Analysis of HSF-Training event data

Jim Pivarski[1], Rocky Bala Garg[2], Richa Sharma[3], Alexander Moreno[4]

Princeton U.[1], Stanford U.[2], U. Puerto Rico, Mayaguez[3], U. Antonio Nariño[4]

October 19, 2024

Why are we analyzing student data?

To understand our students! And make better decisions about what events to plan, how, and when.

Sources of data

- Indico registrations: 24 events since 2019

Sources of data

- Indico registrations: 24 events since 2019
- Pre- and post-event surveys: students are required to fill these out (Zoom password is at the end of the pre-event survey), 17 events since 2021

Sources of data

- Indico registrations: 24 events since 2019
- Pre- and post-event surveys: students are required to fill these out (Zoom password is at the end of the pre-event survey), 17 events since 2021

  Surveys have been analyzed before, but not across all events.

Sources of data

- Indico registrations: 24 events since 2019
- Pre- and post-event surveys: students are required to fill these out (Zoom password is at the end of the pre-event survey), 17 events since 2021

  Surveys have been analyzed before, but not across all events.
- Website traffic: anonymized, but can identify unique visitors *each day*, since May 2023

Sources of data

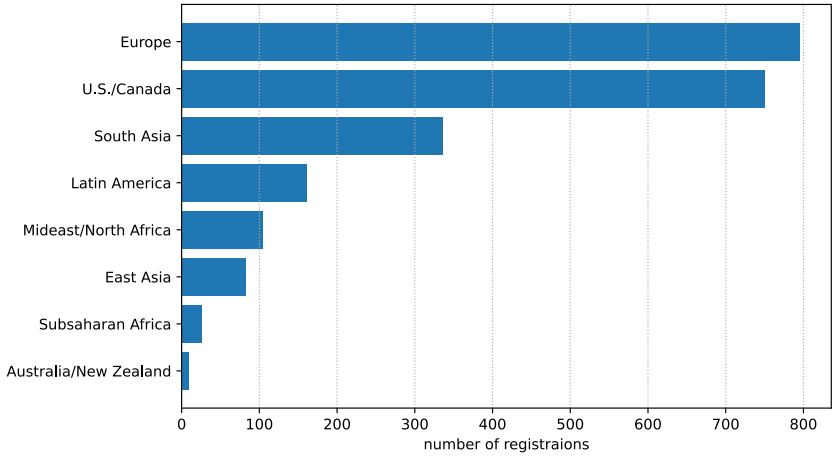
- Indico registrations: 24 events since 2019
- Pre- and post-event surveys: students are required to fill these out (Zoom password is at the end of the pre-event survey), 17 events since 2021

  Surveys have been analyzed before, but not across all events.
- Website traffic: anonymized, but can identify unique visitors *each day*, since May 2023
- Zoom and Slack activity for the February 2024 "Analysis Pipelines" event (at least)

(Indico) Where do the registrants come from?

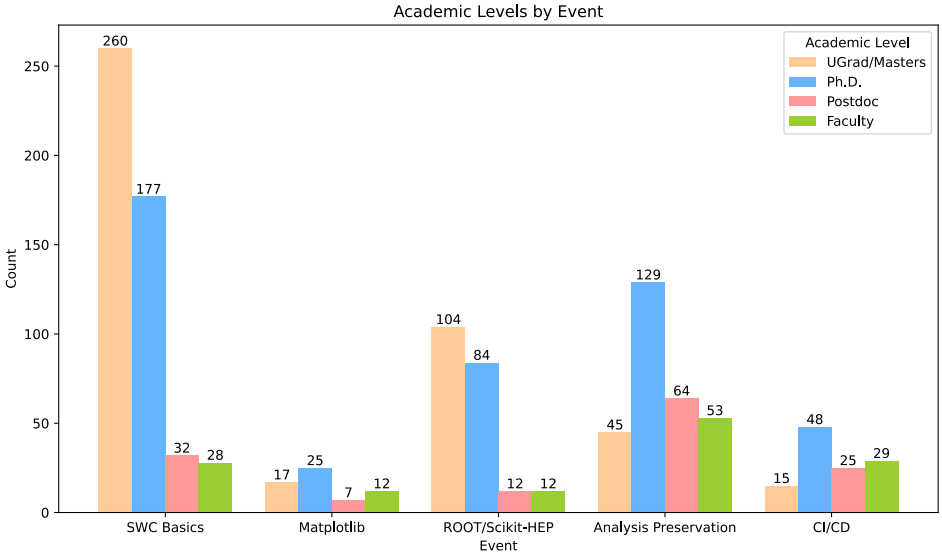
(Note: only a fraction of the students who register actually attend.)

(Indico) Where do the registrants come from?

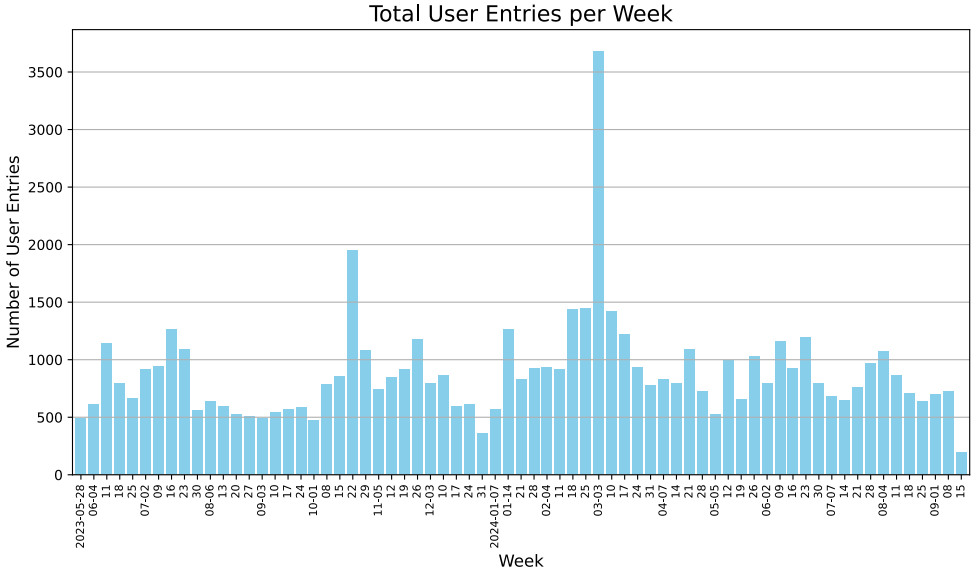


(Note: only a fraction of the students who register actually attend.)

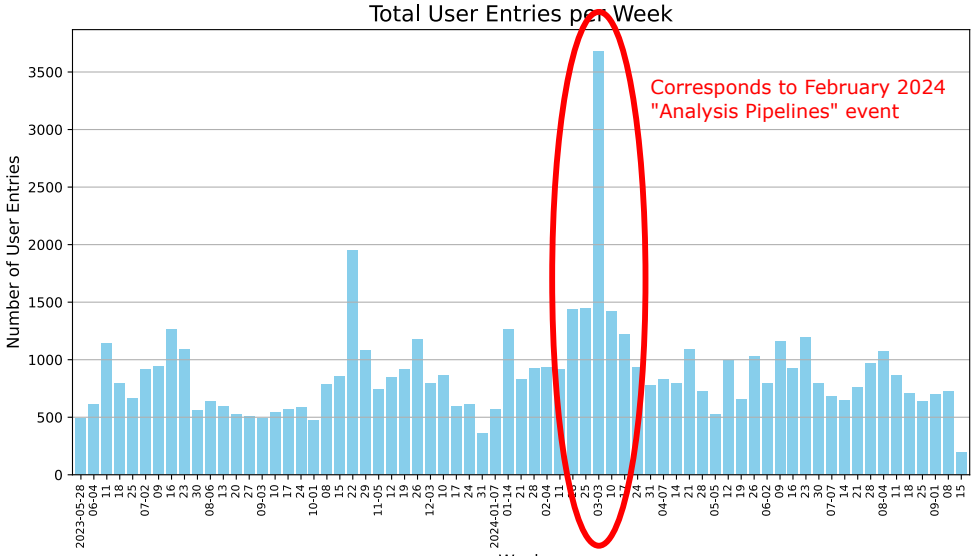
(Surveys) What kinds of students are these?



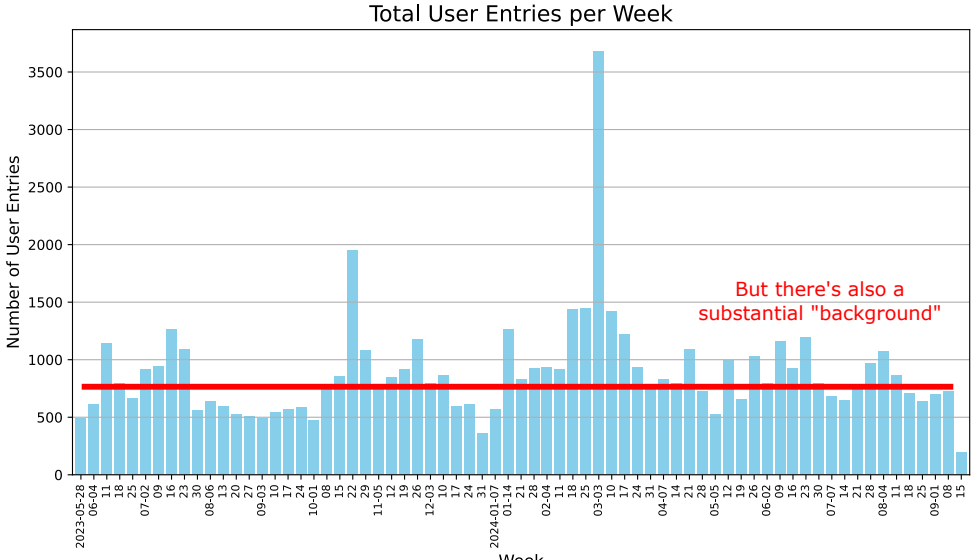
(Website) Who visits the website?



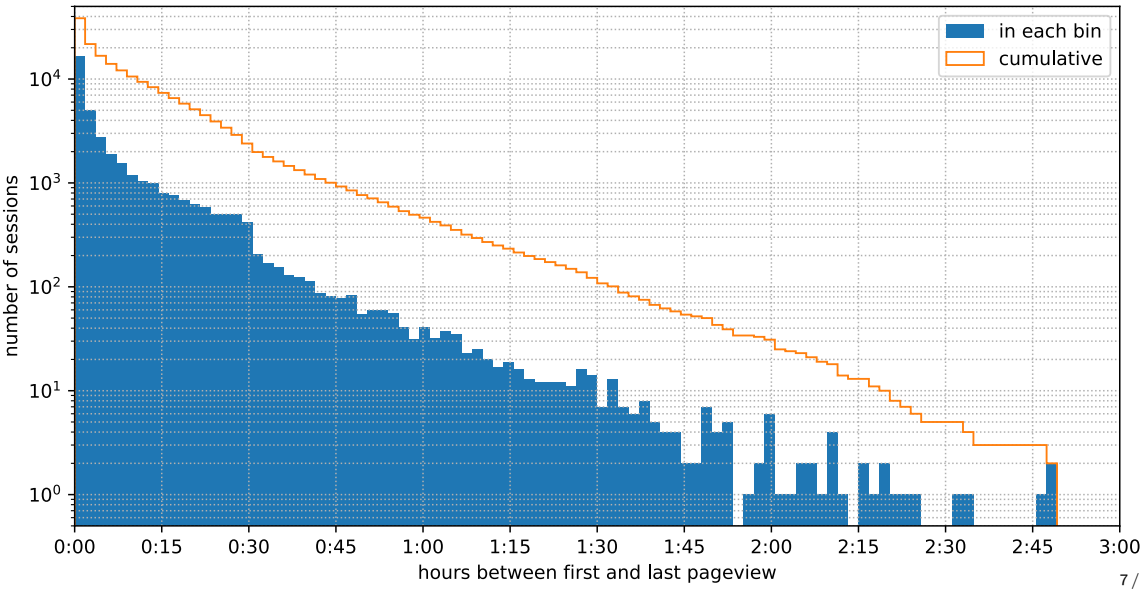
(Website) Who visits the website?



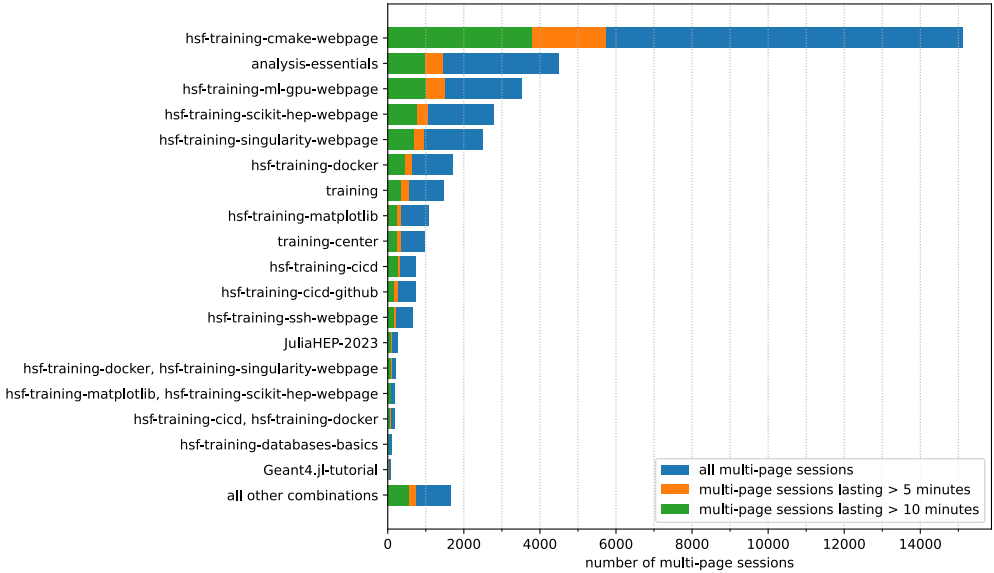
(Website) Who visits the website?



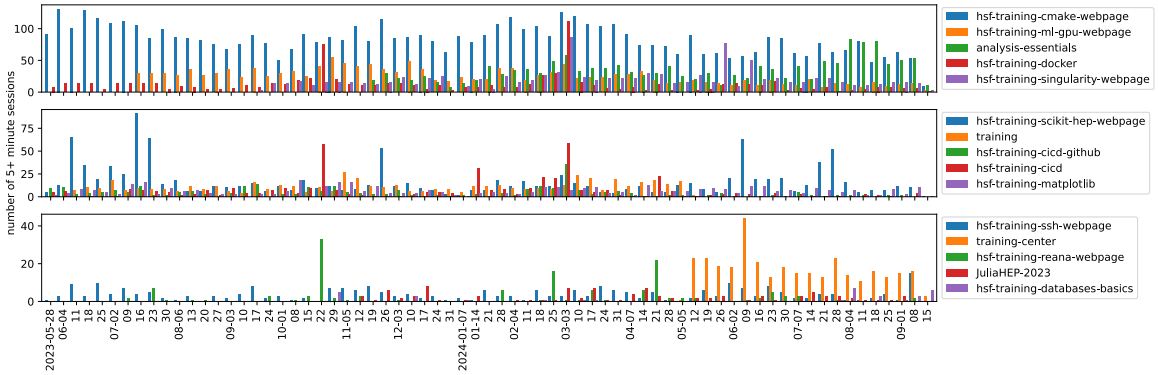
(Website) Are they studying the material or brief glance?



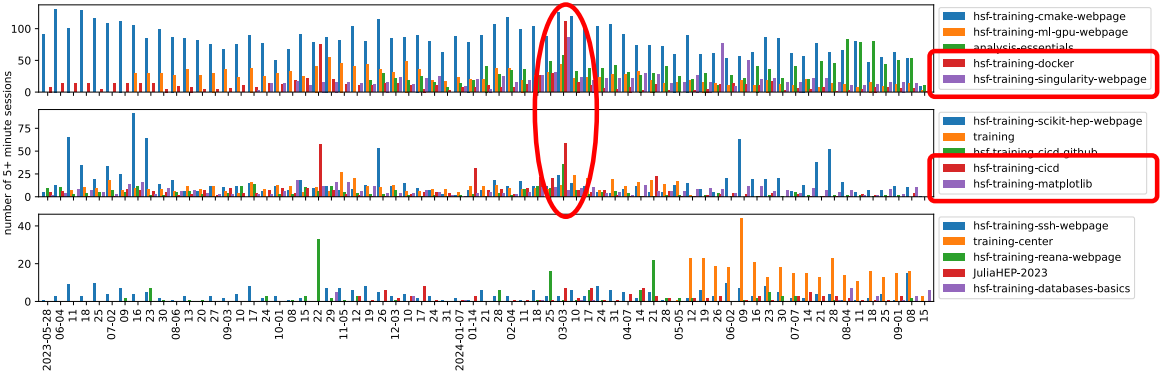
(Website) What are the most popular modules?



(Website) Long-term website visits broken down by module

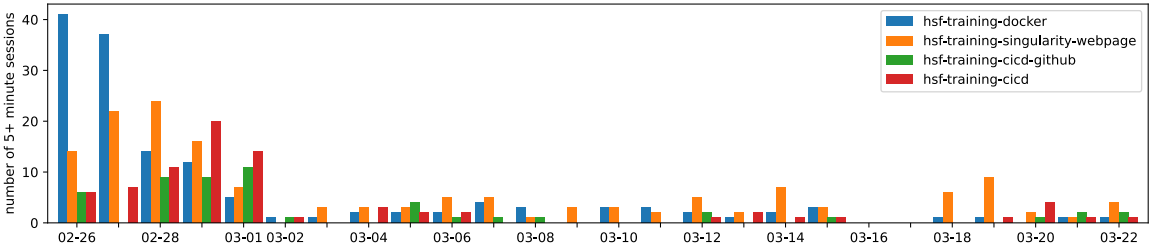


(Website) Long-term website visits broken down by module

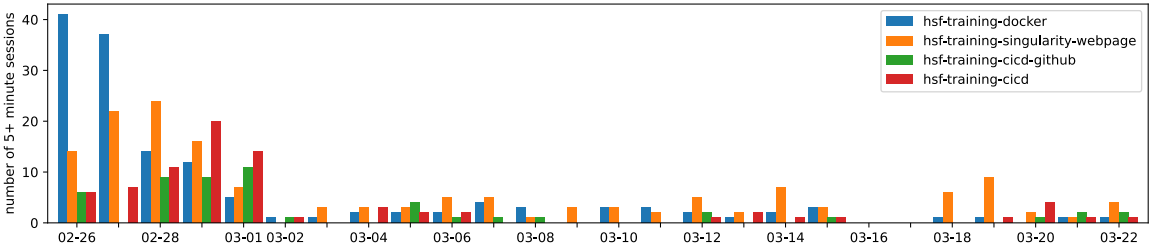


This is the same training event we saw before. Let's zoom into it.

(Website) Modules in the training event, that week + 3 more weeks

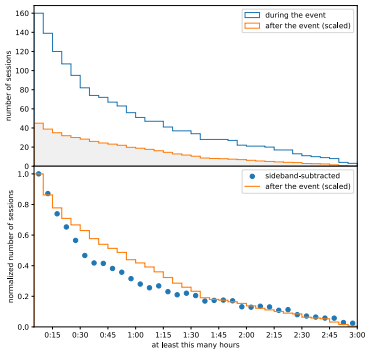


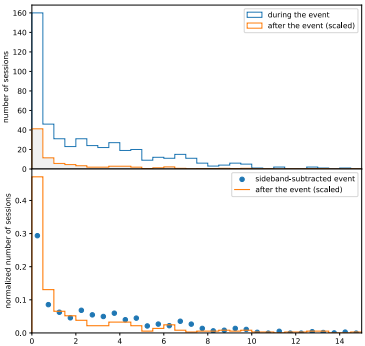
(Website) Modules in the training event, that week + 3 more weeks

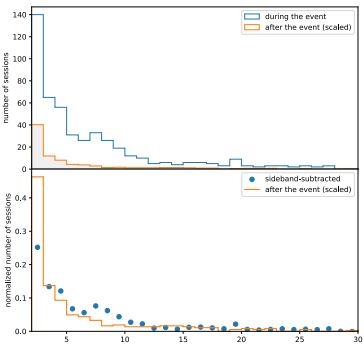


We can do a sideband subtraction: "excess" visitors during the event minus (properly scaled) data from after the event.

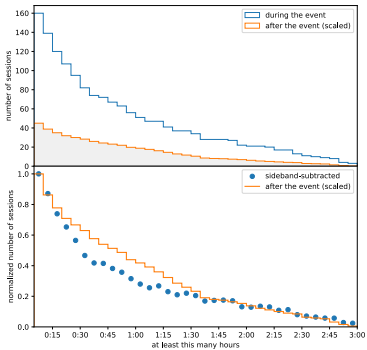
(Website) Properties of the visitors, with sideband subtraction

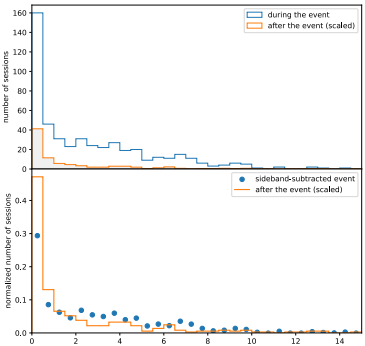


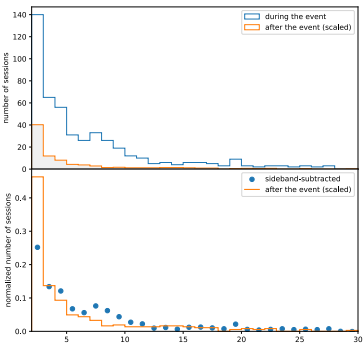




(Website) Properties of the visitors, with sideband subtraction

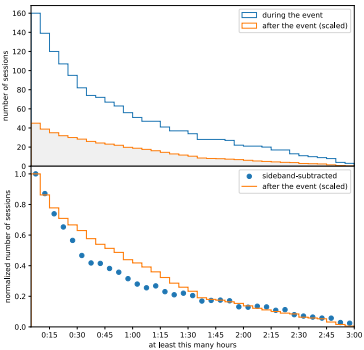


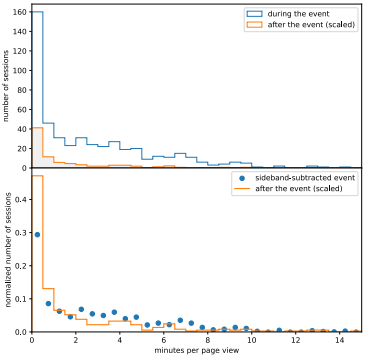


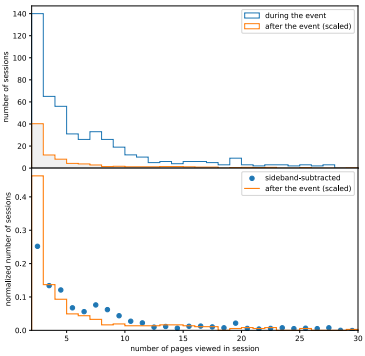


The distributions look *a little* different, but not much different.

(Website) Properties of the visitors, with sideband subtraction



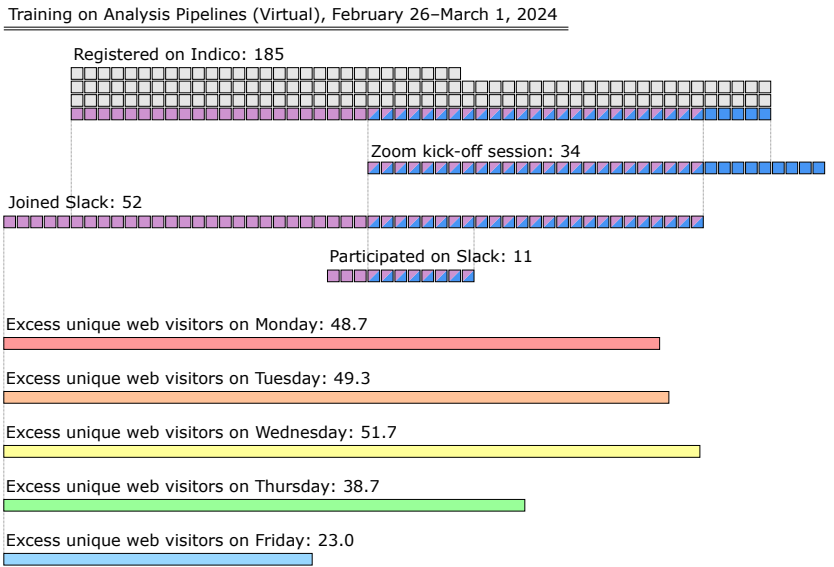




The distributions look *a little* different, but not much different.

Website visitors weeks after the event are as studious as during.

(All sources) Breakdown of February 2024 “Analysis Pipelines”



Conclusions

- First conclusion: we have many website visitors who are just as studious as those during an organized training event.

  About 50/day due to the event, but 18.5/day on typical weekdays.

Conclusions

- First conclusion: we have many website visitors who are just as studious as those during an organized training event.

  About 50/day due to the event, but 18.5/day on typical weekdays.

  - Can we make the website more useful for offline studying?

Conclusions

- First conclusion: we have many website visitors who are just as studious as those during an organized training event.

  About 50/day due to the event, but 18.5/day on typical weekdays.

  - Can we make the website more useful for offline studying?
  - Can we scale back the number of organized events per year?

Conclusions

- First conclusion: we have many website visitors who are just as studious as those during an organized training event.

  About 50/day due to the event, but 18.5/day on typical weekdays.

  - Can we make the website more useful for offline studying?
  - Can we scale back the number of organized events per year?

- The CMake module is especially popular, and we don't even organize a CMake event.

Conclusions

- First conclusion: we have many website visitors who are just as studious as those during an organized training event.

  About 50/day due to the event, but 18.5/day on typical weekdays.

  - Can we make the website more useful for offline studying?
  - Can we scale back the number of organized events per year?
- The CMake module is especially popular, and we don't even organize a CMake event.
- Wildly different distribution of undergrad/grad/postdoc/faculty by event, even events of the same type (not shown in this talk).

Conclusions

- First conclusion: we have many website visitors who are just as studious as those during an organized training event.

  About 50/day due to the event, but 18.5/day on typical weekdays.

  - Can we make the website more useful for offline studying?
  - Can we scale back the number of organized events per year?
- The CMake module is especially popular, and we don't even organize a CMake event.
- Wildly different distribution of undergrad/grad/postdoc/faculty by event, even events of the same type (not shown in this talk).
- There's quite a lot more in the surveys that hasn't been analyzed yet.

Conclusions

- First conclusion: we have many website visitors who are just as studious as those during an organized training event.

  About 50/day due to the event, but 18.5/day on typical weekdays.

  - Can we make the website more useful for offline studying?
  - Can we scale back the number of organized events per year?
- The CMake module is especially popular, and we don't even organize a CMake event.
- Wildly different distribution of undergrad/grad/postdoc/faculty by event, even events of the same type (not shown in this talk).
- There's quite a lot more in the surveys that hasn't been analyzed yet.
- How many Spanish-speaking countries in the registrants/website visitors?
  - Alex is studying that, to gauge interest in a Spanish translation.