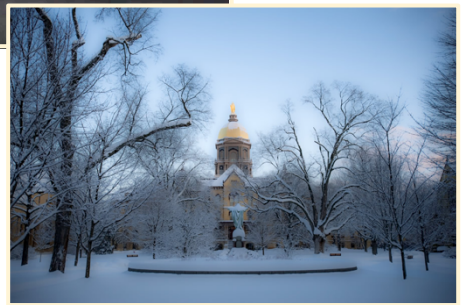
A Deep Dive into Learning Analytics at Notre Dame

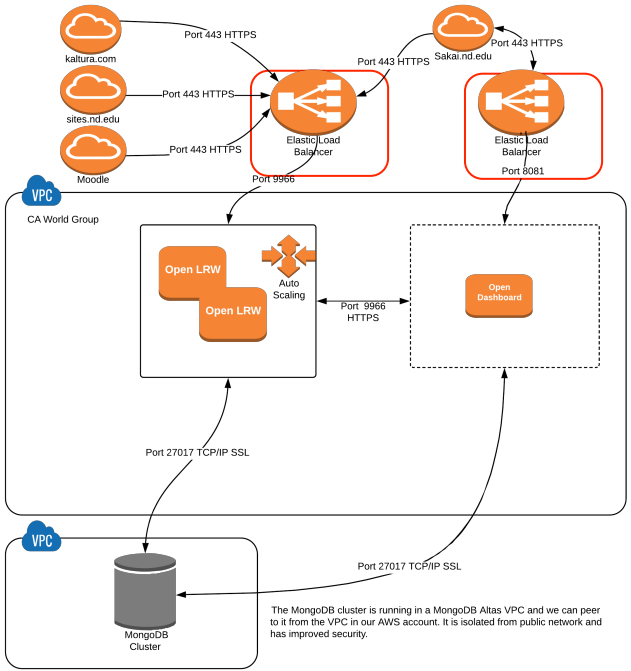
Xiaojing Duan, MS
Technical Lead



Credit: Photo by Loïc Romer

apereo

ND Learning Analytics Infrastructure

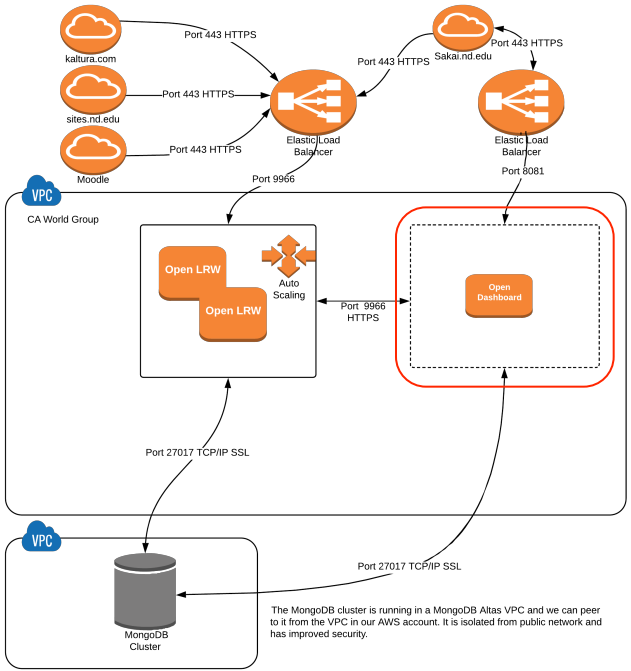


- One Application Load Balancer for balancing the traffic load to Open LRW
- One Application Load Balancer for balancing the traffic load to Open Dashboard

Open Dashboard Application

- Open source Java-based web application framework for visualizing student activity in a course
- LTI compatible
- Instructor facing
- Installation requirement:
  - Recommended operating system is GNU/Linux (Debian/Ubuntu/RHEL)
  - Git
  - Java Development Kit 8
  - Maven 3
  - MongoDB 2.6+
- Open Dashboard Github Repository

ND Open Dashboard Implementation

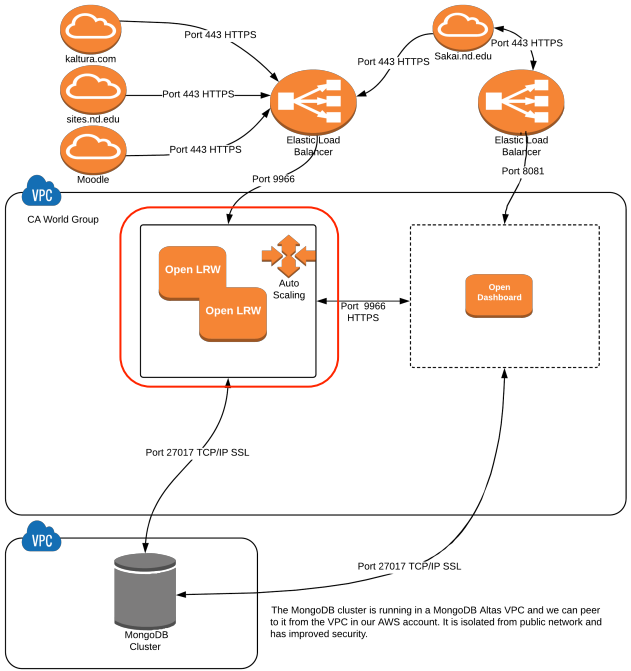


- Open Dashboard (test):
  - One t2.medium instance running Amazon Linux AMI
  - T2 instances are low cost with burstable CPU performance abilities
  - T2.medium offers 2 vCPUs and 4GB of RAM
  - Amazon Linux AMI is supported and maintained by AWS
  - Future prod environment: deploy Auto Scaling group policy

Open LRW Application

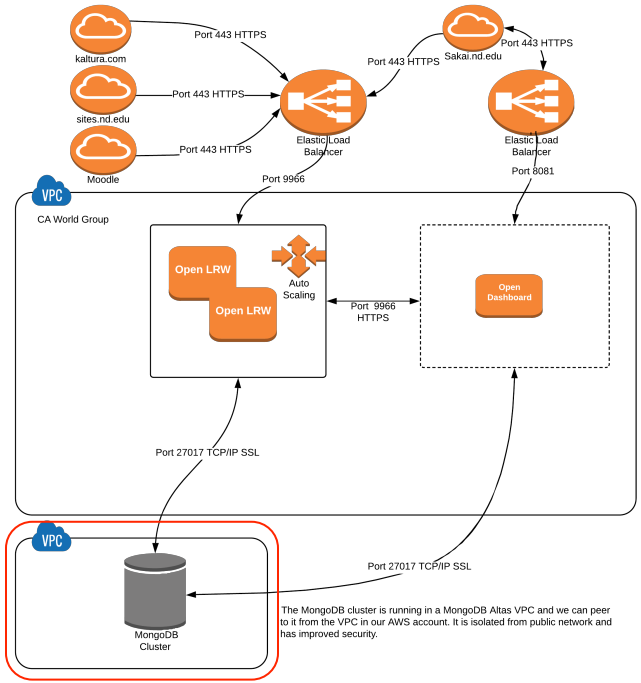
- Open source Java-based Learning Record Warehouse for event capturing
- xAPI / Caliper compatible
- Ability to capture events from multiple sources
- Installation requirement:
  - Recommended operating system is GNU/Linux (Debian/Ubuntu/RHEL)
  - Git
  - Java Development Kit 8
  - Maven 3
  - MongoDB 2.6+
- Open LRW Github Repository

ND Open LRW Implementation



- Open LRW (prod) with auto scaling:
  - Minimum: one t2.medium instance running Amazon Linux AMI
  - Maximum: two cross zones t2.medium instances running Amazon Linux AMI
  - Scaleup: add 1 instance when 70%<= CPU Utilization < +infinity
  - Scaledown: remove 1 instance when -infinity <= CPU Utilization < 20%
  - Load testing: 175 requests/second

ND MongoDB Implementation



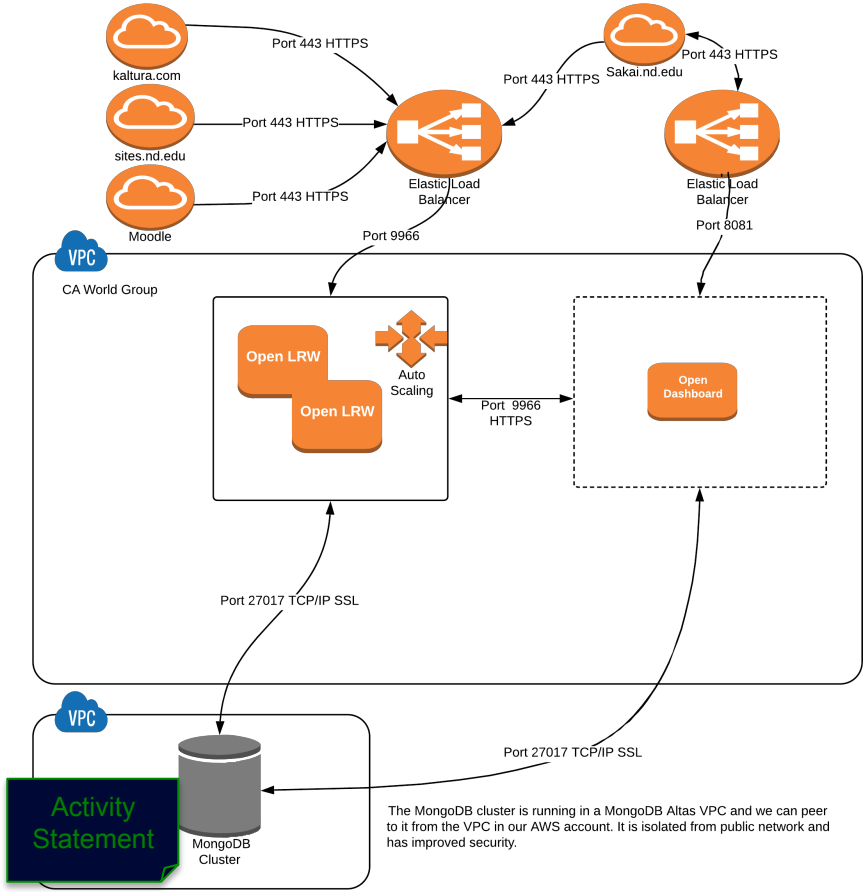
- MongoDB Cluster
  - One primary and two secondary nodes
  - Three cross region nodes for high availability
  - Each node has two vCPUs, 4GB RAM, and 20GB Storage
  - MongoDB 3.6
  - Running in MongoDB Atlas VPC and isolated from public network
  - Current storage usage: 468.24MB for 19,268,731 statements

Sakai xAPI Provider Application

- Provides support for Sakai to work with Open LRW
- Registers activity statements in Sakai
- Sakai then sends the xAPI statements to the connected Open LRW
- Allows you to configure from which tool you want to capture the activities into Open LRW
- Sakai xAPI Provider Github Repository

After implementing Open LRW, MongoDB, and Sakai xAPI Provider

Ability to capture student online learning activities from Sakai in real-time

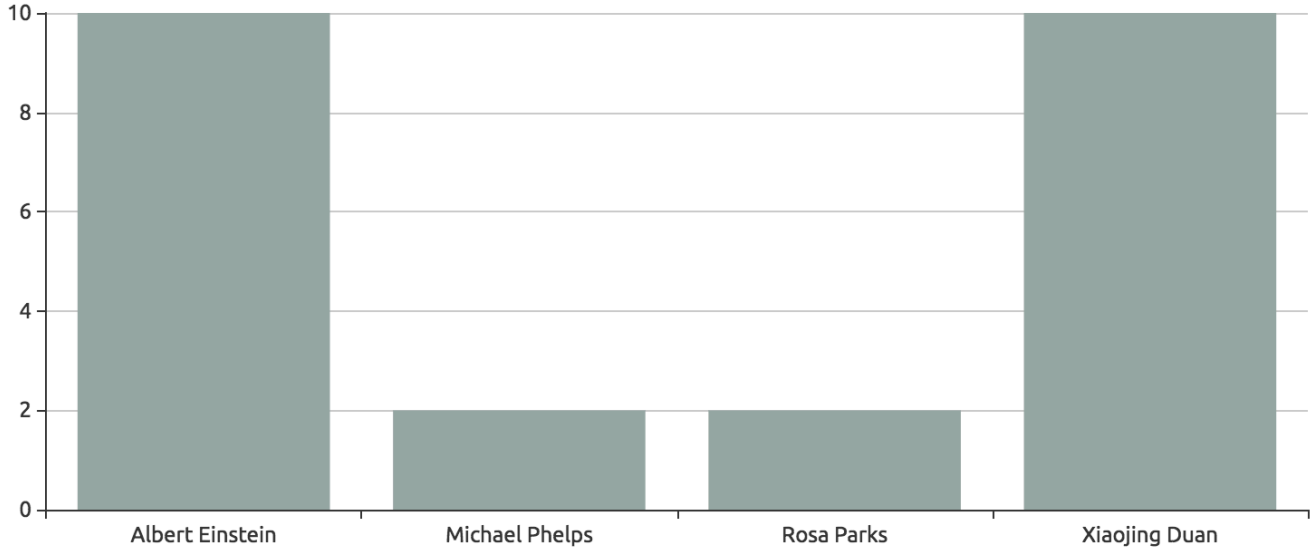

kaltura.com
Port 443 HTTPS
sites.nd.edu
Port 443 HTTPS
Moodle
Port 443 HTTPS
Elastic Load Balancer
Port 443 HTTPS
Sakai.nd.edu
Port 443 HTTPS
Elastic Load Balancer
Port 9966
Port 8081
VPC
CA World Group
Open LRW
Open LRW
Auto Scaling
Port 9966 HTTPS
Open Dashboard
Port 27017 TCP/IP SSL
VPC
Activity Statement
MongoDB Cluster
Port 27017 TCP/IP SSL
The MongoDB cluster is running in a MongoDB Altas VPC and we can peer to it from the VPC in our AWS account. It is isolated from public network and has improved security.

Sample Activity Statement from Sakai


```
{ "_class": "unicon.matthews.caliper.service.repository.MongoEvent",
  "userId": "xduan",
  "organizationId": "58d963b069f7c11e7d90e598",
  "tenantId": "58d963b069f7c11e7d90e597",
  "event": { "_id": "bc167446-d3c6-4afe-8e63-a871ea6956e6",
             "context": "http://purl.imsglobal.org/caliper/v1/Event",
             "agent": { "_id": "xduan",
                        "type": "https://github.com/adlnet/xAPI-Spec/blob/master/xAPI.md#agentaccount",
                        "context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
                        "name": "Xiaojing Duan",
                        "extensions": { "HOMEPAGE": "https://sandbox-sakai.nd.edu" }
                      },
             "action": "http://www.adlnet.gov/expapi/verbs/scored",
             "object": { "_id": "https://sandbox-sakai.nd.edu/portal/assessment",
                         "type": "http://adlnet.gov/expapi/activities/received-grade-assessment",
                         "context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
                         "name": "User received a grade",
                         "description": "User received a grade for their assessment: Introduction; Submitted: on time"
                       },
             "generated": { "_id": "da2f482b-78e1-423b-86f9-74b31f0f0df4",
                            "type": "http://purl.imsglobal.org/caliper/v1/Result",
                            "context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
                            "actor": { "_id": "xduan",
                                       "type": "https://github.com/adlnet/xAPI-Spec/blob/master/xAPI.md#agentaccount",
                                       "context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
                                       "name": "Xiaojing Duan",
                                       "extensions": { "HOMEPAGE": "https://sandbox-sakai.nd.edu" }
                                       },
                            "totalScore": 10.0
                            },
             "eventTime": "2017-06-26T13:54:53Z"
             }
}
```


This statement describes that “Xiaojing Duan” submitted and scored “10.0” points on the “Introduction” assessment at 13:54:53 on 2017-06-26.

Sample Visualization of Data from Sakai



Sample bar chart shows different students' scores on the "Introduction" quiz.

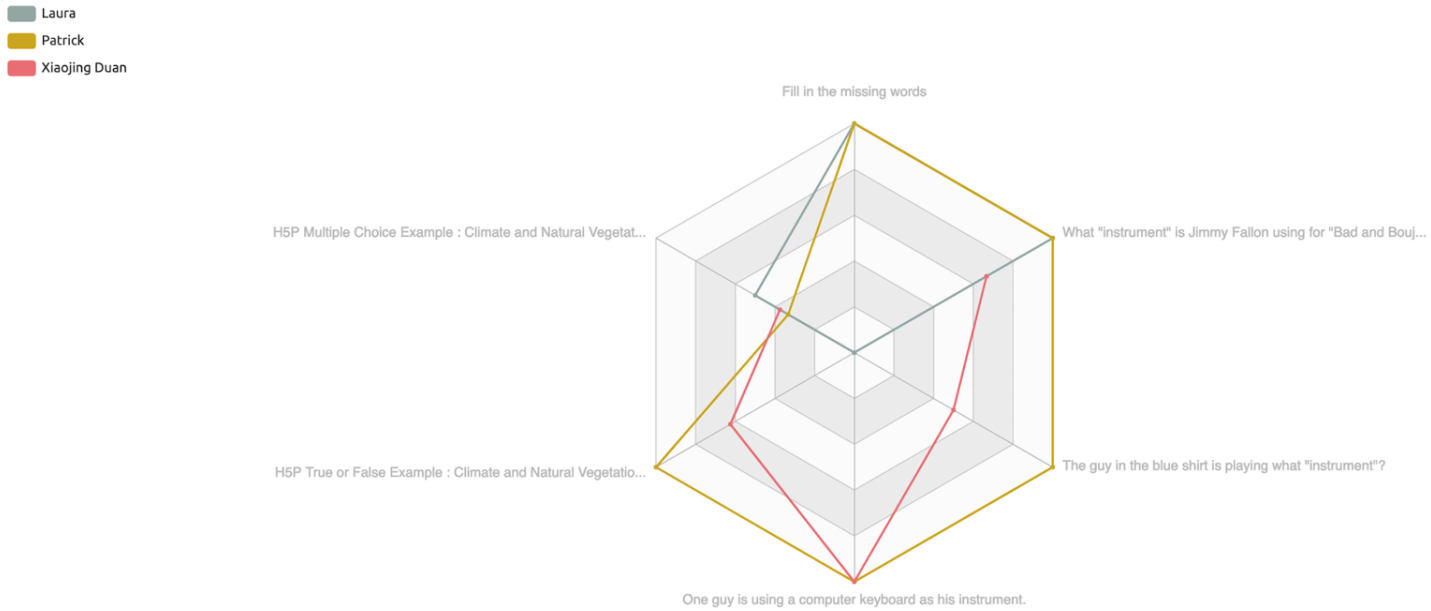
Sample Activity Statement from H5P


```
{ "_class": "unicon.matthews.caliper.service.repository.MongoEvent",
  "userId": "mailto:Xiaojing.Duan@nd.edu",
  "organizationId": "58d963b069f7c11e7d90e598",
  "tenantId": "58d963b069f7c11e7d90e597",
  "event": { "_id": "a4f8d2b5-8e1c-415e-ac9c-bb0afc36af9c",
             "context": "http://purl.imsglobal.org/caliper/v1/Event",
             "agent": { "_id": "mailto:Xiaojing.Duan@nd.edu",
                        "type": "foaf:mbox",
                        "context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
                        "name": "Xiaojing Duan",
                        "extensions": {  }
                      },
             "action": "http://adlnet.gov/expapi/verbs/answered",
             "object": { "_id": "https://sites.nd.edu/kabbott/wp-admin/admin-ajax.php?action=h5p_embed&id=2",
                         "type": "http://adlnet.gov/expapi/activities/cmi.interaction",
                         "context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
                         "name": "H5P True or False Example : Climate and Natural Vegetatio...",
                         "description": "Pine and cacti are typical plants found in the hot desert.\n",
                         "extensions": { "http://h5p.org/x-api/h5p-local-content-id": "2" }
                       },
             "generated": { "_id": "ed56288c-a425-4979-b05d-5e0a1ae4b47a",
                            "type": "http://purl.imsglobal.org/caliper/v1/Result",
                            "context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
                            "actor": { "_id": "mailto:Xiaojing.Duan@nd.edu",
                                       "type": "foaf:mbox",
                                       "context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
                                       "name": "Xiaojing Duan",
                                       "extensions": {  }
                                     },
                            "totalScore": 1.0
                          },
             "group": { "_id": "http://sites.nd.edu/kabbott/h5p-true-or-false-example/",
                        "type": "http://activitystrea.ms/schema/1.0/page",
                        "context": "http://purl.imsglobal.org/ctx/caliper/v1/Context",
                        "name": " | H5P True or False Example"
                      },
             "eventTime": "2017-05-26T18:47:45.586Z"
           }
}
```


This statement describes that “Xiaojing Duan” answered the “H5P True or False Example: Climate and Natural Vegetation” question and scored “1.0” point at 18:47:45 on 2017-05-26.

(H5P is a tool for creating and delivering rich and interactive learning objects.)

Sample Visualization of Data from H5P



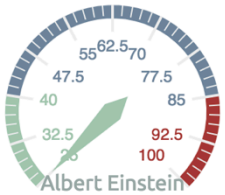
Sample radar chart shows different students' scores on different H5P questions.

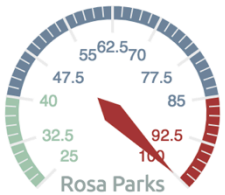
Sample Activity Statement from Kaltura

This statement describes that “Xiaojing Duan” watched 50% of the “Video1-Introduction” at 16:12:16 on 2017-08-15.

(Kaltura is a video platform for delivering and storing live and on-demand videos.)

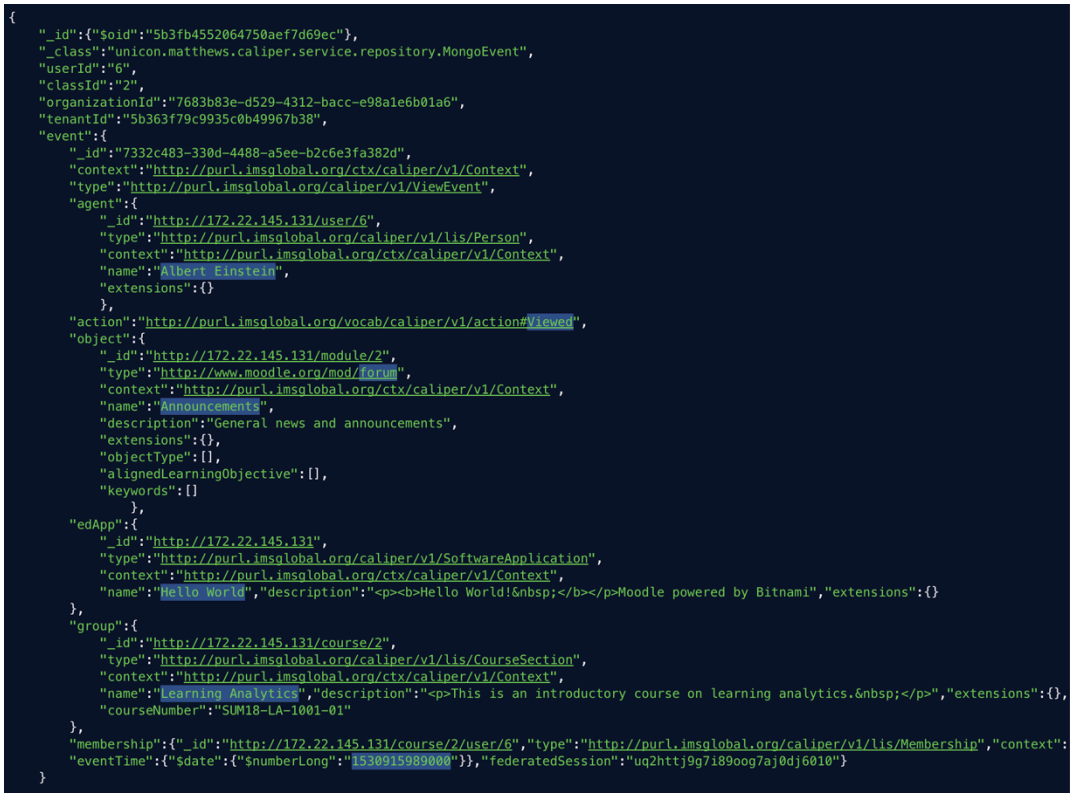
Sample Visualization of Data from Kaltura





Sample gauge chart shows the maximum length of the “Video1-Introduction” watched by different students.

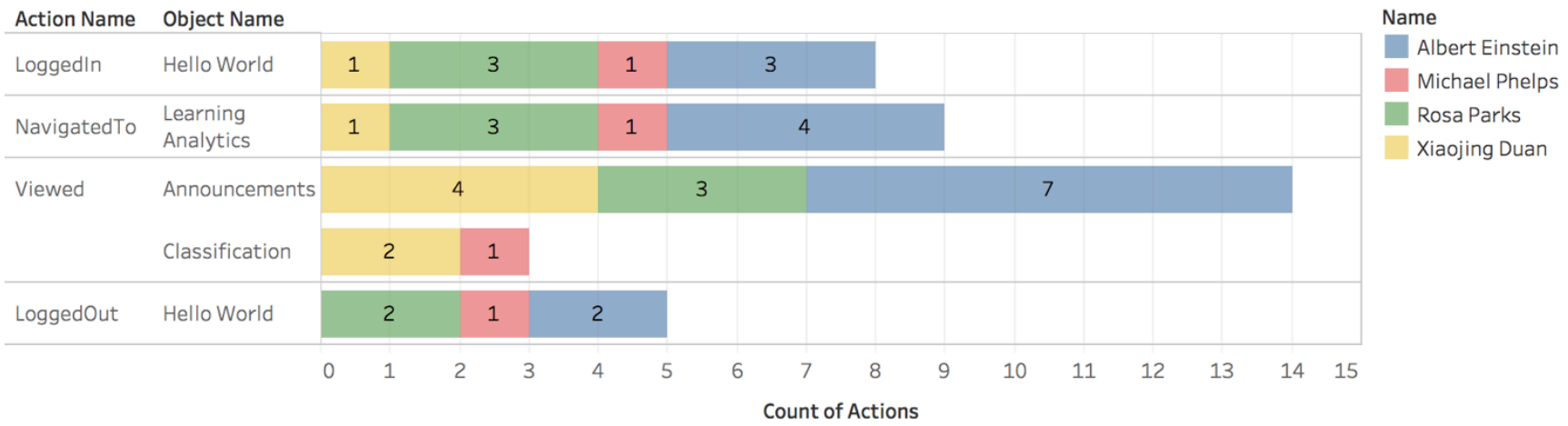
Sample Activity Statement from Moodle



This statement describes that “Albert Einstein” viewed the “Announcements” forum in the “Learning Analytics” course on the “Hello World” Moodle site at 1530915989000 (18:26:29 on 2018-07-06).

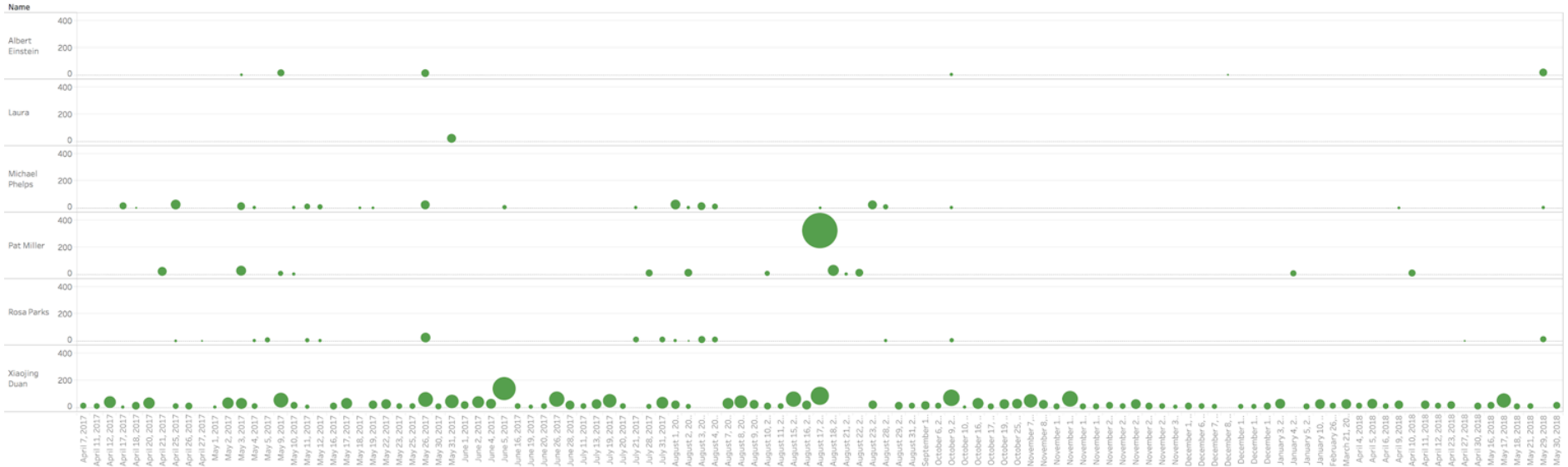
(Like Sakai, Moodle is an open source learning management system.)

Sample Visualization of Data from Moodle



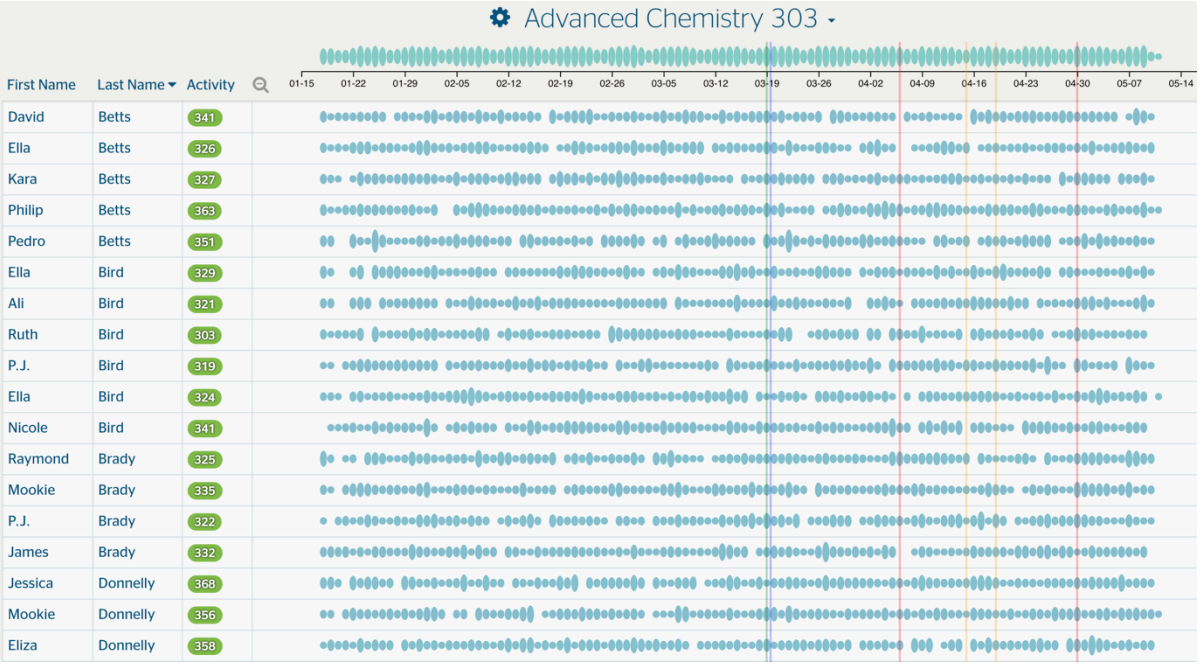
Sample bar chart shows the types and number of activities taken by different students in Moodle.

Sample Visualization of Combined Data



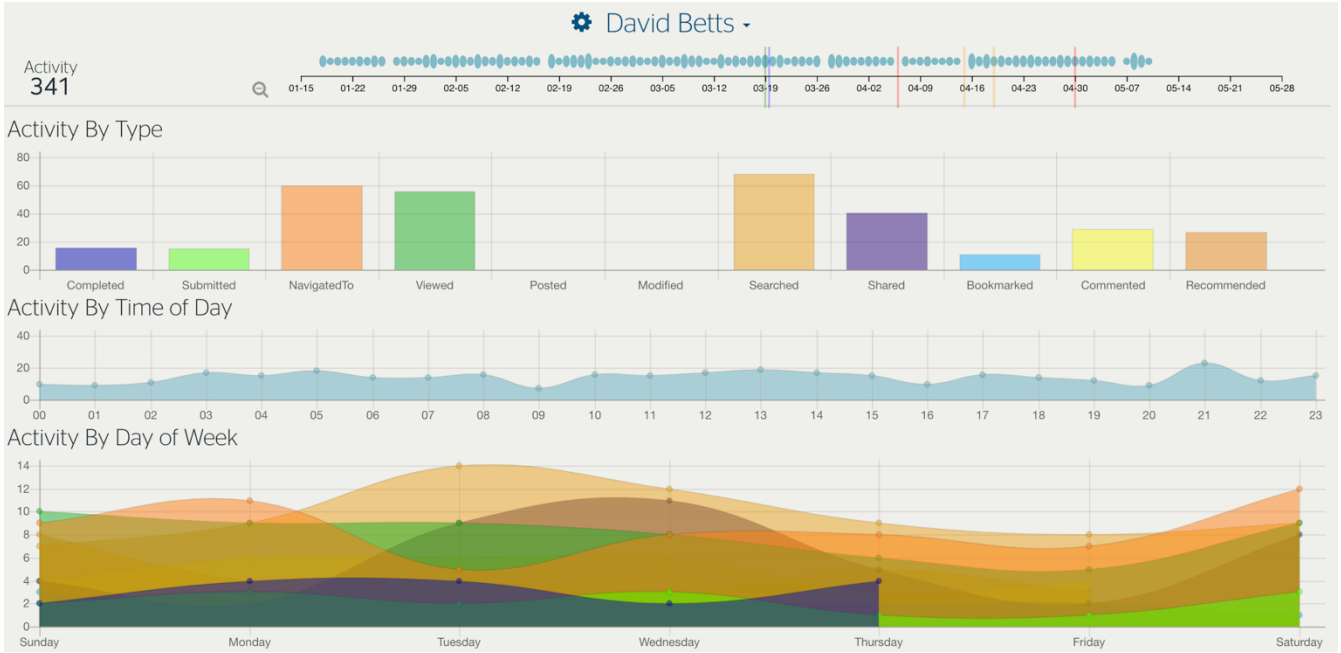
Sample bubble chart shows the total number of activities taken by different students on various learning platforms over time.

Open Dashboard Demo



This aggregated view shows the total number of activities each student had over time.

Open Dashboard Demo



This detailed view shows a student's activity by type, time of day, and day of week.

Our Initial Project: Moreau Analytics

Challenge: How can we ensure everyone has the opportunity to thrive?

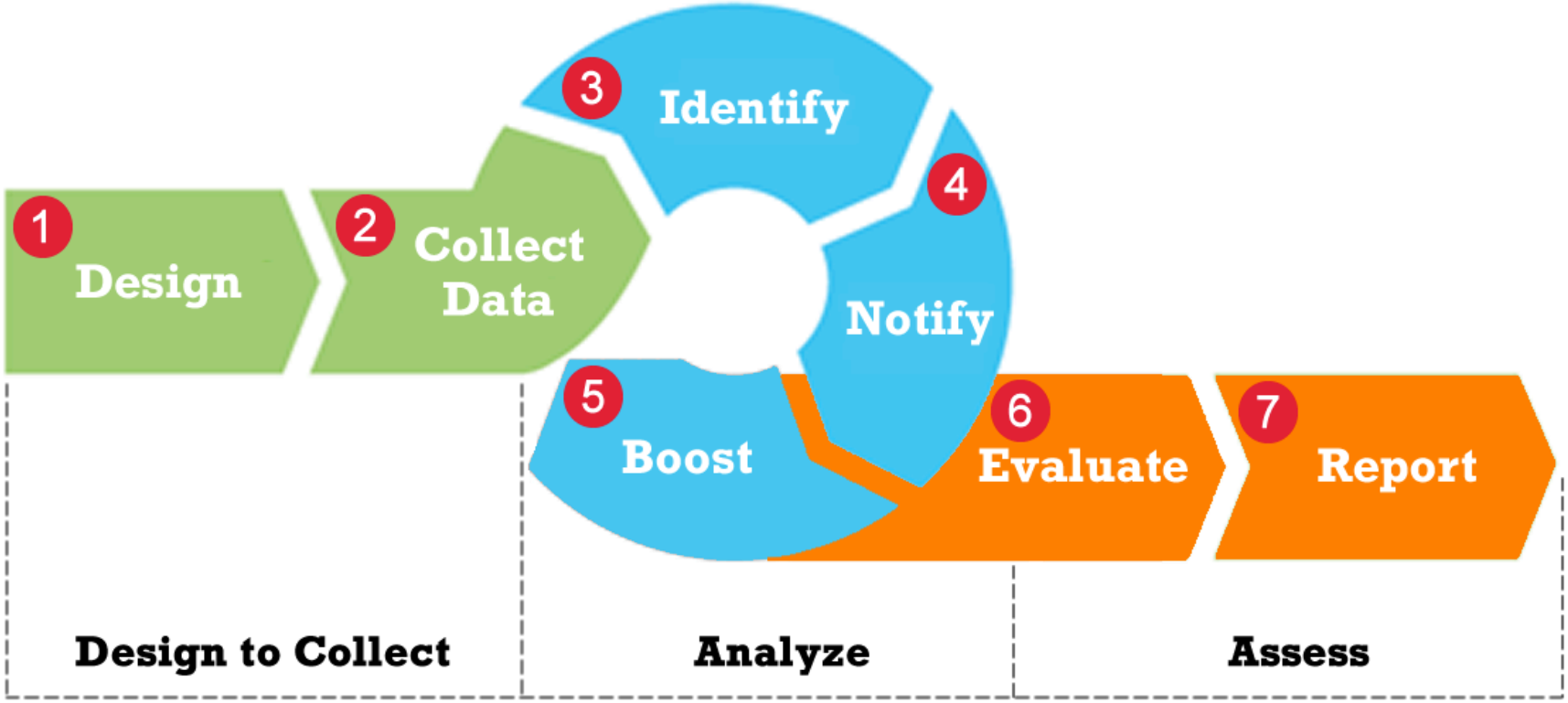
- 98% of our first year students are retained to the second year.
- How can we identify and help the 2% (~40 students) who don't continue?

Goal: Liberate the data with our Learning Analytics process to boost every students' potential to thrive (without risk of harm) in the Moreau First Year Experience (FYE) course.

Moreau

MOREAU FIRST YEAR EXPERIENCE

The Process We Followed



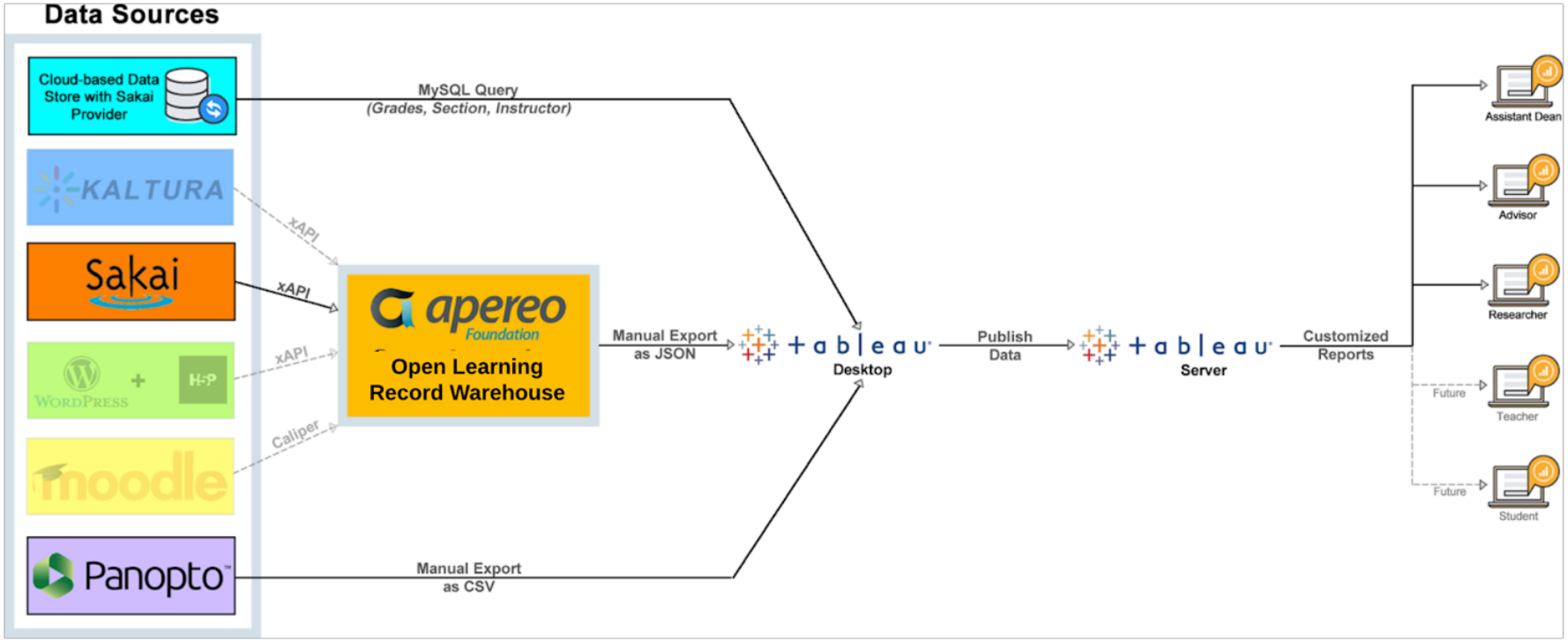
1 Design

Course Design in Sakai

- Enabled Sakai as the single hub for all the students' work
- Created device-friendly Coursework page
- Used visual and language cues on course sites
- Centralized content in "Super Site" for easy maintenance and click recording
- Created custom role for instructors
- Embedded data collection

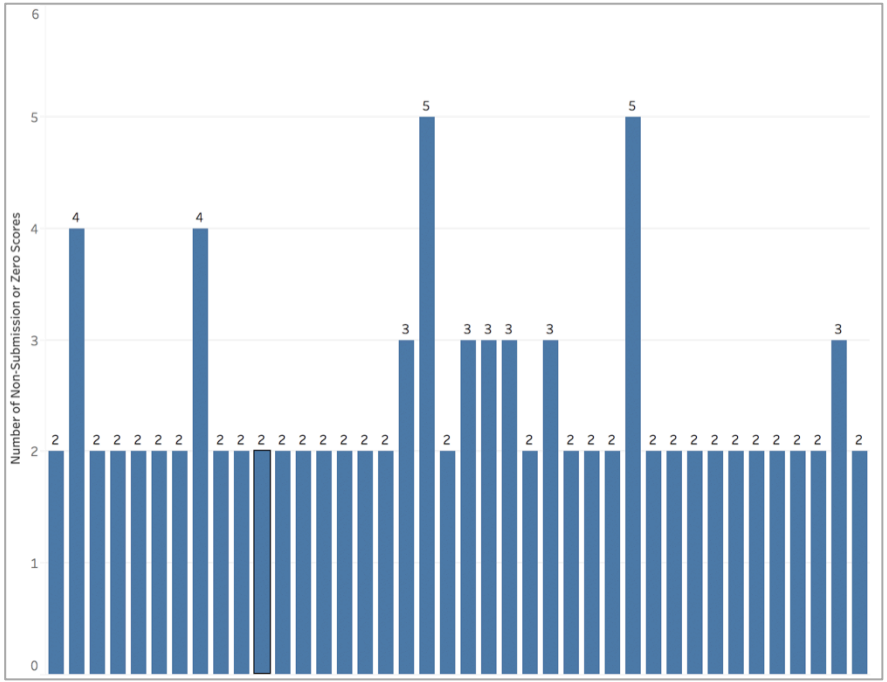
2 Collect Data

Open LRW In Action



3 Identify

Early Identification: 39 students
Weeks 1-6 Homework Deficiency Report



Midterm Identification: 52 students
Grades Comparison Report

Since 16 students appeared on both reports, our final set of at risk student was 75.

Early Notification

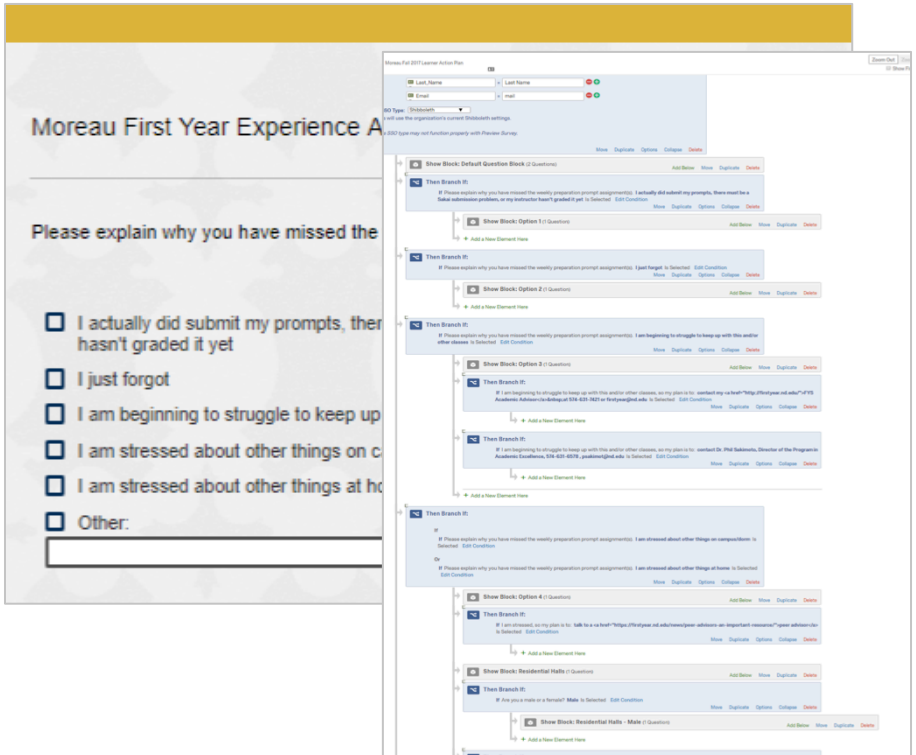
Bottom Up Student Notification

Midterm Notification

Top Down Advisor Requests Appointment

5 Boost

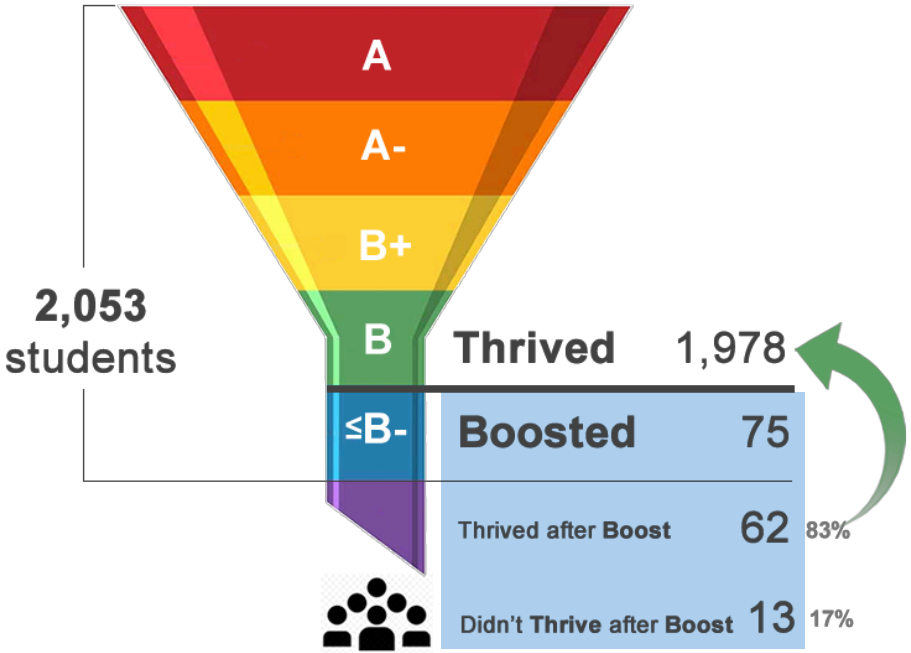
Early Boost: Personalized Learner Action Plan

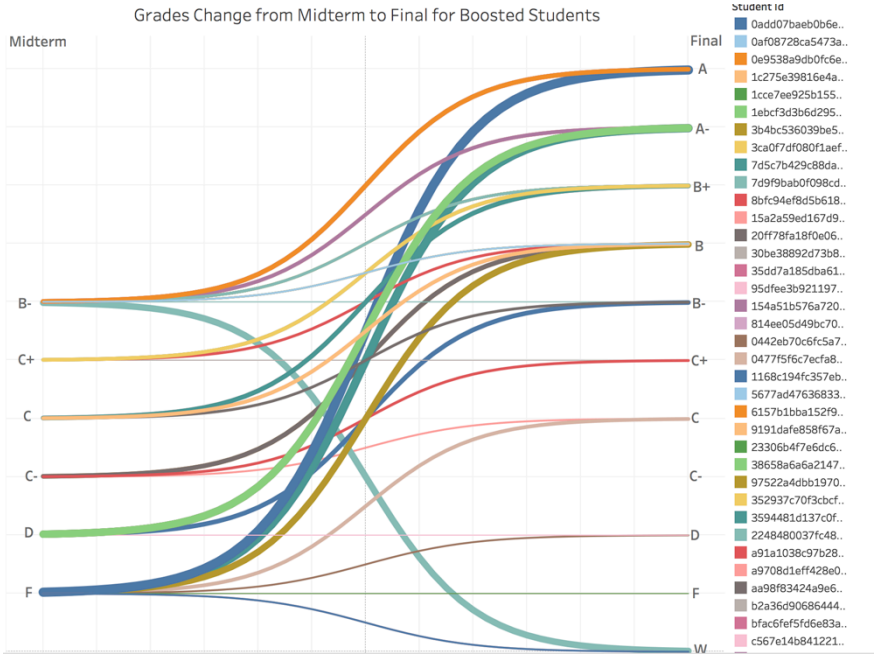


Midterm Boost: Advisor Face-to-Face Data Informed Discussion

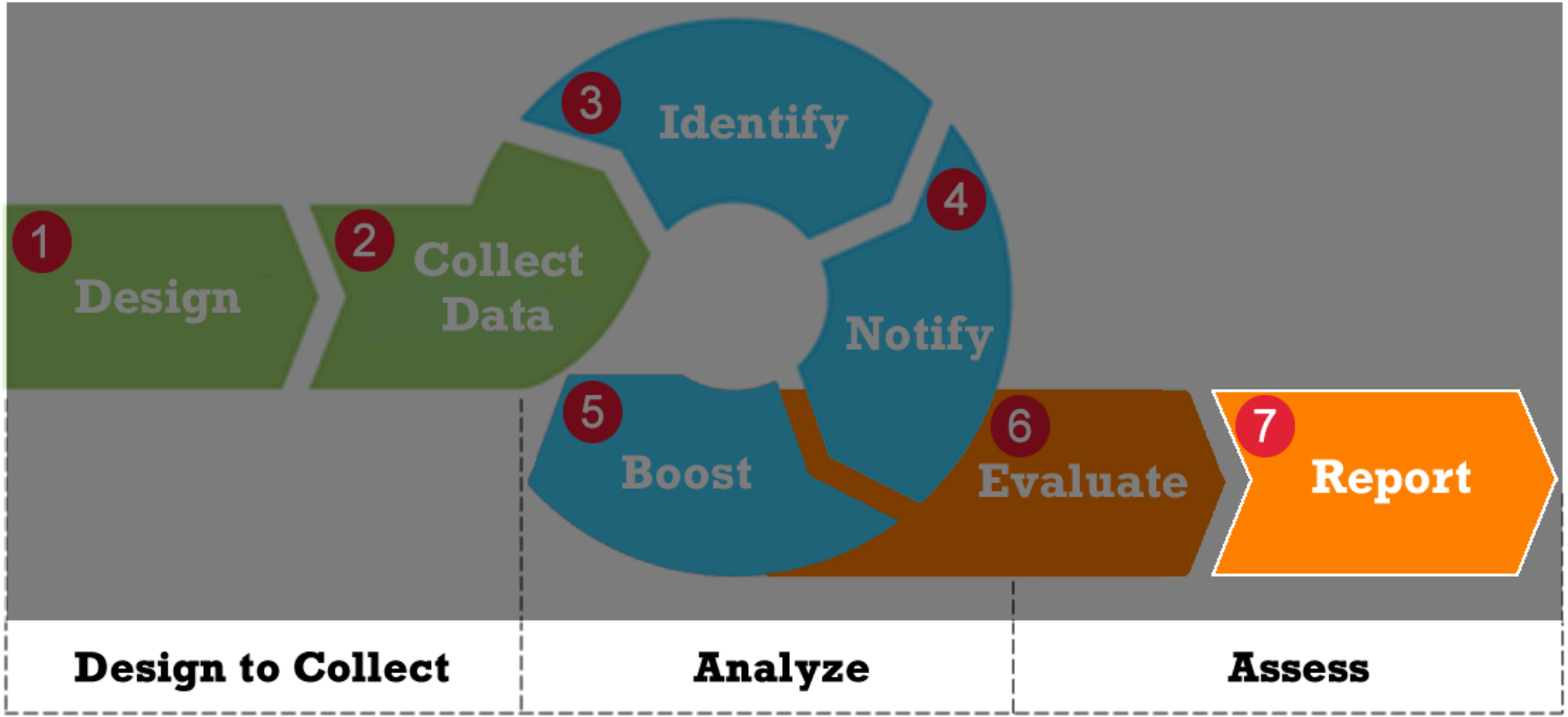
6 Evaluate

Students Thrived after Boosting





Report


1
Design
2
Collect Data
3
Identify
4
Notify
5
Boost
6
Evaluate
7
Report
Design to Collect
Analyze
Assess

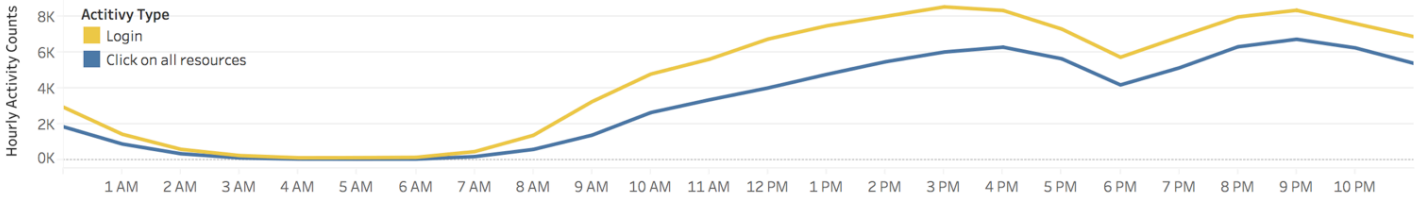
7 Report

Scenario #1

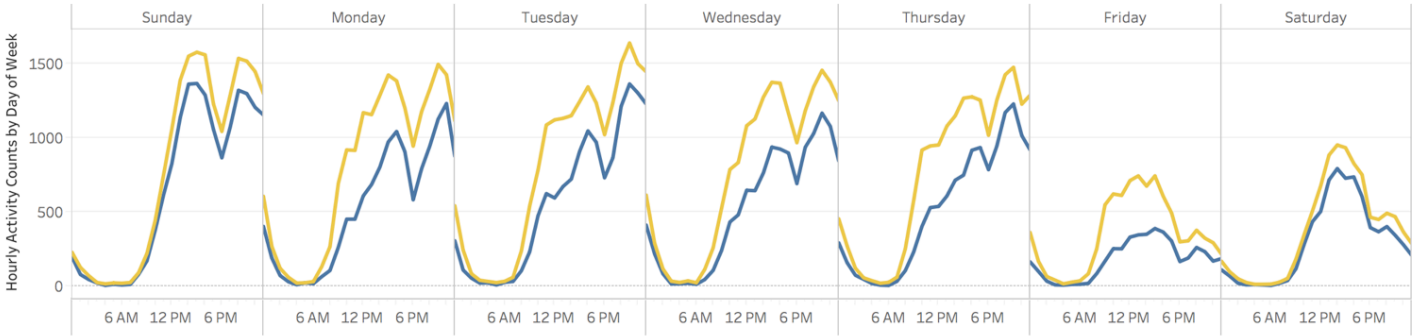
Program director, Instructors, Course designers: When do students work on the course online?

Program director, Instructors, Course designers: What are the course workload rhythms for the day/week/semester schedule?

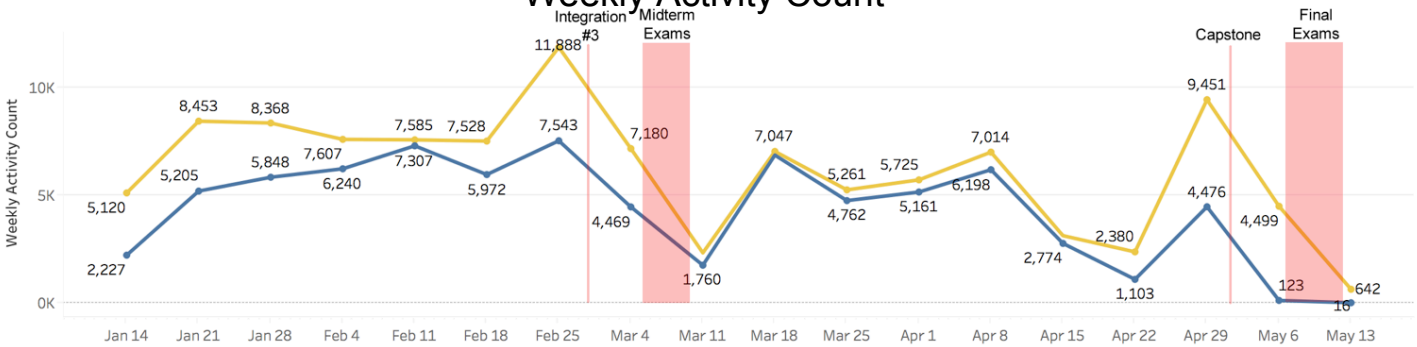
Hourly Activity Count



Hourly Activity Count by Day of Week



Weekly Activity Count



7 Report

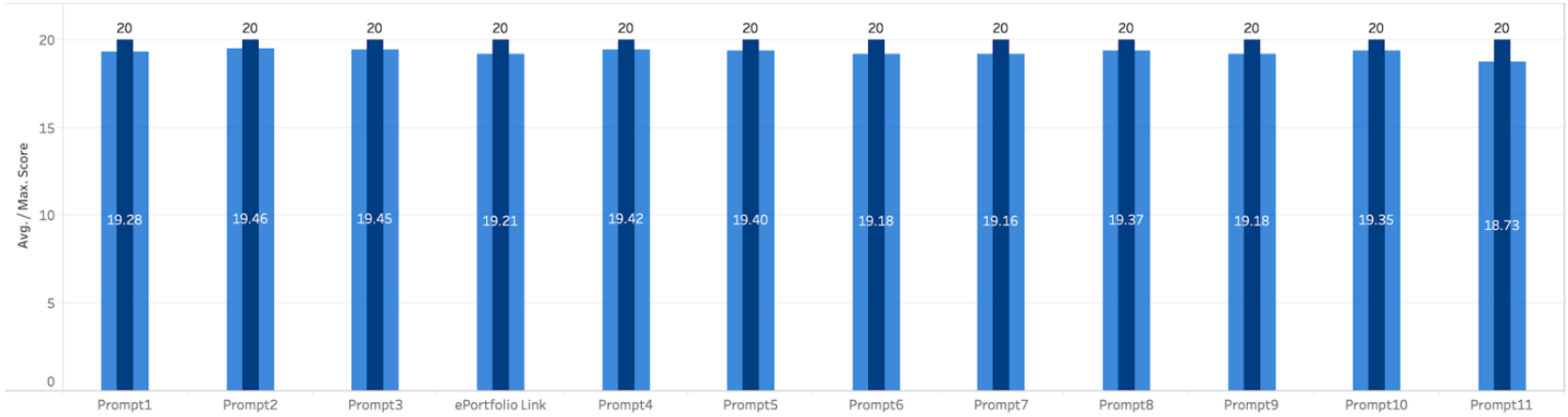
Scenario #2

Program director: Are instructors following the rubrics for grading?

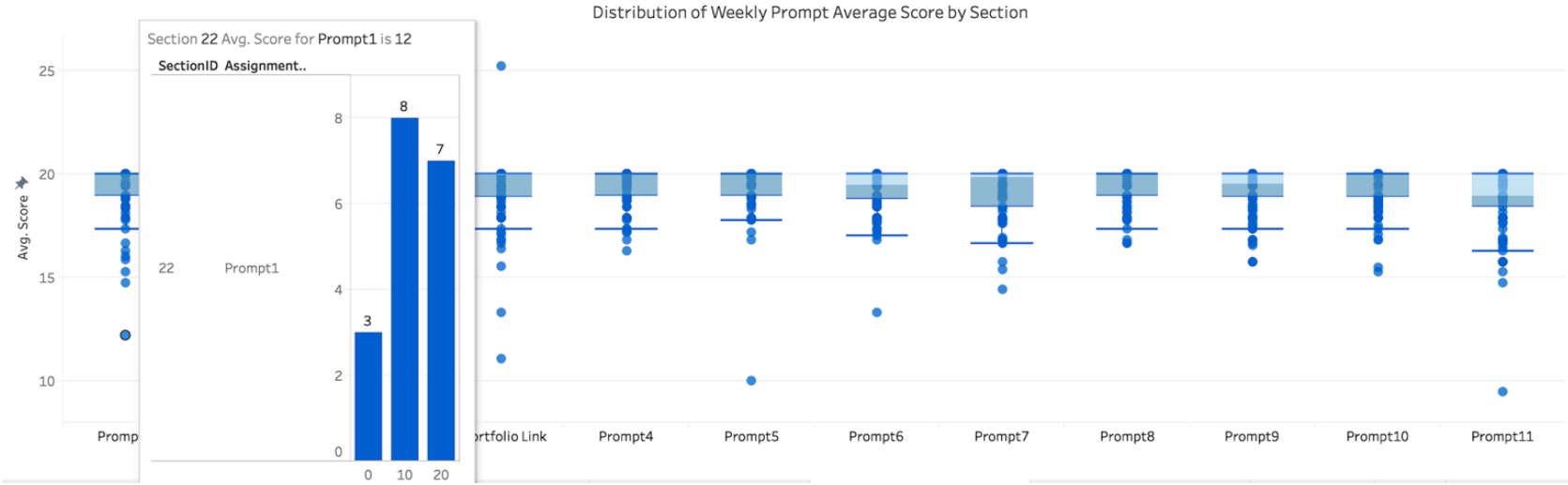
Program director: Is the grading consistent across different sections?

Student: Is my instructor grading fairly?

Weekly Prompt Average Score of All FYE Students



Distribution of Weekly Prompt Average Score by Section



7 Report

Scenario #3

Program director: Which sections had the most missed assignment submission ?

Course designer: Which assignment had the most missed submission?

Missing Submissions by Section

Section Number	Total Missing	Capstone ePortfolio	Prompt11	Prompt10	Prompt9	Prompt8	Prompt7	Prompt6	Integration #3 ePortfolio	Prompt5	Prompt4	ePortfolio Link	Prompt3	Prompt2	Prompt1
1	12	1	1	1	1	2		1		1		1		1	2
46	8		6		2										
21	7		3			2	1			1					
9	6		2	1	2			1							
70	5	1	1	1		1							1		
98	5		1	1	1	1				1					
22	4		4												
23	4								2	2					
31	4		1	2		1									
36	4					1	2	1							
42	4		3	1											
72	4		1	1						1		1			
105	4		4												
3	3				1					1					1
33	3		1		1		1								
40	3		1							1			1		
58	3					1					1				1
59	3		1		1										1
113	3				1		1	1							
5	2		1	1											
11	2				2										
12	2				1	1									
19	2									1			1		
20	2				1		1								
29	2	2													
32	2		1					1							
52	2		1							1					
57	2			1			1								
78	2				1			1							
114	2		1			1									
15	1					1									
16	1						1								
17	1				1										
39	1					1									
48	1	1													
79	1								1						
107	1			1											
109	1								1						
Totals	119*	5	34	11	16	13	8	6	4	10	1	2	3	1	5

*Only 119 missed submissions out of 28,742 possible submissions for the entire course

7 Report

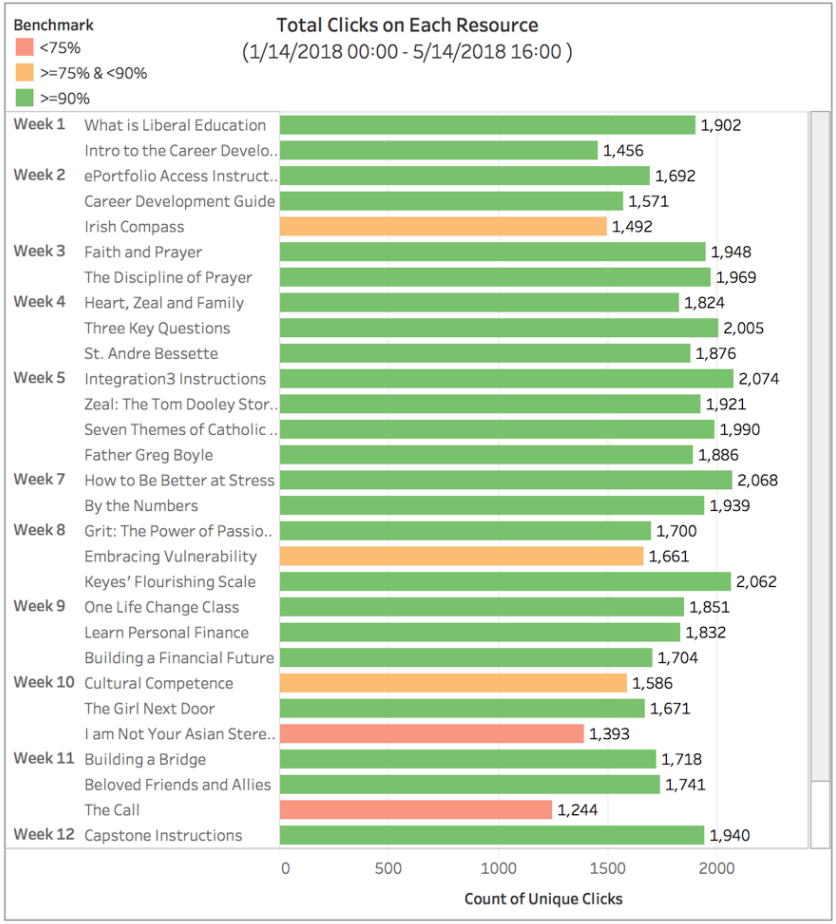
Scenario #4

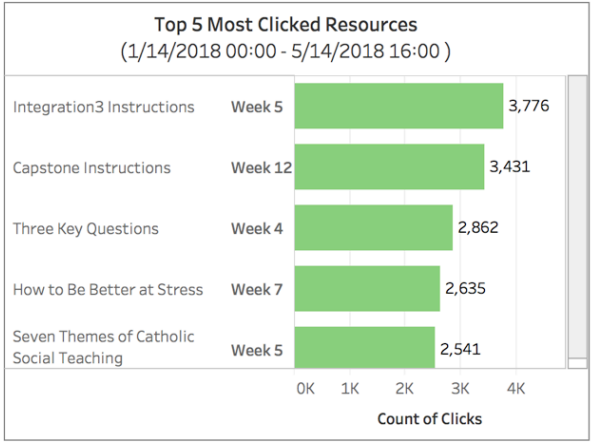
Program director: Are students engaging with the course materials?

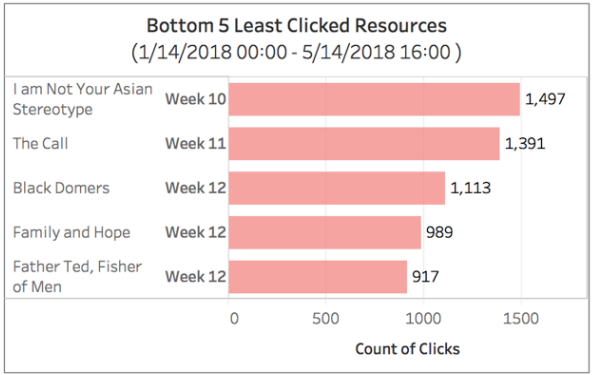
Program director, Course designers: Which course materials are most engaging?

Program Director, Course designers: What is the optimal timing for a selected course content?

Clicks on Content Resources







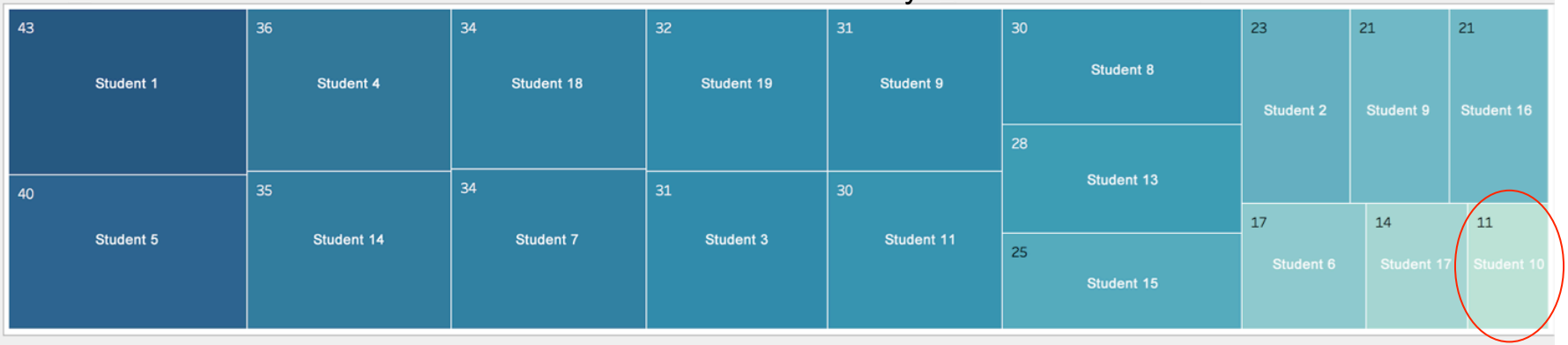
7 Report

Scenario #5

Instructor: Who didn't do the course readings before class?

Instructor: Does students' class participation match their online preparation?

Total Content Resource Clicks by Student



Detailed Content Resource Clicks by Student

	Week 1		Week 2			Week 3		Week 4			Week 5			Week 7		Week 8			Week 9			Week 10			Week 11			Week 12		
	Intro t..	What i..	Caree..	ePortf..	Irish C..	Faith ..	The Di..	Heart,..	St. An..	Three ..	Fathe..	Integr..	Seven ..	Zeal: T..	By the..	How t..	Embra..	Grit: T..	Keyes'..	Buildi..	Learn ..	One Li..	Cultur..	I am N..	The Gi..	Belove..	Buildi..	The Call	Black ..	Capst..	Family..	Fathe..
Student 1	1	1	1			1	1	2	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		4		
Student 2		1		1		1	1				1	2	1	1	1	1	1	1	1			1							2		
Student 3		1		1	1	1	1	1	1	2	1	2	1	2	1	1	2	1	1	1	1	1	1	1	1	1			1		
Student 4		1	1		2	1	2	1	1	1	1		1	1	2	2	2	2	1	1	1	1	1	1				1		1	1
Student 5		1		1		1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		2		
Student 6		1				1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1		1	1			1		
Student 7						1	1		1	1	1	1	1	1		1			1	1	1			1					1		
Student 8	1	1		1		1	1	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	2	1	
Student 9	1	2	1	2	1	1	1	1	1	1	1	3	1	1		1	1	1	1	1	1	1	1	1	2	2	1	1		1	1
Student 10		1	1	1	1	1	1	1				1				1			1										1		
Student 11		1				1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1				1			1		
Student 12		1	1	1	1	1	1	1	2	1	1	1	2	1	1	1		1	1		1	1					1	1			
Student 13	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1
Student 14	2	2	2	1	1	1	1	1	1	1	2	5	2	1	2	1	1	1	1	1	1	1	1	1	2	1	1	1	4		
Student 15		1	1	2	1	1	2	2	2	2	1	3	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2		3		
Student 16	1	1	1	1	1	1	1	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1		
Student 17	1	1	1	1		1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1		
Student 18	1		2		1	1	1			1	1	3	2		1	1	1	1	1	1	1	1	1	1					1		
Student 19	1	1	1		1	2	2	1	1	1	1	2	1	1	2	2	1	1	1	1	1	1			2	1	1	1		4		

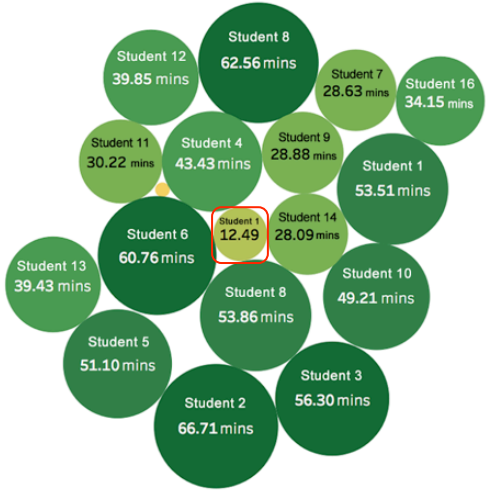
7 Report

Scenario #6

Instructor: Did my students watch all the required videos?

Instructor: How much of the video did they watch?

Minutes of Video Watched by Student



	Week3 Two Notre Dames: Your Holy Cro.. 13.27 minutes in length	Week5 Mind: The Word of Life 6.17 minutes in length	Week5 Notre Dame as An Intellectual an.. 7.53 minutes in length	Week6 Hope: How to Fail Well 10.87 minutes in length	Week6 The Integrity of the Family 13.40 minutes in length
Student 1					
Student 2					
Student 3					
Student 4					
Student 5					
Student 6					
Student 7					
Student 8					
Student 9					
Student 10					
Student 11					
Student 12					
Student 13					
Student 14					
Student 15					
Student 16					
Student 17					
Student 18					

What We Learned

- Our work is scalable to other large credit-granting courses, especially the first-year gateway courses
- The Sakai/Open Learning Record Warehouse/Tableau connection has allowed us to easily collect and analyze data and produce useful customized reports for administrators, instructors, advisors, and researchers
- The design of the course is crucial in removing student barriers to learning, collecting good data, and enabling analysis
- We can help our students thrive by utilizing learning analytics
- Patience is required as your knowledge and expertise grow in this area

Our Current Project: STEM Analytics

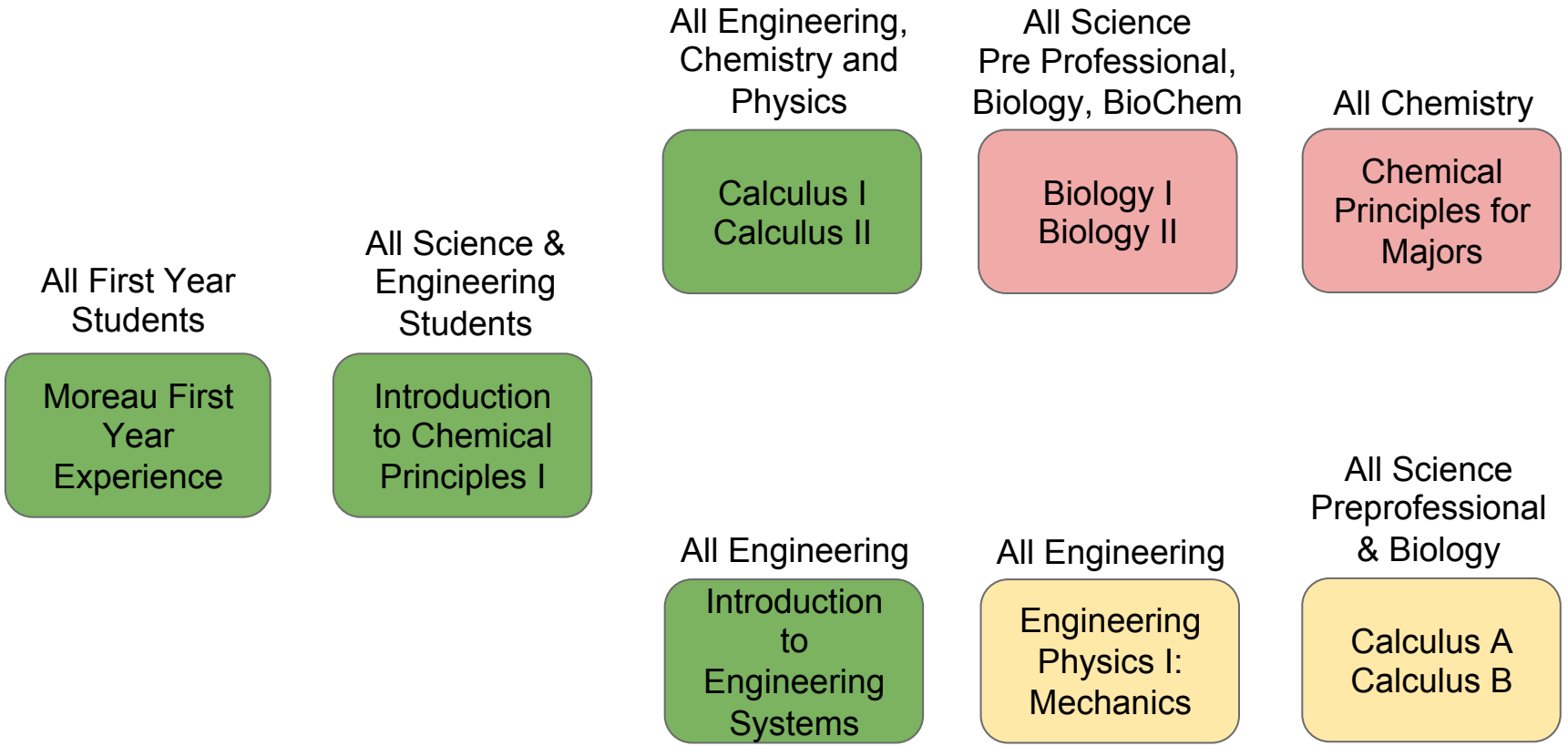
Challenge: “Can we help the underserved/underprepared special population (Science & Engineering Scholars) and also support the general population?”

Margaret Dobrowolska, Associate Dean for Undergraduate Studies

Goal: identify underperforming Science Technology Engineering Mathematics (STEM) students as early as possible, boost their potential to thrive, and improve the overall STEM retention rate

Method: scale the Moreau Analytics process both horizontally and vertically

Horizontal Scaling: Multiple STEM Courses

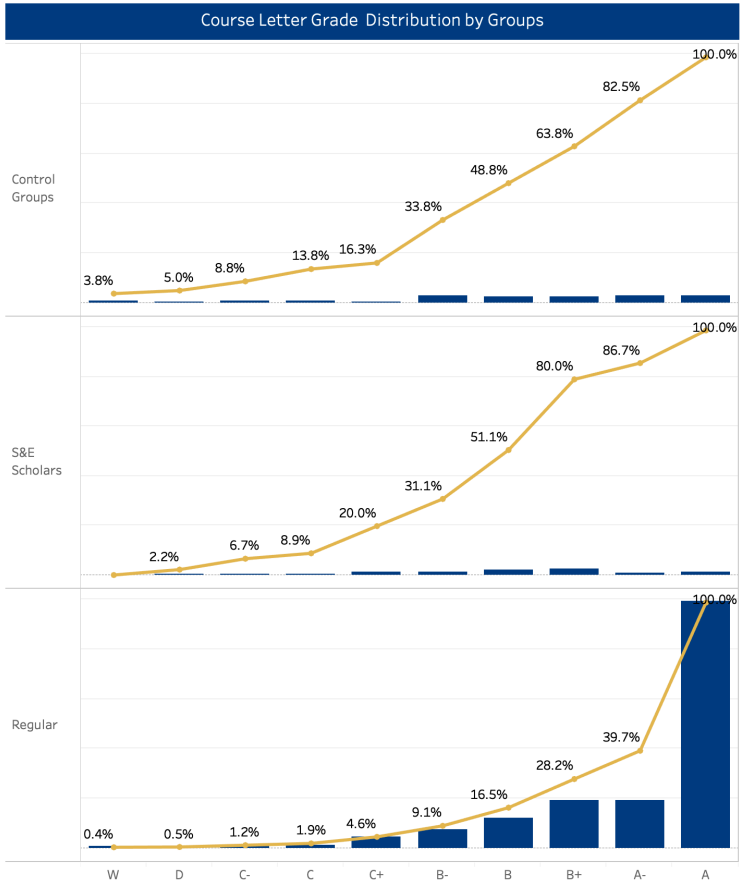


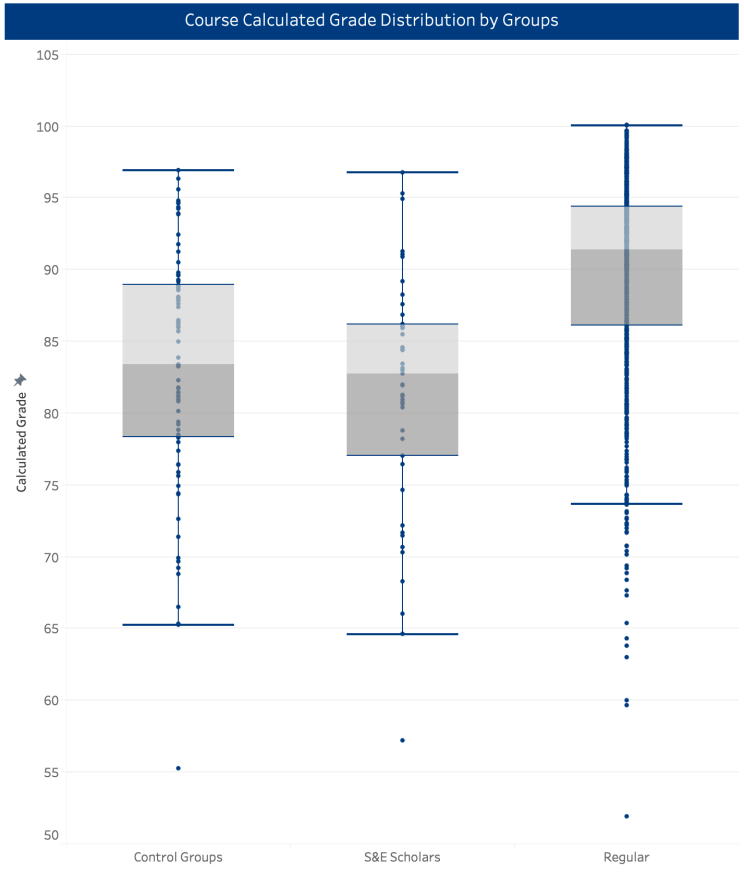
Vertical Scaling: Deeper Analysis

1. Overall course performance analysis
2. Individual assessment performance analysis
3. Assessment Item Analysis

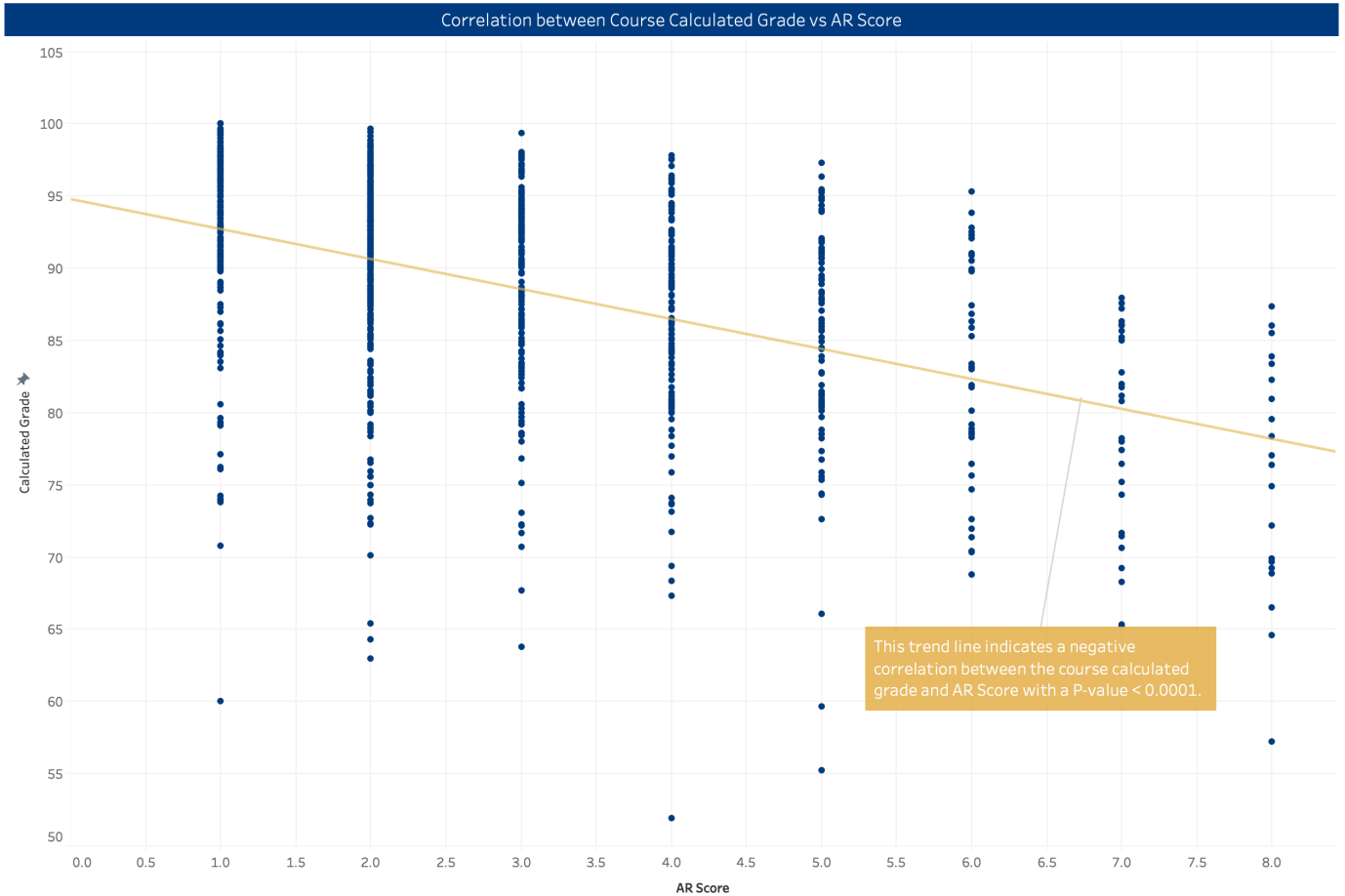
Overall Course Performance Analysis

- Is there an achievement gap between the Science & Engineering (S&E) Scholars and the rest of the student population in the Introduction to Chemical Principle course?
- Did the treatment program boost the S&E Scholars' performance?
- What's the best feature for identifying the underserved and underprepared S&E Scholars?





This dashboard compares the overall course performance between different groups of students.

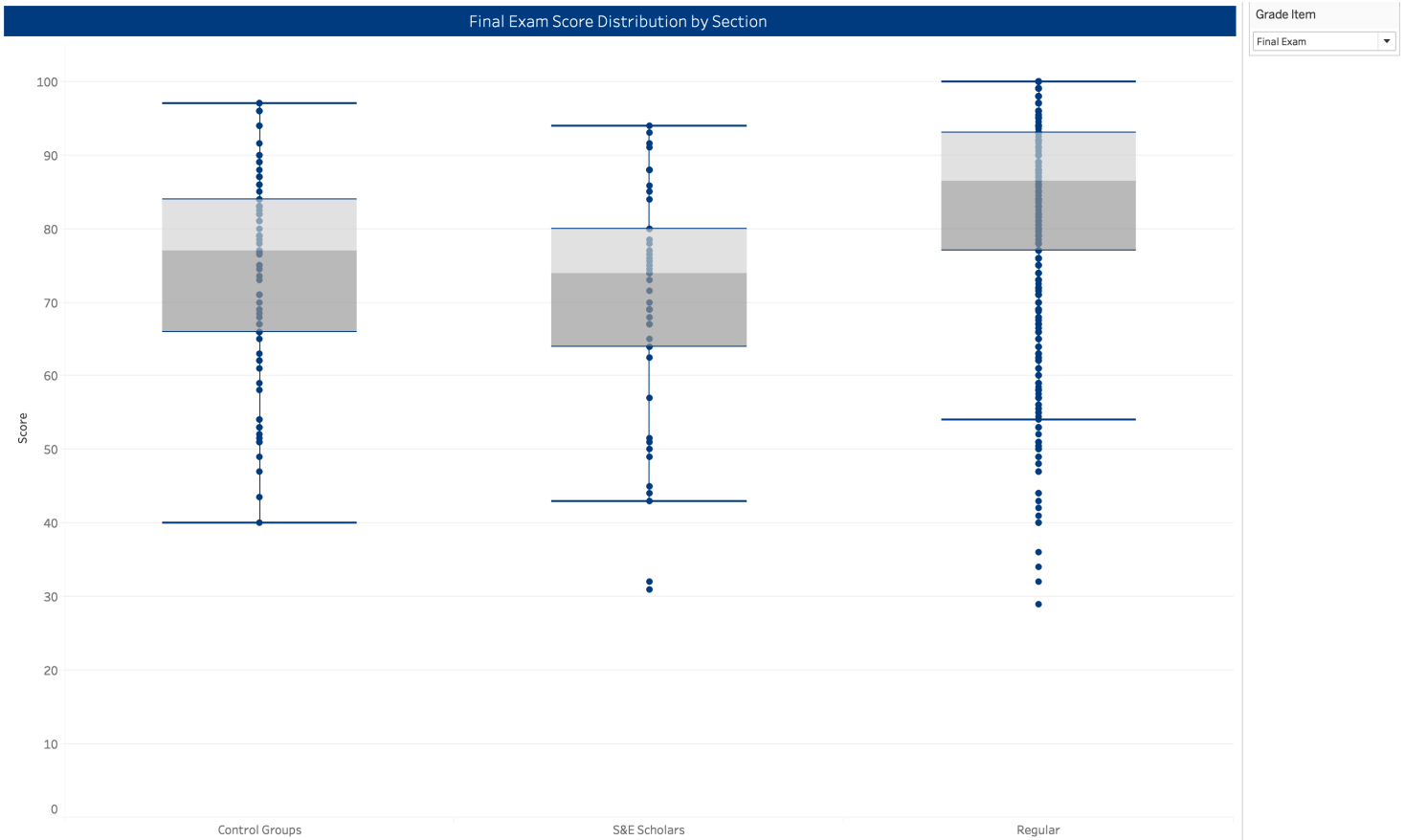


This chart indicates a negative correlation between the Academic Readiness (AR) score and the course calculated grade.

Individual Assessment Performance Analysis

- How did the S&E Scholars perform on individual assessment in the Introduction to Chemical Principle course?

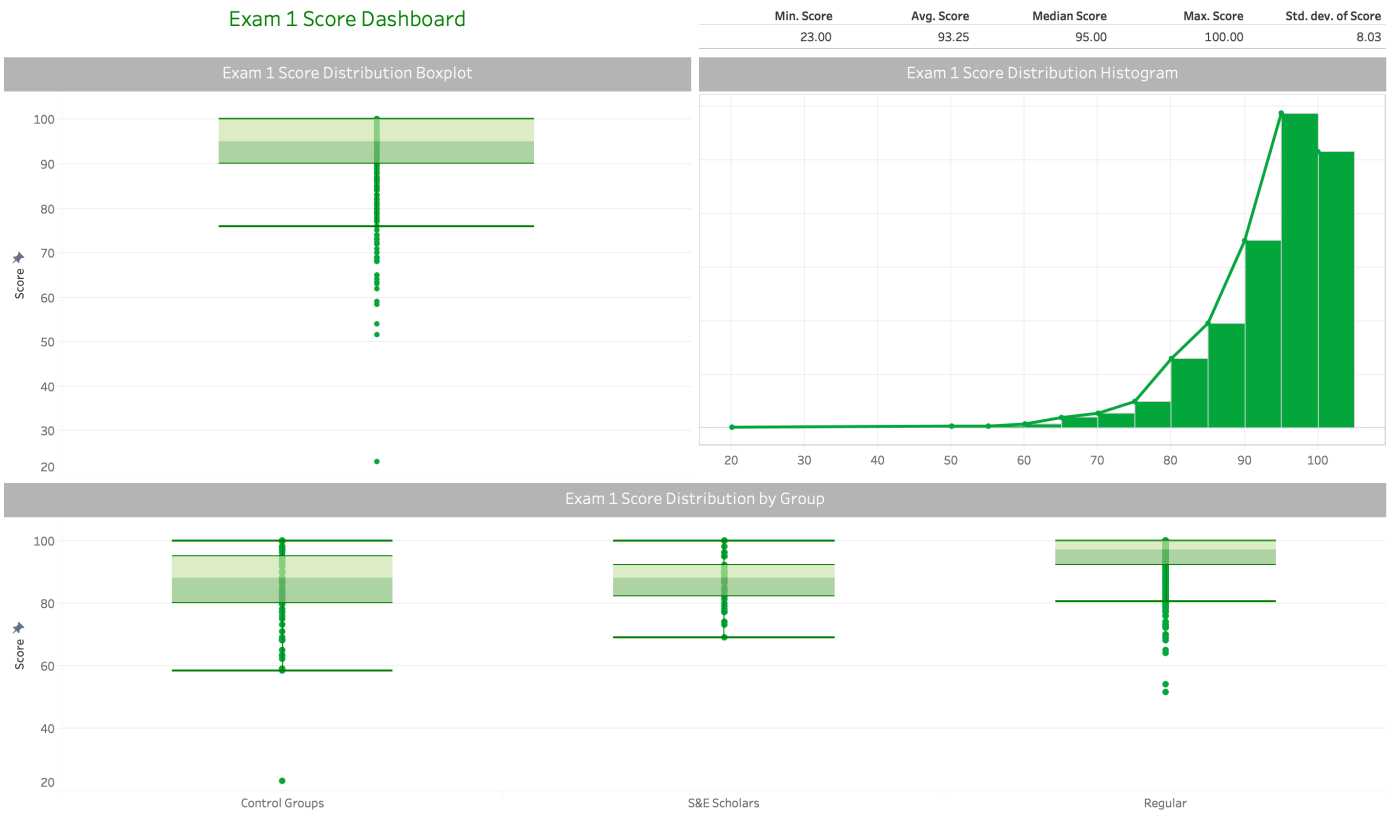
- Compared to other groups, which assessment did the S&E Scholars demonstrate a performance gap?



This chart compares the performance gap between different groups on Exam 1.

Assessment Item Analysis

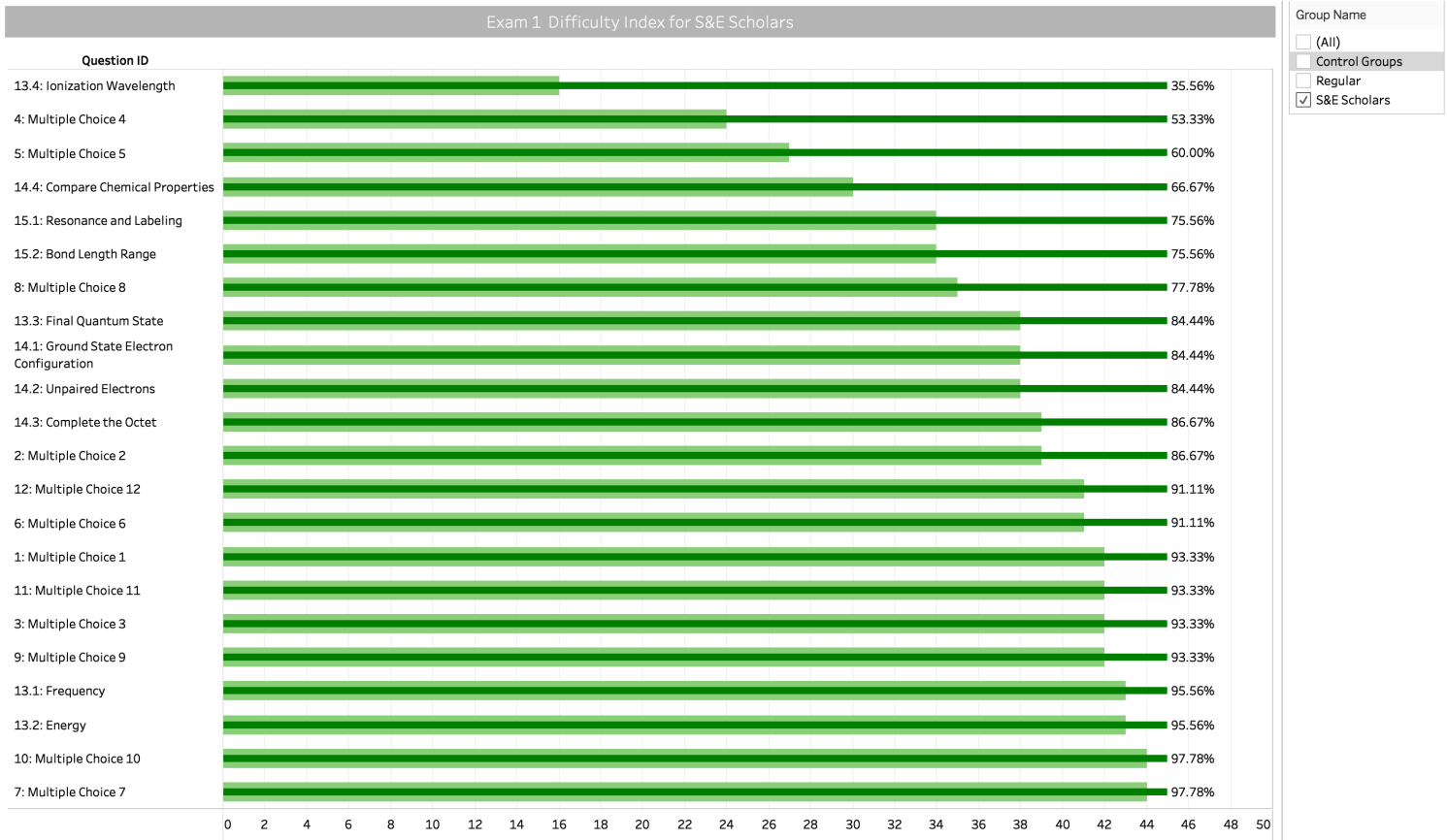
- What is the difficulty index for each question item in an assessment?
- What is the discrimination index for each question item in an assessment?
- Do the questions demonstrate difficulty consistency between different groups of students?



This dashboard shows the stats and distribution of Exam 1 scores.

This graph shows the difficulty and discrimination indexes of each item in Exam 1.

[Item Analysis Reference]



This chart shows the difficulty index of each item in Exam 1 for the S&E Scholars.

This dashboard shows the total attempts and correct counts for each item in Homework set 9.

THANK YOU!

QUESTIONS?

Contact information for the team:

- Patrick Miller (pmiller@nd.edu)

- Xiaojing Duan (xduan@nd.edu)

- Alex Ambrose (gambrose@nd.edu)

- Kevin Abbott (kabbott@nd.edu)

Credit: Photo by Loïc Romer

apereo