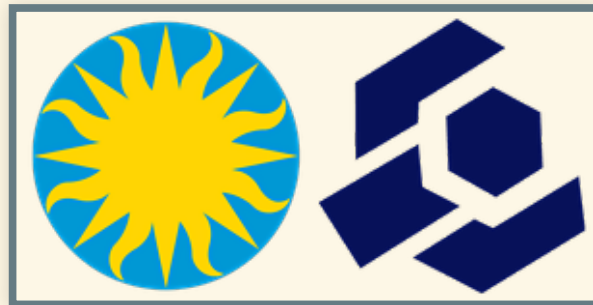


SPRING 2020

QUARTERLY DATA

CARPENTRY WORKSHOP



Before we start...

CODE OF CONDUCT

- We are dedicated to providing a welcoming and supportive environment for all people, regardless of background or identity.

CODE OF CONDUCT

https://docs.carpentries.org/topic_folders/policies/code-of-conduct.html

CODE OF CONDUCT

- Go to Instructors or Helpers if you feel the Code of Conduct is being violated.
- There is also a big blue "Report a Code of Conduct Incident" on the course website if you feel the Instructors are not handling appropriately.

INTRODUCTIONS

INSTRUCTORS AND HELPERS

INTRODUCTIONS

LEARNERS

WHAT IS THE CARPENTRIES?

<https://carpentries.org/>

The Carpentries is a global volunteer-driven organization dedicated to teaching computational skills to researchers and anyone who works with data.

CARPENTRIES AT THE SMITHSONIAN

<https://datascience.si.edu/carpentries>

The Smithsonian Carpentries Team is made up of Smithsonian staff and fellows from around the institution.

CARPENTRIES WAY OF TEACHING

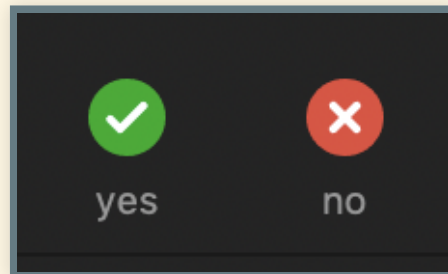
Code along, rather than lecturing from PowerPoint, or
just demonstrating

Odds of running into an error are high, but don't panic –
this is a great opportunity to learn

WE WON'T LEAVE YOU BEHIND

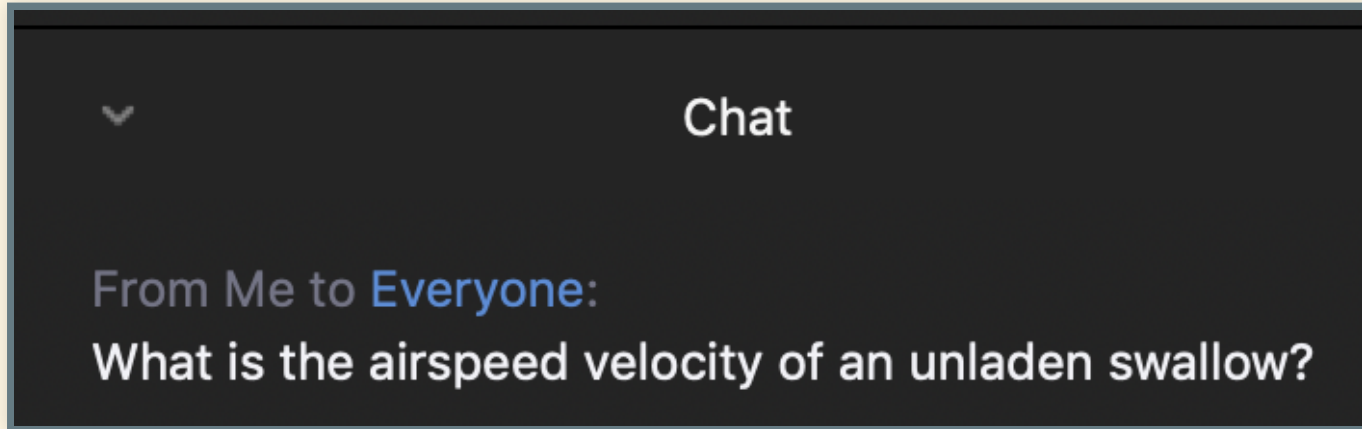
HOW TO MAKE SURE WE DON'T LEAVE YOU BEHIND

In the real world, we used to use Post-It notes to indicate progress, but for this workshop we'll use the "yes" / "no" buttons in the Zoom Participants panel.



HOW TO ASK QUESTIONS

The Instructors will try to build in time between lessons to ask questions, but use the Zoom chat for questions that come up during a lesson.



COURSE GITHUB REPOSITORY

<https://github.com/SmithsonianWorkshops/2020-05-19-smithsonian>

COURSE ETHERPAD

<https://pad.carpentries.org/2020-05-19-smithsonian>

A collaborative document – think of it as a stripped-down version of Google Docs

YOUR FIRST CHALLENGE

- Go to the EtherPad
- Add your name, unit, and your favorite emoji (😊) under Attendance Day 1

WHAT WE WILL BE TEACHING IN THIS WORKSHOP

We will be teaching you data organization, cleaning, and visualization skills, so that your analyses can be easily understood and reproduced by your colleagues ... and your future self.

DAY ONE (TODAY)

TIDY DATA PRINCIPLES IN SPREADSHEETS

OPENREFINE

DAY TWO (TOMORROW)

SQL

DAY THREE (NEXT TUE)

R AND RSTUDIO

WORKING WITH DATA FILES

DAY FOUR (NEXT WED)

R AND RSTUDIO

VISUALIZING DATA WITH GGPLOT2

We will be teaching you a lot of valuable skills on a demo dataset, but you won't be able to add these tools to your toolbox until you work through how to apply them to your own data.

How to actually learn any new programming concept



Essential

Changing Stuff and
Seeing What Happens

O RLY?

@ThePracticalDev