

AMSTATNEWS

The Membership Magazine of the American Statistical Association

[Printed Issues](#)
[Practical Significance Podcast](#)
[Additional Features](#)
[Columns](#)
[Member News](#)
[Departments](#)

SEARCH

Search the site ...

MORE TO SEE



2026 SDSS Registration Open; Lightning Talk Submissions Sought

January 1, 2026 By [Meg Ruyle](#)



Meet New ASA Members: Spotlight on William Brown

January 1, 2026 By [Meg Ruyle](#)



Connection, Contribution, and Evidence

January 1, 2026 By [Meg Ruyle](#)



Q&A with ASA Founder Mark Glickman

January 5, 2026 By [Meg Ruyle](#)



You are here: [Home](#) / [Additional Features](#) / Mutual Benefits of Data Sharing: Insights from CourseKata and Savills

Mutual Benefits of Data Sharing: Insights from CourseKata and Savills

September 2, 2025 — [Leave a Comment](#)

Who provides the data behind student projects—and what do they gain in return? In the following pieces, donors share their motivations and takeaways.

CourseKata's Experience as a DataFest Donor

Ji Y. Son, Claudia C. Sutter, and James W. Stigler

CourseKata is a research and development project devoted to improving how students learn statistics and data science. We design innovative online curricula for use in high school and college classrooms, and we partner with learning scientists to figure out how to improve the teaching and learning experience by improving the curriculum.

We've long argued that education needs an R&D arm because the same kind of data-driven insights that fuel innovation in tech, medicine, and policy could (and should) improve how people learn. But here's the catch: Education data is notoriously hard to access, interpret, and act on. We're trying to change that.

When we first served as judges at the University of California at Los Angeles site of DataFest, we were blown away by the quality of student thinking. These weren't just shallow dashboards or flashy visuals. They were thoughtful, creative, and rigorous analyses of authentic data. Each team flashed the same data set from a different angle, like the proverbial blind men encountering different parts of an elephant. Then, in just two slides and five minutes, each team presented their part of the story. And by the end of DataFest, a fuller picture of the data had emerged.

So, in 2024, CourseKata became a DataFest donor.

What We Donated

We shared an anonymized [data set](#) of student behavior and learning outcomes from 1,625 college students across 48 courses at 11 institutions, all using our online curriculum in 2023. This data set included engagement logs (e.g., active time on page), responses to formative assessment questions, and measures of psychological constructs related to motivation, mindset, and more. All data was de-identified per institutional review board protocols and publicly documented.

What We Got Back

The student teams brought a range of perspectives and found plenty of surprises. One team investigated cheating and discovered the best source for getting the right answers to CourseKata questions ... was the CourseKata textbook, itself.

Other teams tackled the question of whether instructional videos helped or hindered learning. Their answer? Both. Some found watching videos was associated with better performance, while others found not enough people watched the videos to make them worth it.

Several groups applied clustering techniques to explore patterns across motivation, engagement, and learning. These groups found students tended to fall into three clusters: those who engaged a little and got a lot right, those who engaged a lot and got a lot right, and those who engaged a little and got a lot wrong. These distinctions are helping us rethink how we interpret "time on task" and how to support different learner profiles.

One group that especially impressed us, Team Abercrombie, analyzed the behaviors of the highest-performing students and found they were the students most likely to revisit earlier sections of the textbook when working on new problems. Based on this insight, the team designed a system called ABER (Automated Built-in Engine for Review). ABER used vector embeddings to represent all textbook paragraphs and questions, then used similarity matching to recommend the most relevant sections of the text for review after a student answered a question.

We were so intrigued by their work that we invited several members of the team—Justin Gong, Hairan Liang, and Lukas Hager, all from UCLA—to join us as interns. They helped us refine a prototype of ABER, explored ways to use large language models to match new formative assessment questions to our learning goals, and even worked on generating new questions for concepts that were under-measured in our current curriculum.

Other DataFest participants have since joined us as interns, continuing to bring their creativity and curiosity to real-world projects. Over the past year, they've contributed to initiatives such as prototyping researcher dashboards, visualizing motivation and engagement across different textbook versions, and tracking changes in these measures across academic years.

For many of these students, it was their first opportunity to work directly with educational data in a research and development setting. They told us how much they appreciated the chance to work on projects with meaningful (and sometimes real-time) impact.



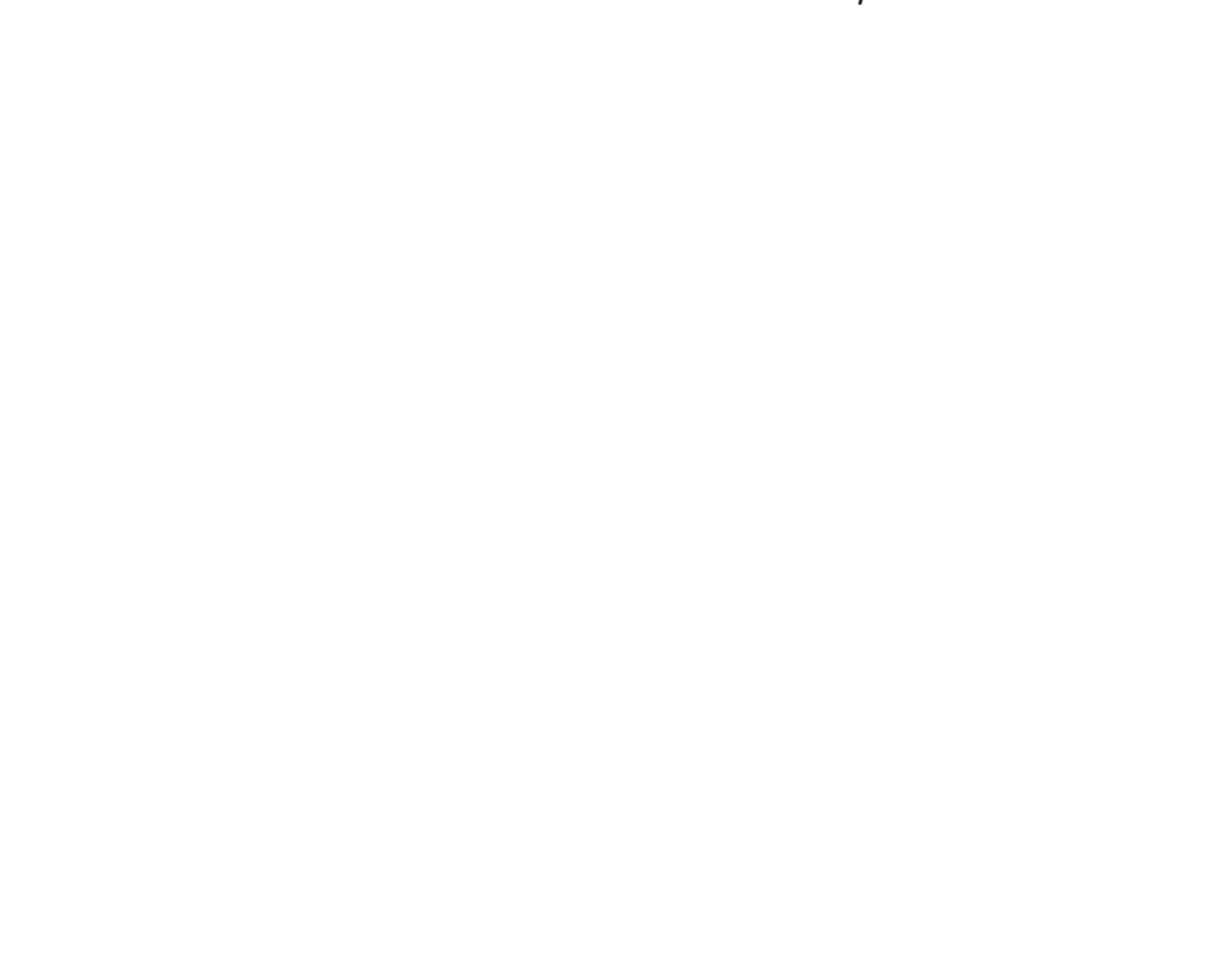
University of Toronto students watch the CourseKata introduction video (Jim Stigler and Ji Yun are pictured in the video).

possible when curious minds are unleashed on real data. Because these student researchers are closer in age and experience to the learners in our data set, they brought a much-needed perspective we don't get from seasoned researchers or curriculum developers. For CourseKata, it was a rare opportunity to be both generous and selfish at once: to support the next generation of data scientists while also learning more about how students learn.

If we want education to be truly evidence-based, we need to be thoughtful (and even a little shrewd) about how we use data to drive impact in a world in which data is too often used just to sell ads, products, and services. That means sharing it, contextualizing it, and being open to what others might find. DataFest gave us all that, and it reminded us that sometimes the best analysts are the ones just starting out.

Savills Workplace Studio

John Rissmiller, Associate Director, Savills Workplace Studio



John Rissmiller is associate director of Savills Workplace Studio

both qualitative and quantitative data analysis to help clients reimagine their workplaces.

In my current role at Savills Workplace Studio, I help build out a database of benchmarking data. I also get data from our clients and analyze it to help them make decisions. The ability to distill lots of data down to a quick and effective story was something I first began practicing at DataFest and is a skill I employ every day. Finding my love for data and complexity at DataFest has driven my career and had a lot to do with why I chose my current organization.

My organization uses data in everything we do, from helping with leases and designing spaces to managing the build-out of spaces. Data and being active in our communities is in our DNA. So, the opportunity to use our data to give back and give opportunity to students around the world was something everyone at Savills was excited to do.

Our data set considered a wide variety of corporate real estate data, from rents and locations to sizes and time. It was critical to provide data that could be looked at in a variety of ways and complexities. This allows students of all skill sets to engage and have fun during what is an intense few days.

“I highly recommend any organization interested in being a data donor to do it and go to a competition. It's truly inspiring and a great way to give back to the next generation of budding data specialists.

As a former participant, it was great to see students engage in our data and think about it from a totally different perspective than I or my team might. More importantly, after a few events, I got LinkedIn messages from students who wanted to tell me about how much fun they had and how interested they are in joining the industry. It's great to see that our work can be inspiring for students today.

There is nothing quite like DataFest for students. Even coming from a school that focused on project-based learning, the mix of the time crunch, the large data, working with a team—all of it combines into a really fun weekend in which you learn so much, not only about statistics, but about yourself, your skills, and where you want to grow and develop.

I highly recommend any organization interested in being a data donor to do it and go to a competition. It's truly inspiring and a great way to give back to the next generation of budding data specialists.

Filed Under: [Additional Features](#), [DATAFEST](#), [Special Features](#)
Tagged With: [benchmarking data](#), [Claudia C. Sutter](#), [communities](#), [CourseKata](#), [James W. Stigler](#), [Ji Y. Son](#), [John Rissmiller](#), [Savills Workplace Studio](#)

Leave a Reply

Your email address will not be published. Required fields are marked *

Comment *

Name *

Email *

Website

☐ Save my name, email, and website in this browser for the next time I comment.

POST COMMENT

ASA HOME

[American Statistical Association](#)

[Communications from the Executive Director](#)

[ASA Leader Hub](#)

[ASA Career Connect](#)

ADVERTISERS

Stata

ARCHIVES

Select Month

CATEGORIES

Select Category

EDITORIAL STAFF

Managing Editor
Megan Murphy

Graphic Designers / Production Coordinators
Olivia Brown
Meg Ruyle

Communications Strategist
Val Nirala

Advertising Manager
Christina Bonner

Contributing Staff Members
Emily Fekete, Naomi Friedman and Kim Gilliam

American Statistical Association • 732 North Washington Street • Alexandria, VA 22314