If you build it, they will come
Establishing a data discovery curriculum across a medical campus

Young-Joo Lee
Data Informationist, Welch Medical Library

MLA '22 NEW ORLEANS MAY 3-6 2022 #MLANET22 MLANET.ORG RECONNECT, RENEW, REFLECT

JOHNS HOPKINS UNIVERSITY & MEDICINE

background

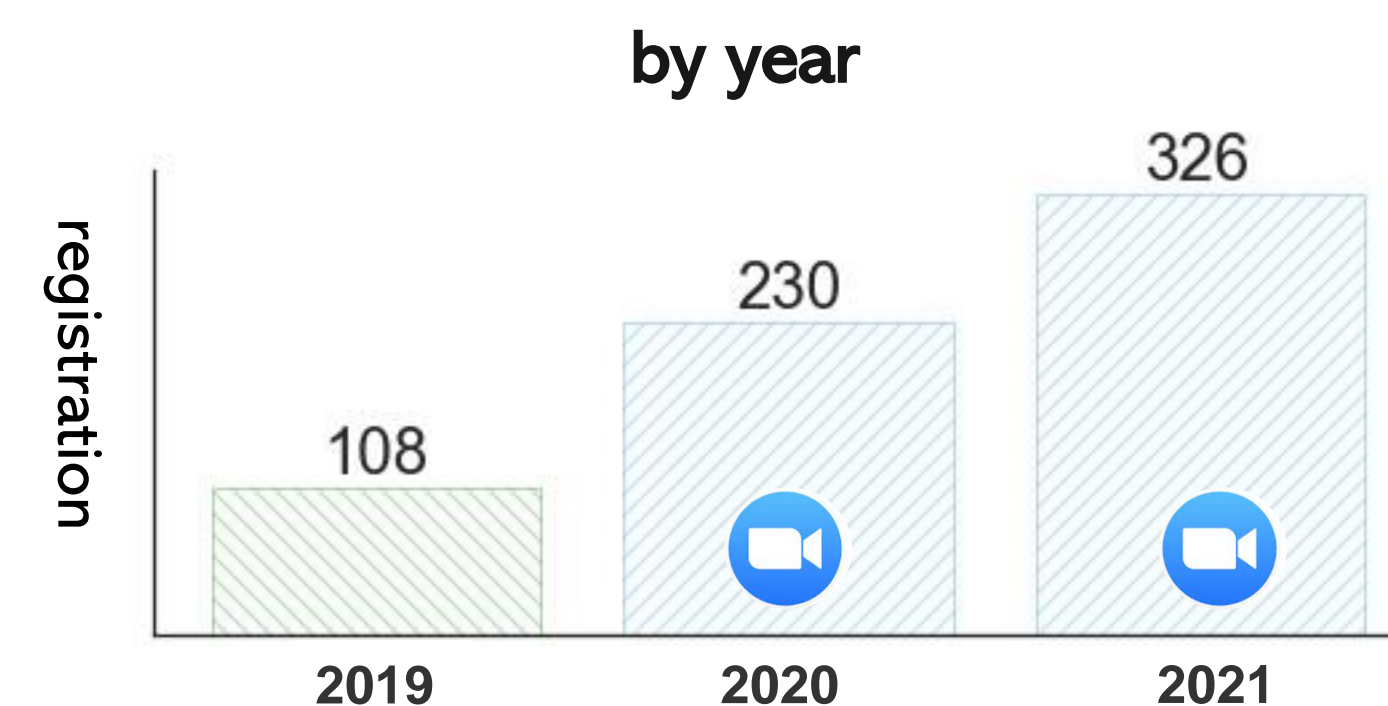Need for data discovery education
- Strong interest in secondary analysis across campus
- Wide range of data available thanks to data sharing mandates & open science movement
- Lack of searching skills for data (As shown in the study of Gregory, 2020)
- Challenges in data searching – too many data catalogs, repositories, no federated search engine, etc.
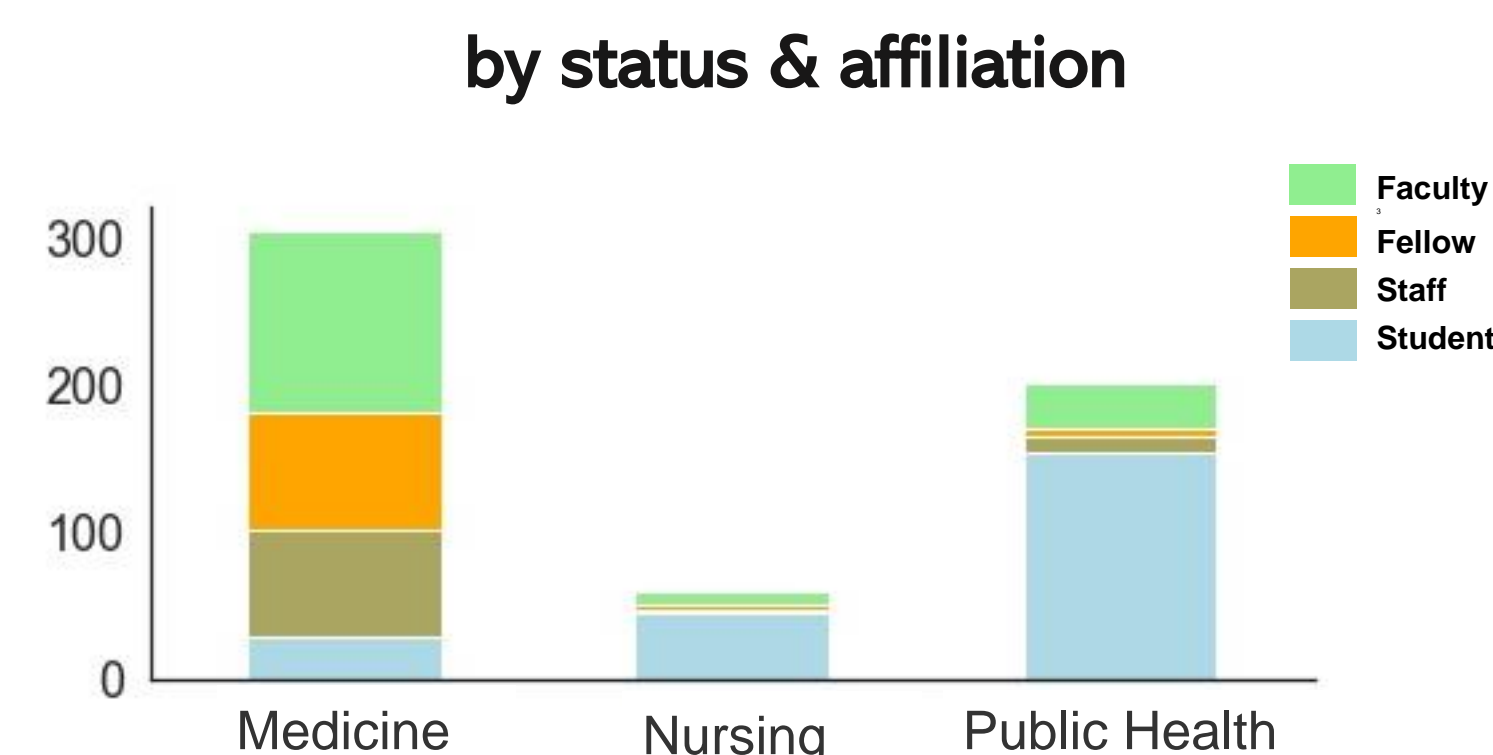
curriculum building

Data discovery curriculum launched in **2019**
- Led by the data informationist in collaboration with basic science informationist, faculty, and other guest speakers
- **Phase 1**
  Introductory data discovery class (1.5 hours) covering:
  - sponsors/producers of health data
  - major data catalogs & repositories
  - data searching strategies for PubMed & Google
- **Phase 2**
  Classes (1.5 hours) on specific topics such as:
  - mental health data
  - genomic data
  - environmental health data
- **Phase 3 (upcoming)**
  new classes being developed
  - discovery classes on different topics such as cancer data, covid data, etc.
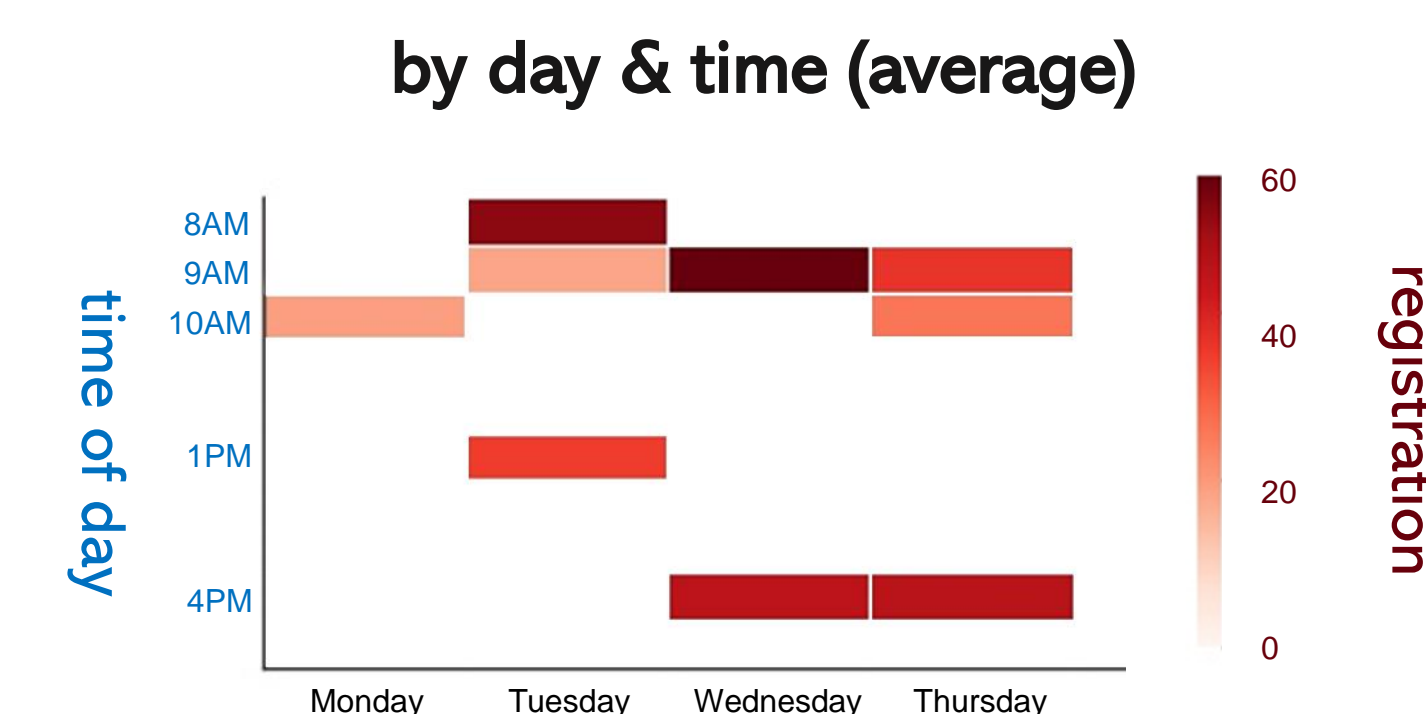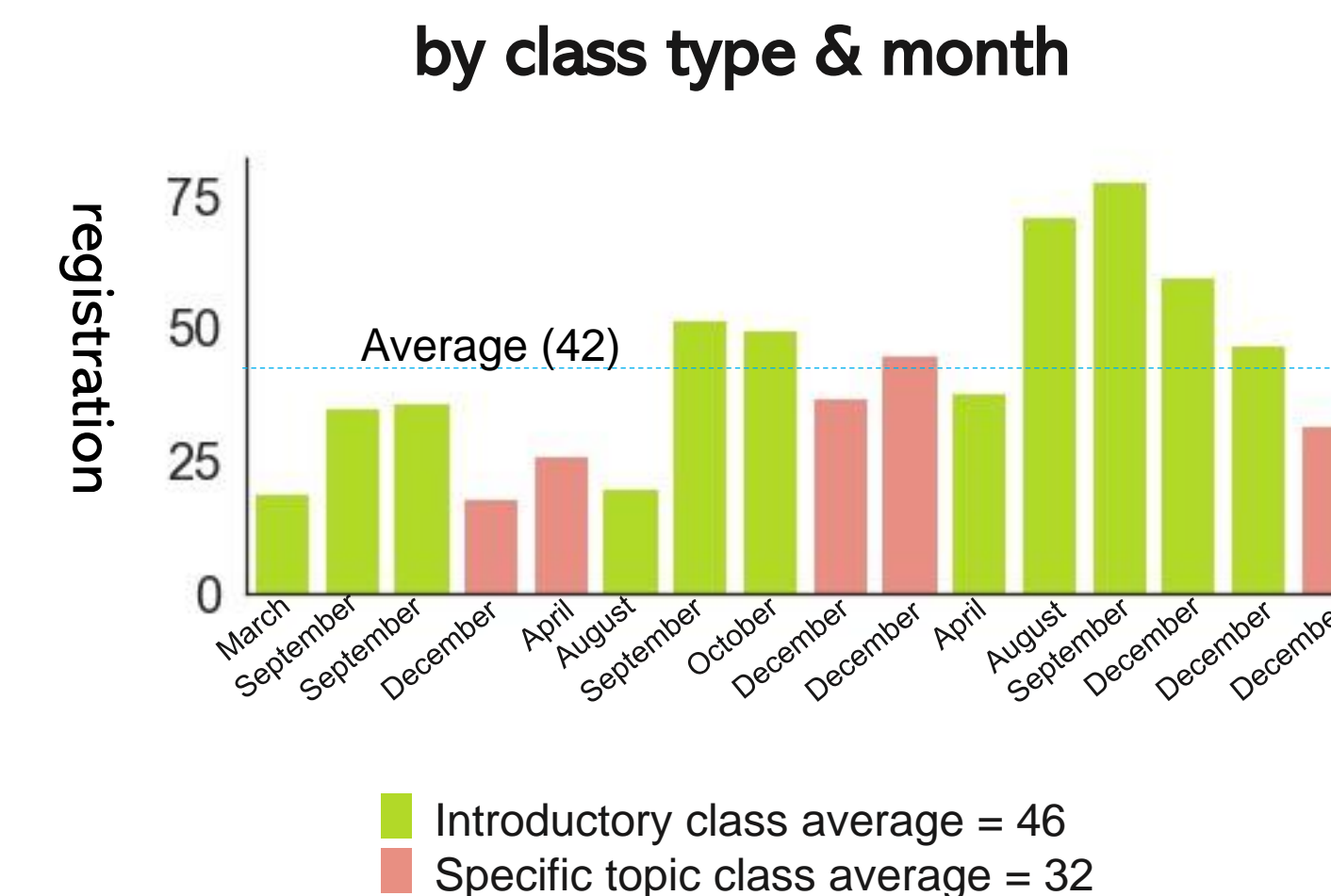 - In-depth workshop on different databases

registration data

by year

- Total registration = 664
- Total registrants = 556
  - recurring registrants (>1 registration) = 91

by class type & month

Average (42)

- Introductory class average = 46
- Specific topic class average = 32

by status & affiliation

Faculty / Fellow / Staff / Student

Medicine / Nursing / Public Health

- Top groups
  - Public Health students = 154
  - Medical faculty = 124

by day & time (average)

- Popular times
  - early mornings (8am, 9am) and late afternoon (4pm), particularly for clinicians
  - no unpopular time

conclusion

If you build it, they will come and come
- The number of registrations proves strong interest in data discovery
- Other evidence of strong interest:
  - Requests to teach data discovery classes by clinical departments
  - Integrated to multiple curricula (schools of public health and nursing)
  - Increasing data discovery consultations
- Promising future for online learning
  - Particularly for clinicians

Outreach is essential for success
- Strong correlation between outreach and registration
  - Cases of Oncology (Medicine) and Mental Health (Public Health)

references

- Gregory, K et al. (2020). Understanding data search as a socio-technical practice. *Journal of Information Science*, *46*(4), 459–475.
- Kellam, LM (2011), *Numeric data services and sources for the general reference librarian*. Chandos Publishing: Oxford.

acknowledgement

ylee191@ jhmi.edu @thatyoungjoolee