

Learning Analytics

INFO 4100

INFO 5101





Prof. Rene Kizilcec
Information Science



Complete the survey this week


104 responses thus far



Need more for later HW


 Course / Week 1: Introduction / About the Course 

[< Previous](#)





[Next >](#)

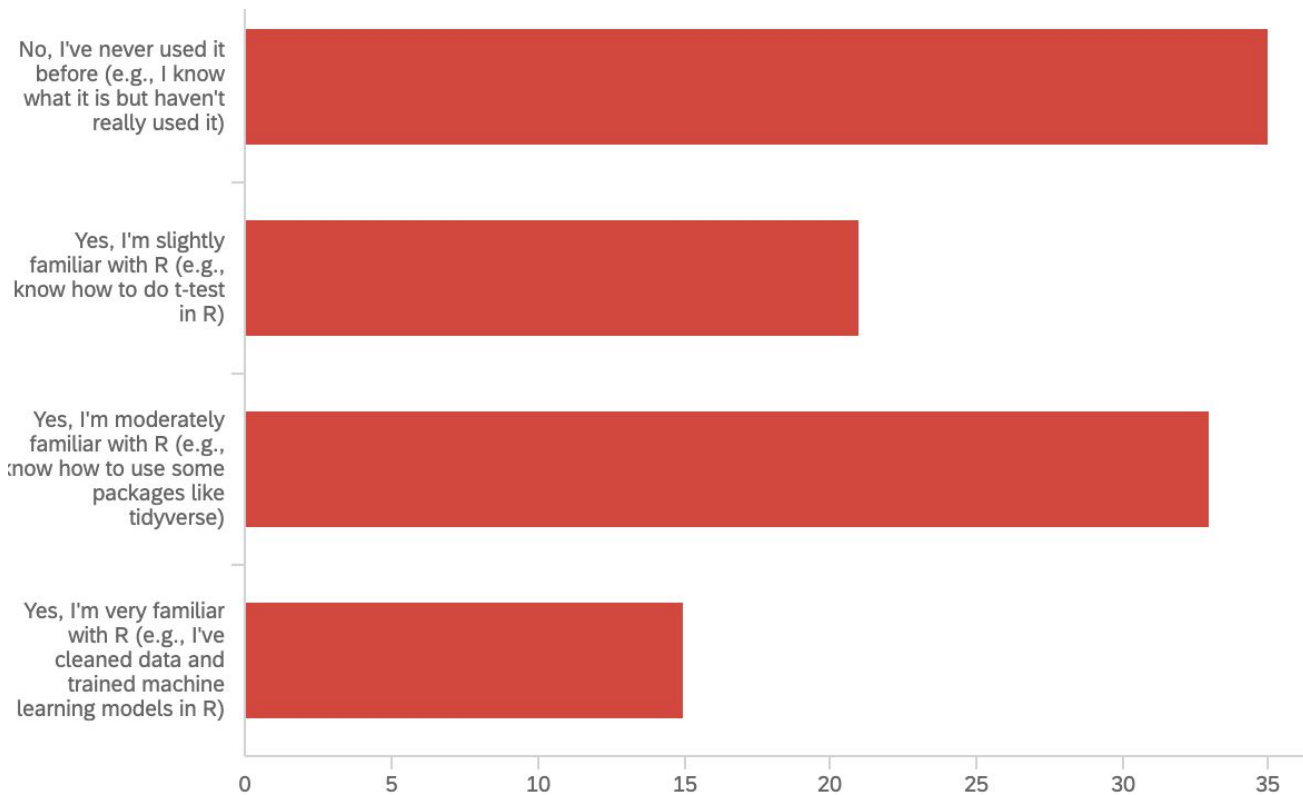
Start Survey (required)
[Bookmark this page](#)

Required Survey for Learning Analytics (INFO 4100/5101)

This 20-minute survey has two important goals:

1. to understand who is in the course and calibrate content to your level of background knowledge
2. to generate authentic data that you will analyze later in the course for homework assignments

You *R* not alone ...



Course flow & first R homework

- Lab sections are for working with R / homework support
- First HW is challenging intentionally
- You have two weeks to practice before the next R HW
- Remember to submit everything on edX
- GenAI help is allowed (LLM-based TA almost ready)

Learning to Make Data Science Decisions



Today at 9:19 AM

I found some entries in the dataset where attempts are 0, but the correct in the first attempt is 1 and seconds > 0 . Do we need to filter out these incorrect entries before answering all the questions? (edited)



1



Real data is messy,
and I didn't clean it up

Understand the issues
and make justifiable
decisions

Careers Connected to Learning Analytics

- Lots of interest in learning about / pursuing LA careers
- Invited practitioners to share first-hand experiences

Videos interspersed in the course materials

Careers Connected to Learning Analytics



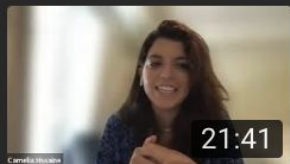
Interview with John Willse (Udemy) - Cornell INFO 4100 Learning Analytics

John Willse, Senior Data Science Manager at Udemy, talks about his path from academia to the education technology industry when he observed a change in the industry. John shares how Udemey created an assessment product from...



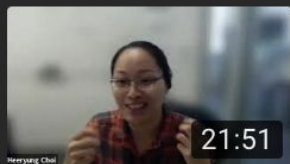
Interview with Alan Hickey (Coursera) - Cornell INFO 4100 Learning Analytics

Alan Hickey, Senior Manager of Machine Learning at Coursera, talks about what drew him to work in machine learning for education, different kinds of machine learning projects, and what it is like to start off in a technology company. This is an...



Interview with Camélia Hssaine (Codecademy) - Cornell INFO 4100 Learning Analytics

Camélia Hssaine shares insights about working in the learning analytics industry from her role as Head of Data Science at Codecademy. This is an interview with Dr. Rene Kizilcec for the Cornell course on Learning Analytics.



Interview with Heeryung Choi - Cornell INFO 4100 Learning Analytics

Dr. Heeryung Choi shares insights from her work on self-regulated learning in online courses and how to measure it with surveys and behavioral data. This is an interview with Dr. Rene Kizilcec for a course on Learning Analytics.

Careers Connected to Learning Analytics

- Lots of interest in learning about / pursuing LA careers
- Invited practitioners to share first-hand experiences
- Feel free to ask me about this in OH
- Reach out to Cornell Career Services
- I'll reach out to YouTube Edu team

Emergence of Big Data in Education

What is driving the emergence of big data in education?

Emergence of Big Data in Education

What is driving the emergence of big data in education?

1. Recording & storing of institutional data became digitized
 - a. Student Information Systems (SIS)

Emergence of Big Data in Education

What is driving the emergence of big data in education?

1. Recording & storing of institutional data became digitized
 - a. Student Information Systems (SIS)
2. Recording student/classroom interactions became easier
 - a. Learning Management Systems (LMS)

Emergence of Big Data in Education

What is driving the emergence of big data in education?

1. Recording & storing of institutional data became digitized
 - a. Student Information Systems (SIS)
2. Recording student/classroom interactions became easier
 - a. Learning Management Systems (LMS)

Anything else you think is important?

Levels of Big Data in Education

Microlevel data.

Mesolevel data.

Macrolevel data.

Levels of Big Data in Education

Microlevel data. Clickstream, timestamped interactions, etc.

Mesolevel data. Submission content, social interaction, responses, affective states, etc.

Macrolevel data. Institutional records, enrollments, admissions, schedules, etc.

Applications of Big Data in Education

Microlevel Data

Identifying Knowledge Components

Identifying Metacognitive and SRL Skills

Identifying Affective States

Evaluating Student Knowledge

Using Data for Actionable Knowledge

Clustering Student Profiles and
Discovering How to Personalize

Applications of Big Data in Education

Microlevel Data

- Identifying Knowledge Components
- Identifying Metacognitive and SRL Skills
- Identifying Affective States
- Evaluating Student Knowledge
- Using Data for Actionable Knowledge
- Clustering Student Profiles and
Discovering How to Personalize

Mesolevel Data

- Supporting and Evaluating Cognitive Functioning
- Supporting and Examining Social Processes
- Detecting Behavioral Engagement
- Examining Affective Constructs

Applications of Big Data in Education

Microlevel Data

Identifying Knowledge Components

Identifying Metacognitive and SRL Skills

Identifying Affective States

Evaluating Student Knowledge

Using Data for Actionable Knowledge

Clustering Student Profiles and
Discovering How to Personalize

Mesolevel Data

Supporting and Evaluating Cognitive Functioning

Supporting and Examining Social Processes

Detecting Behavioral Engagement

Examining Affective Constructs

Macrolevel Data

Early-Warning Systems

Course Guidance and Information Systems

Administration-Facing Data Analytics

Three Themes from Student Reflections

Practical Value and Scientific Insight

Data Ethics and Privacy

Methods and Analysis

Group Activity (4–5 students)

1. Pick a level (macro, meso, micro) and application from the reading (link: <http://tiny.cc/1obyuz>)
2. Select and examine ONE (or 2 if strongly connected) of the papers highlighted for that application (click citation, get PDF)
3. Identify and discuss the paper with respect to three themes
 - Practical Value and Scientific Insight
 - Data Ethics and Privacy
 - Methods and Analysis

Group Activity Guiding Worksheet

Level: _____

Application: _____

Study title and authors: _____

This study is about _____

The data it uses is _____ from _____

This use of data provides value for ____ (practical) and ____ (scientific)

The methods used to analyze the data are _____

Here we are concerned about _____ (stakeholder – privacy issue)

Group Activity Share-out

Share your findings in class, one for each level

Post your findings on [#in-class-activity](#) channel in Slack

You can start working with big edu data now

Datasets made available by education companies available online, including Duolingo, ASSISTments, eedi Cognitive tutors, and more

Public competitions (some with pretty big prizes)

XPrize

Kaggle



You can start

Datasets made available including Duolingo,

Public competitions

XPrize
Kaggle



Results

Recently Launched ▾



Feedback Prize - English Language Learning

Evaluating language knowledge of ELL students from grades 8-12
Featured · Code Competition · 49 Teams · 3 months to go

\$55,000

...



Feedback Prize - Predicting Effective Arguments

Rate the effectiveness of argumentative writing elements from students gr...
Featured · Code Competition · 1557 Teams · 8 days ago

\$55,000

...



NBME - Score Clinical Patient Notes

Identify Key Phrases in Patient Notes from Medical Licensing Exams
Featured · Code Competition · 1471 Teams · 4 months ago

\$50,000

...



Feedback Prize - Evaluating Student Writing

Analyze argumentative writing elements from students grade 6-12
Featured · Code Competition · 2058 Teams · 6 months ago

\$160,000

...



LearnPlatform COVID-19 Impact on Digital Learning

Use digital learning data to analyze the impact of COVID-19 on student lear...
Analytics · 0 Teams · a year ago

\$20,000

...

