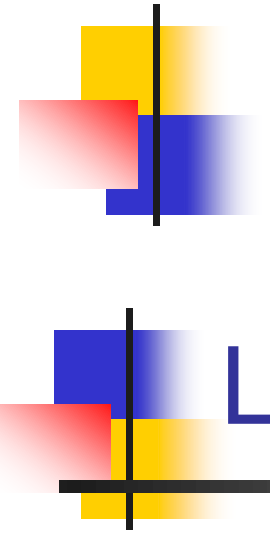
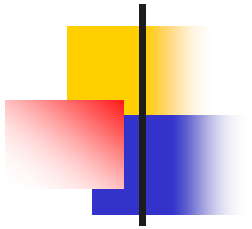


Learning Analytics

Dr. Terry James

2017



- A few questions as thematic anchor

Analytic Insight

- You see which chapter the class finds difficult.
- You drill down to see which topic in the chapter the class finds difficult.
- You drill down and see which students in the class find a topic difficult.
- You drill down to which type of question in a topic a particular student finds difficult.

Do we need tests?

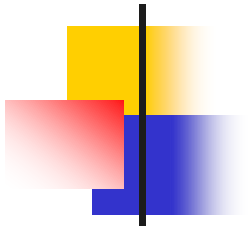
- You can tell BEFORE the test
  - Where the class will struggle,
  - the students that will succeed,
  - the topic some students need help on,
  - and the level of difficulty for a topic that needs to be re-taught.
- Do we actually need tests?
  - Skip the test and just hand out grades?

Skip the test?

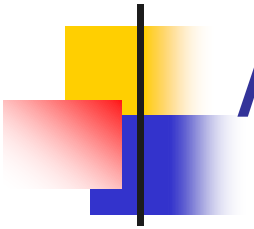
- You cannot skip the test, because people like to believe they have freewill, they are not predictable, …
- Should you check for cheating if a student test result does not match the analytic predictive model?
  - We can constantly add improvements to the analytic equation/model.

Could this change teaching?

- Sort results for difficulty ratio by topic
  - If you improve materials, target your work on the topics that helps the most
- Offer to help students who need it
  - Some students don't ask for help
- How do we teach without analytic data?
  - It is like having no map or GPS.



Basics



Analytics

- Analytics is used in business to improve performance by find important hidden patterns in business data.
  - Every transaction creates data.
  - Data is increasingly pervasive with computers.
  - How can we use data to improve?
  - What gold is in data awaiting discovery?

Learning analytics

- If we capture more data in education, can we use analytics to improve performance in education?
- Some examples:
  - Predict individual performance
  - Personalize for individual student needs
  - Improve accountability
  - Improve assessment feedback
  - Recommend resources
  - Improve student success (Papamitsiou & Economides, 2014)

Background to level set

- 2010 –launched a new Statistics course.
- Textbook, 60 video lectures, and millions of practice online questions.
- Every view of the textbook, lectures, or questions is logged for every student.
- Can we find pedagogical gold using learning analytics in this data?

Traditional course data

- For traditional courses, the weight and emphasis in math courses is usually *tests* and the final *course grade*.
- Educators assume the course grade reflects student learning from reading, lectures, and homework questions.

Learning analytics

- In learning analytics, we do not assume a course grade reflects textbook reading and/or homework questions.
- In analytics, we measure reading, video lectures, or homework and use the data to predict grades (ElAtia, Ipercel, & Hammad, 2012).
- We build equations to predict performance

Learning analytics granularity

- When using learning analytics, you must pick the level of data granularity?
- We use *topic* as the level of granularity.
  - A chapter consists of many topics.
  - Every time a student reads the book, watches a lecture, or does a question, we log the count, day, and the topic.

Adaptive learning

- We use adaptive learning.
- The difficulty of questions automatically adjusts to student ability.
- The pace of the course is *personalized* to student ability.
- Difficulty varies from 1 to 7 levels depending on the complexity of the topic.
- If you answer 4 questions correctly in-a-row at the most difficult level for the topic, you *complete* the topic and can move ahead.

Data capture for assessment

- For each student question, we log:
  - Topic
  - Level of difficulty
  - Correct or Wrong answer
  - Count of correct and wrong answers in a row
  - Topic Complete (yes/no)

    (Computer = 3 correct in a row at high level for topic)
  - Total number of questions for level
  - Ease index for student for level

    = number correct / total questions

High Level Results

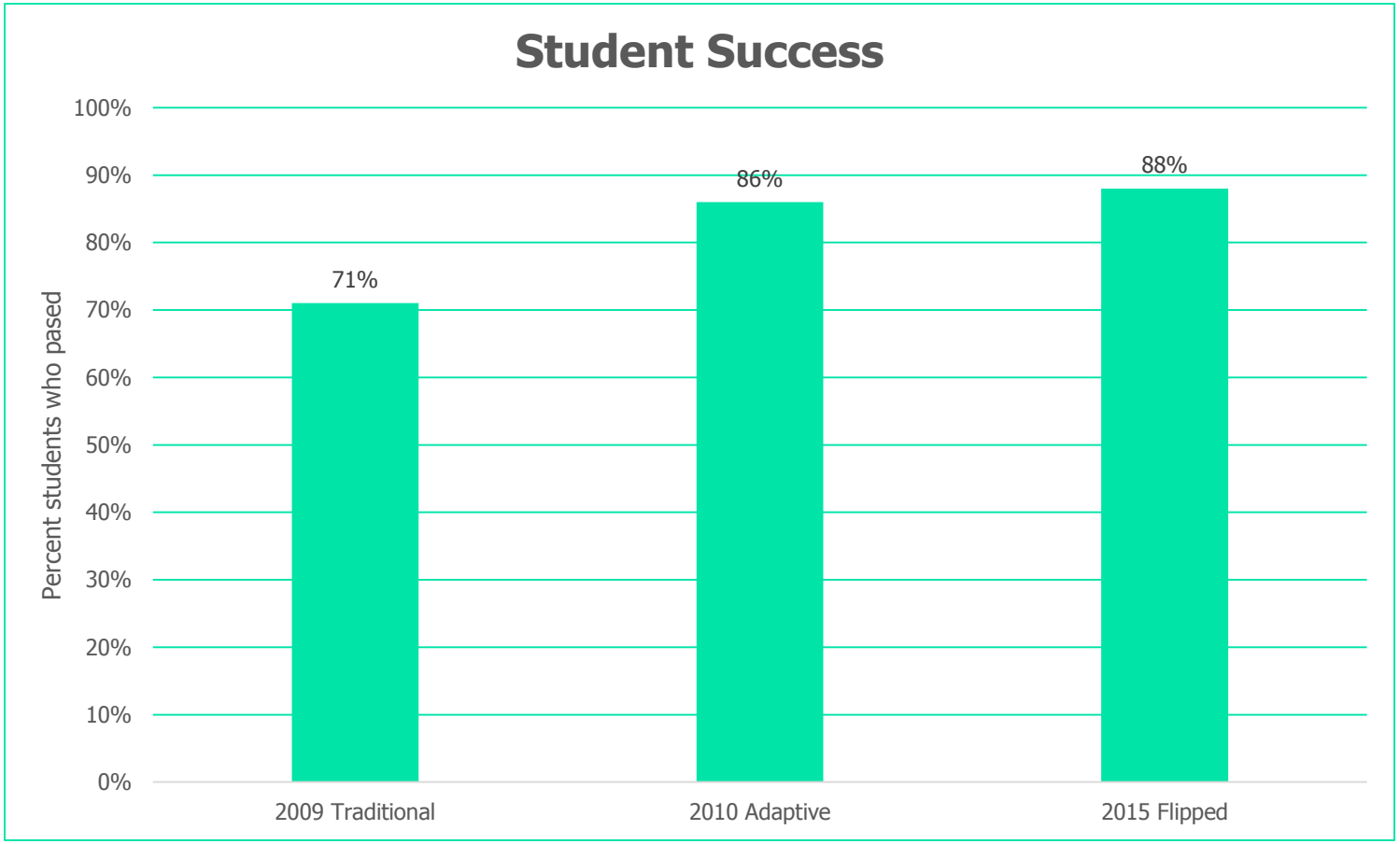


Chart 1: Percentage of students who passed the course using traditional lecture, adaptive learning, and flipped classroom.

Overview

- 2009 – Traditional
  - This is the benchmark
  - Standard lecture format
  - Sample size (n) = 143 students
- 2010 – Adaptive
  - Online questions. Personalized. Self-paced
  - n= 85
- 2015 and 2016
  - Adaptive learning plus online textbook and video lectures
  - High score, alert system for students in trouble
  - Flipped classroom, blended learning, hybrid learning
  - n = 249

Basic totals: Flipped classroom

- 249 Students in sample.
- Engagement
  - *Reading* -12,821 topics read
  - *Lectures* -10,444 video lectures seen
  - *Practice* -85,999 practice problems

- Students are engaged.
- Students found value in the resources.

Basic statistics – flipped classroom

- Grades

Mean	67% C+
Median	68%
Standard Deviation	17%
Maximum	97%
Minimum	0%

- The key result is that 88% of students passed the course which is very high for a statistics course.

Basic statistics

- Textbook – average topic count per student

Mean	47
Median	35
Standard deviation	44
Maximum	206
Minimum	0

- We measured 31 topics of the textbook using 249 students.

  Students are reading topics multiple times.
  - reviewing textbook examples when working on questions, reviewing before a quiz is common.

Basic statistics

- Video lectures -average topic count

Mean	41
Median	29
Standard deviation	42
Maximum	185
Minimum	0

- We measured 31 topics of the lectures with 249 students.

  Students are watching lectures multiple times.
- Some students learn using the book, others prefer video lectures, or both. Students can learn using any resource.

Basic statistics

- Practice questions

Mean	319
Median	293
Standard deviation	144
Maximum	800
Minimum	2

- One reason for improved student success is the large number of practice questions.
- 293 unique questions per student using paper requires a professor grade 72,957 questions (249 x 293) by hand. Too much for a human. We need a computer.

Analytics - median per student by grade

Grade	Book	Lecture	Practice	Complete	Average Grade
A	53	24	365	33	85%
C	49	44	358	32	65%
F	21	19	173	20	38%

- Students with higher grades generally read more, practiced more, and completed more topics.
- Notice video lectures declines in the A grade.
- To *complete* a topic, the student must do 3 questions correct in-a-row at the highest level of difficulty.
- We measured 33 topics above.

Ease ratio

- We calculates the ease ratio for every topic and level of difficulty for every student. (High score = easier).

Topic	Average ease ratio by topic
Hypothesis	.62
Percentile	.62
Independence	.66
Discrete Variance	.69
Binomial distribution	.73

- We can use this analytic measure to improve instruction materials and time allocated for more difficult topics.
- Note: the literature calls (number correct/total) a difficulty ratio but some feel wrong/total shows difficulty and correct/total shows ease (Frey, 2006).

Ease ratio

- If we look at ease ratio within difficulty levels, insight is gained.
- Harder questions (levels of difficulty) may be easier for students (high ease ratios are easier)!

Topic	Level 1	Level 2	Level 3
Percentile	.52	.73	
Binomial	.76	.67	.73
Hypothesis Proportion	.53	.69	.66

- For percentile or hypothesis, it seems grasping the initial concept is more difficult than moving to a more complex question.

Predictive Analytics -Practice

- Does over-practice on questions improve success in statistics?
- Correlation shows a statistically significant relationship between number of *practice* questions and final *grade* at the 95% confidence level. (P-value = 0.0000)
- R^2 indicates 18% of the final grade was explained by number of practice questions.

Predictive Analytics -complete

- To complete a topic, the student must correctly answer questions at all levels of difficulty.
- Correlation shows a statistically significant relationship between *completing* topics and *final grade* at the 95% confidence level. (P-value = 0.0000)
- R^2 indicates 43% of the final grade was explained by number of topics completed.
- Clearly the difficulty and variety of questions completed is more important than the number of practice questions.

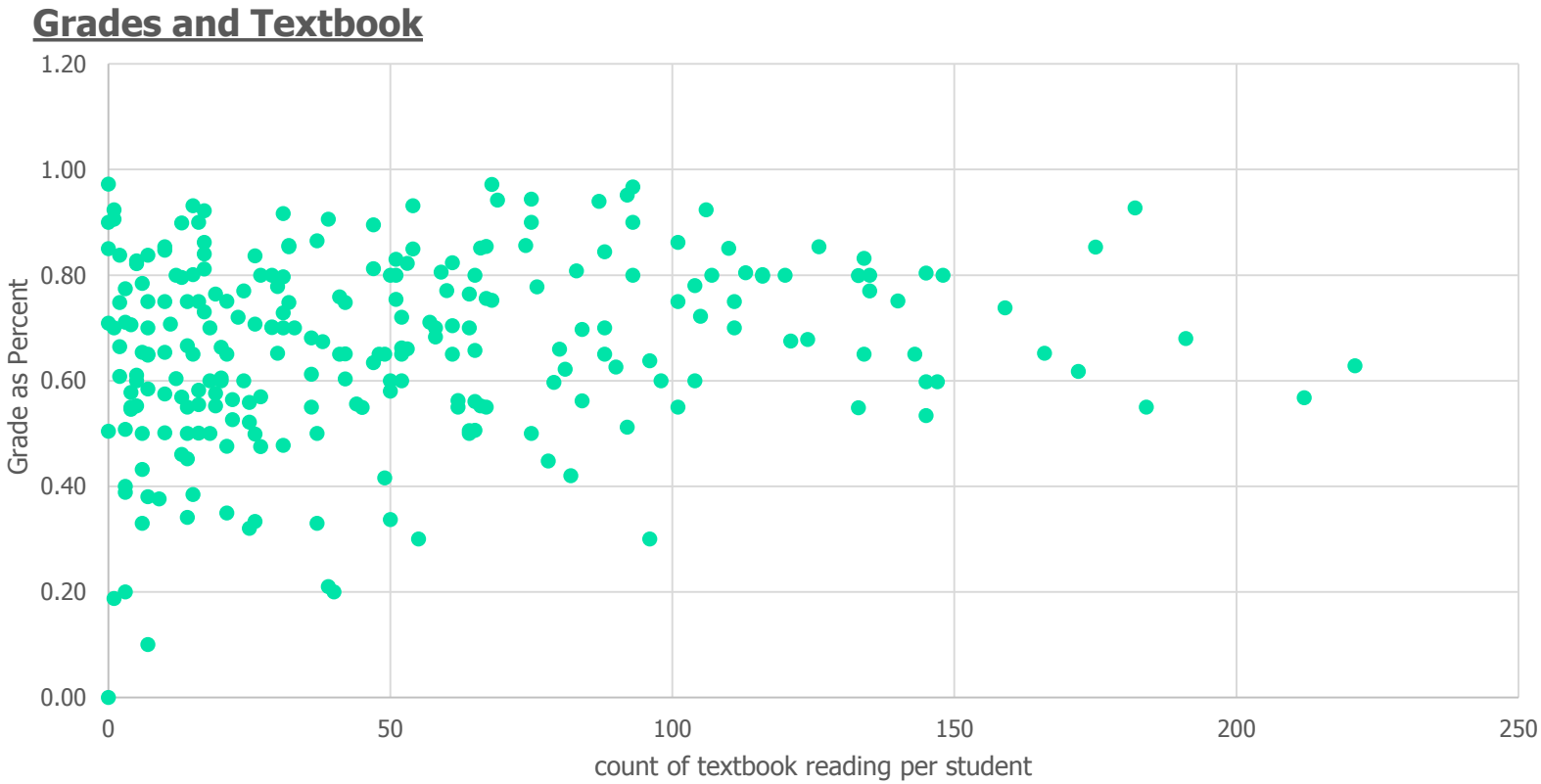
Predictive equation

- Final *grade* (percent) = 36.8 + *2*.26*(complete*)
- For each topic completed, final grade will increase by 2.26%
- Multiple regression of both *complete* and *practice* variables to predict *grade* did not improve the R^2 of .43 value (P-value = 0.000) provided by *complete* alone, so we can simplify the equation to one variable.

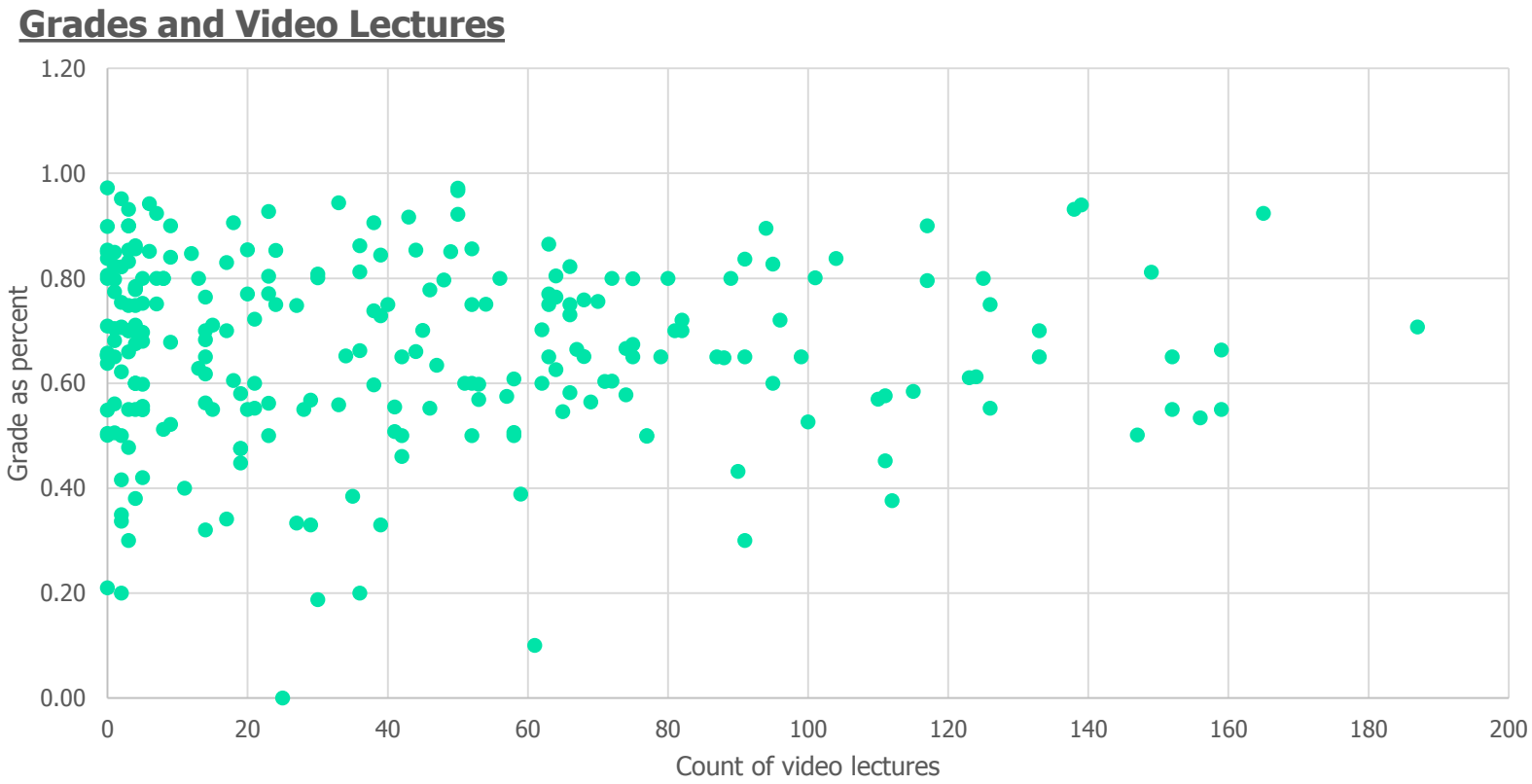
Book and Video lectures

- As a professor, my expectation is any view of the textbook or video lectures will increase understanding and eventually the grade.
- Analytic regression analysis showed no relationship between the count of textbook or video access and final grade.
- A graph of textbook and grade, or video and grade, does not show a clear linear or non-linear relationship. The chart shows no relationship.

Grades and textbook (r^2 = .03)



Grades and video ($r^2 = 0.00$)



Book and video analytics

- We looked at this lack of relationship between textbook and grade as follows:
  - Multiple regression
  - Logistic regression using pass-fail grade
  - Trimmed 10% of outliers
  - Created a video+book index
  - Regression by grade cluster: A,B,C,D,F.
- No statistically significant effect ratio was found for book/video and final grade.

Book/Video and Grade

- My thought is too many uncontrolled variables.
  - Prior learning, help from professor, help from friends, access to Wikipedia, …
- We need a controlled experiment for the book and video variables to prove a relationship.
  - Pre-assessment
  - Read one topic.
  - Post-assessment.
  - Use a controlled environment.
- My prior bias was so strong I expected that under any condition you could see a relationship.
- Does reading matter?

Controlled experiment and Adaptive learning

- Adaptive learning personalizes learning to individual needs.
- Controlled experiments standardizes learning to control for all variables except the intervention variable.
- There is a conflict of goals between individual adaption and experimental control.

Student survey on Video

- **Question: Why you want the video lectures?**
  Mark ALL answers that apply.
  - If I miss a class because of work responsibilities 63%
  - If I miss a class due to illness 56%
  - If I am late to a class and miss material 63%
  - Videos are helpful as English is a 2nd language. 56%
  - Video help if you sleep in or work on an important assignment instead of coming to class. 38%
  - I like the videos to check for mistakes in the practice questions. 81%

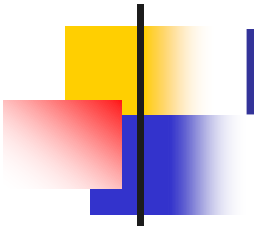
Dashboard (www.growingknowing.com, 2015)

Statistics | How To | About Us | Help

Dashboard for Professors showing Practice Progress by Topic

Section: TTT
Date: 16-November-2016
Active Students: 3
Inactive Students: 0

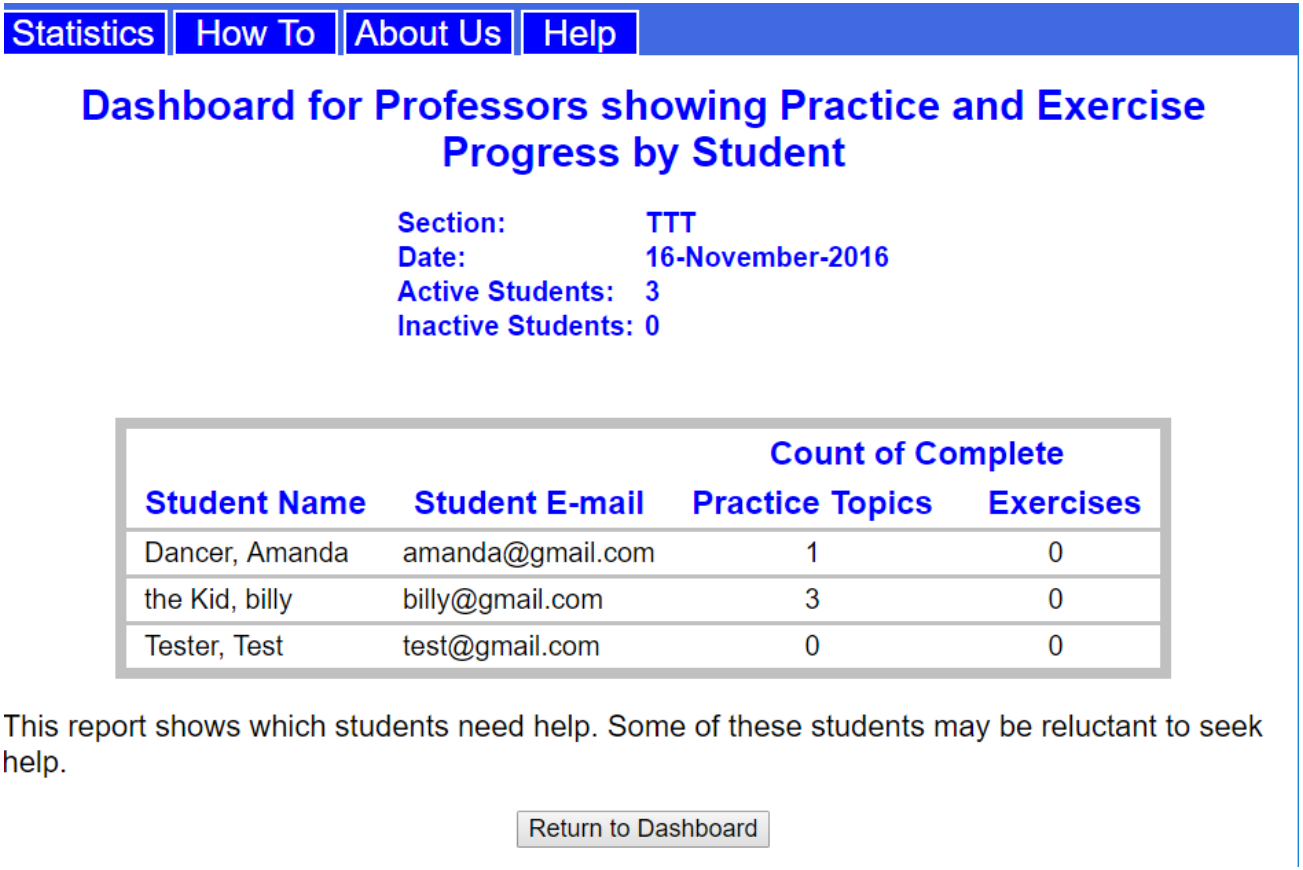
Topic Name	Number of Students Incomplete
Mean	2
Median	2
Mode	2
Range	3
Variance	3
Standard Deviation	3
Skewness	3
Coefficient Variation	3
Empirical Rule	3



Dashboard

- With the dashboard, a professor can see which topics the class finds difficult.
- The professor can re-teach or provide more time for difficult topics.

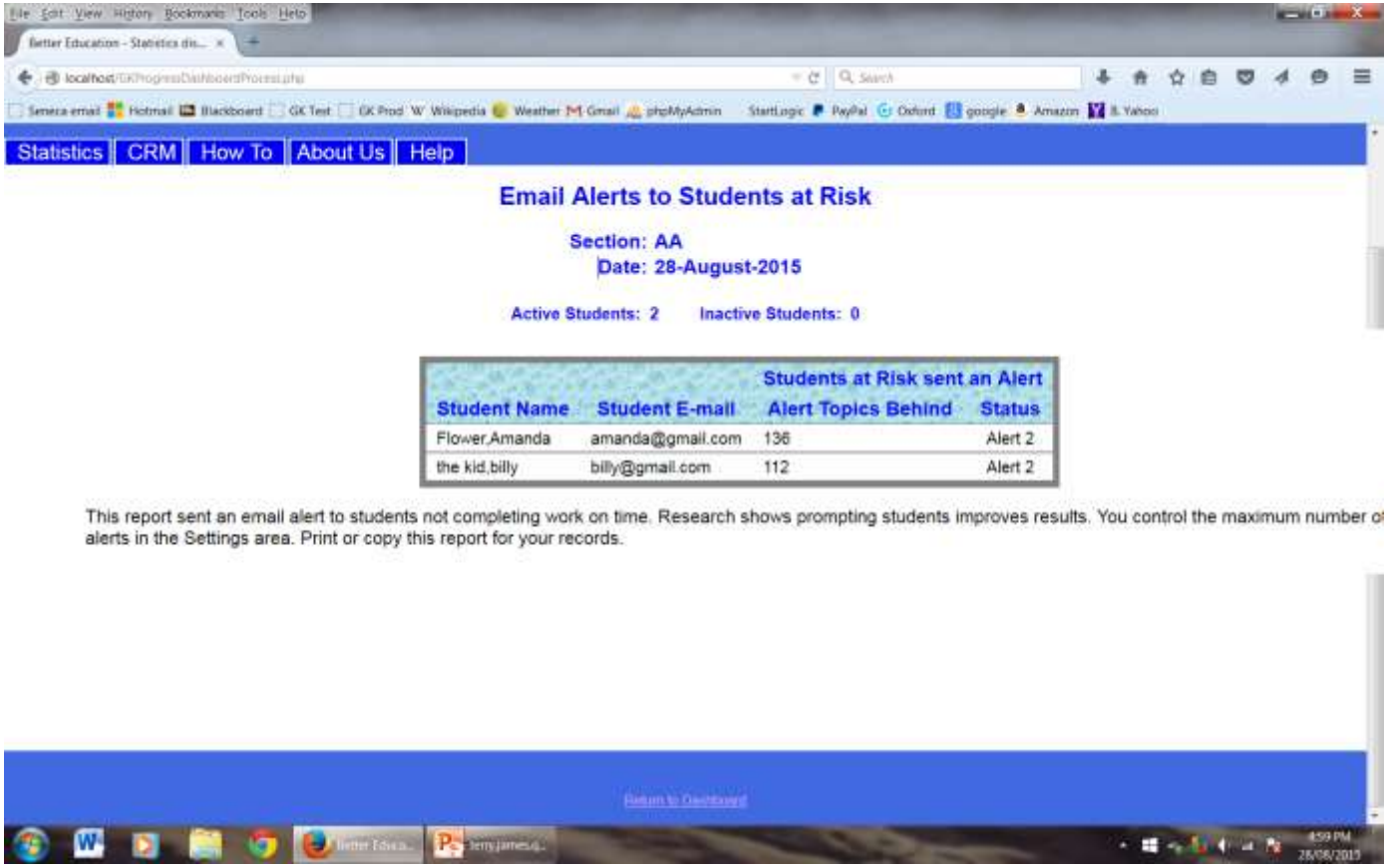
Dashboard – drilldown



Dashboard - drilldown

- A professor can drilldown to find who in a class finds the topic difficult.
- The professor can then do an intervention.
- Some students are reluctant to ask for help.
- A professor can predict how the class will do and who will succeed BEFORE the test.
- Note: the dashboard reports provided and student names show fake results to protect student privacy. (www.growingknowing.com, 2015)

Alert (www.growingknowing.com, 2015)



Alerts

- The system tracks performance so multiple levels of automated alert can be sent to students.
  - Does the student need attention?
  - Help? Encouragement? A plan?
- We do not have a large enough sample to have statistical significance (not enough people fail!)
- Alerts can be tailored by each professor.
- About half the students who get an alert make a commitment to catch-up. They pass the course.
  - Currently a first alert is sent if you are 25% behind schedule

High score –highly motivating

Summary – Learning Analytics

- We improved student success significantly.
- We can predict performance using an equation.
- We see who needs intervention before the test.
- We can automatically alert students before it is too late.
- Analytics allows us to sort topics by difficulty for the class, student, or question.
  - We can allocate time when most needed
- More research is need into textbook and video lecture analytics.

References

ElAtia, S., Ipperciel, D., & Hammad, A. (2012). Implications and challenges to using data mining in educational research in the Canadian context. *Canadian Journal of Education, 35*, 2, 101-119.

Frey, B. (2006). *Statistical hacks: Tips and tools for measuring the world and beating the odds.* Sebastopol, CA: O'Reilly.

Papamitsiou, Z. & Economides, A. (2014). Learning Analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society, 17* (4), 49-64.