Research Workshop Attendance, Marketing, and Satisfaction

C.M. Curry (STEM Services) and B. Narr (Metrics Coordinator)

2025-11-13

Table of contents

Executive Summary

Part I

Overall analysis

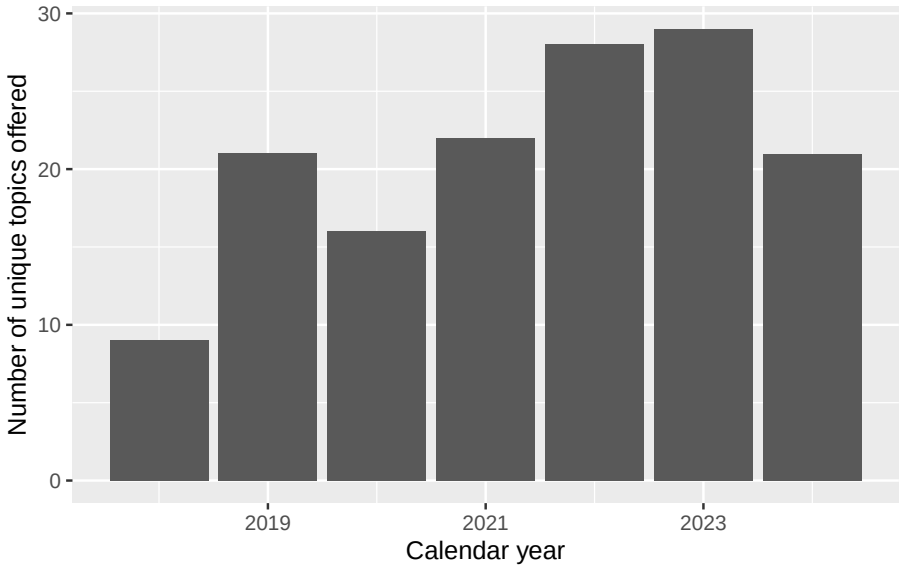
This section covers all workshops combined.

Chapter 1

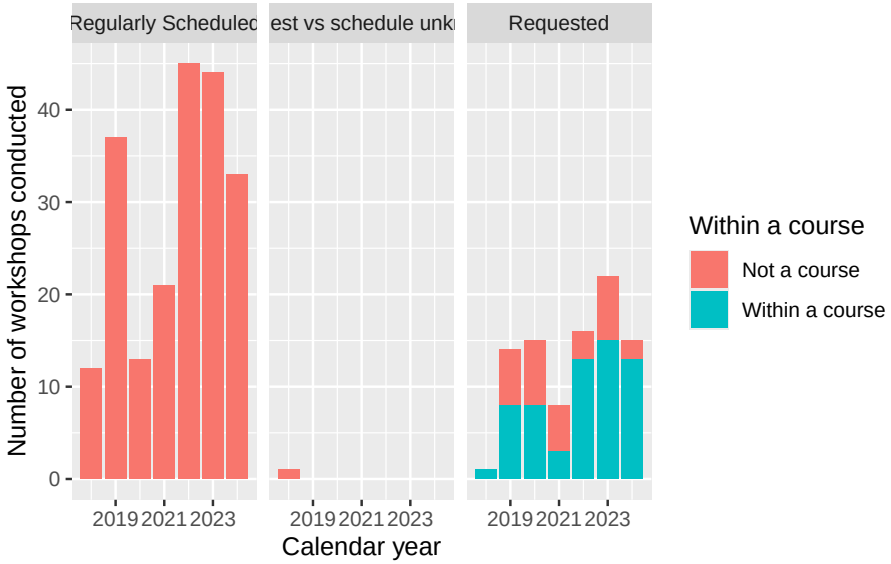
Offerings

1.1 Workshop topics offered

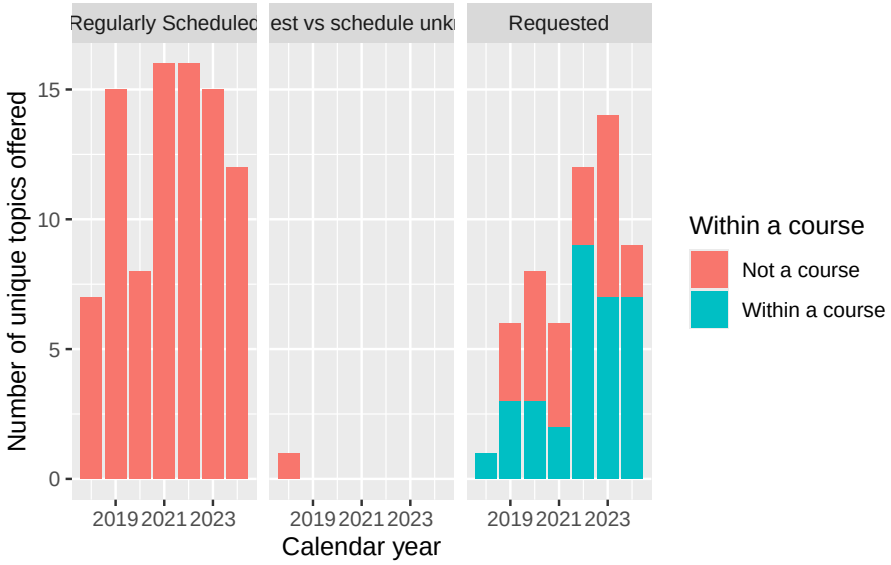
The number of distinct topics has declined since 2022, likely due to loss of digital scholarship specialist expertise.



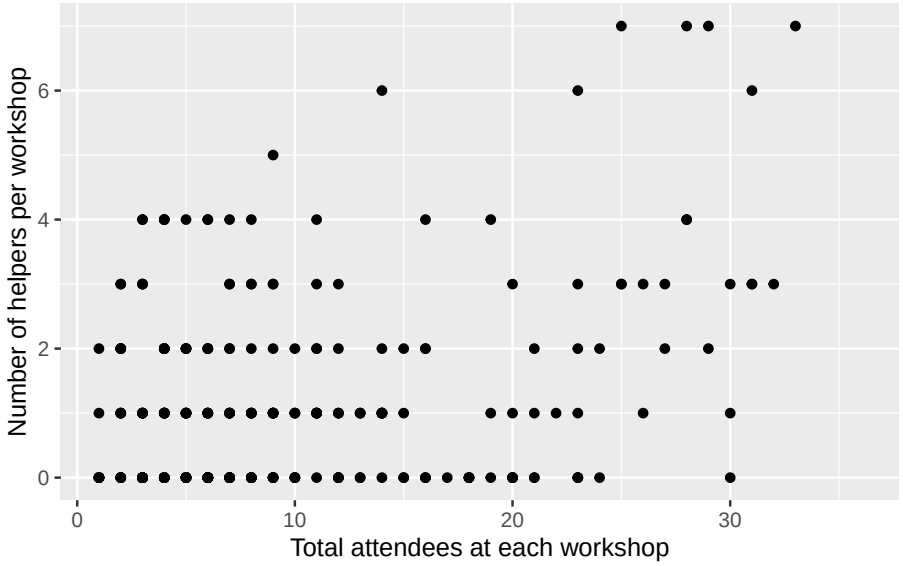
While we offer many scheduled workshops each semester (upper left panel, "Not requested/Regularly scheduled"), we continue to increase our course-based "workshops on request" visits (lower right panel) as well as a low but steady number of out-of-class requested workshops (upper right panel).

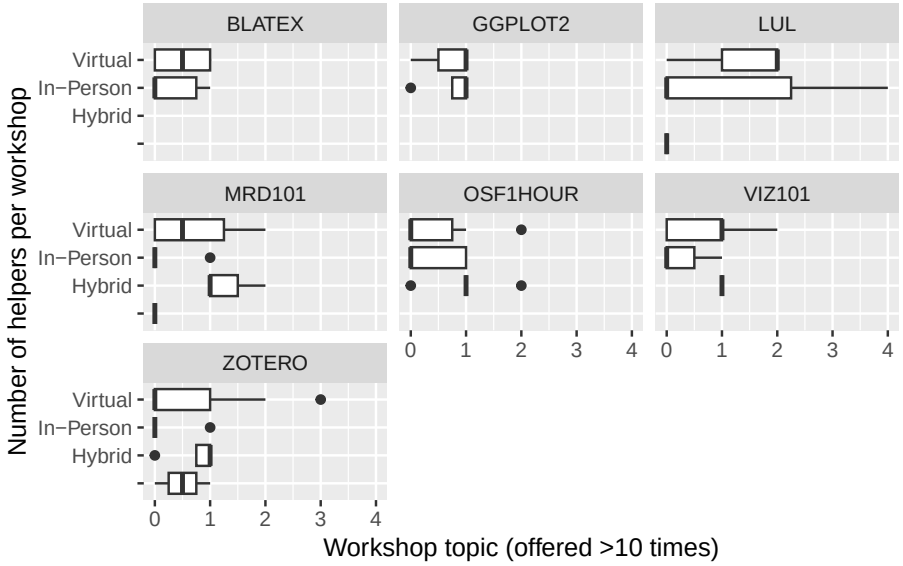


We have increased the number of topics we've brought "on request" to groups and courses.



1.2 Workshop staffing needs





Chapter 2

Marketing

2.1 Executive summary

Our overall most effective marketing methods are emails, the OU Libraries website, word-of-mouth, and outreach to instructors.

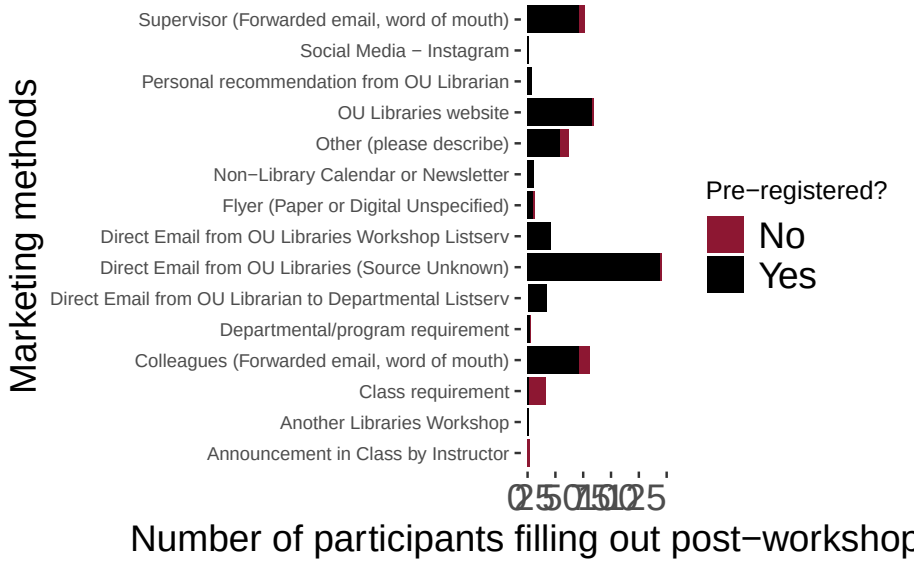
This determination is made by combining what methods bring in high attendee counts (1) and what methods bring the highest ratio of attendance to registration (2).

1. The top three marketing methods that bring in the highest counts of attendees who fill out our post-workshop surveys (hereafter, "attendees" and "attendance") are emails, word-of-mouth, and the OU Libraries website
2. The marketing methods with the highest response "intensities" (relative proportion of attendees) were class requirements (making liaison outreach to instructors critical), librarians emailing a departmental listserv directly (we have these permissions for at least two STEM departments to my knowledge), and the OU Libraries website.

2.2 Absolute effectiveness (counts of attendees) by marketing method

2.2.1 Overall counts of attendees by marketing method and registration status

These data for people who filled out a post-workshop survey (ie are confirmed to have attended a workshop) and answer the question "How did you hear about this workshop?". They are grouped by pre-registered or not pre-registered (passerby walk-ins, class workshops where registration wasn't required). We do not get 100% completion of surveys.

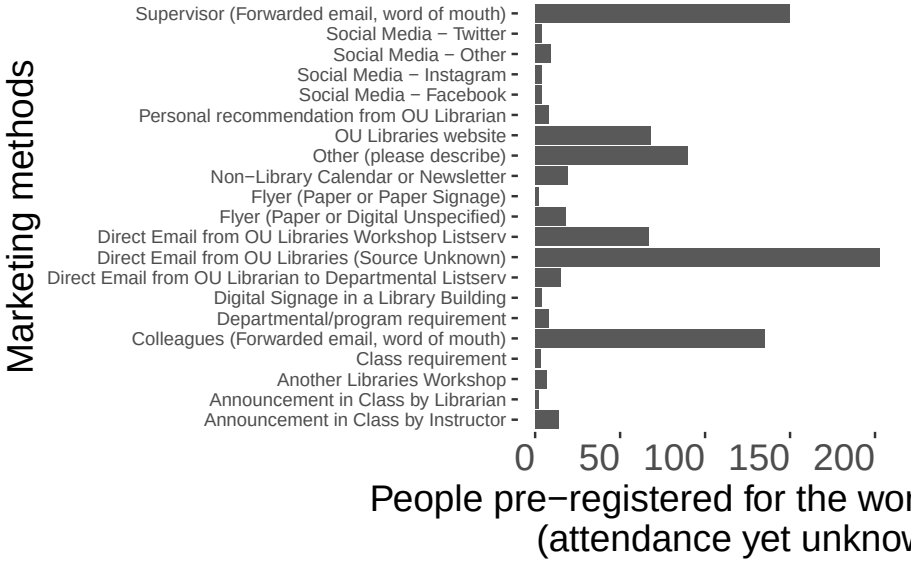


2.2.2 Marketing methods for selection "Other (please describe)" with at least two words

These words suggest that professors, librarians, and previous workshops were included in methods that reached these attendees. Later data cleaning to incorporate emails/professors/word-of-mouth into the existing categories could be useful.

2.3 Marketing sources for people who registered for workshops (attendance yet unknown)

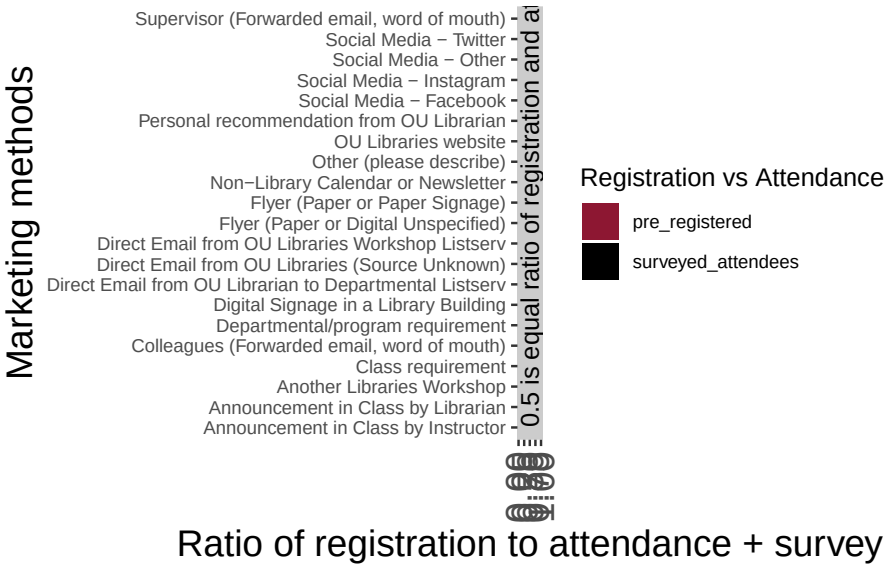
These data are from people who filled out the pre-registration form to attend a workshop. When they filled out this form, we do not know if they will go on to actually attend the workshop. Most of the people in the previous section will have registered, so there is overlap between the datasets, but the post-workshop surveys are anonymous so we cannot connect them directly. The next section shows a broad-level view of the relative effectiveness of methods for people who register vs people who follow through and attend.



2.4 Relative effectiveness in marketing methods

We do not get 100% survey responses at our workshops. However, I wanted to see to see if some marketing methods are proportionately more effective in getting pre-registrants to actually show up at the workshop. The differences are overall statistically significantly different (Chi-squared = 71.5, df =- 20, p < 0.0001).

Below is a chart that represents this visually as a ratio of attendance to pre-registrant counts. The vertical beige bar shows approximately where about the same proportion (i.e., 0.5 out of 1, or 50/100) of people who pre-registered (red) ended up actually attending and filling out a survey (black).



A completely red bar indicates that none of the people who completed the pre-registration survey later completed a post-workshop survey. It is possible some of those attendees showed up but did not complete the offered post-workshop survey. People who pre-registered and heard about the workshop via Twitter, Facebook, Other Social Media Not Specified, Announcement in Class by Librarian, and paper fliers have this result, suggesting those media may not be resulting in attendees or not resulting in attendees willing to provide feedback.

A completely black bar indicates that pre-registration did not occur but people attended anyways (you see these attendees as the red caps to the black bars in the previous section). This occurred more with people who attended a workshop given in a class ("Class Requirement"; that category suggests either we visited the class, which does not involve LibCal registration, or that the instructor asked people to attend a workshop outside of class). With online workshops requiring pre-registration, we don't necessarily expect any bar to be completely

black (all walk-ins). However, in-person workshops can accept passer-by walk-ins who didn't pre-register or people who saw the event via any other marketing method (again, see the previous section's chart) and just showed up without pre-registering. Zoom workshops could have other registrants also forward the Zoom link.

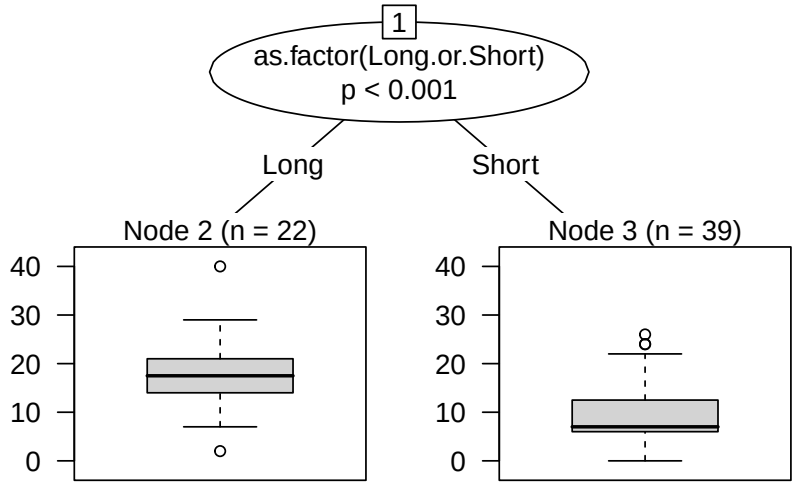
Thus, any method with some black has some attendance, and methods with larger amounts of black indicate relatively stronger responses. Thus, the highest response "intensity" are for class requirements, librarians emailing a departmental listserv directly (we have these permissions for at least two STEM departments to my knowledge), and people who find events by examining the OU Libraries website.

2.5 Marketing implementation details

2.5.1 Email detail level and timing

I used registrants instead of confirmed attendees, because that felt more relevant to the emails, and it was more consistently documented. The factors that were considered here were lead time (how much time between the email and the event) and length of email (long = one that was typed by one of you, short = automated libzoom email). Email length was the deciding factor for number of registrants!

Thus, the time between registration opening and workshop did not affect attendance. As such, CMC proposed we open all workshops at the start of the semester. This change was approved by the committee in slack in Fall 2023 (2023/10/03) to be implemented for the Spring 2024 workshops. We discussed adding a second reminder for already registered participants, but only one reminder is possible using LibCal automated emails. We will now post once advertising the full schedule, and then continue doing the 3-weeks-advertising to remind people again.


1
as.factor(Long.or.Short)
p < 0.001
Long
Short
Node 2 (n = 22)
Node 3 (n = 39)
40
30
20
10
0
40
30
20
10
0

Chapter 3

Attendance

3.1 Executive summary

Counts of workshop attendance are only influenced by whether a workshop was presented as part of course instruction. Format (virtual, in-person, hybrid) and marketing did not change attendance counts. Even topics (Chapter 10) didn't impact attendance counts.

3.2 Per-workshop attendance by format and request

3.2.1 Statistical analysis

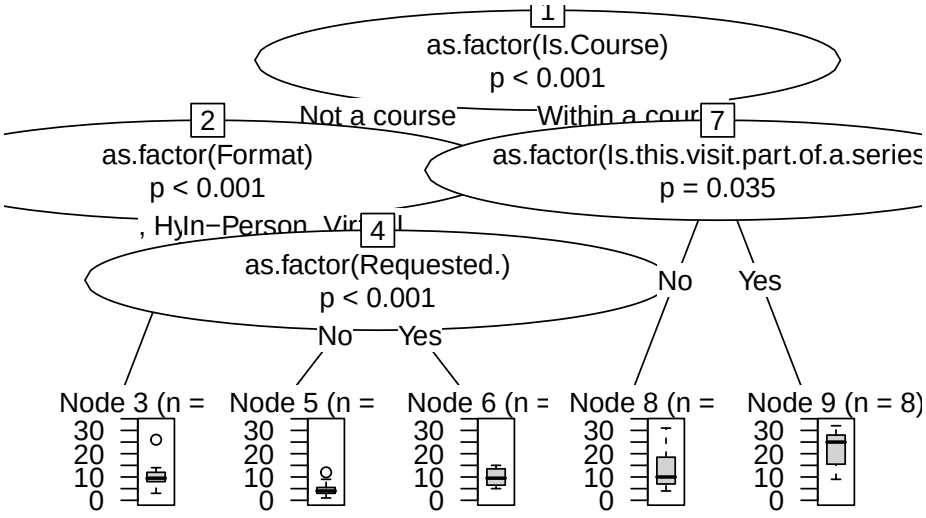
For scheduled workshops, format and year impact attendance numbers in a conditional inference tree analysis. Virtual and hybrid workshops get generally higher numbers. Within in-person workshops, the most recent two years (2023-2024) have lowest numbers attending per workshop.

- format (in person vs virtual vs hybrid)
- calendar year (starts in January)
- semester (spring, summer, or fall)
- workshop topic
- multi-day vs single day scheduling

When we examine all workshops (included on-request visits to classes and on-request workshops), we add the following variables:

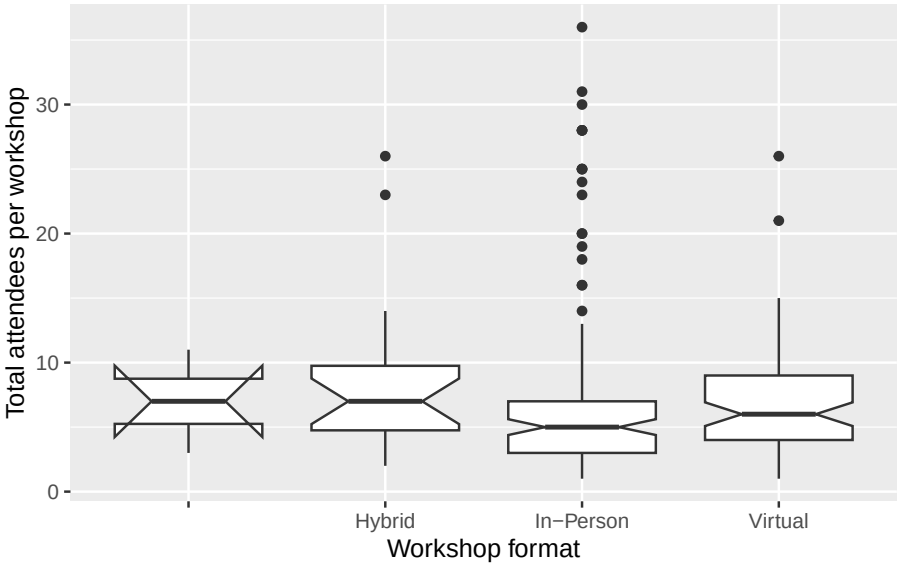
- workshop was in a course
- requested or not

Whether a workshop was in a course (where students are presumably required to attend) was the most dramatic factor in increased attendance numbers. For workshops not in a course, virtual and hybrid workshops have the highest attendance.



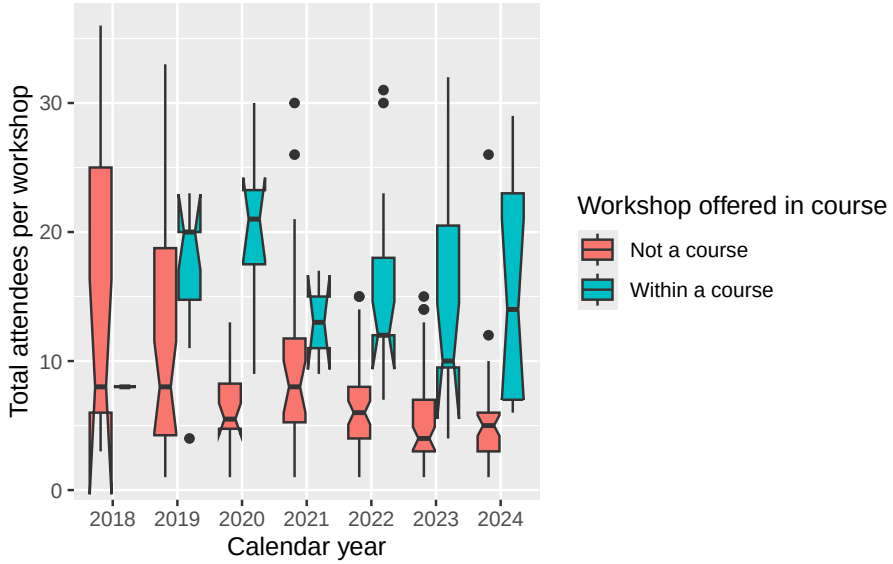
3.2.2 Format

Ignoring all other factors, the workshop format was not significantly different. You can also see this with a different visual analysis using notched box plots. The overlapping "notches" here show that there is no difference between median attendance for scheduled (not on request, not in a class) workshops by format.



3.2.3 Workshop attendance over time

The median number of people at each scheduled workshop has declined slightly per calendar year from 2021-2024. Workshop attendance at courses is higher, presumably because most classes have to reach a certain enrollment to "make".

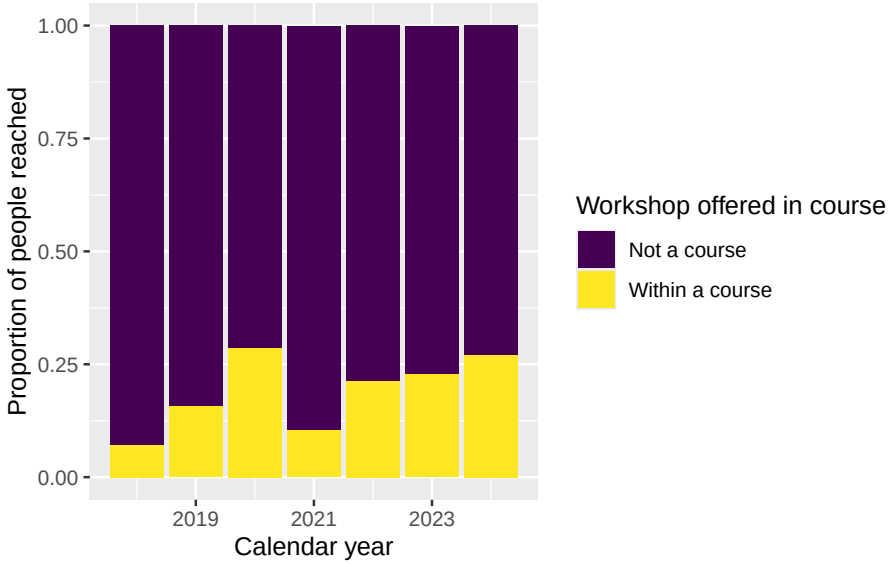


3.3 Total people reached

The total number of people reached has no strong trends since 2021.

```
# A tibble: 3 x 8
  Is.Course       `2018` `2019` `2020` `2021` `2022` `2023` `2024`
  <chr>            <int>  <int>  <int>  <int>  <int>  <int>  <int>
1 Not a course       190    496    128    257    311    293    188
2 Within a course      8    134    162     39    205    233    201
3 Total              198    630    290    296    516    526    389
```

The proportion of total people reached each calendar year in courses peaked in 2020, declined, and then has increased each calendar year since 2021.


1.00
0.75
0.50
0.25
0.00
Proportion of people reached
2019
2021
2023
Calendar year
Workshop offered in course
Not a course
Within a course

Chapter 4

Feedback

4.1 Executive summary

Feedback data show no difference in perception of workshop value by learners among formats (in person, hybrid, or virtual). Qualitative (word-cloud illustrations) feedback are generally positive.

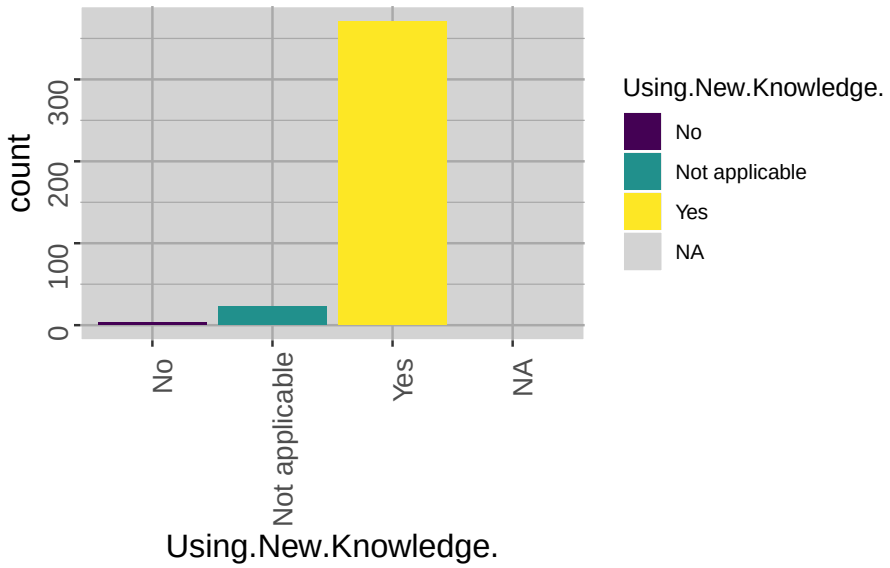
4.2 Quantitative questions

For each question, we examine statistically and graphically (if differences were statistically significant) the rankings divided by instruction format (DSI request 2023/07), request/not, and course/outreach variables. In the future, we may check rankings vs length of course in hours.

Chapter 5

TODO need to clean out NA for variables we test, as it errors out in the statistical test (graphs work and just automatically remove)

5.0.1 Do you anticipate that this new knowledge can be applied towards your work?

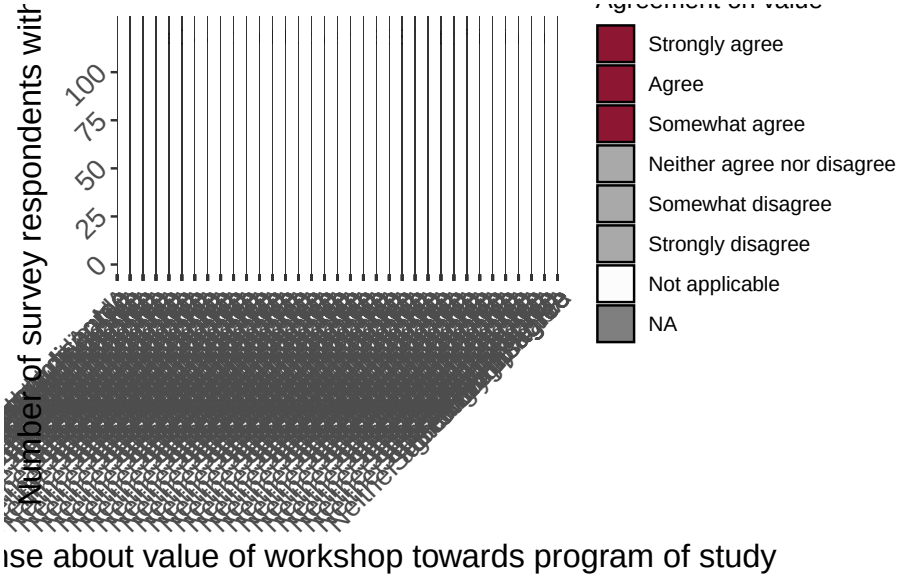


There was no significant difference in perception of learning new knowledge from the workshop among learning formats.

```
[1] FALSE
```

5.0.2 This workshop was valuable towards your program of study.

Most participants perceived the workshop as valuable towards their program of study. Interestingly, "disagree" is never chosen by any of our several hundred respondents.



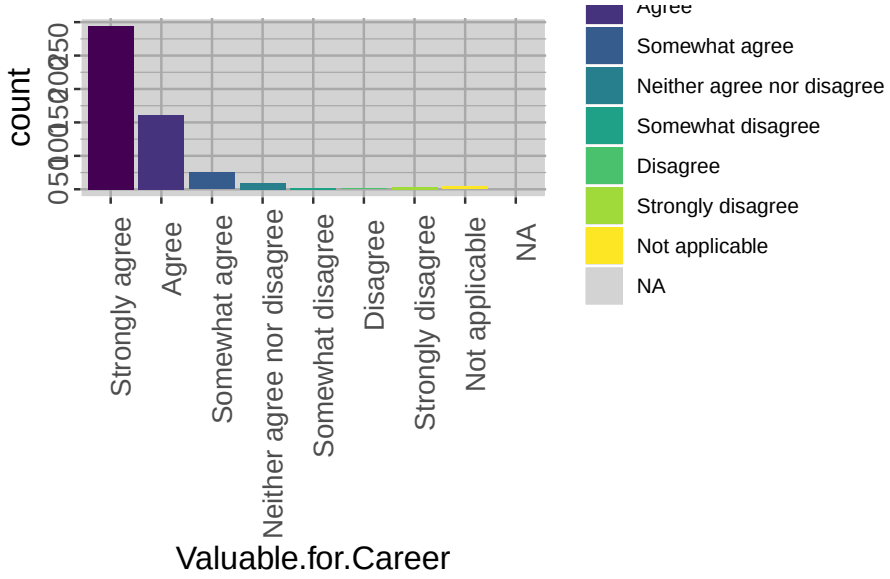
ıse about value of workshop towards program of study

There was no difference in perception of the material as "valuable to my program of study" between workshop formats (hybrid, online, or in person).

```
[1] FALSE
```

5.0.3 This workshop was valuable towards your career.

Most participants perceived the workshop as valuable towards their career.



There was no difference in perception of the material as "valuable to my career" between workshop formats (hybrid, online, or in person).

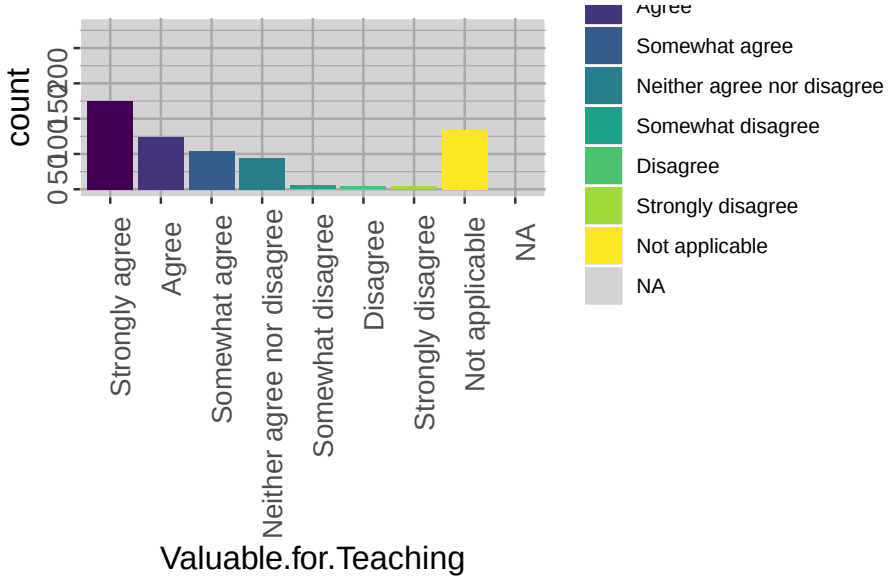
```
[1] FALSE
```

There was no difference among career levels (classifications) for perception of value towards their careers.

```
[1] FALSE
```

5.0.4 This workshop was valuable towards your teaching.

More respondents felt the workshops were not applicable towards their teaching but were otherwise positive. As our workshops are research focused, this is probably reasonable.

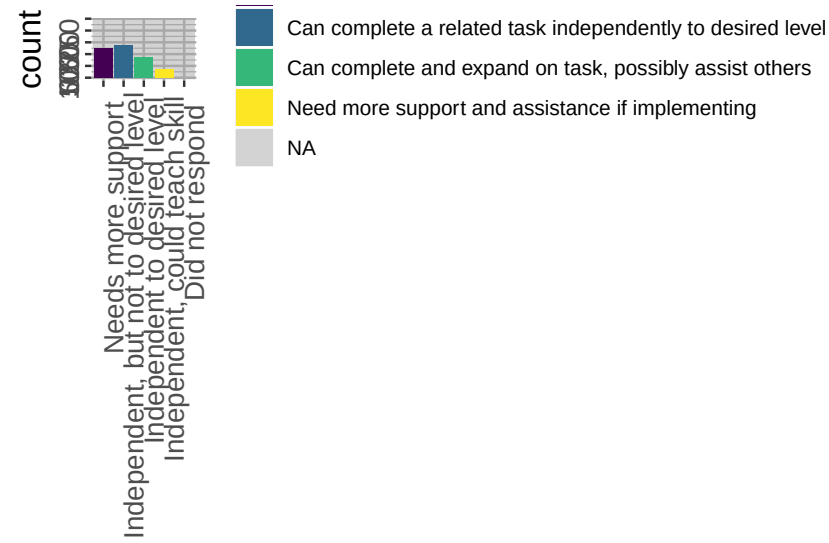


There was no difference in perception of the material as "valuable to my teaching" between workshop formats (hybrid, online, or in person).

```
[1] FALSE
```

5.0.5 How do you rate your skills after learning about this workshop topic?

We have a range of skill ratings after the workshop. This could be an area for improvement but could also be dependent on whether we have targeted the right audience with the right level of materials.



There was no difference in perception of future abilities to use skills from the workshop between formats (hybrid, online, or in person).

```
[1] FALSE
```

5.1 Qualitative feedback (Wordcloud code)

5.1.1 "What did you like about the workshop?"

5.1.2 "What about the workshop needed improvement?"

5.1.3 "What other workshop topics would interest you?"

We can examine this wordcloud to see if we need to advertise existing workshops more in addition to having ideas about new ones.

5.1.4 "Any other comments?"

Chapter 6

Name one aspect of the workshop that you liked.

Chapter 7

Name one aspect of the workshop that could be improved.

Chapter 8

What other workshop topics would interest you?

Chapter 9

Please let us know any other comments you have about the workshop.

Part II

Topics

These reports are created on request by the instructor for each workshop. Contact Claire or B. if you want one for your workshop.

Chapter 10

Topics

For aggregate workshop metrics, please read Attendance (Chapter 3), Marketing (Chapter 2) or Feedback (Chapter 4). For an overall description of how many topics and frequencies of offerings, see Offerings (Chapter 1).

10.1 Data sources

We use only scheduled workshops (i.e., not in a course, not on request outside of a course) in these analyses as it has a statistically significant impact on attendance (Chapter 3). We include only workshops with more than four sessions (two years).

Chapter 11

Concepts of Data-Driven Visualization

The materials for this workshop are available in OSF and GitHub.

For aggregate workshop metrics, please read Attendance (Chapter 3), Marketing (Chapter 2) or Feedback (**?@sec-feedback**). For an overall description of how many topics and frequencies of offerings, see Offerings (Chapter 1).

11.1 Descriptive numbers

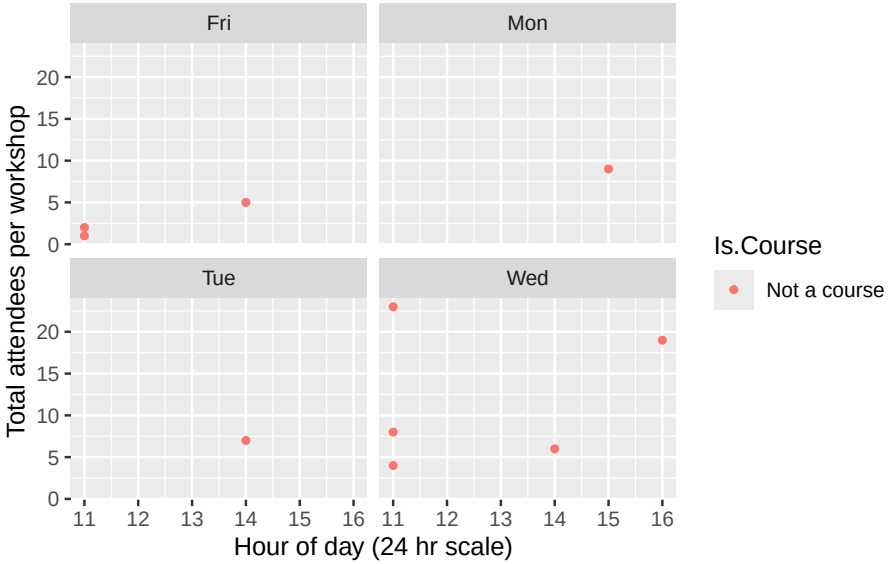
How many times offered per semester, per year. Number of different instructors.

```
     Entered.By  n
1 Curry, Claire 20
```

11.2 Attendance

11.2.1 By Timing

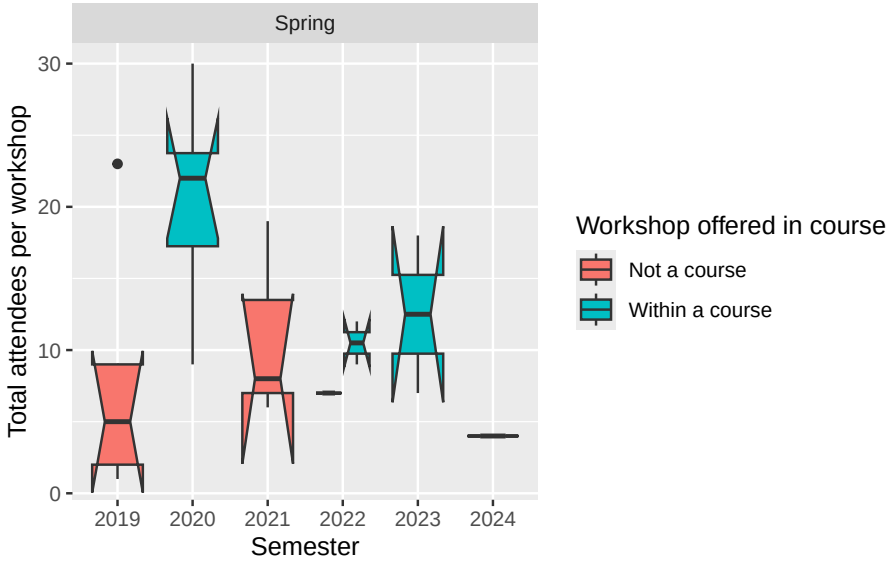
There are no obvious trends based on time of day or weekday with the current sample size. In-class workshops do not record time of day and thus are not included.



11.2.2 By semester and year

This workshop is offered only in spring semesters. More attendees come to in-class offerings.

```
Notch went outside hinges
i Do you want `notch = FALSE`?
Notch went outside hinges
i Do you want `notch = FALSE`?
Notch went outside hinges
i Do you want `notch = FALSE`?
Notch went outside hinges
i Do you want `notch = FALSE`?
Notch went outside hinges
i Do you want `notch = FALSE`?
```



11.3 Marketing

11.3.1 Over time

11.3.2 By semester

11.4 Satisfaction

11.4.1 Over time

11.4.2 By semester

Technical details

Data sources and processing

If no data update needed

Data download from OSF

Data for this report are gathered from three sources (created by processing the four datasets below in "update data"):

- pre-registration questionnaire when people register
- attendance numbers entered by workshop instructors in LibInsights
- post-workshop surveys.

Registration numbers include learner names and emails and so all three datasets are kept on a private OSF repository and must be downloaded there. Do not include these in github.

```
############
# Load data from OSF
############

# https://www.statology.org/r-check-if-file-exists/

# Authenticate
# Your OSF PAT should be in .REnviron file following these instructions:
# https://docs.ropensci.org/osfr/reference/osf_auth.html

if(file.exists("raw_data/README.md")){
  print("OSF data already downloaded.")
} else {
  print("Download it!")

library(osfr)

# ## Retrieve project
```

```
osf_node <- osf_retrieve_node("qjkvf")

## List files in project
osf_files <- osf_ls_files(osf_node)

## Download the files
osf_download(x = osf_files,
             path = NULL, #default save to local working directory
             recurse = TRUE, #download all nested files
             conflicts = "overwrite", #OSF is the canonical version
             progress = TRUE #show progress bar
 )
}
```

```
[1] "OSF data already downloaded."
```

If data update needed

Raw data downloads

This consists of four datasets.

- Registration numbers (pre-workshop interest) from LibCal
  - Downloaded per calendar then concatenated (?)
  - Registered <- read.csv("raw_data/processed_Registered.csv")
- Attendance from LibInsights for Course-integrated instruction
  - How to get this data via download from "Course Integrated Library Instruction"
    - Method 1
      - Filter by Custom Date Range (July 1, 2018, to present)
      - Add Additional Report Filter "Name of Workshop" "Is Not" "Null", then click "Add this filter".
      - Click "generate report".
    - Method 2
      - Use saved filter "Workshop Filters (CMC)"
    - Method 2
      - Use saved filter "Workshops in Courses (CMC)"
    - After either method, click "Reports" tab and then green button "Export Data to CSV"
    - Save the file in OSF as /raw_data/attendance_courses.csv
- Attendance from LibInsights for Outreach
  - How to get this data via download from "Outreach and Programming Form Analysis"
    - Method 1
      - Filter by Custom Date Range (July 1, 2018, to present)
      - Add Additional Report Filter "Type of Outreach" "Is" "Workshop", then click "Add this filter".

        - Click “generate report”.
    - Method 2
        - Use saved filter “Workshop Filters (CMC)”
  - After either method, click “Reports” tab and then green button “Export Data to CSV”
    - Save the file in OSF as /raw_data/attendance_outreach.csv
- Post-workshop surveys (feedback and satisfaction) from LibWizard
  - postworkshopsurveys_named <- read.csv(“raw_data/processed_postworkshopsurveys_named.csv”)

Raw data processing to write three working files

Titles have varied over time with marketing experiments, typos, and accidental changes. This code assigns a standardized code between each topic.

```
library(dplyr)
```

```
Warning: package 'dplyr' was built under R version 4.4.1
```

```

Attaching package: 'dplyr'
```

```
The following objects are masked from 'package:stats':

    filter, lag
```

```
The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
topics_complete_variations <- read.csv("./workshop_metadata.csv")

# remove any accidental duplicates of name variations (otherwise will create a many-to-many join
topics_complete_variations_cleaned <-topics_complete_variations %>%
  dplyr::distinct(WorkshopNameVariations, WorkshopCode)

# Must attach standard titles to all three datasets (registrations, attendance, and post-workshop
```

The four raw datasets are combined and cleaned into three working datasets.

- Registration numbers
- Attendance
  - Course and Outreach datasets are each given a column for Is.Course (Y = Course, N = Outreach) before being combined.

```
library(lubridate)
```

```
Warning: package 'lubridate' was built under R version 4.4.3
```

```

Attaching package: 'lubridate'
```

```
The following objects are masked from 'package:base':

    date, intersect, setdiff, union
```

```
############
# Attendance data is created by combining Course Instruction attendance data from LibI
############

# Outreach
attendance_outreach <- read.csv("raw_data/attendance_outreach.csv",
                                header=TRUE,
                                stringsAsFactors=FALSE)
attendance_outreach$Is.Course <- "Not a course"
attendance_outreach$Requested.[is.na(attendance_outreach$Requested.)] <- "Unknown"
attendance_outreach$Requested.[attendance_outreach$Requested.==""] <- "Unknown"

# Courses
attendance_courses <- read.csv("raw_data/attendance_courses.csv",
                                header=TRUE,
                                stringsAsFactors=FALSE)
attendance_courses$Is.Course <- "Within a course"
attendance_courses$Requested. <- "Yes"
attendance_courses$Total.Attendees <- attendance_courses$Total.Attendance


############
# We clean dates out by weekday, year, month, day, and times for later analysis and da
############

# Outreach

# Month, day, year, hour, minute, seconds
attendance_outreach$Event.Date.and.Time_2 <- ymd_hms(attendance_outreach$Event.Date.an
                                                    truncated = 3)


# separate year, month, date, times
attendance_outreach$year <- year(attendance_outreach$Event.Date.and.Time_2)
attendance_outreach$month <- month(attendance_outreach$Event.Date.and.Time_2)
attendance_outreach$day <- day(attendance_outreach$Event.Date.and.Time_2)
attendance_outreach$wday <- wday(attendance_outreach$Event.Date.and.Time_2,
                                 week_start = 1, #1 = Monday,
                                 label = TRUE    # days of weeks as characters
                                 )
attendance_outreach$hour <-as.numeric(
  format(attendance_outreach$Event.Date.and.Time_2, "%H"))
```

```
attendance_outreach$Date <-
  format.Date(attendance_outreach$Event.Date.and.Time_2, "%m/%d/%Y")

# convert 00 hour to NA
# attendance_outreach[attendance_outreach$hour==00, "hour"] <- NA

attendance_outreach$hour <- as.numeric(attendance_outreach$hour)

# Semester defined as spring (month 1-5), summer (month 6-7), fall (month 8-12)

attendance_outreach$Semester[attendance_outreach$month>0&
                             attendance_outreach$month<6] <- "Spring"
attendance_outreach$Semester[attendance_outreach$month>5&
                             attendance_outreach$month<8] <- "Spring"
attendance_outreach$Semester[attendance_outreach$month>7&
                             attendance_outreach$month<=12] <- "Spring"


# Courses
# Month, day, year, hour, minute, seconds
attendance_courses$Event.Date.and.Time_2 <-
  ymd_hms(attendance_courses$Session.Date,
          truncated = 3)


# separate year, month, date, times
attendance_courses$year <- year(attendance_courses$Event.Date.and.Time_2)
attendance_courses$month <- month(attendance_courses$Event.Date.and.Time_2)
attendance_courses$day <- day(attendance_courses$Event.Date.and.Time_2)
attendance_courses$wday <- wday(attendance_courses$Event.Date.and.Time_2,
                                    week_start = 1, #1 = Monday,
                                    label = TRUE    # days of weeks as characters
                                    )
attendance_courses$hour <-as.numeric(
  format(attendance_courses$Event.Date.and.Time_2, "%H"))

attendance_courses$Date <-
  format.Date(attendance_courses$Event.Date.and.Time_2, "%m/%d/%Y")

# convert 00 hour to NA
# attendance_courses[attendance_courses$hour==00, "hour"] <- NA

attendance_courses$hour <- as.numeric(attendance_courses$hour)

# Semester defined as spring (month 1-5), summer (month 6-7), fall (month 8-12)
```

```
attendance_courses$Semester[attendance_courses$month>0&
                            attendance_courses$month<6] <- "Spring"
attendance_courses$Semester[attendance_courses$month>5&
                            attendance_courses$month<8] <- "Spring"
attendance_courses$Semester[attendance_courses$month>7&
                            attendance_courses$month<=12] <- "Spring"



############
# Join outreach and course datasets
############

# using dplyr's bind_rows instead of cbind ensures all columns are kept, including thos

attendance_named <- dplyr::bind_rows(attendance_outreach,
                                     attendance_courses)


############

## attendance
attendance_named <- left_join(attendance_named,
                              topics_complete_variations_cleaned,
                              by =  c("Name.of.Workshop" = "WorkshopNameVariations")
  dplyr::filter(!is.na(id))




############
# Write to file for use in qmds.
############
# Attendance
write.csv(attendance_named,
          file = "raw_data/processed_attendance_named.csv",
          row.names = FALSE)
```

- Post-workshop surveys
- Deal with all these data (attendance data are in ok now)

```
############
# Registration data has to be imported per-room from LibCal.
############
```

```
# Event registration
Room339 <- read.csv("./raw_data/preworkshop/Room339_lc_events_20230105043223.csv")
RoomLL123 <- read.csv("./raw_data/preworkshop/LL123_lc_events_20230105043507.csv")
RoomGeneral <- read.csv("./raw_data/preworkshop/General_lc_events_20230105043258.csv")
RoomLearningLabClassroom <- read.csv("./raw_data/preworkshop/LearningLabClassroom_lc_events_20230

# Merge
Registered <- rbind(Room339,
                    RoomLL123,
                    RoomGeneral,
                    RoomLearningLabClassroom
)




###########
# Post-workshop survey data are from a subset of learners who attended and also filled out a surv
###########

# Post-workshop surveys from libwizard only
## Used for both feedback AND marketing
FeedbackActuallyAttended <- read.csv("./raw_data/postworkshop/report.csv",
                                     na.strings=c("","NA"))


library(tidyr)
```

```
Warning: package 'tidyr' was built under R version 4.4.1
```

```
library(stringr)
```

```
Warning: package 'stringr' was built under R version 4.4.1
```

```
FeedbackActuallyAttended2 <- FeedbackActuallyAttended %>%
  separate_longer_delim(This.workshop.was...,
                        delim = ", ...") %>%
  separate_wider_delim(cols = This.workshop.was...,
                       delim = ". - ",
                       names = c("This_workshop_was",
                                 "Ranking"),
                       too_few = "align_start")

FeedbackActuallyAttended2$This_workshop_was <-
  str_remove(FeedbackActuallyAttended2$This_workshop_was,
             pattern = fixed("..."))
```

```
postworkshopsurveys_wide <- FeedbackActuallyAttended2 %>%
  pivot_wider(names_from = This_workshop_was,
              values_from = Ranking) %>%
  dplyr::select(-`NA`)



#rows <- as.numeric(count(postworkshopsurvey))

#postworkshopsurvey <- na.omit(postworkshopsurvey)




###########
# Changes in categories recorded over time have resulted in the need for some automated
###########

# Creating and splitting current levels for marketing
marketing_current_categories <- c("Announcement in Class by Instructor",
                                  "Announcement in Class by Librarian",
                                  "Non-Library Calendar or Newsletter",
                                  "Another Libraries Workshop",
                                  "Walk-in",
                                  "Digital Signage in Non-Library Building",
                                  "Digital Signage in a Library Building",
                                  "Social Media - Other",
                                  "Social Media - Instagram",
                                  "Social Media - Facebook",
                                  "Social Media - Twitter",
                                  "Personal recommendation from OU Librarian",
                                  "Class requirement",
                                  "Departmental/program requirement",
                                  "OU Libraries website",
                                  "Direct Email from OU Librarian to Departmental Lists",
                                  "Direct Email from OU Libraries Workshop Listserv",
                                  "Direct Email from OU Libraries (Source Unknown)",
                                  "Flyer (Paper or Digital Unspecified)",
                                  "Flyer (Paper or Paper Signage)",
                                  "Supervisor (Forwarded email, word of mouth)",
                                  "Colleagues (Forwarded email, word of mouth)")
```

```
## Convert old categories to equivalent new
Registered$How.did.you.hear.about.this.workshop.[Registered$How.did.you.hear.about.this.workshop.
Registered$How.did.you.hear.about.this.workshop.[Registered$How.did.you.hear.about.this.workshop.
Registered$How.did.you.hear.about.this.workshop.[Registered$How.did.you.hear.about.this.workshop.
Registered$How.did.you.hear.about.this.workshop.[Registered$How.did.you.hear.about.this.workshop.
Registered$How.did.you.hear.about.this.workshop.[Registered$How.did.you.hear.about.this.workshop.
Registered$How.did.you.hear.about.this.workshop.[Registered$How.did.you.hear.about.this.workshop.

Registered$ReducedCategories <- Registered$How.did.you.hear.about.this.workshop.
Registered$ReducedCategories[!(Registered$How.did.you.hear.about.this.workshop. %in% marketing_cu
Registered$OtherDescriptions[!(Registered$How.did.you.hear.about.this.workshop. %in% marketing_cu

## Convert old categories to equivalent new in feedback/post-workshop surveyd ata too.
postworkshopsurveys_wide$How.did.you.hear.about.this.workshop.[postworkshopsurveys_wide$How.did.y
postworkshopsurveys_wide$How.did.you.hear.about.this.workshop.[postworkshopsurveys_wide$How.did.y
postworkshopsurveys_wide$How.did.you.hear.about.this.workshop.[postworkshopsurveys_wide$How.did.y
postworkshopsurveys_wide$How.did.you.hear.about.this.workshop.[postworkshopsurveys_wide$How.did.y

postworkshopsurveys_wide$ReducedCategories <- postworkshopsurveys_wide$How.did.you.hear.about.thi
postworkshopsurveys_wide$ReducedCategories[!(postworkshopsurveys_wide$How.did.you.hear.about.this
postworkshopsurveys_wide$OtherDescriptions[!(postworkshopsurveys_wide$How.did.you.hear.about.this
```

```
Warning: Unknown or uninitialised column: `OtherDescriptions`.
```

```
## registrations
registrations_named <- left_join(Registered,
                                 topics_complete_variations_cleaned,
                                 by =  c("Title" = "WorkshopNameVariations"))


## postworkshop surveys
postworkshopsurveys_named <- left_join(postworkshopsurveys_wide,
                                       topics_complete_variations_cleaned,
                                       by =  c("What.workshop.did.you.attend." = "WorkshopNameVar









############
# Month, day, year, hour, minute, seconds
############
postworkshopsurveys_named$Date_2 <- ymd_hms(postworkshopsurveys_named$What.date.was.the.workshop.
                                                          truncated = 3)
```

```
# separate year, month, date, times
postworkshopsurveys_named$year <- year(postworkshopsurveys_named$Date_2)
postworkshopsurveys_named$month <- month(postworkshopsurveys_named$Date_2)
postworkshopsurveys_named$day <- day(postworkshopsurveys_named$Date_2)
postworkshopsurveys_named$wday <- wday(postworkshopsurveys_named$Date_2,
                                       week_start = 1, #1 = Monday,
                                       label = TRUE    # days of weeks as characters
                                       )

# Overwrite original column with formatted date
postworkshopsurveys_named$Date <-
  format.Date(postworkshopsurveys_named$Date_2, "%m/%d/%Y")


############
# Ordered scales require using R levels to order categories for later graphing or analy
############

#This will need to be repeated for all three "valuable for" questions, so the levels ar

# First, rename difficult columns
# rename(df, newname = oldname)
postworkshopsurveys_named <- postworkshopsurveys_named %>%
  dplyr::rename(Valuable.for.Program = `valuable towards your program of study`,
                Valuable.for.Teaching  = `valuable towards your teaching`,
                Valuable.for.Career = `valuable towards your career`,
                Using.New.Knowledge. = `Do.you.anticipate.that.this.new.knowledge.can.h

## Next, order the levels for all three.
postworkshopsurveys_named$Valuable.for.Program <- factor(x = postworkshopsurveys_named$
                                                         levels = c("No answer",
                                                                    "Not applicable",
                                                                    "Strongly agree",
                                                                    "Agree",
                                                                    "Somewhat agree",
                                                                    "Neither agree nor disagre
                                                                    "Somewhat disagree",
                                                                    "Disagree",
                                                                    "Strongly disagree"

                                                         ))


############
# Finally, the tidied feedback and attendance data are combined to compare feedback wit
```

```
###########

joined_post <- full_join(postworkshopsurveys_named,
                         attendance_named,
                         by = c("Date", "WorkshopCode"))
```

```
Warning in full_join(postworkshopsurveys_named, attendance_named, by = c("Date", : Detected an un
i Row 91 of `x` matches multiple rows in `y`.
i Row 91 of `y` matches multiple rows in `x`.
i If a many-to-many relationship is expected, set `relationship =
  "many-to-many"` to silence this warning.
```

```
# It is necessary to join by both date and workshop code as occasionally 2 or more people will te
# needs to be a full join because there are some courses not entered and vice versa.

# Filtering by NA is necessary because not everyone has entered their workshops into LibInsights
joined_attendance_post_clean <- joined_post %>%
  dplyr::filter(!is.na(WorkshopCode))

# Later data analysis will remove observations with no feedback automatically.


###########
# Write to files for use in qmds.
###########


# Satisfaction
write.csv(joined_attendance_post_clean,
          file = "raw_data/processed_joined_attendance_post_clean.csv",
          row.names = FALSE)

# Marketing

#not sure why this one is different
write.csv(postworkshopsurveys_named,
          file = "raw_data/processed_postworkshopsurveys_named.csv",
          row.names = FALSE)

write.csv(Registered,
          file = "raw_data/processed_Registered.csv",
          row.names = FALSE)
```

Website info

This document was created are using the **quarto** package [**?**] to format this book. Each page can be updated by editing the .qmd file and pressing "render". Use Build > Render Book to make whole website again.