UNIVERSITY OF KENTUCKY

see blue.

Student Success Analytics @ UK

Descriptive, Predictive and Prescriptive

Vince Kellen, Ph.D.
Senior Vice Provost
Analytics and Technologies
University of Kentucky
vkellen@uky.edu

April 4, 2015

Topics

- Definitions
- General philosophy
- Organization and architecture
- Community
- Examples
- Learning analytics

Definitions

- Descriptive, predictive analytics
  - Descriptive analytics describes what was. Predictive analytics describes what will be
  - Statistical methods can be used in both categories. It is what people do with the analytic method that makes them descriptive or predictive
  - People instinctually predict using descriptive data and descriptive data can be highly predictive (example later)
- Prescriptive analytics
  - Analytics that give directions, directives, recommendations, “nudges” or reminders
- Student success analytics
  - Analysis of data that evaluates student readiness, involvement, retention and progress towards the degree
- Student learning analytics
  - Analysis of data inside the class that evaluates student engagement with and mastery over learning outcomes and concepts in the class

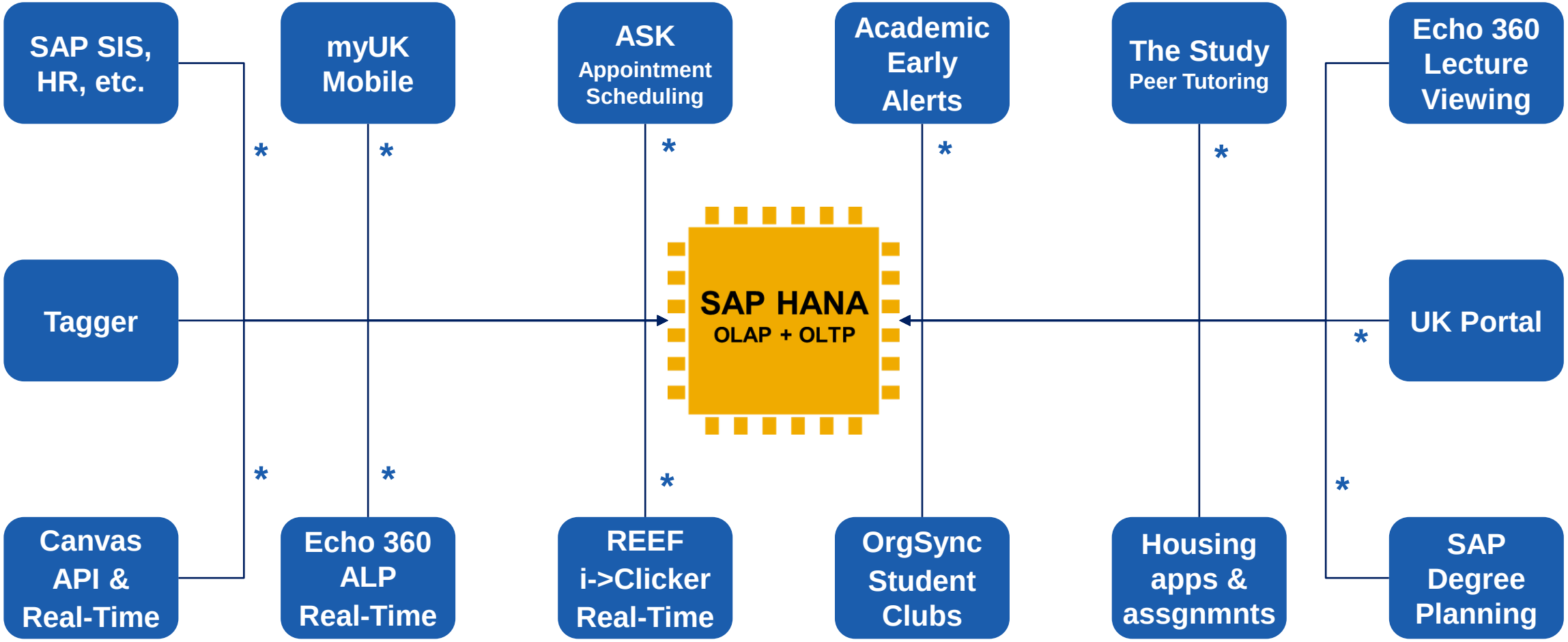
General philosophy

- Express maximum semantic complexity
  - Integrate all data from all relevant systems
  - Bring in all fields and attributes possible
  - Maintain data at the lowest level of detail
  - Respect security and privacy
- Integrate the myUK Mobile application
  - Use the myUK Mobile as the primary means for prescriptive analytics
  - Utilize micro-surveys to design interactions and learn about students
- Develop a strong community of practice
  - Empanel students (through student government) to collect feedback, provide input
  - Encourage analysis of data in colleges and support units (40+ active members, 100+ members)
  - ‘Community-sourced’ analytics

Analytic models

Model	Description
Enrollment	Enrollment in a class, midterm and final grades, credit hours attempted and earned, instructor teaching the class, class level, room building, time and meeting pattern
Student retention and graduation	Student demographics and cohort identification (e.g., John Doe is in the 2009 entering first-year student cohort), fall-to-spring, 1, 2, 3, 4, 5 year retention rates; 4, 5, 6-year graduation rates, incremental retention rates
Student demographics	Demographics, such as age, high school GPA, entrance test scores (SAT, ACT) and subcomponent scores. Also, in a secure location with limited access, demographic details such as name, address, email, financial aid data
Student performance	Presents the enrollment data in such a way as to easily show the student's performance for each term, including credit hours earned, term GPA, cumulative GPA for that term, etc., midterm and final grades, academic progress
Student academic career	List of the majors and minors for each student and degrees awarded, major and minor switching history, retention rates by program. Class correlation matrix (grades, entrance test scores, etc.). Details on students who transfer in and out, including transfer institution, credit hours transferred in, etc.
Productivity	Number of student, enrollments, credit hours and sections taught per term by faculty, department, program, college. Research expenditures per term and grant proposals submitted and won
Micro-surveys, other surveys	Capture questions and answers from the My UK Mobile micro-survey feature, annual freshman survey of non-cognitive factors, other ad-hoc surveys
Student involvement	Interaction history with various student software including myUK Mobile, UK's student portal, the learning management system, clickers, course capture and playback, academic alerts, student organizations, advising interactions, Tagger events, early LMS assignment grades
Facilities	Classroom capacity and enrollment (utilization) at 5-minute increments, how far students walk to/from classes

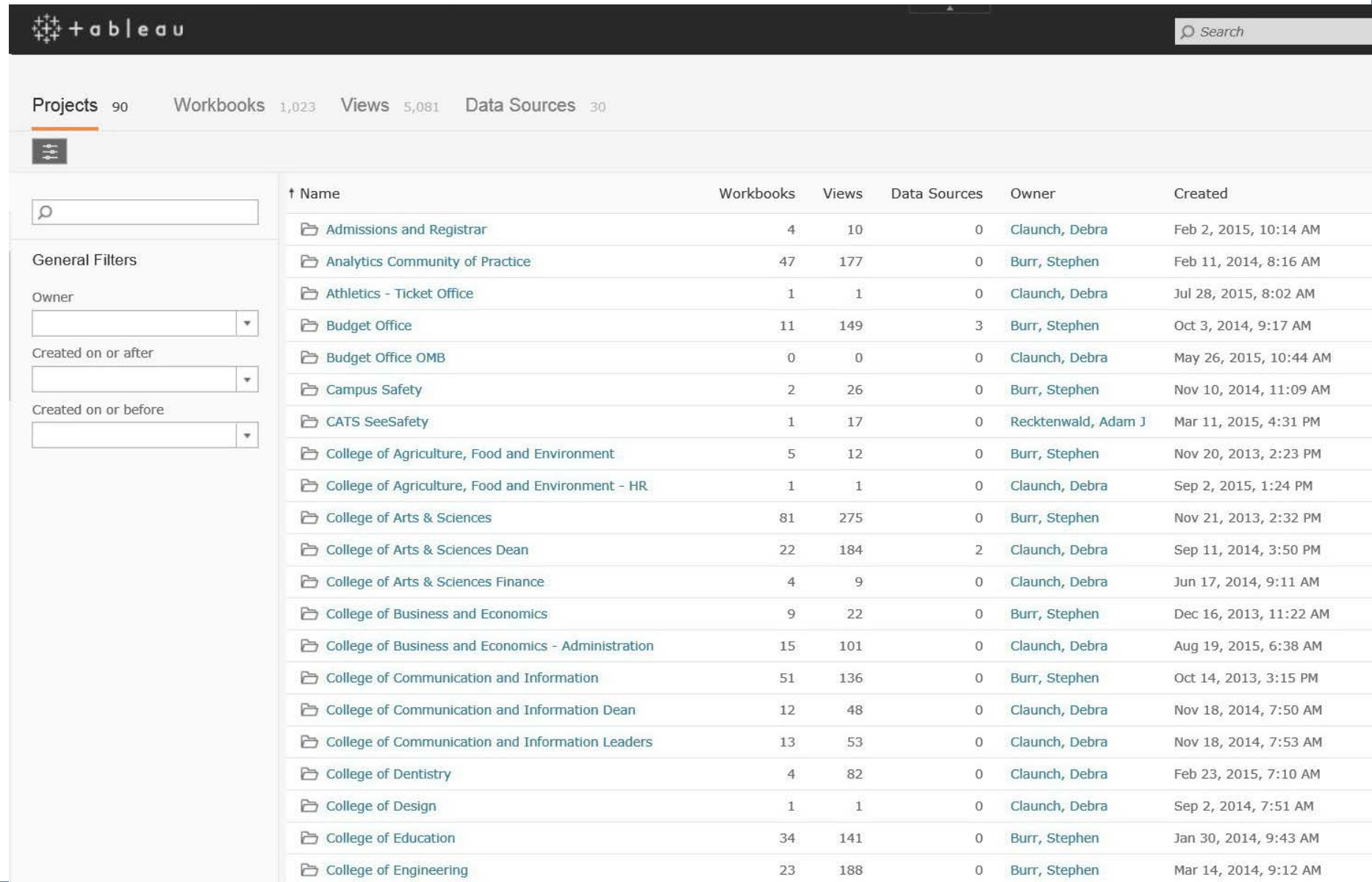
Source systems



** Real-time data feeds. All others are nightly or multiple times a day*

Tableau Web Server and our Data Analytics Community

- 90 folders for each college, business unit, etc.
- 1,023 workbooks with most created by the community
- 5,081 tabs in the workbooks
- High use of key workbooks (enrollments, retention dashboard, etc.)



Name	Workbooks	Views	Data Sources	Owner	Created
Admissions and Registrar	4	10	0	Claunch, Debra	Feb 2, 2015, 10:14 AM
Analytics Community of Practice	47	177	0	Burr, Stephen	Feb 11, 2014, 8:16 AM
Athletics - Ticket Office	1	1	0	Claunch, Debra	Jul 28, 2015, 8:02 AM
Budget Office	11	149	3	Burr, Stephen	Oct 3, 2014, 9:17 AM
Budget Office OMB	0	0	0	Claunch, Debra	May 26, 2015, 10:44 AM
Campus Safety	2	26	0	Burr, Stephen	Nov 10, 2014, 11:09 AM
CATS SeeSafety	1	17	0	Recktenwald, Adam J	Mar 11, 2015, 4:31 PM
College of Agriculture, Food and Environment	5	12	0	Burr, Stephen	Nov 20, 2013, 2:23 PM
College of Agriculture, Food and Environment - HR	1	1	0	Claunch, Debra	Sep 2, 2015, 1:24 PM
College of Arts & Sciences	81	275	0	Burr, Stephen	Nov 21, 2013, 2:32 PM
College of Arts & Sciences Dean	22	184	2	Claunch, Debra	Sep 11, 2014, 3:50 PM
College of Arts & Sciences Finance	4	9	0	Claunch, Debra	Jun 17, 2014, 9:11 AM
College of Business and Economics	9	22	0	Burr, Stephen	Dec 16, 2013, 11:22 AM
College of Business and Economics - Administration	15	101	0	Claunch, Debra	Aug 19, 2015, 6:38 AM
College of Communication and Information	51	136	0	Burr, Stephen	Oct 14, 2013, 3:15 PM
College of Communication and Information Dean	12	48	0	Claunch, Debra	Nov 18, 2014, 7:50 AM
College of Communication and Information Leaders	13	53	0	Claunch, Debra	Nov 18, 2014, 7:53 AM
College of Dentistry	4	82	0	Claunch, Debra	Feb 23, 2015, 7:10 AM
College of Design	1	1	0	Claunch, Debra	Sep 2, 2014, 7:51 AM
College of Education	34	141	0	Burr, Stephen	Jan 30, 2014, 9:43 AM
College of Engineering	23	188	0	Burr, Stephen	Mar 14, 2014, 9:12 AM

Analytics Community of Practice (CoP) Principles

1) **Be safe and secure**. Respect the acceptable use of information policies and guidelines the university has in place. Please have good passwords and secure your laptop, desktop and other devices appropriately. Treat private student and UK information appropriately.

2) **Be collegial.** University data is a community asset and a community of people steward the data. Use and share the data with the best interests of the university community in mind. Since parts of our data analysis environment is designed to allow for greater transparency, analysis will potentially be able to see other unit data. While we will make private to a unit what absolutely needs to be private, the way the university runs its business often involves multiple colleges and units at the same time requiring broad data access. Don't use your access to take unfair advantage of another unit.

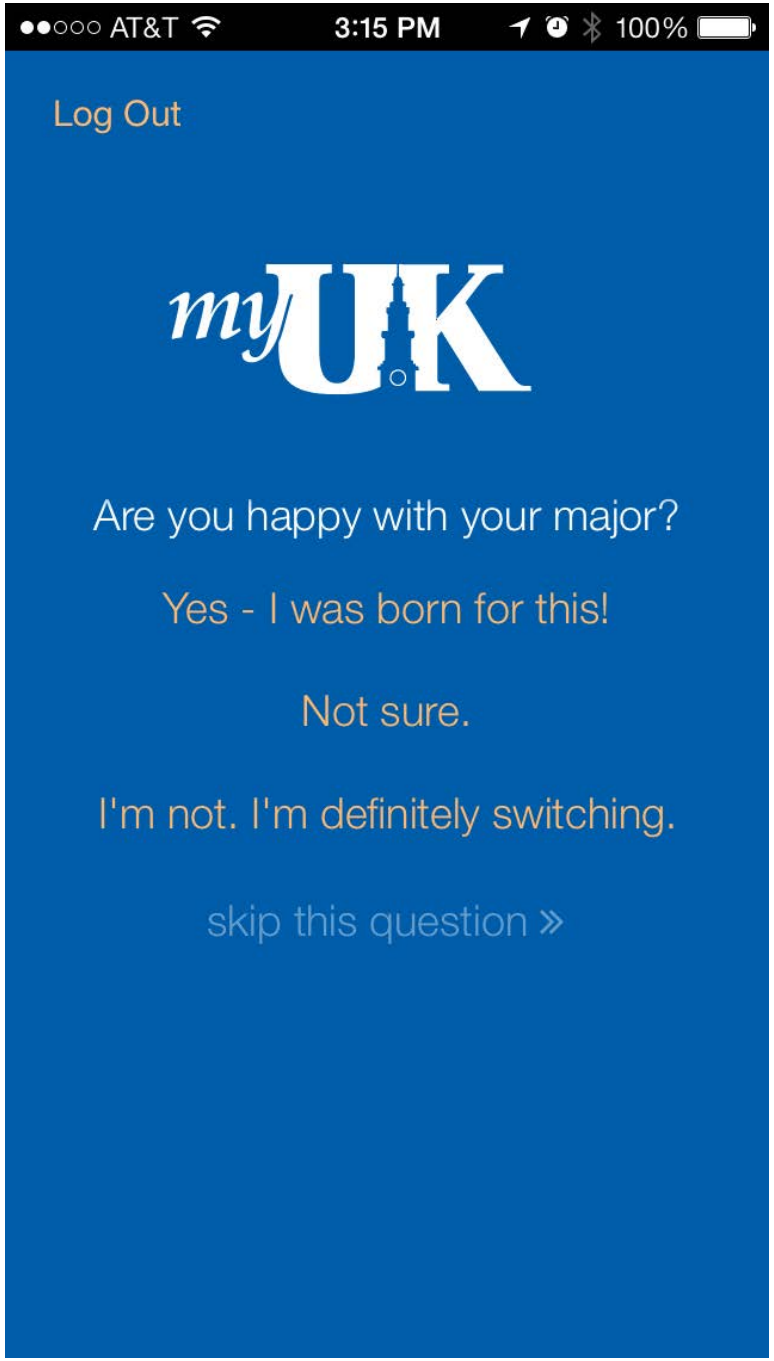
3) **Help improve data quality.** If you see data that doesn't appear to be correct, let someone know. We have a team of staff dedicated to helping improve data quality. This team can work with colleges and units on any data entry and data management processes that might need to be changed to improve data quality.

4) **Be open-minded and inquisitive.** Data can be represented in multiple ways at the same time. While the teams are taking great care to enable multiple views of the data to support the community, you might have a valid and unique perspective. In time, we can accommodate more ways of looking at the same data while not interfering with other views or taxonomies.

5) **Share.** The main benefit from open analytics is the power of a community of analysts learning from each other rather than a few select individuals hoarding knowledge or access. As the community improves its knowledge and skill with the data, the university can improve accordingly.

Mobile & micro-surveys

- We connected analytics to our mobile app to support personalized messaging
- We deliver simple 1-question surveys with the ability to trigger message responses and interactions
- In 24 months, we have gotten 181,285 surveys, averaging about 40-45% response rates



myUK Mobile Dashboard

Last updated 8/26/2015 6:08:41 AM

TOTAL APP USERS	TOTAL DEVICES	PUSH-ENABLED USERS	MICRO-SURVEY RESPONSES
29,803	52,355	21,794	181,285

5404 log-ins to the app today

1772 micro-survey responses recorded today

Users spend around **a minute and 2 seconds** in the app

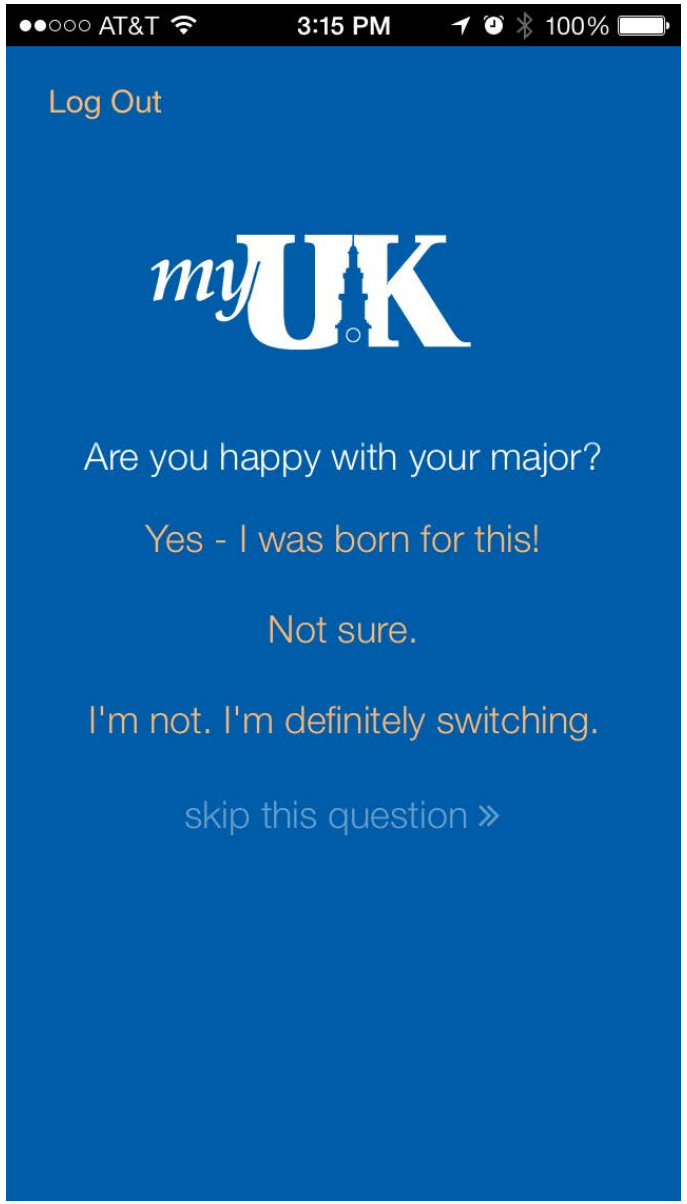
Mobile OS Breakdown

APPLE VS. ANDROID

92%

8%

Questions: Simple, Approachable

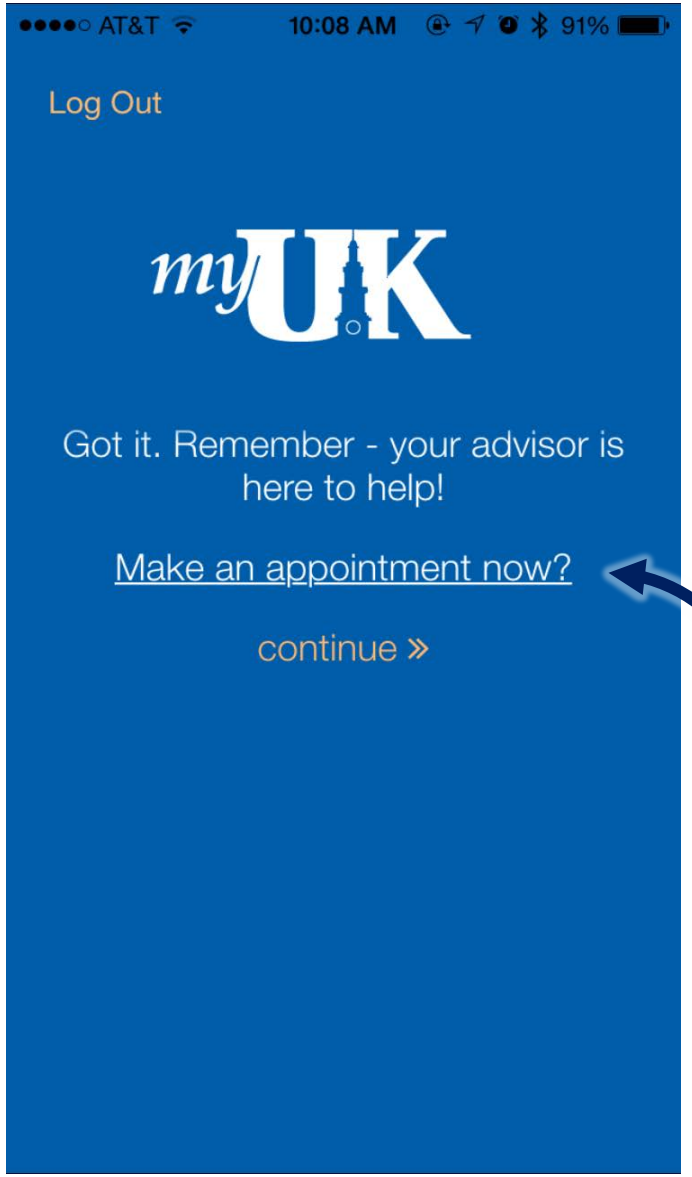


Upon login to myUK Mobile, a user has a chance (~ 1 in 4) of seeing a question that is designed for a simple read/respond reaction.

The question's 'voice' is carefully crafted to make the question approachable and the response options natural.

◁ Select an option…

Responses: Telling, Useful

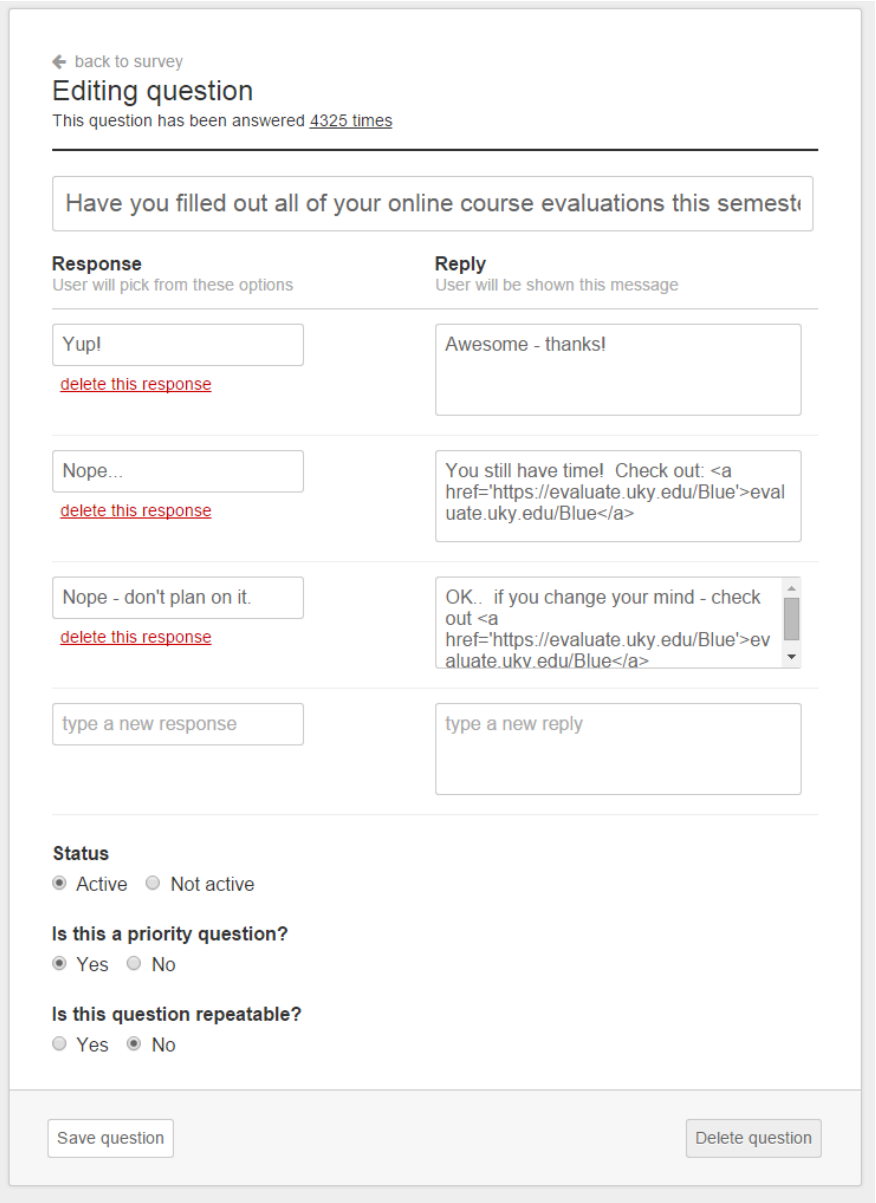


Each ***response-option*** has a full page ***reply.***

The **replies** are designed to be appreciative, understood quickly and *get-out-of-the-way*.

Responses can contain personalized actions!

- Make an appointment with advisor / counselor
- Link to more information
- Signup for an event
- - Anything -



Creating Questions

A administration tool is used to create **questions**, **response-options** and corresponding **replies**.

A reply can include links or other personalized elements.

Questions can be flagged as:

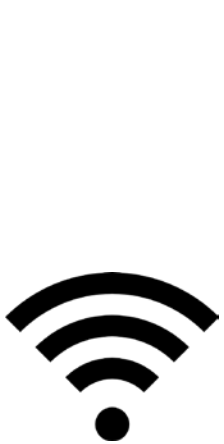
*High Priority**: It will be presented upon the next login

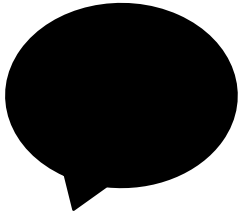
Repeatable: Asking the same question over a period of time (e.g., “Are you stressed?” each 60 days)

Grouped: Presented to a small segment of the population (default: ‘all students’)

*push notifications can also be used to encourage login to the app – very useful for High Priority questions.

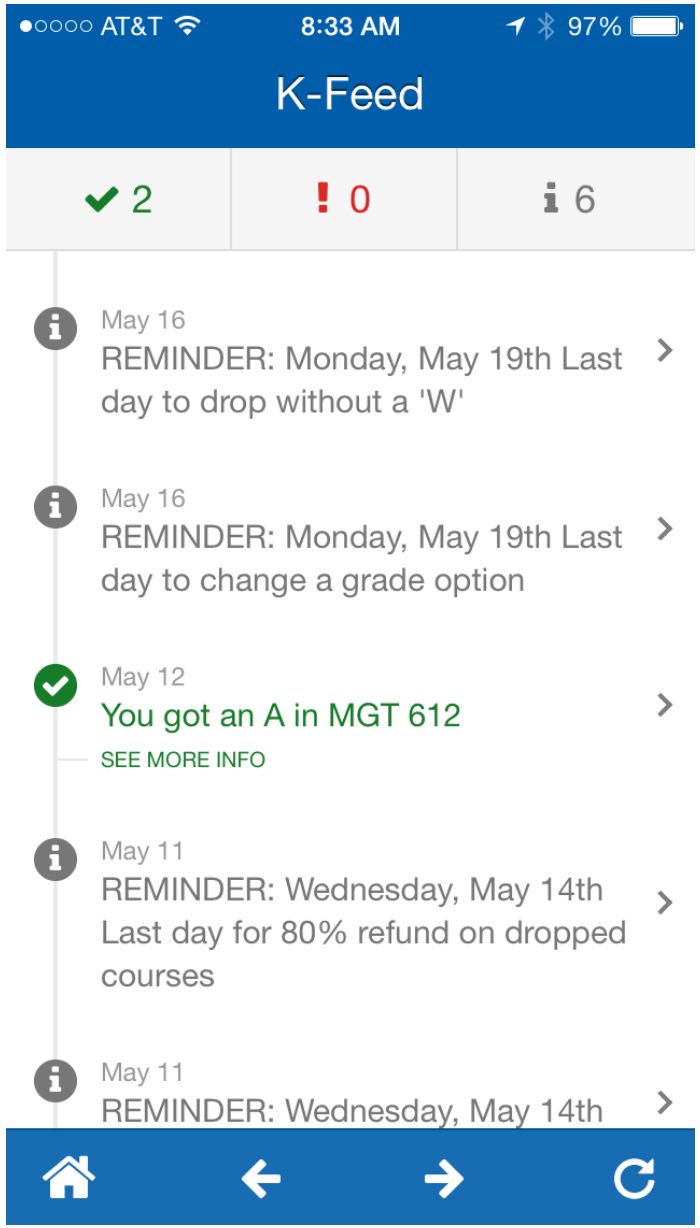
Tagger: General purpose event attendance

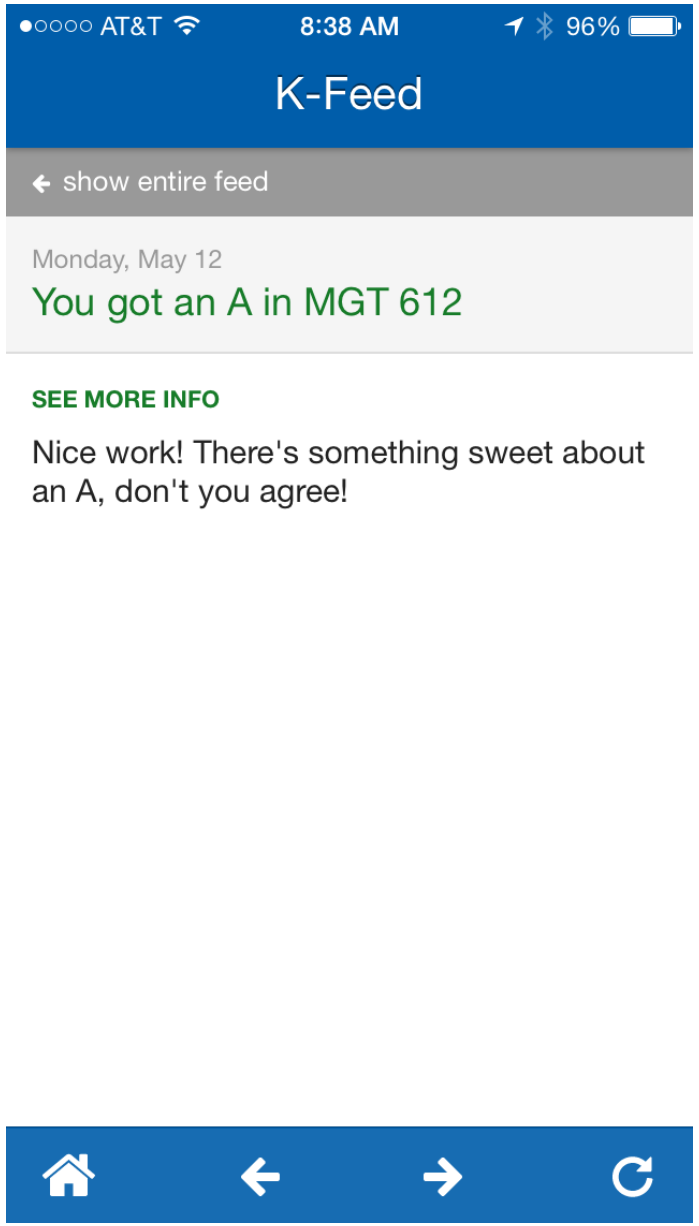




K-Feed

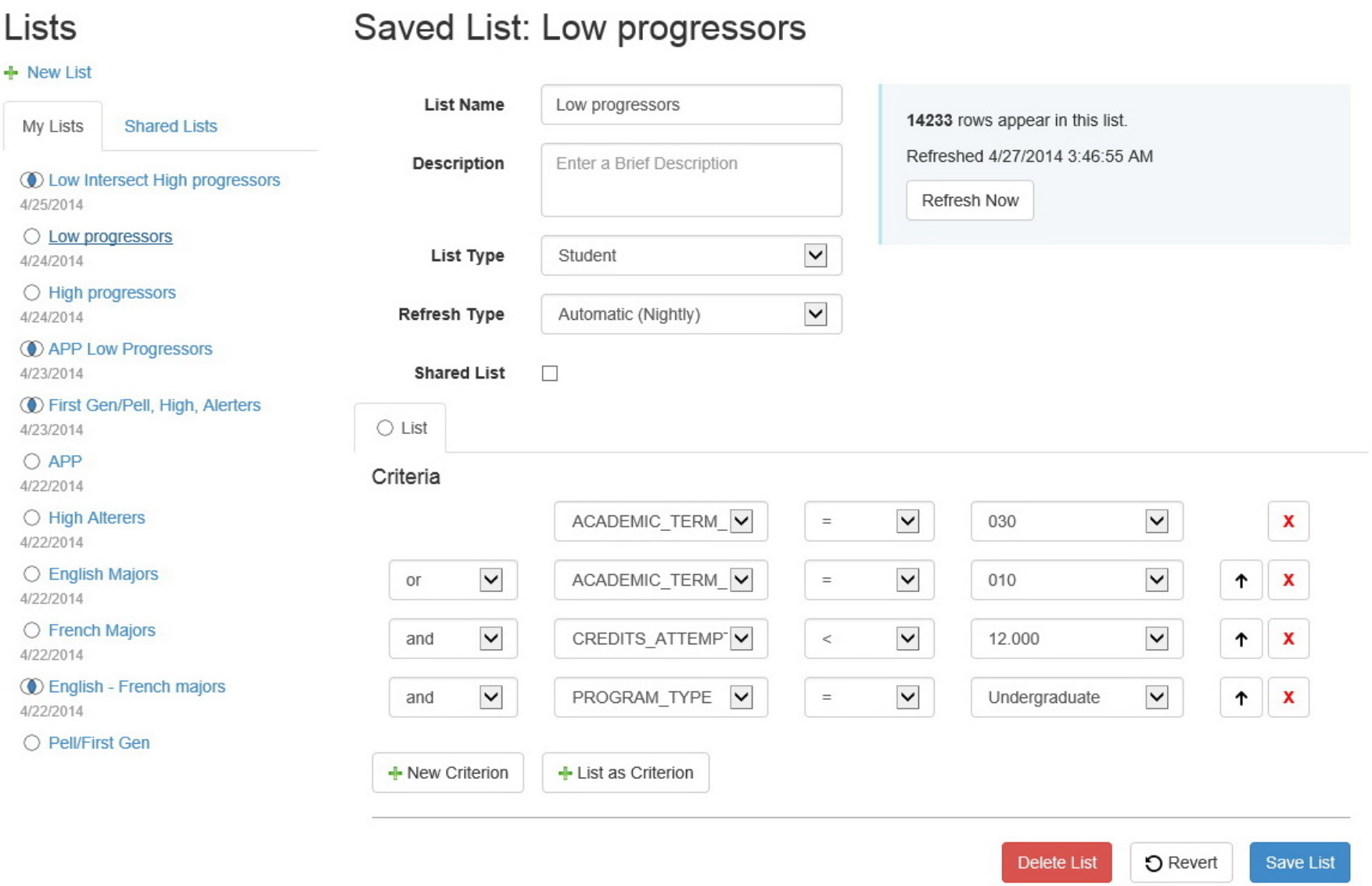
Intelligent, personalized alerts, news, reminders





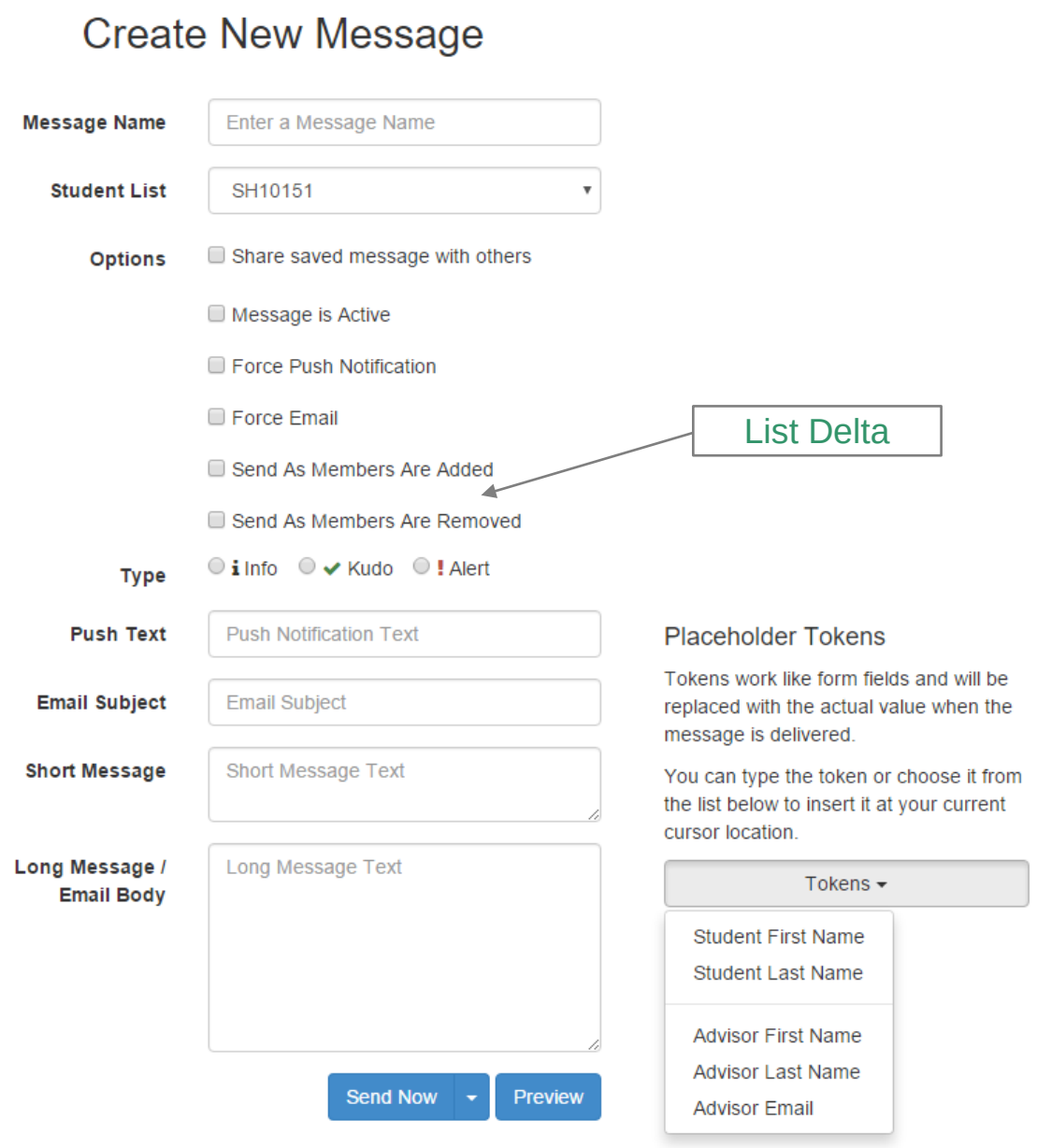
List builder

- Iteratively query any/all fields of your choosing, linking in using AND or OR Boolean logic
- Combine different lists using SET operations
- Refresh lists regularly (nightly or more frequently, called List Delta)
- This provides advanced filtering and combining that can be shared across multiple dashboards, visualization tools
- Since lists are refreshed nightly (or more frequently), we can keep track of each time a student (or other entity) is added or removed from a list, triggering personalized message



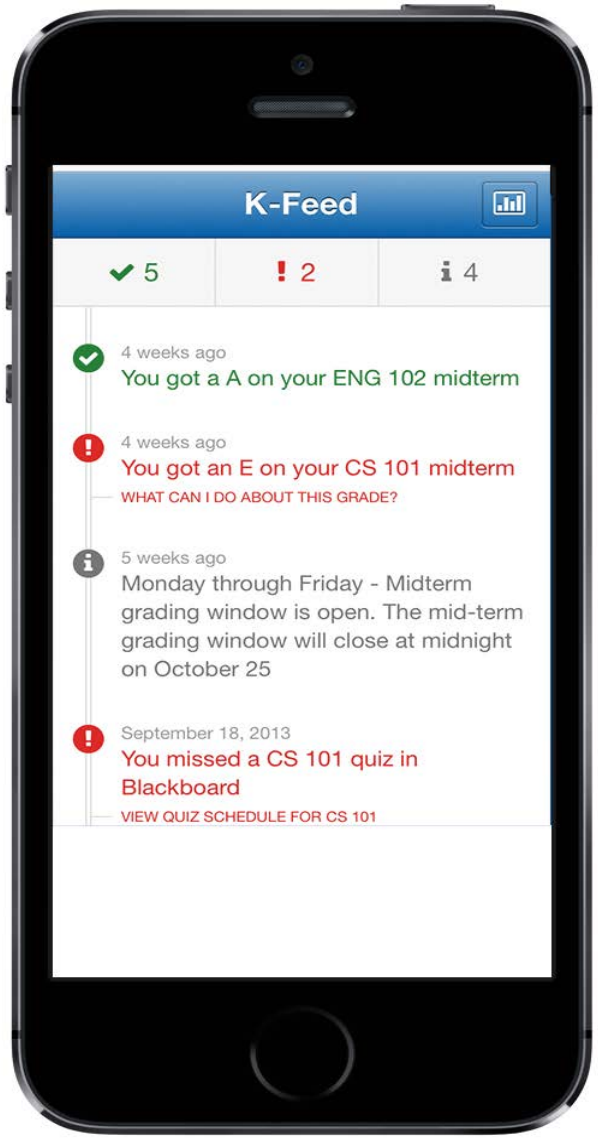
Message Builder

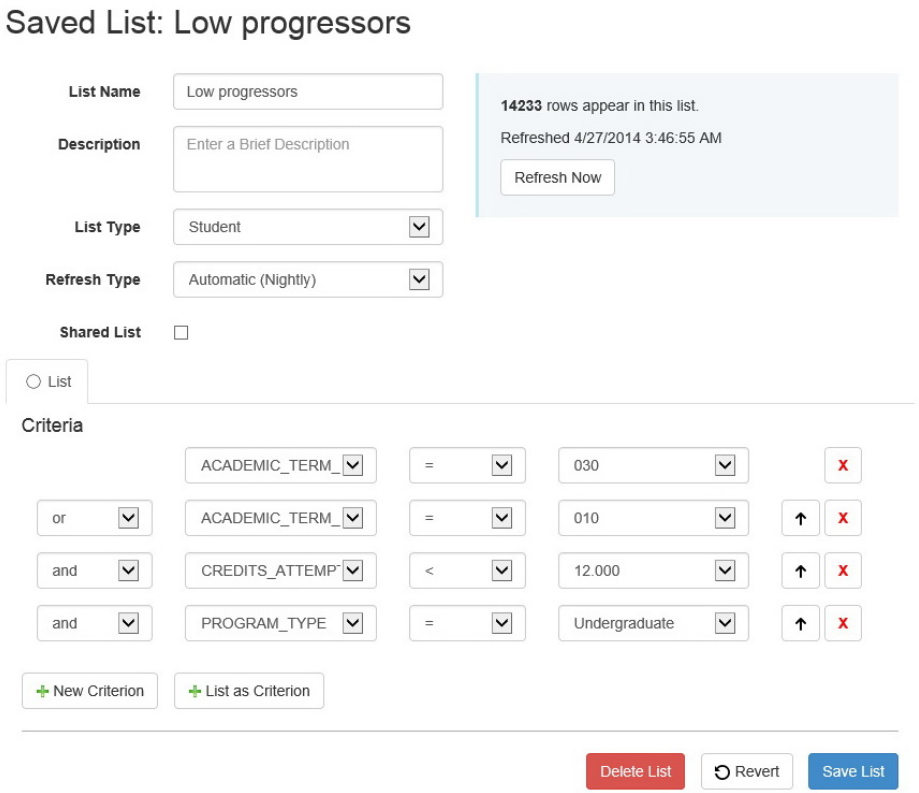
- Using lists created, enable personalized messages to be sent to students, faculty, staff
  - Email
  - Mobile app K-Keed
  - Mobile push notification
  - Other applications
    - Learning management system
    - Active learning systems
    - Web sites
- Using List Delta functionality
  - Send messages as soon as the query rule “fires” and adds or removes someone from the list
- List Builder, List Delta and Message Builder can respond to changes in ANY attribute ANYWHERE in the system
  - Data-driven universal workflow working at the platform layer, not the application layer



We would like to ‘nudge’ students along

- End-user micro-segmentation of students
- Set up automated, “trigger based” messages
- Triggers are completely user-defined based on combinations of 800 variables (ListBuilder / ListDelta)
- Messages to the mobile app (push notifications, K-Feed item) are push to micro-segments of students based on ListBuilder / ListDelta





Cohort registration

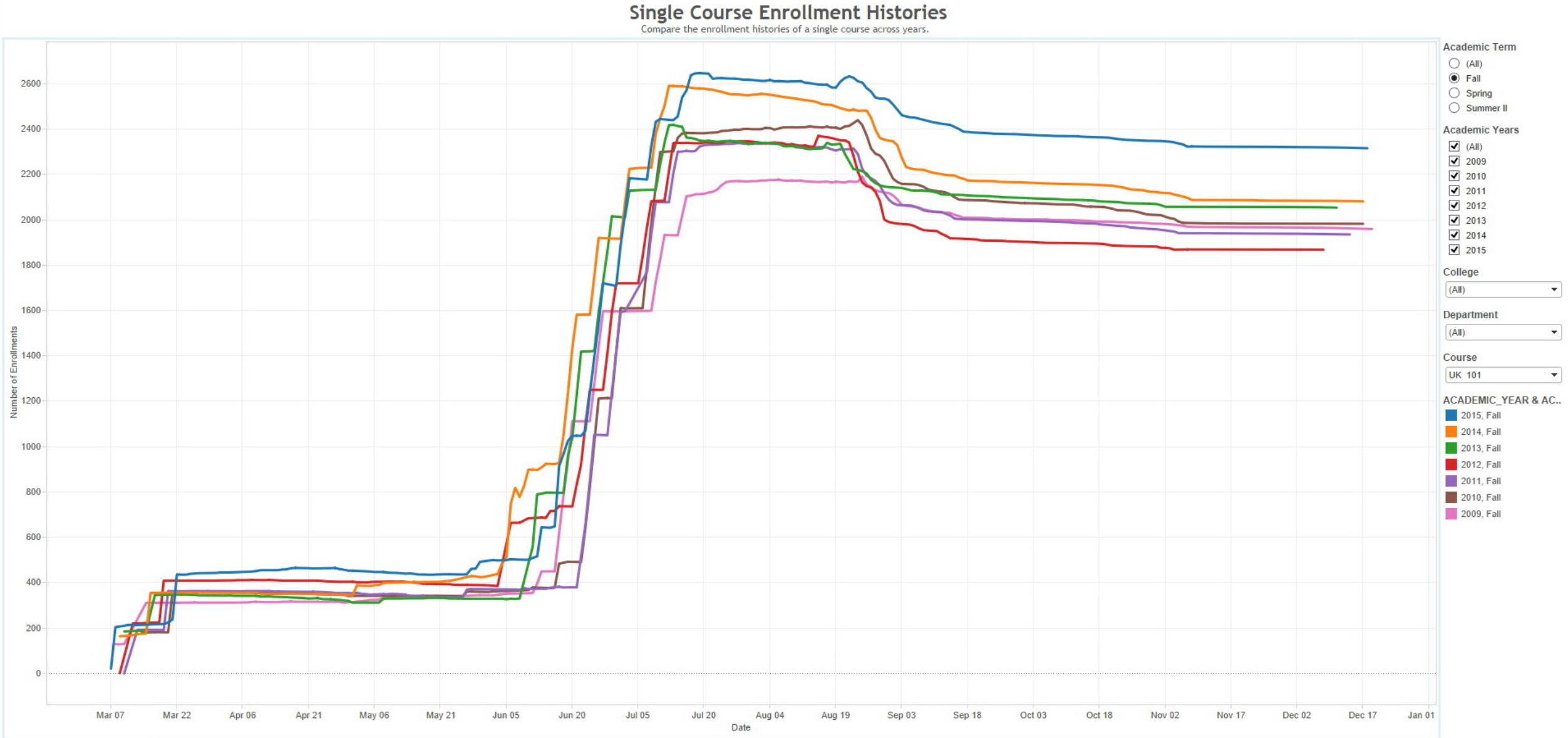
- Cohort Registration is the idea that we'd like to group students and classes to make as many 'common paths' as possible. Last semester with tried this with a small group of students – by hand… a painful process for humans to work out. One of the most tedious issues is to find combinations of classes and sections that do not conflict with others. Doing this with two or even three large classes isn't terribly difficult – but five or six with variability in capacity and suddenly the problem is much more complex.

- We're using Hana to compute all possible combinations between grouping of classes (20 million rows) and recording only those that 'fit' together. This 'fit' includes time and walkability and soon will include student preference (morning, mid-day, afternoon), student-instructor fit, enhanced peer-grouping and other student and class factors.

- The seemingly simple process of finding those non-conflicting classes for a few hundred students is going from weeks to 3 seconds. It took one developer a couple of days to get the core process working. We avoided purchase and implementation of a third party software package

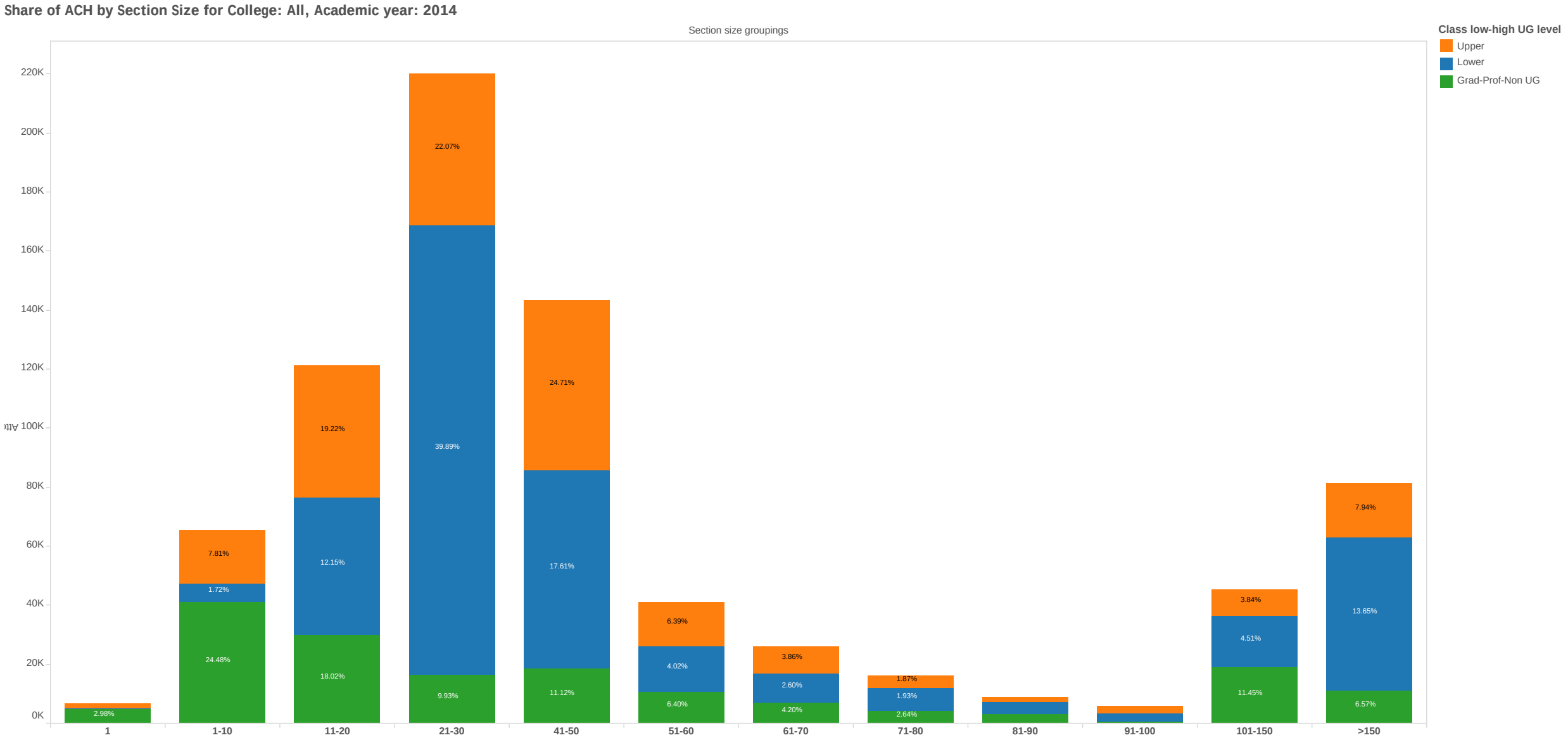
- In the future, this will also tap into our planned n-dimensional performance correlation matrix…

Course enrollment “curves” (aka, Arnie’s S-Curve)



UNIVERSITY OF KENTUCKY® see blue.

Share of credit hours by section size



Student readiness and frequency of the use of “The Study” and 2nd year retention rates

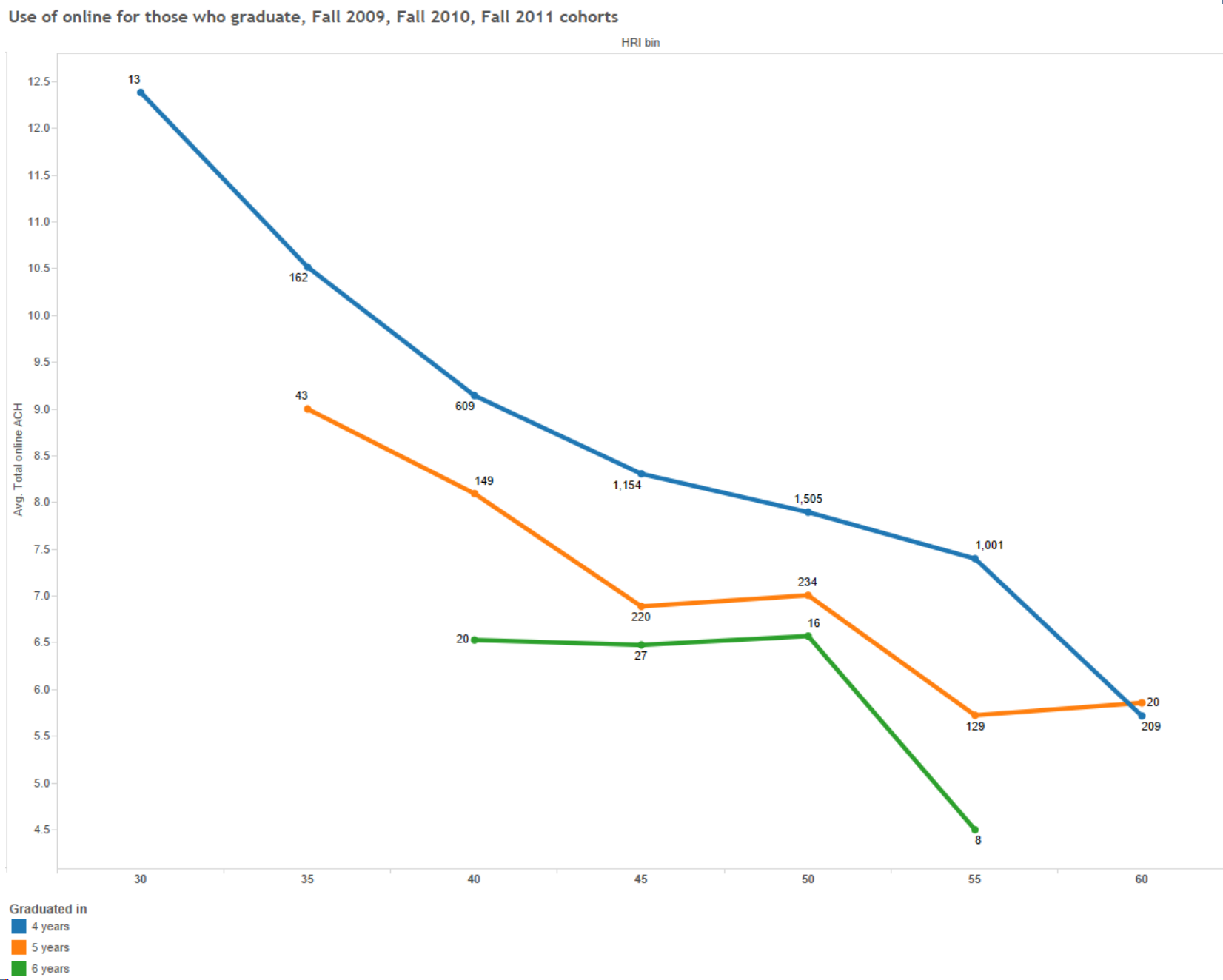
2nd Fall Retention by HS Readiness and Study Attendance

Readiness groups

- < 23
- 23-27
- > 27

Use of online for those who graduate

- Students who graduate in 4 years take more online classes than students who graduate in 5 or 6 years
- Students who have higher readiness (HRI) tend to take less online classes than students with lower readiness



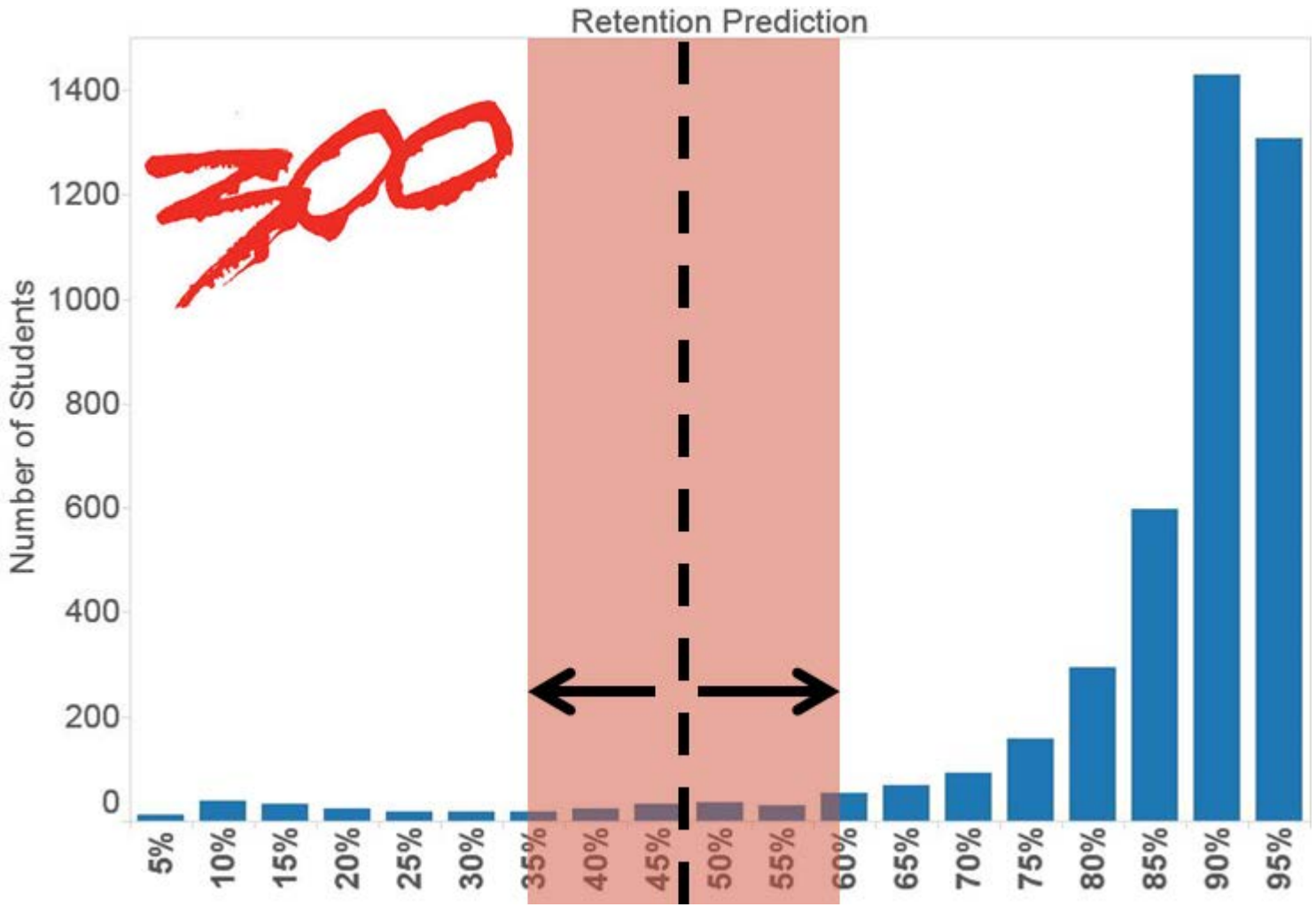
Targeted student interactions

In addition to our work on difficult student cases, we needed to find a way to reach a ‘murky middle’ group of students

We identified students who are just as likely to come back as they are not

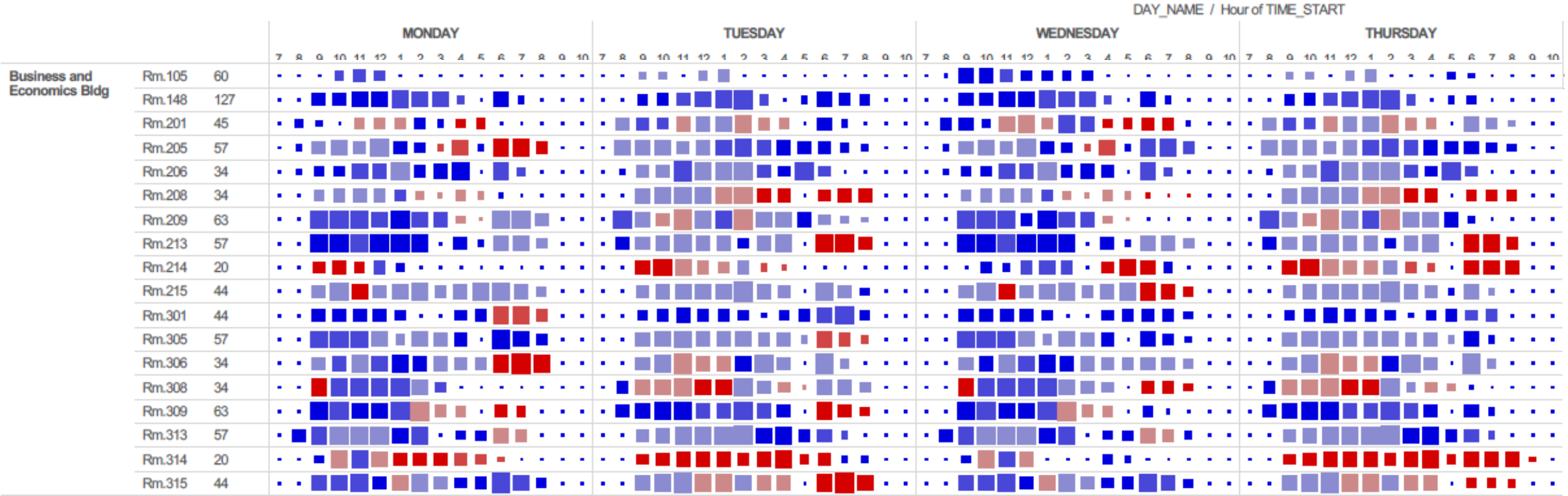
The predicted reenrollment was about 50%

After interventions, the actual enrollment was about 65%

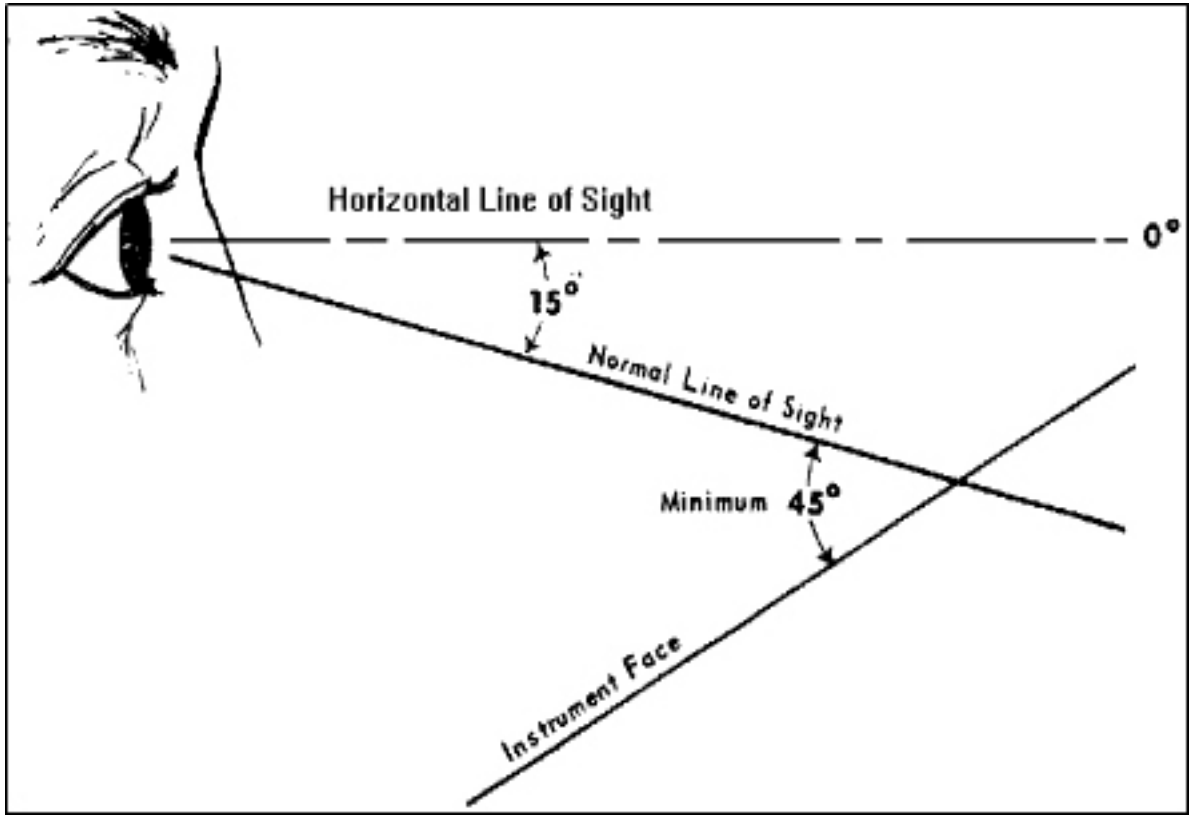


Classroom utilization

Room detail - one week



Learning Analytics: We need ‘line of sight’ to the student



- Right now I can see you nodding off in the back of the room and can adjust my approach, real-time and in a personalized manner, calling you out and giving you a strong verbal ‘nudge’
- Just because we go digital or we use tools from vendors does not mean we should lose line of sight to the student or that we can’t respond within seconds
- Also, teachers have to know what works and what does not in the digital realm. The learning analytics should be a public good, not a private one, IMHO

Uber Tutor: Confusion detector

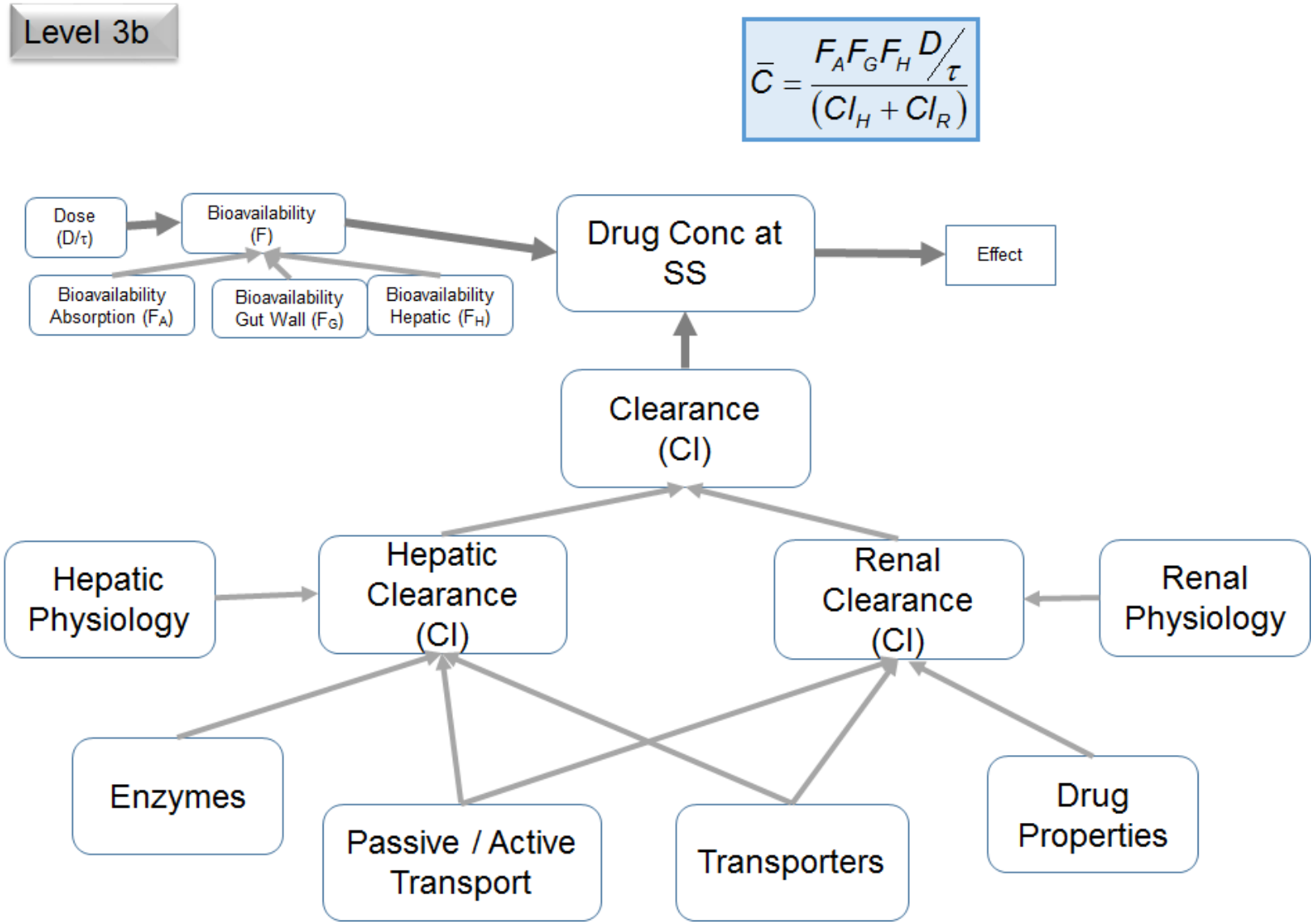
In the middle of working online, a student searches for “dorsolateral prefrontal cortex” in the online course content, replays the associated segment in the video lecture and has gotten 4 out of 5 related quiz questions incorrect.

1. Through content analysis, real-time analytics detect the confusion regardless of the type of content and notifies the peer tutor.
2. The real-time analytics service lets the student know that a peer tutor is available and will be calling
3. The peer-tutor and the student discuss the concept “dorsolateral prefrontal cortex” and arrange to meet later to explain
4. The real-time analytics service places a entry in the students advising log and notifies the faculty

Can we help learners build proper mental models?

Natural language processing (NLP) and neural-networks. Can these classes of technology:

- Extract concepts from audio and text?
- Find highly relevant articles, scientific papers and discussions regarding these concepts?
- Help learners construct and validate their knowledge?



What learning analytics innovations are others working on?

- Unizin (consortium focused on Instructure Canvas value-add services)
  - Confusion detection
- UC Berkeley
  - Real-time “engagement score” for in-class gamification, Canvas add-on component
- University of Texas
  - NLP tool to detect when a discussion thread is going off the rails and automatically intervene
  - Biometric data for evaluation of affect in student learning
- University of Michigan
  - Identify students for follow-on faculty 1:1 engagement
- University of Kentucky
  - Instructor defined in-class triggers to send ‘nudges’ to students via mobile app
  - Real-time clicker data collection within the real-time analytic environment

Analytic maturity scale: Rate your organization's abilities

Capability \ Level	Level 1 Disorganized	Level 2 Managed	Level 3 Autonomic
Technical	• IT silos • One-time analyses • Non-integrated data • Brittle • Problematic data quality • Slow analytic turnaround	• IT priority • Reusable analytic components • Centralized data warehouses • Modular BI components • Managed data quality • Ongoing analytic projects • Enterprise analytic architecture	• Analytics inside IT services • Real-time analytic platform • Common use analytic models • Federated analysis • Heterogeneous access • Fast data quality cycles • Enterprise integration architecture
Organizational	• Unit silos • Between-group infighting • Non-transparent • Reliance on a few analysts • Analysis role ambiguity	• Mandated dashboards • Unit-level measurement • Between-group competition • Central analysis services • Defined analysis roles • Defined unit/function incentives • Production orientation • Numeric preference	• Variegated dashboards • Capacity-level measurement • Between-group collaboration • Community-sourced analytics • Maximal transparency • Distributed analysis roles • Intrinsic motivation • Consumption orientation • Visualization preference
Strategic	• Shallow metrics • Inward orientation • Low quantitative orientation • Fact-free decision-making	• Balanced scorecards • Performance orientation • Benchmarking • Data-centric strategic planning • Periodic strategic review	• Dynamic, complex metrics • Value creation orientation • Competitive intelligence • Shorter planning timeframes • Strong operation-strategy link • Quick top-down • Flexible bottom-up

There's the rub!

- Disagreements between units on campus about control over data and technology
  - Data is a community property, knowledge is a community product
  - Data silos make this type of analysis shared here nearly impossible
- Individual and organizational defensiveness
  - People often have trouble dealing with disconfirming data. The need to maintain a 'popular narrative' even when the data disagree is a universal problem
  - Transparency! Sharing assumptions and the reasoning process for making conclusions, staying close to other independent research and allowing for individuals to independently replicate results (the scientific method) are all important
- Students should do this on their own! We shouldn't be looking at this data…
  - Make clear your intent (student success). Students want institutions to help them, in a caring way
  - Use transparency to keep multiple eyes on the data and the usage of the data
  - The perspective differences between administrators and students is real and problematic
  - Respect all privacy policies

Questions?