Building Oceanhackweek: a week of data science, hacking, collaboration and more

Wu-Jung Lee, Amanda Tan, Valentina Staneva, Filipe Fernandes, Joseph Gum, Emilio Mayorga, Timothy Crone, Ryan Abernathy

Oceanhackweek is a workshop focused on educating ocean scientists with **modern data science skills** and building an **inclusive and cohesive user community**. After three workshops and in the planning phase of another here are our **key elements for success**:

Participant selection
- **Beyond "Hello World" ability** in target programming language, experience equivalent to a Carpentries class preferred
- **Basic domain knowledge** (oceanography)
- Recruit highly **skilled individuals as TAs** to lead projects and group discussions

Curriculum Planning
- Mixture of **hands-on tutorials** and **presentations**
- **Equal time** for group projects and tutorials
- Tutorials publicly **available online** (ex. Github) during and after the event
- Tutorials focus on **data and computational techniques used in oceanography today**

Project Facilitation
- **Self-organized** groups after project brainstorming on the first day
- Instructors function as **proactive "roamers"** and accessible **help desks** to field questions
- TAs work in their own groups and/or answer questions

Cyberinfrastructure support
- All tutorials run in a **cloud-based JupyterHub** to minimize overhead in computing environment setup
- Participants bring their **own computers** for learning and project work

Critical lessons learned
- Focus on **tools and techniques** that participants can take with them
- Establish a **code of conduct** so all feel empowered to learn

# A **workshop** teaching **data science** skills with **collaborative projects** in oceanography: **Oceanhackweek**

What topics do you want to learn about? 1st: 2nd: 3rd:

Big Data / Cloud Computing (Pangeo, Xarray, Dask)	**Machine Learning** (Scikit-learn)	**Visualization** (Holoviews, Cartopy)	**Open Science**
Data Preparation	**Automated Pipelines**	**Structured Data** (Standard Vocabularies, netCDF)	**Reproducible Science** (FAIR)
Large Project Datasets (Argo, OOI, GO-SHIP)	**Field Setup / Software Infrastructure**	**Software Engineering Best Practices**	**Raw Sensor Data / Sensor Interfaces**

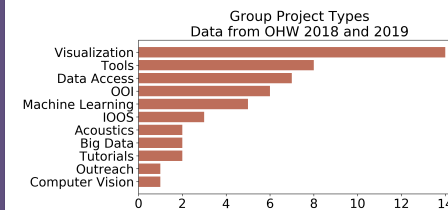Want to participate? Visit: **oceanhackweek.github.io**
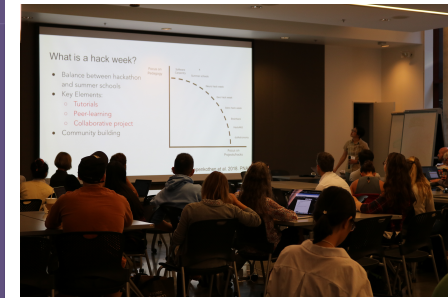
"The workshop provided an excellent introduction to the existing infrastructure of oceanographic data-gathering resources. Also provided was an introduction to the open-source evolving tools for accessing and utilizing large data sets. The technical challenges in data-intensive research are daunting, and here this workshop's model and enactment of collaborative work was particularly valuable to me." ~ Rachel Jackson, San Francisco State University

"It is sometimes hard to predict whether a given conference, training or hack-a-thon might be worth your time. [...] The real challenge, of course, is to find a group of like minded people that want to expand their horizons and learn together, and that is just what I found when I attended Ocean Hackweek." ~ Christian Saranson

How much do you think you learned? Data from OHW 2019
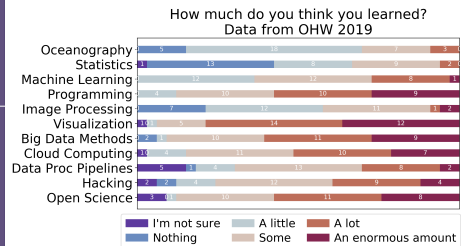
Quick Stats:
- 150+ Participants
- 22 Projects
- 23 Tutorials on GitHub
- 3 Hackweeks to date
- 2 Hackweeks in 2020: East and West

Group Project Types Data from OHW 2018 and 2019