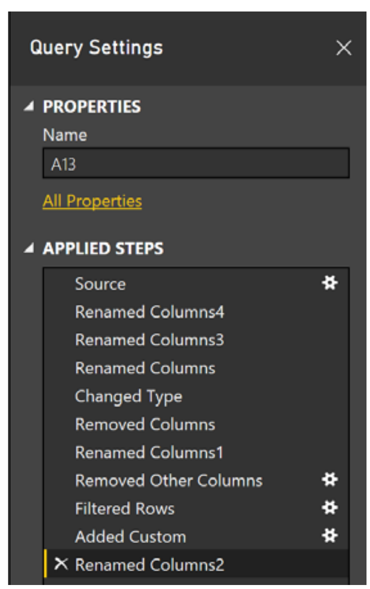
- Challenge – each row contained individual data elements
- Columns by individual numbered question or numbered answer
- Report needs to be updated as students continue to take survey
- Report needs to be disaggregated by school and program
- Comments need to be organized

```sql
SELECT TOP
      ,[Q11Quality]
      ,[A11QualityScore]
      ,[A28QualityComment]
      ,[Q12Values]
      ,[A12ValuesScore]
      ,[A30ValuesComment]
      ,[Q13Balance]
      ,[A13BalanceScore]
      ,[A31BalanceComment]
      ,[Q14Journey]
      ,[A14Journey]
      ,[Q15Experience]
      ,[A15Experience]
      ,[Q16Interprofessional]
      ,[A16InterprofessionalScore]
      ,[A32InterprofessionalComment]
      ,[Q17Respected]
      ,[A17RespectedScore]
      ,[A33RespectedComment]
      ,[Q18Overall]
      ,[A18OverallScore]
      ,[A34OverallComment]
      ,[Q21Recommend]
      ,[A21RecommendScore]
      ,[A35RecommendComment]
      ,[Q22Courses]
      ,[A22Courses]
      ,[Q24Socially]
      ,[A24Socially]
      ,[Q25Academically]
      ,[A25Academically]
      ,[Q26Professionally]
      ,[A26Professionally]
      ,[Q27Faculty]
      ,[A27FacultyScore]
      ,[A36FacultyComment]
  FROM [commonTablesV2].[dbo].[exitSurvey]
```

57

Power BI Solution for Qualtrics Survey Data

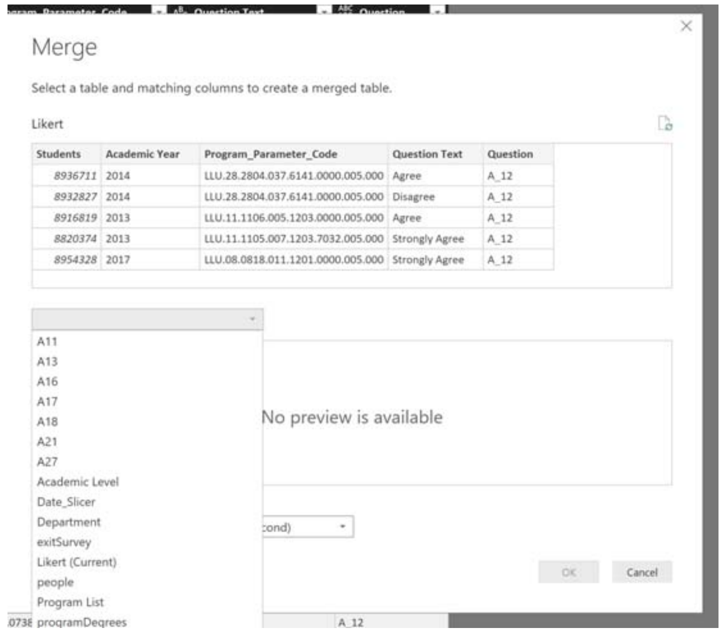
- Dataset is duplicated for each question
- Extra columns (other answers) deleted
- New columns for filters are added
- Columns renamed for clarity
- Each step is recorded, can be "played back" and edited at any time



58

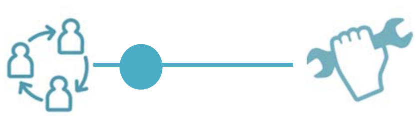
Power BI Solution for Qualtrics Survey Data

- Each copied table is then merged
- Power BI has options that will automate this each time source table is updated
- Resultant merged table now ready to be used for developing visualization



Analytics Application #4

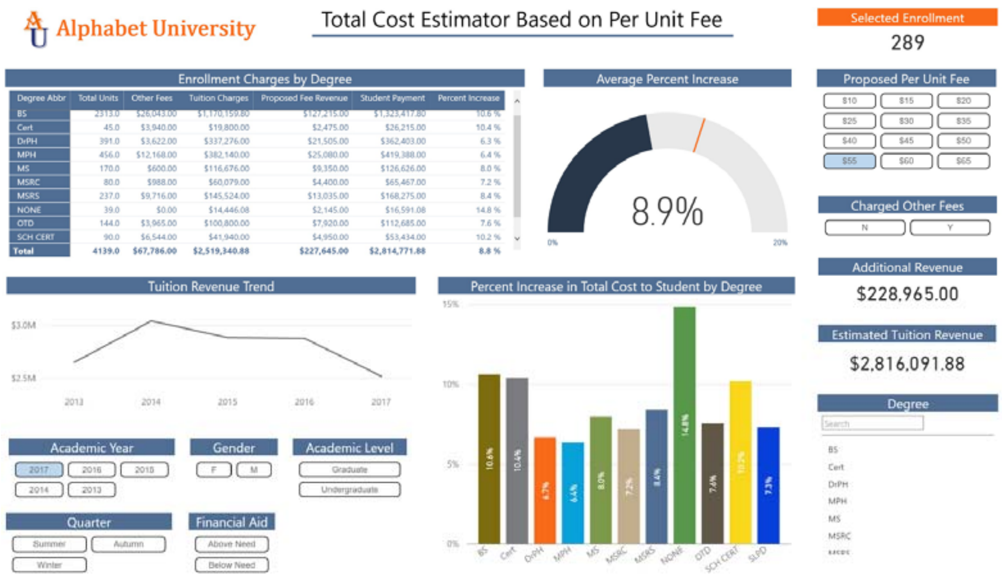
What-if Analysis



FUNCTIONAL TECHNICAL

60

Tuition revenue what-if planner



Degree Abbr	Total Units	Other Fees	Tuition Charges	Proposed Fee Revenue	Student Payment	Percent Increase
BS	2313.0	$26,043.00	$1,170,159.80	$127,215.00	$1,323,417.80	10.6 %
Cert	45.0	$3,940.00	$19,800.00	$2,475.00	$26,215.00	10.4 %
DrPH	391.0	$3,622.00	$337,276.00	$21,505.00	$362,403.00	6.3 %
MPH	456.0	$12,168.00	$382,140.00	$25,080.00	$419,388.00	6.4 %
MS	170.0	$600.00	$116,676.00	$9,350.00	$126,626.00	8.0 %
MSRC	80.0	$988.00	$60,079.00	$4,400.00	$65,467.00	7.2 %
MSRS	237.0	$9,716.00	$145,524.00	$13,035.00	$168,275.00	8.4 %
NONE	39.0	$0.00	$14,446.08	$2,145.00	$16,591.08	14.8 %
OTD	144.0	$3,965.00	$100,800.00	$7,920.00	$112,685.00	7.6 %
SCH CERT	90.0	$6,544.00	$41,940.00	$4,950.00	$53,434.00	10.2 %
Total	**4139.0**	**$67,786.00**	**$2,519,340.88**	**$227,645.00**	**$2,814,771.88**	**8.8 %**

61

What-if activity

Final Activity

Read *Cost Estimator Worksheet. . .*, see dashboard, and then

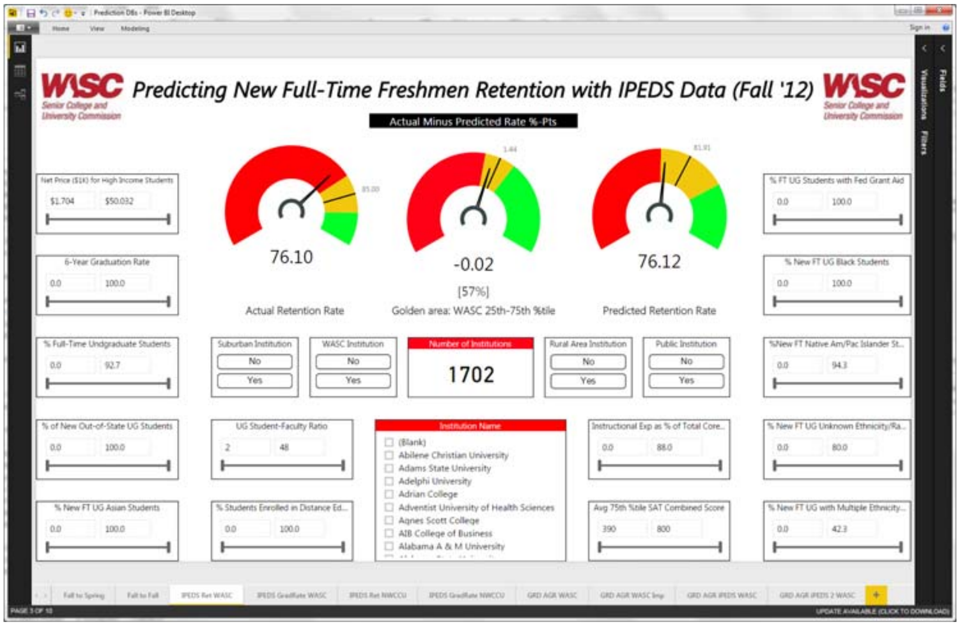
A. Generate a research question to answer using the parameters in the dashboard.

B. Discuss which parameters need to be checked and unchecked in order to visualize the answers.

C. Share with a neighboring table.

62

Institutional benchmarking using what-if



New WASC comparative dashboard under development. Compare actual versus predicted rates and perform what-if analyses.

ARC Session: Thursday, April 20, 2:15-3:15

63

Thank You

John Stanley
Director, Institutional Research
University of Hawai'i – West Oahu
jstanley@hawaii.edu

Kenneth Nelson, MD, MS
Associate Director, Educational Effectiveness
Loma Linda University
wknelson@llu.edu

64