LBNL CompBoot @ IRIS-HEP Roadmap

Link to the actual workshop material

Sam Meehan
(obo the instructors and participants who gave feedback)
13 September 2019

Goal : What they tell you

- Analyses (often students/contacts) show workflows like this
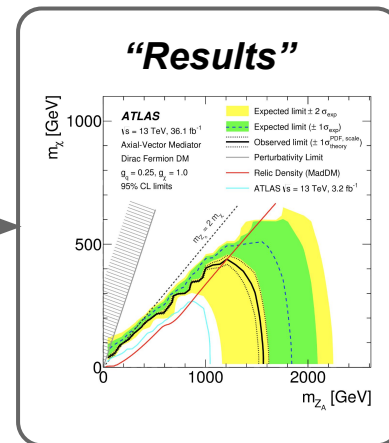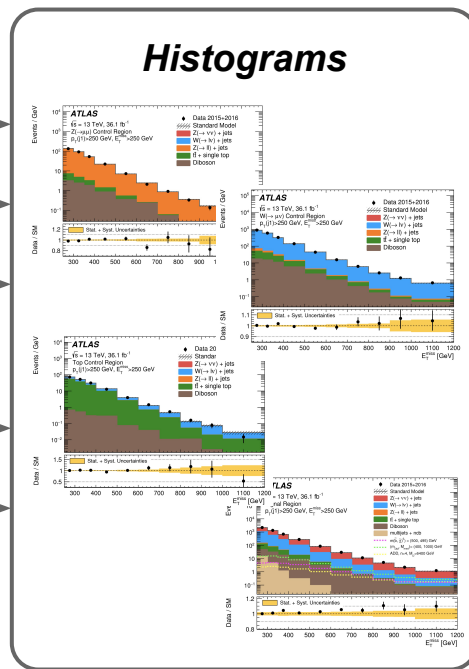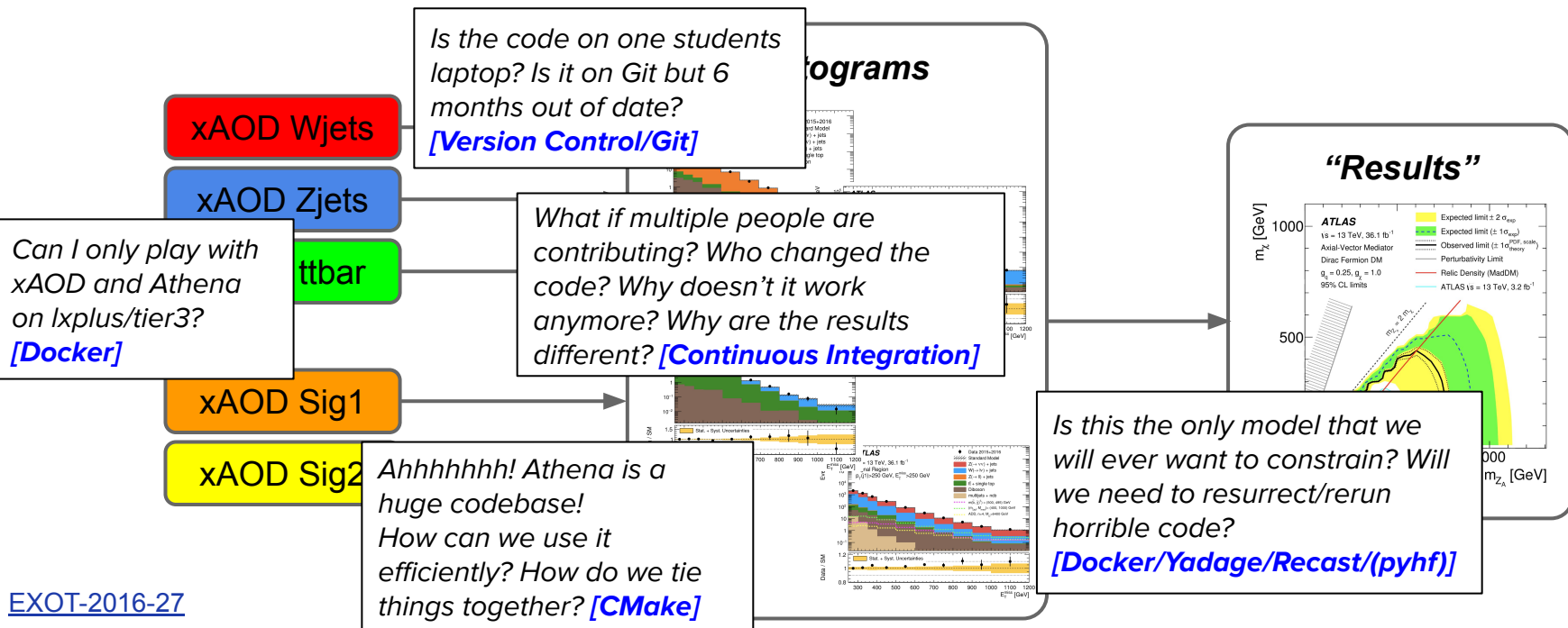 - This looks very nice and makes it feel "straightforward" to do analysis
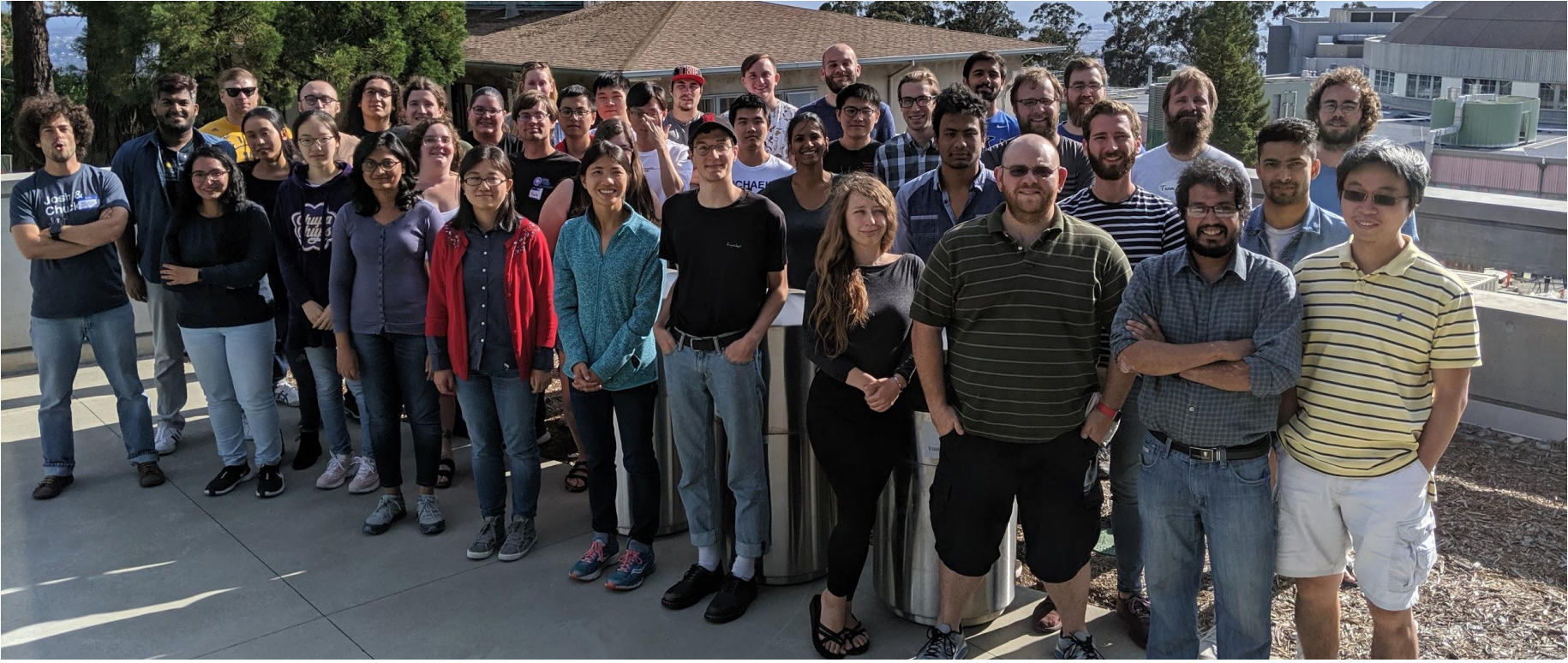

EXOT-2016-27

Goal : What they (don't) tell you

- ... but they don't highlight all of the frustrating aspects that are rarely taught
  - cause (unnecessary) holdups/inefficiency/frustration
- We aimed to mitigate these inefficiencies/frustrations via formal education
  - Thank you to those who support this and encouraged/allows their students/post-docs to attend

Is the code on one students laptop? Is it on Git but 6 months out of date?
[Version Control/Git]

Can I only play with xAOD and Athena on lxplus/tier3?
[Docker]

What if multiple people are contributing? Who changed the code? Why doesn't it work anymore? Why are the results different? ***[Continuous Integration]***

Ahhhhhhh! Athena is a huge codebase! How can we use it efficiently? How do we tie things together? ***[CMake]***

Is this the only model that we will ever want to constrain? Will we need to resurrect/rerun horrible code?
[Docker/Yadage/Recast/(pyhf)]

xAOD Wjets

xAOD Zjets

ttbar

xAOD Sig1

xAOD Sig2

...*ograms*

"Results"

Who : Participants

Who came?

- We had a rather diverse audience
  - {gender, age, native location/language}

Age
29 responses

Gender
30 responses

What is your native language?
31 responses
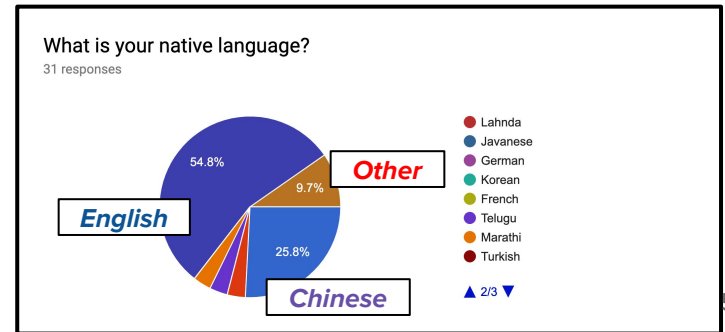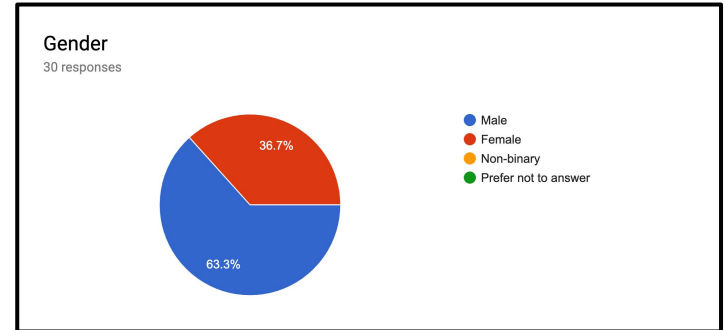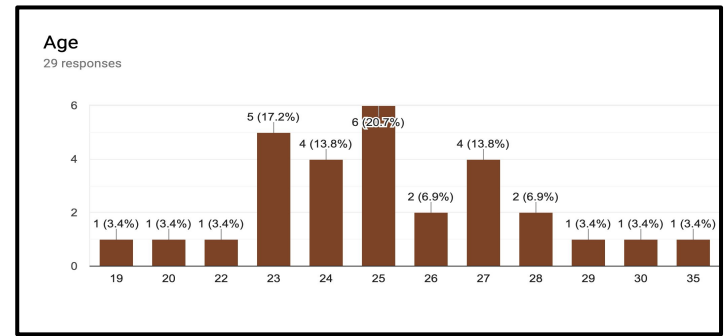
Who came? - Career Stage

- We reached the intended target!
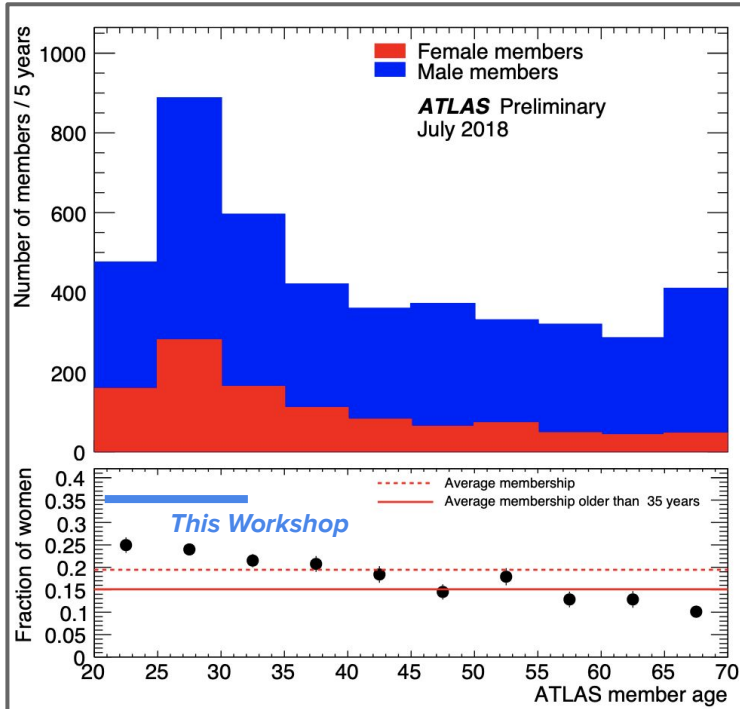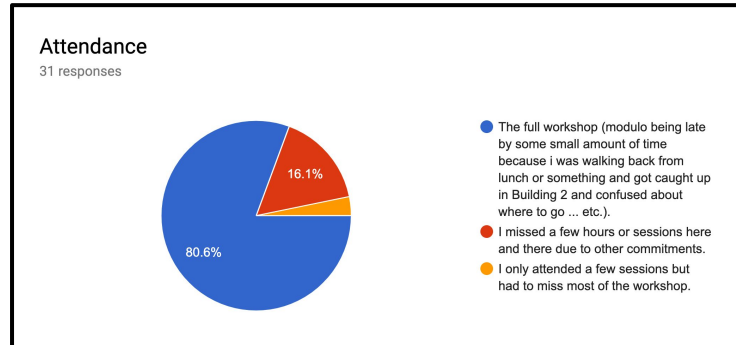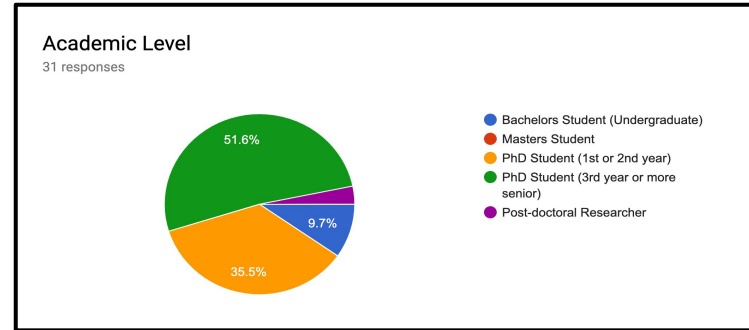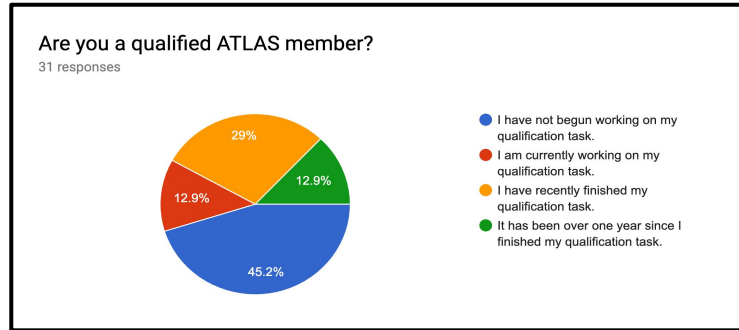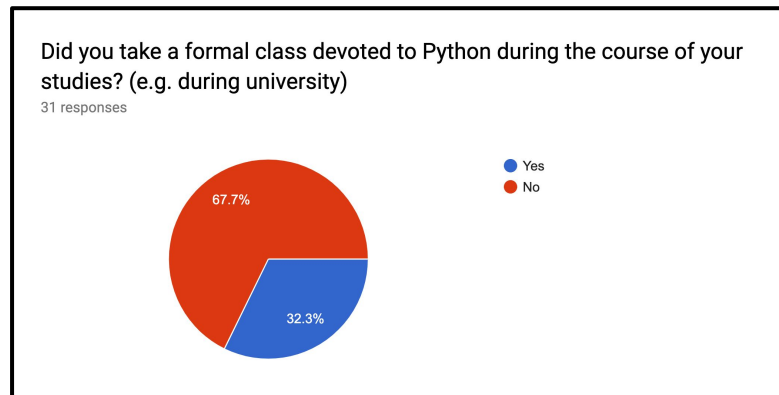 - Who : graduate students who can put these skills to work
- And they came for the entire workshop! ➜ Did not waste anyones time!

Are you a qualified ATLAS member?
31 responses

- I have not begun working on my qualification task.
- I am currently working on my qualification task.
- I have recently finished my qualification task.
- It has been over one year since I finished my qualification task.

45.2%, 12.9%, 29%, 12.9%

Academic Level
31 responses

- Bachelors Student (Undergraduate)
- Masters Student
- PhD Student (1st or 2nd year)
- PhD Student (3rd year or more senior)
- Post-doctoral Researcher

51.6%, 9.7%, 35.5%

Attendance
31 responses

- The full workshop (modulo being late by some small amount of time because i was walking back from lunch or something and got caught up in Building 2 and confused about where to go ... etc.).
- I missed a few hours or sessions here and there due to other commitments.
- I only attended a few sessions but had to miss most of the workshop.

80.6%, 16.1%

What did they know?

- The ***most common tools currently in ATLAS***
  - AnalysisBase and Athena are >99% written in ***c++***
  - We steer our code (i.e. jobOptions) with ***python***
  - We build our code with **CMake**

- Is this allowing us to reach our full potential

Did you take a formal class devoted to c++ during the course of your studies? (e.g. during university)
31 responses

45.2%
54.8%

- Yes
- No

I know how to use the CMake utility to build a c++ project.
31 responses

64.5%
32.3%

- True
- False
- Not exactly but partially

Did you take a formal class devoted to Python during the course of your studies? (e.g. during university)
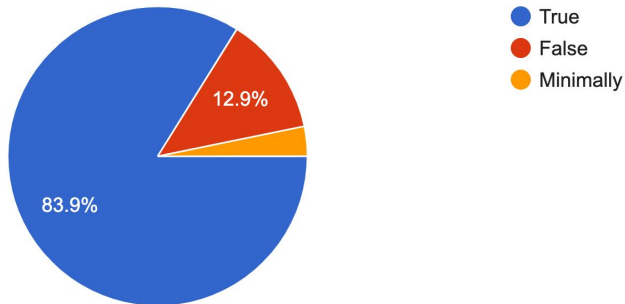31 responses

67.7%
32.3%

- Yes
- No
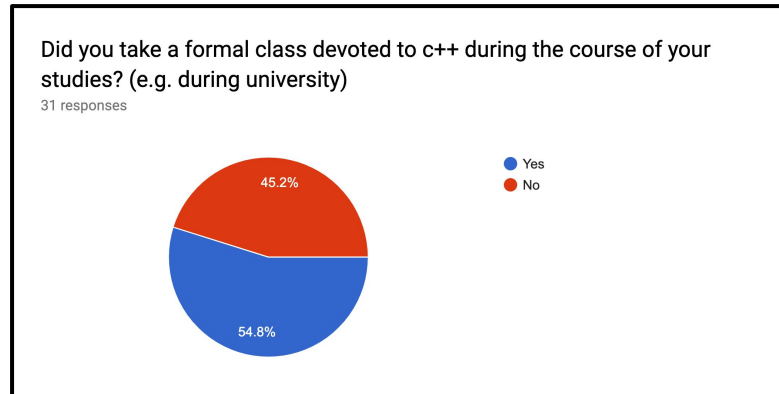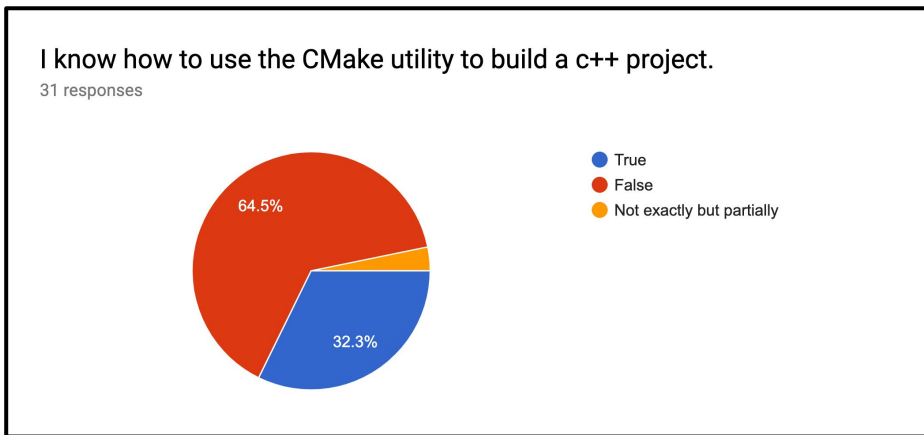
What did they know?

- Git is something that is being used
  - Basic skills (merge request) are "alright"
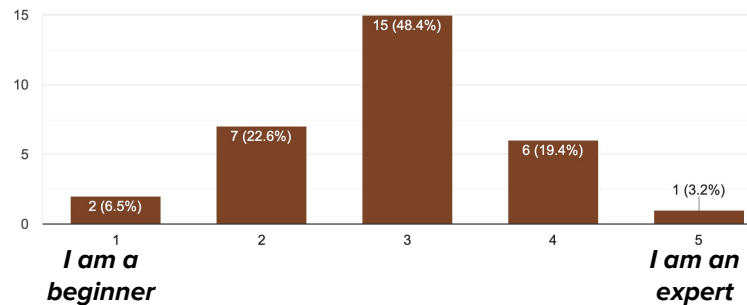 - Advanced skills (conflict res.) are "shaky"

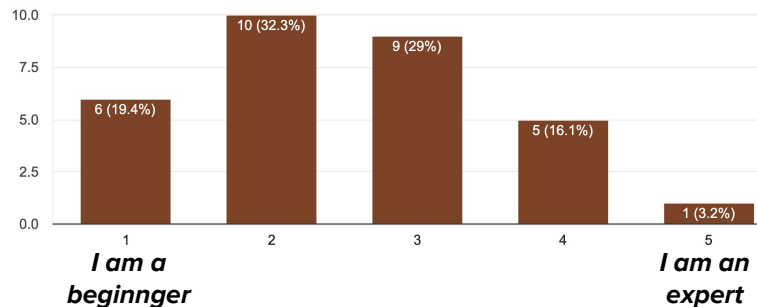I have used version control before.

31 responses

- True
- False
- Minimally

83.9%

12.9%

How comfortable are you when making a merge request (also called a pull request by some) to someone else's repository?

31 responses

15

10

5

0

2 (6.5%)
7 (22.6%)
15 (48.4%)
6 (19.4%)
1 (3.2%)

1 — *I am a beginner*
2
3
4
5 — *I am an expert*

How comfortable are you when resolving conflicts when committing and merging code?

31 responses

10.0

7.5

5.0

2.5

0.0

6 (19.4%)
10 (32.3%)
9 (29%)
5 (16.1%)
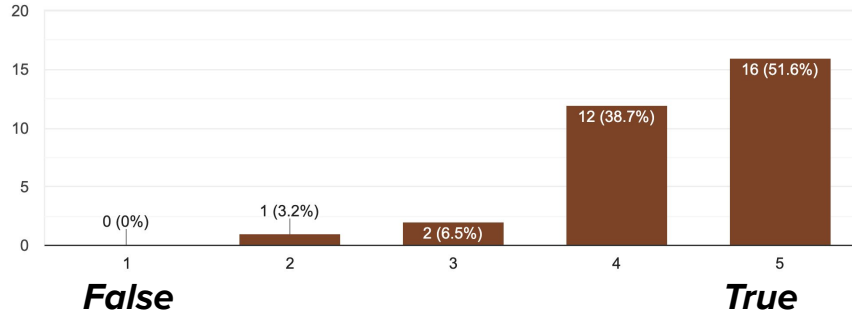1 (3.2%)

1 — *I am a beginnger*
2
3
4
5 — *I am an expert*

What did they know?

- Everyone knows you *should* implement automated tests
- … but noone really knows *how* to do this

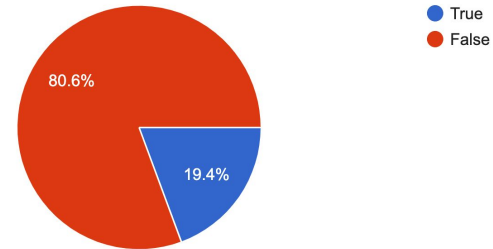Automated tests are a good thing to develop for your code.
31 responses

False ... *True*

I have used continuous integration and automated testing when developing code in the past.
31 responses

- True
- False

80.6%
19.4%

Schedule

Monday	Tuesday	Wednesday	Thursday	Friday

Git [Kelly/Kunal]	CMake [Henry]	Testing/CI [Kelly/Kunal]	Docker [Matthew]	pyhf Fitting / Documentation / Machine Learning / RDataFrames / EventLoop
Lunch	Lunch	Lunch	Lunch	Lunch
ATLAS-Git [Dan]	ATLAS-CMake [Karol]	ATLAS-CI/CD [Giordon]	ATLAS-Recast [Danika]	Free/Worktime
Jupiter Pizza		Cultural Event : Berkeley Bowl	Career Panel *Send questions!*	
			Dinner : 19h30 @ Great China	

Schedule

Monday	Tuesday	Wednesday	Thursday	Friday
Welcome **Surveys!**				pyhf Fitting / Documentation / Machine Learning / RDataFrames / EventLoop
Git [Kelly/Kunal]	CMake [Henry]	Testing/CI [Kelly/Kunal]	Docker [Matthew]	
Lunch	Lunch	Lunch	Lunch	Lunch
ATLAS-Git [Dan]	ATLAS-CMake	ATLAS-CI/CD [on]	ATLAS-Recast [Danika]	Free/Worktime
Jupiter Pizza		vent : Berkeley Bowl	Career Panel **Send questions!**	
			Dinner : 19h30 @ Great China	

software carpentry
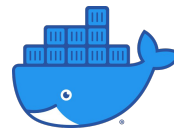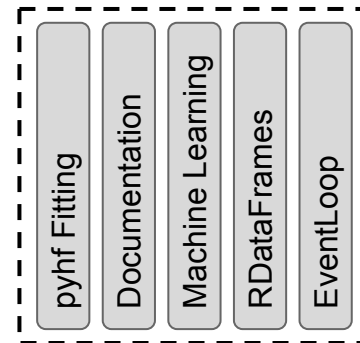
Schedule

	Monday	Tuesday	Wednesday	Thursday	Friday
	Welcome **Surveys!**				
	Git [Kelly/Kunal]	CMake [Henry]	Testing/CI [Kelly/Kunal]	Docker [Matthew]	pyhf Fitting · Documentation · Machine Learning · RDataFrames · EventLoop
	Lunch	Lunch	Lunch	Lunch	Lunch
	ATLAS-Git [Dan]	ATLAS-CMake [Karol]	ATLAS-CI/CD [Giordon]		Free/Worktime
	Jupiter Pizza		Cultural Event Berkeley Bowl	Dinner : 19h30 @ Great China	

iris hep

Schedule

	Monday	Tuesday	Wednesday	Thursday	Friday
	Welcome **Surveys!**	CMake [Henry]	Testing/CI [Kelly/Kunal]	Docker [Matthew]	pyhf Fitting / Documentation / Machine Learning / RDataFrames / EventLoop
	Git [Kelly/Kunal]				
	Lunch	Lunch	Lunch	Lunch	Lunch
	ATLAS-Git [Dan]	ATLAS-CMake [Karol]	ATLAS-CI/CD [Giordon]	ATLAS-Recast [Danika]	Free/Worktime
	Jupiter Pizza		Cultural Event : Berkeley Bowl	Career Panel **Send questions!**	
				Dinner : 19h30 @ Great China	
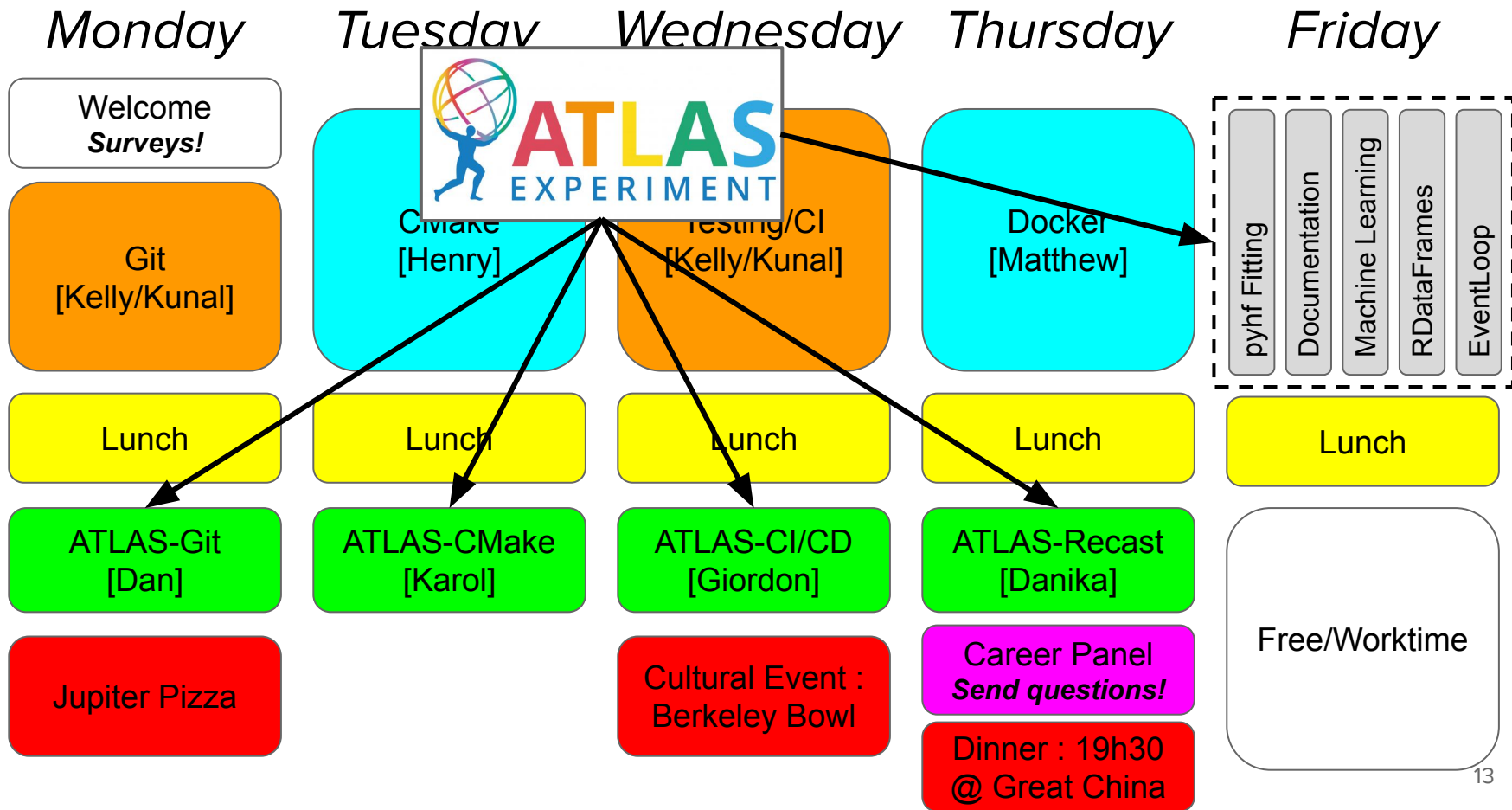
Who : Instructors

Danika MacDonell
Univ. of Victoria

Nils Krumnack
Iowa State

Matthew Feickert
Univ. of Illinois

Giordon Stark
UC Santa Cruz

Henry Schreiner
Princeton

Karol Krizka
LBNL

Dan Guest
UC Irvine

Adam Parker
Cal State East Bay

Sam Meehan
Univ. Washington/CERN

Who : Support Elsewhere

Zach Marshall
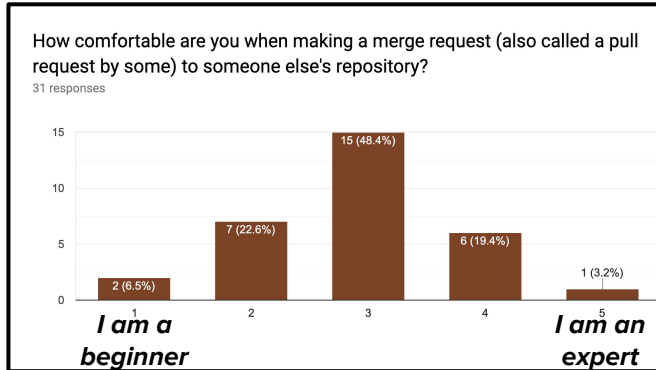LBNL

Lukas Heinrich
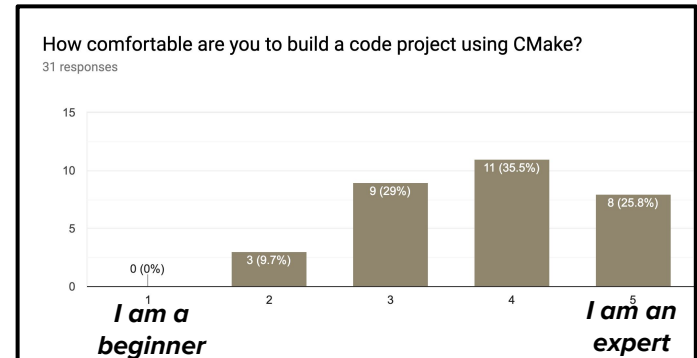New York Univ. / CERN

Shih Chieh Hsu
Univ. of Washington

Survey Feedback
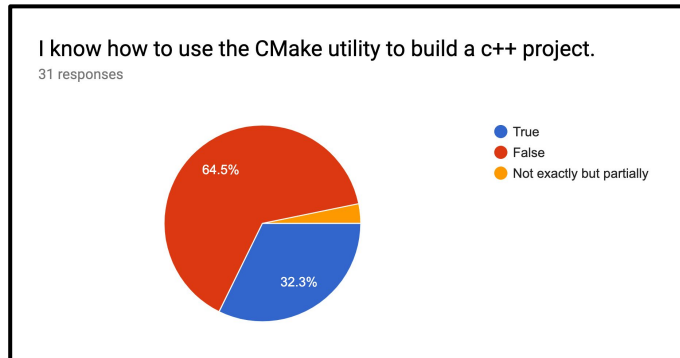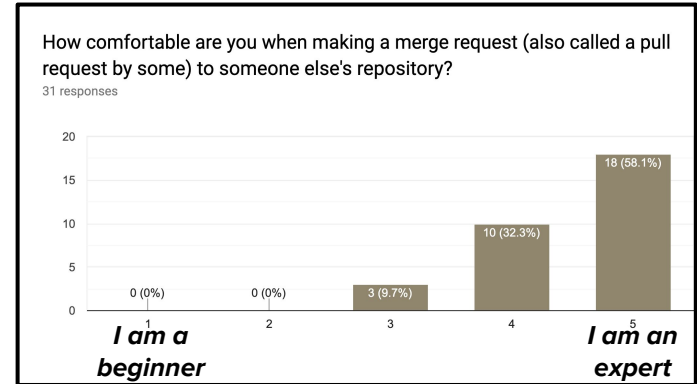
- Pre/Post workshop surveys allow for a derivative : #success

Before

How comfortable are you when making a merge request (also called a pull request by some) to someone else's repository?
31 responses

2 (6.5%) · 7 (22.6%) · 15 (48.4%) · 6 (19.4%) · 1 (3.2%)

I am a beginner ... *I am an expert*

After

How comfortable are you when making a merge request (also called a pull request by some) to someone else's repository?
31 responses

0 (0%) · 0 (0%) · 3 (9.7%) · 10 (32.3%) · 18 (58.1%)

I am a beginner ... *I am an expert*

I know how to use the CMake utility to build a c++ project.
31 responses

- True
- False
- Not exactly but partially

64.5% · 32.3%

How comfortable are you to build a code project using CMake?
31 responses

0 (0%) · 3 (9.7%) · 9 (29%) · 11 (35.5%) · 8 (25.8%)

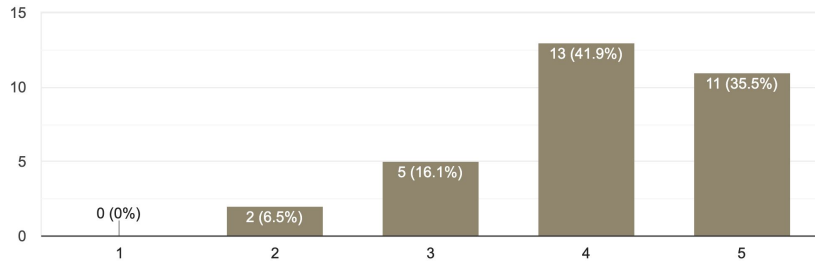*I am a beginner* ... *I am an expert*
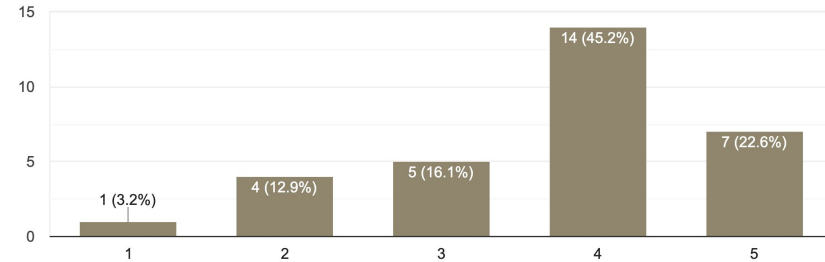
Survey Feedback

Other good stuff!

How comfortable are you with implementing Continuous Integration for your analysis in GitLab?
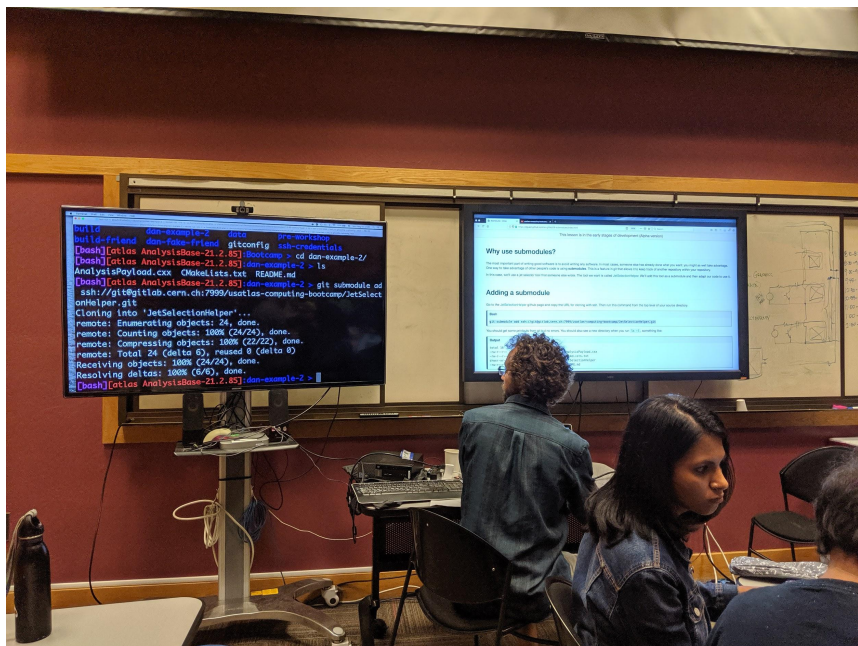
31 responses

How comfortable are you when working with docker images/containers?
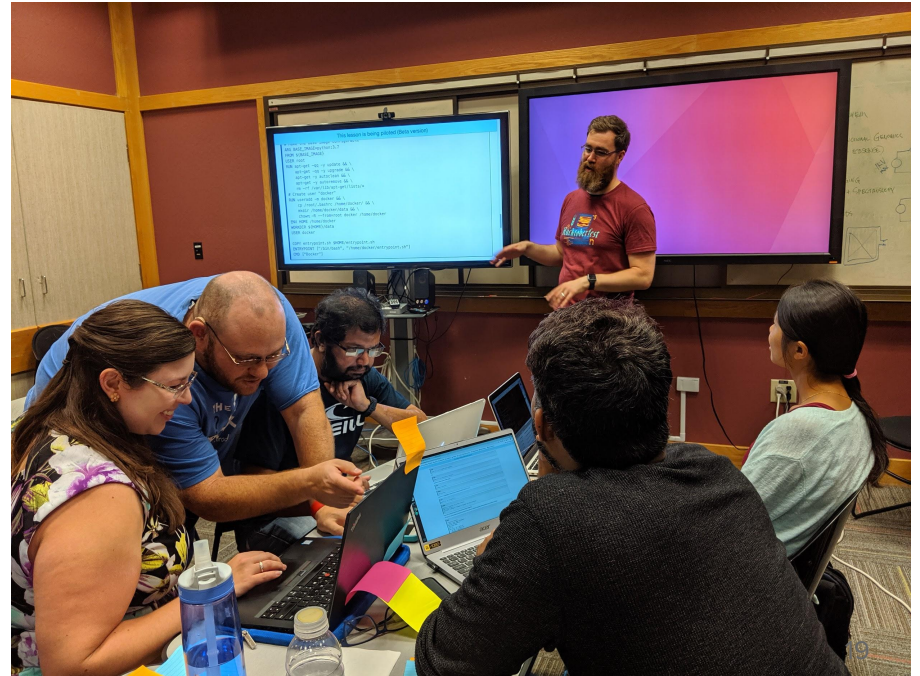
31 responses

We did real work!

- Hands on relied on two aspects
  - [1] : Scroll through Software Carpentry formatted website - NOT SLIDES!!!!!!!!!!
  - [2] : Work through exercise in parallel - "coding on the fly"

We did real work!

- Hands on relied on two aspects
  - [1] : Scroll through Software Carpentry formatted website - NOT SLIDES!!!!!!!!!!
  - [2] : Work through exercise in parallel - "coding on the fly"

We did real work

- Ratio of Student : Instructor = 4
  - This was *essential* to allow for the "hands on" aspect of the workshop to be successful
- Large time commitment on behalf of the instructors
  - Can't just "do your talk" and then leave

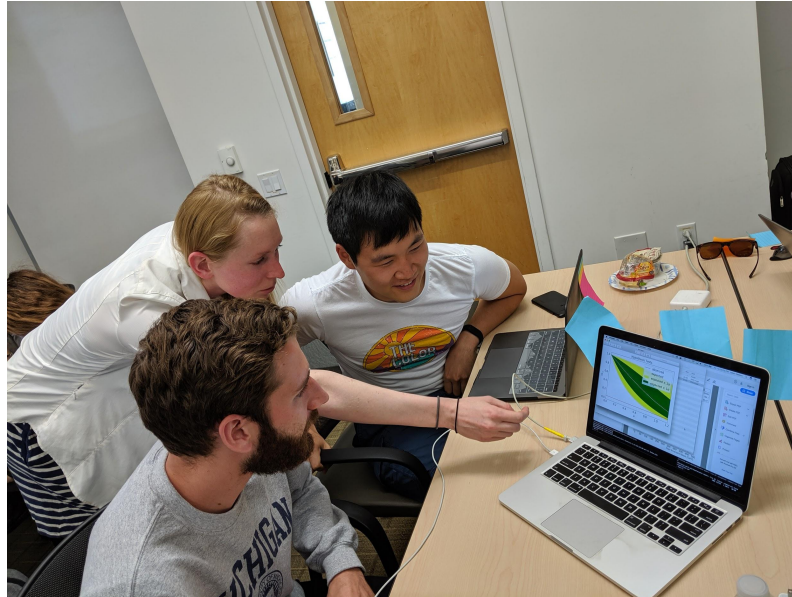**Zach : "I'm confused that ..."**   **Zach : "Yeah, I already tried that ..."**   **Zach : "Ahhhh, that makes sense!"**
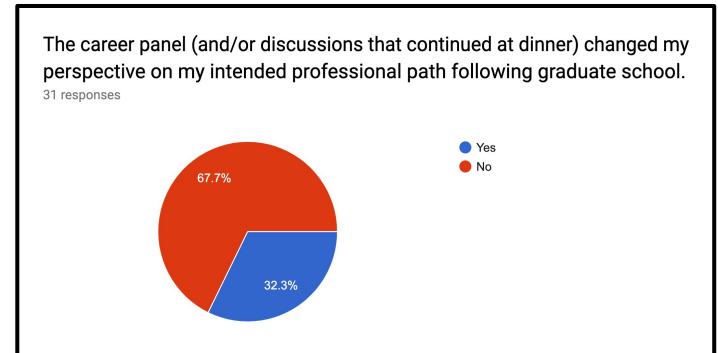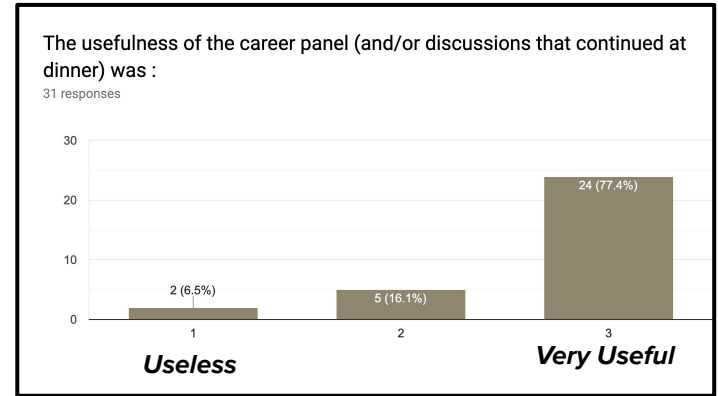
We got results!

- Produced a mu-scan of CLs to reinterpret a (toy) di-bjet bumphunt using a **docker image** produced using **continuous integration** that **builds the analysis with CMake** contained within a **GitLab repository**

Career Reflections

- Thursday afternoon/evening panel discussion with former ATLAS/CERN folks
- Discussion continued at the bootcamp dinner afterwards

Kathy Copic, Michaela Paganini, Luke de Oliveira, Francesco Rubbo, Mariana Rihl, Andrew Hard

The usefulness of the career panel (and/or discussions that continued at dinner) was :

31 responses

Useless — *Very Useful*

- 1: 2 (6.5%)
- 2: 5 (16.1%)
- 3: 24 (77.4%)

The career panel (and/or discussions that continued at dinner) changed my perspective on my intended professional path following graduate school.

31 responses

- Yes: 32.3%
- No: 67.7%

Funding : How Much

- 20k was awarded for the original ATC proposal - THANK YOU!
- Peter Loch suggested/helped re-purpose the HFSF leftover funds (3.5k)
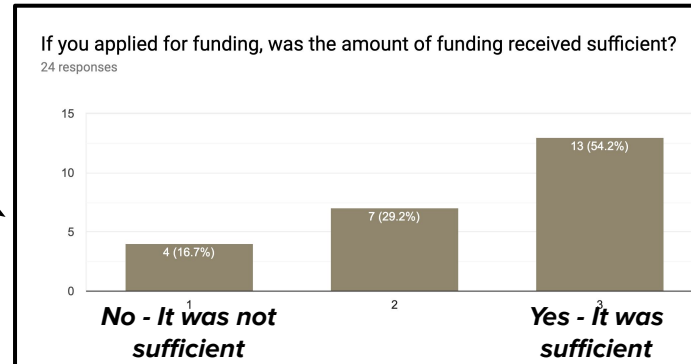- Partnering with FIRST-HEP greatly helped facilitate success!

Snacks [600]

Dinners [3400]

Instructor Travel [2700]

Instructor Housing [2000]

Allotted 20000

Participant Travel [5500]

Participant Housing [7500]

+

Software Carpentry (FIRST-HEP) [2000]

Henry (FIRST-HEP) [1200]

Adam/Nils (ATC) [1800]

Total Cost = 27000
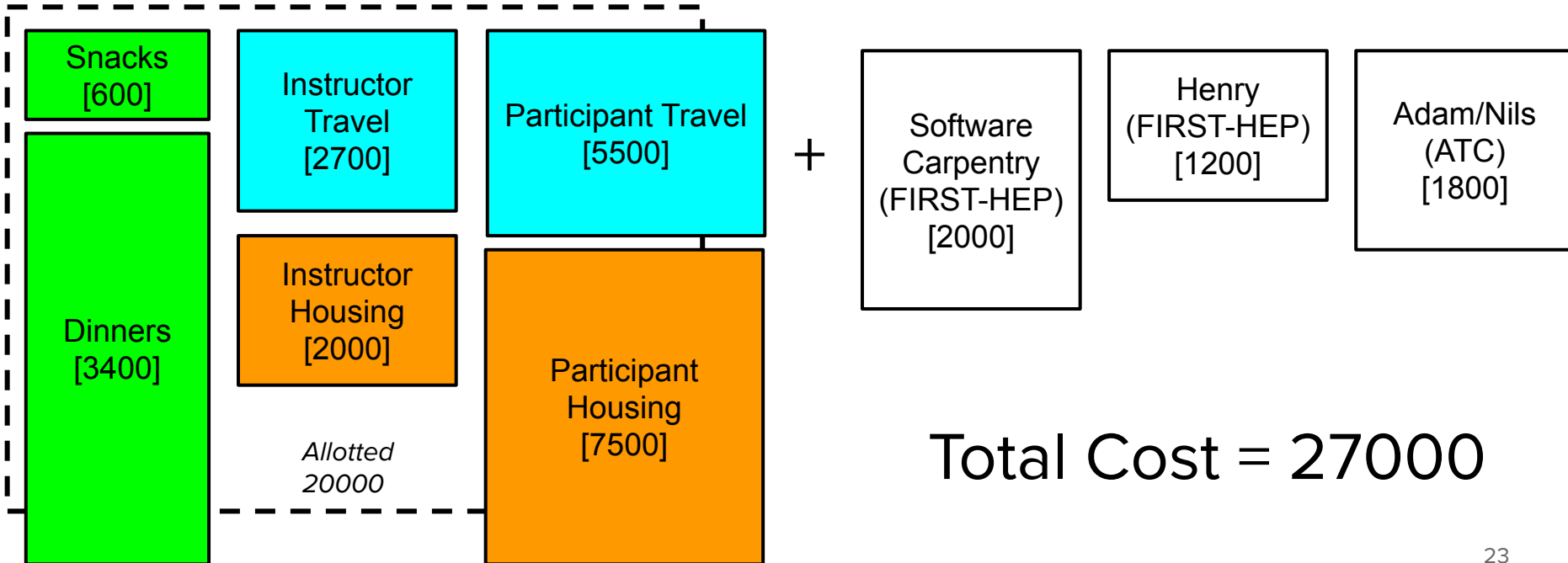
Funding : Was it "enough"?

- We wanted to assure that any student who *wants* to learn *can* learn
  - Within "uncertainties" I think that our approach to assisting participants worked well

Did you apply for travel funding?
31 responses

- Yes
- No
- No - but in retrospect I wish I had.
- What? I could have applied for funding?

67.7%
29%

If you applied for funding, did you/do you expect to also receive support from your institution?
28 responses

- Yes
- No
- Not sure

28.6%
39.3%
32.1%

If you applied for funding, was the amount of funding received sufficient?
24 responses

4 (16.7%)
7 (29.2%)
13 (54.2%)

No - It was not sufficient

Yes - It was sufficient

Future : Location

- Success of the workshop is highly dependent on the location
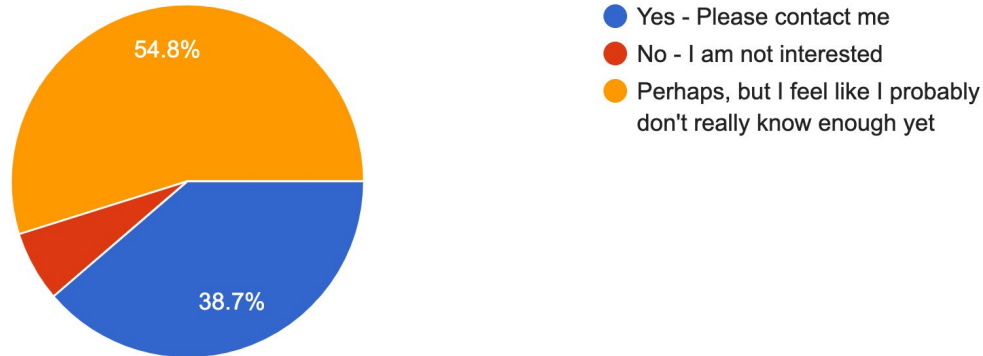 - Is this event "vidyo-able" and can be held remotely? - No [my opinion]

Future : Instructors

- We didn't scare them all off!
- We should reach out and ask if these people want to be involved!

Would you be interested in being involved in the USATLAS Computing Bootcamp (this thing) next year in 202...who are enthusiastic experts to help.]

31 responses

- Yes - Please contact me
- No - I am not interested
- Perhaps, but I feel like I probably don't really know enough yet

54.8%

38.7%

Opinions

- [1] We need to have particular locations for hands-on events
  - Needs forethought : >6 months at LBNL apparently (but there are fantastic rooms there if we could have gotten them
- [2] We should use the SWC approach to preserve/reuse our work
  - Fantastic interface and result for tutorial - systematized to provide a uniform interface for students
  - Preserved in Git! (Practice what we preach)
  - Can feed back *into* the open source community
- [3] We have plenty of in-house enthusiasm to help prepare these
  - I get the sense that many of these people/students/post-docs have not felt encouraged by their superiors to invest time in training ➜ who changes this? [Us, I suppose]
- [4] Technicalities
  - e.g. Use a white terminal with dark text
  - Can mitigate with Software Carpentry Training