WHOI Data Science Training Camp syllabus 2020

https://tinyurl.com/WHOI-Data-Science-syllabus

[link to last year’s syllabus]

22 & 23 January 2020, afternoons (12:30 to 4:30pm), WHOI Quissett Campus, Clark 5-07

Please note start time 12:30

Lead instructor and contact: Stace Beaulieu, Redfield 1-12, x3536 (stace@whoi.edu)

Description: With data and software increasingly recognized as scholarly research products, and with current momentum towards computational reproducibility and open science (see references below), it is imperative to have practical training in data management and software development. WHOI’s “Navigating an Ocean of Data” Training Camp provides an opportunity for students, postdocs, and others at WHOI to learn essential practices and skills for data management and research computing, including practices specific to the ocean sciences and exposure to WHOI’s institutional resources. The training is guided by WHOI technical specialists during two afternoons, each with two sessions.

Requirements: There are no pre-reqs. Participants need to bring their laptops; each session includes hands-on activities.

Registration limited to ~40 participants: Registration is required (there is no registration fee). Registration will open to JP students on 6 January 2020, to WHOI postdocs on 9 January 2020, and to all other WHOI staff and students on 13 January 2020.

Sponsored by: WHOI Doherty Chair in Education Award through Academic Programs and co-sponsored by WHOI Ocean Informatics initiative, MBLWHOI Library, and WHOI Information Services.

Please read ahead of time to prepare for the workshop:


- The Carpentries Code of Conduct (sections at the top re: welcoming and supportive environment): https://docs.carpentries.org/topic_folders/policies/code-of-conduct.html
### Weds. 22 January 2020 (please bring your laptop)

<table>
<thead>
<tr>
<th>12:30 - 14:15</th>
<th>Module 1: “Good enough practices in scientific computing”</th>
<th>This module sets the stage for the other modules, touching on foundational practices for data, software, and project management and collaboration. These practices are popularized by The Carpentries, in which WHOI is a member organization. Link to slides</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Instructor: Stace Beaulieu (Coordinator, WHOI Ocean Informatics initiative)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Helpers: Shannon Rauch, Karen Soenen, and Amber York (BCO-DMO), Jaxine Wolfe (NES-LTER)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14:15 Break 15 min</td>
<td></td>
</tr>
<tr>
<td>14:30 - 16:15</td>
<td>Module 2: Data management</td>
<td>This module will cover the research data life cycle, general practices for data management and citation, and an introduction to WHOI institutional resources available through the Data Library. Link to slides</td>
</tr>
<tr>
<td></td>
<td>Instructors: Lisa Raymond and Audrey Mickle (WHOI Data Library)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16:15 Describe how to prepare for tomorrow:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Choose a journal article relevant to your research and determine if you can access the data and/or software that support the results.</td>
<td></td>
</tr>
</tbody>
</table>

### Thurs. 23 January 2020 (please bring your laptop)

<table>
<thead>
<tr>
<th>12:30 - 14:15</th>
<th>Module 3: Software development and research computing</th>
<th>This module will cover general practices for software development, demo several tools including Jupyter Notebooks and WHOIGit, and introduce WHOI IS resources including our high-performance computing (HPC) cluster. Link to slides Part 1 and Part 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Instructors: Joe Futrelle, Nick Symmonds, Roberta Mazzoli, Rich Brey (WHOI Information Services)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14:15 Break 15 min</td>
<td></td>
</tr>
</tbody>
</table>
14:30 - 16:15

**Module 4: Best practices in the ocean sciences**

Instructors: Stace Beaulieu (Information Manager, NES-LTER) and Danie Kinkade (Director, BCO-DMO)

This module will cover practices specific to the ocean sciences for sharing data and software and will introduce national and international organizations with resources for further learning and engagement. [Link to slides](#)

16:15 Poll for Data Science topics for 2020 Summer Series; Vote for Software Carpentry and Data Carpentry Workshops in 2020

---

**References:**


Software installation:

Please note due to limited time we are not going to be able to include software installation for this workshop. We will be offering Software and Data Carpentry Workshops in 2020 so you will have more opportunities for hands-on coding.