

From Analyzing Abundant Data to Identifying Actionable Steps: A closer look at library student data

Kate Peterson
Mariya Gyendina
Jan Fransen



Can you associate the library with student success?

Yes! Library use is positively associated with...



GPA



Retention



Academic engagement



Scholarship



Academic skills



2011-2019 - focused on
10,000 foot level with “big
data”

GPA (2012)

Compared to their first-year peers who did not use the library, students who use the library at least once have

- Significantly higher **cumulative grade point averages** in their fall semester and first year
- Significantly higher **retention** to their second semester and second year of enrollment

Note: These findings held controlling sex, race/ethnicity, first-generation status, ACT scores, number of AP credits, and other collegiate experiences.

Retention (2016) (Used propensity score matching)

- Using the library at least one time in the first year of enrollment significantly increased the odds that students would graduate in four years OR remain enrolled after four years
- First year students who used electronic resources and books had significantly improved odds of graduation in four years

[Read more](#)

This got us included in
existing programming
but wasn't "actionable"

Layers of Data

**Office of Institutional Research Performance Data
Term and Cumulative GPA, Retention**

**Office of Institutional Research Demographics Data
College, Level, Major, Gender, Ethnicity, Age**

**Libraries Data (13 Access Points)
Circulation, Digital, Instruction, Reference, and Workstation**

No new data: 2015-2016

Impact size
Individual
experience

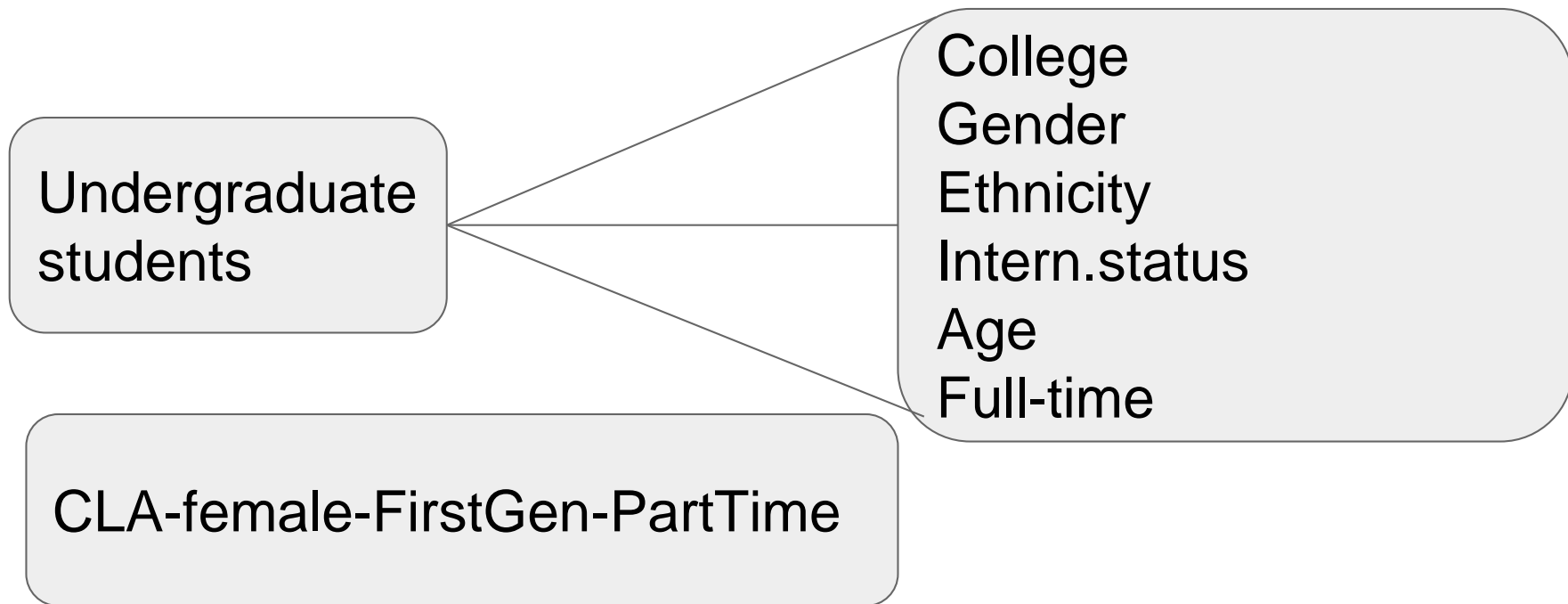


Building
smaller
datasets



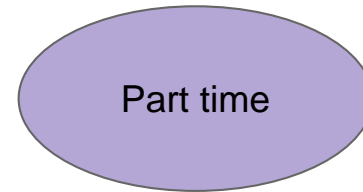
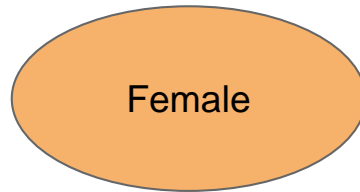
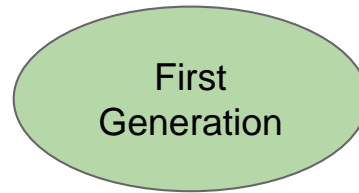
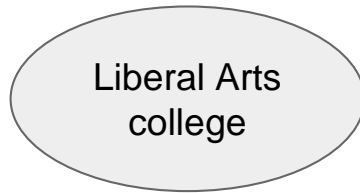
Actionable
steps

Deeper data analysis



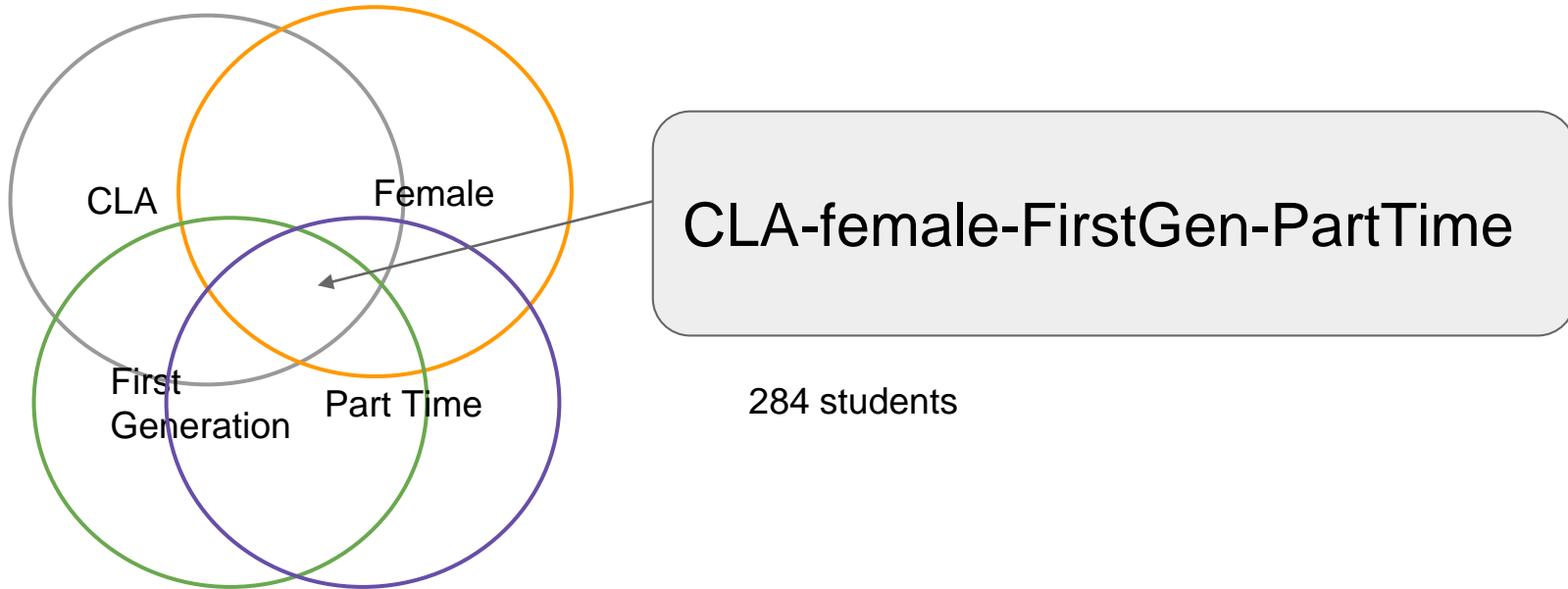
Level 1

Set to analyze = All students that fit ONE criterion



Level 2-4

Set to analyze = Every possible combination of 2, 3, or 4 criterion



Writing code to write code

- Using all the values for all the characteristics and combining two, three, or four of them yields **103,320 sets!**
- Using Microsoft Access, we put each characteristic into its own table with a row for each value, then wrote VBA code to write R code that built each set and conducted the statistical analysis (lots of nested loops!)
- R code wrote out CSV files that we could then review

CLA-female-FirstGen-PartTime

Library use
indicator

GPA of
students who
use Lib

GPA of
students who
do not use Lib

- Statistical significance (p-value)
- Effect size (Cohen's f, eta sq)

Five Undergraduate Colleges

CSE - science&engineering

CEHD - education

CLA - liberal arts

CFANS - food, agriculture&nat.sciences

CBS - biological sciences

Goals - Evidence based decision making

Ideally this data will help the Libraries

- Prioritize work that has the greatest impact
- Identify gaps in the Libraries' impact, specifically related to minority groups.
- Consider new ways to staff and support such work (yes, liaison model but also college teams or new models?)
- Think creatively to reach/teach students
- Measure and sustain its reach to keep new classes of students learning

Why we are looking at instruction

Multiple ways to look at the data

Reach 90+% of students in some cases/by some indicators

Focus on instruction because:

- Even split or fewer students have had it
- Moderate to high effect sizes for the models we built
- Somewhat in our control

Models

- At least 50 students (25 who used the Libraries; 25 who did not)
- Who we do not see:
 - Groups with very high outreach levels, where everybody has contact with the libraries
 - Small groups with fewer than 50/25 students

Examples

Example

- CLA - First Gen
 - Total: 3562, Yes instruction: 562, No instruction: 3000
 - CUM GPA p-value : .044
 - Cohen's f 0.298/Eta sq 0.08
- Goal: Increase instruction and outreach for First Gen CLA students
- Strategy ideas:
 - Talk with CLA student group - [First Generation Students](#)
 - Talk with [CLA President's Emerging Scholars](#)
 - Talk with CLA academic advisors
 - Work to analyze which courses these students are more likely to take - target instruction

Example

- All Colleges - Black, Male, Transfer (notFallNHS)
 - Total: 551, Yes instruction: 56, No instruction: 495
 - CUM GPA p-value: .014
 - Cohen's f 0.37/Eta sq 0.12
- Goal: Increase instruction and outreach
- Strategy ideas:
 - Talk with [Black Student Union](#)
 - Talk with [Coordinator of African American Recruitment](#)
 - Find out where most students are transferring from
 - Outreach to seven historically black Greek letter organizations on campus

Example

- CEHD - Female, First Gen
 - Total: 550, Yes instruction: 102, No instruction: 448
 - CUM GPA p-value: .018
 - Cohen's f 0.354/Eta sq 0.11
- Goal: Increase instruction and outreach for CEHD First Gen
- Strategy ideas:
 - Better integration with First Year Inquiry course (EDHD 1925W)
 - Apply for [First Generation Institute Microgrant](#)
 - More robust outreach with TRIO

**Let's get
going...oh
wait....it's
2020**

Issues with the Data

Data collection

- Missing and incomplete data
 - Course reserves
 - Course integrated instruction (typos, sections, wrong courses, etc.)
 - Gathering users when it doesn't match with traditional course
 - Workshops - registration vs attendance
 - Assuming all students attend all the time

Next Steps

- Identify and pilot targeted outreach/programming
- Follow-up with qualitative work

Questions? Comments?

Let us know!

Library Data and Student Success

- Mariya Gyendina mgyendin@umn.edu
- Kate Peterson katep@umn.edu
- Jan Fransen fransen@umn.edu
- Shane Nackerud snackeru@umn.edu
- Carissa Tomlinson
toml0035@umn.edu
- Kristen Mastel meye0539@umn.edu