

# Introduction

# Welcome to class!

1. Introductions
2. Class overview
3. Getting R up and running



[Photo by [Belinda Fewings](#) on [Unsplash](#)]

**Before we start ..**

Poll: How are you feeling right now?

# About Us

**Carrie Wright (she/her)**

Senior Staff Scientist and Training Lead, Fred Hutchinson Cancer Center

Associate, Department of Biostatistics, JHSPH

PhD in Biomedical Sciences

Email: [cwright2@fredhutch.org](mailto:cwright2@fredhutch.org) Web: <https://carriewright11.github.io>



# About Us

Ava Hoffman (she/her)

Senior Staff Scientist, Fred Hutchinson Cancer Center

Associate, Department of Biostatistics, JHSPH

PhD in Ecology

Email: [ahoffma2@fredhutch.org](mailto:ahoffma2@fredhutch.org) Web: <https://avahoffman.com>



## About Us: TA

**Elizabeth Humphries (she/her)**

Staff Scientist, Fred Hutchinson Cancer Center

PhD in Molecular Epidemiology

Email: [ehumphri@fredhutch.org](mailto:ehumphri@fredhutch.org)

**NOTE** this is not her dog



# About you!

Please introduce yourself on Slack!

[Slack Workspace](#)

# The Learning Curve

Learning a programming language can be very intense and sometimes overwhelming.

We recommend fully diving in and minimizing other commitments to get the most out of this course.

Like learning a spoken language, programming takes **practice**.



# The Learning Curve

Learning R has been career changing for all of us, and we want to share that!

We want you to succeed – We will get through this together!



## What is R?

- R is a language and environment for statistical computing and graphics developed in 1991
- R is both open source and open development



[source: <http://www.r-project.org/>]

# Why R?

- Free (open source)
- High level language designed for statistical computing
- Powerful and flexible - especially for data wrangling and visualization
- Extensive add-on software (packages)
- Strong community



[source: [https://github.com/r-ladies/meetup-presentations\\_baltimore](https://github.com/r-ladies/meetup-presentations_baltimore)]

## Why not R?

- Little centralized support, relies on online community and package developers
- Annoying to update
- Slower, and more memory intensive, than the more traditional programming languages (C, Perl, Python)



[source -School vector created by nizovatina - [www.freepik.com](http://www.freepik.com)]

# Introductions

What do you hope to get out of the class?

Why do you want to use R?



[Photo by [Nick Fewings](#) on [Unsplash](#)]

# Logistics

# Course Website

<https://daseh.org/>

Materials will be uploaded the night before class. We are constantly trying to improve content! Please refresh/download materials before class.



# Learning Objectives

- Understanding basic programming syntax
- Reading data into R
- Recoding and manipulating data
- Using add-on packages (more on what this is soon!)
- Making exploratory plots
- Performing basic statistical tests
- Writing R functions
- **Building intuition**

# Course Format

## ONLINE VIRTUAL COURSE

- Lecture with slides, interactive
- Lab/Practical experience
- Two 10 min breaks each day - timing may vary
- Sep 30 - Oct 10, 10:30am - 2:00pm PST on Zoom

# Course Format

## IN-PERSON CODE-A-THON

- Mostly independent group work
- Frequent check-ins with instructors and other groups
- Some lectures about the practical aspects of coding
- Oct 23 - 25 (in person in Seattle)

**Deadline for arranging travel assistance is October 11!**

# Pulse Check Survey

<https://forms.gle/JAsLbyRDtCtxwgFG9>

Let us know anonymously how you're doing with the material.



[source - Banner vector created by pch.vector - [www.freepik.com](http://www.freepik.com)]

## Homework

While we do have homework assignments on the course schedule, these are **strictly optional!!!**

We encourage you to try the assignments, as the best way to get comfortable with any programming language is through practice.

# Your Setup

If you can, we suggest working virtually with a **large monitor or two screens**.

This setup allows you to follow along on Zoom while also doing the hands-on coding.



[source - [reddit.com](#)]

# Research Survey

## Research Survey

We are collecting data about user experience with our course to learn more about how to improve the data science education experience. This data may ultimately be used for a research publication and reporting to the NIH.

<https://forms.gle/vxp8vxdN2Ck2j7uY6>

Where to find help

## Useful (+ mostly Free) Resources

Found on our website under the Resources tab:

<https://daseh.org/resources.html>

- videos from previous offerings of the class
- cheatsheets for each class

# Help!!!

Error messages can be scary!

- Check out the FAQ/Help page on the website: <https://daseh.org/help.html>
- Ask questions in Slack! Copy+pasting your error messages is really helpful!

We will also dedicate time today to debug any installation issues



## Installing R

- Install the [latest R version](#) (4.4.1 (called 'Race for your Life') as of June 14, 2024)
- [Install RStudio](#)
- [R Tools Issue](#)

More detailed instructions [on the website](#).

RStudio is an **integrated development environment** (IDE) that makes it easier to work with R.

More on that soon!

## Summary

- [Class Website](https://daseh.org/) - <https://daseh.org/> - Logistics, resources, and help!
- [Research Survey](https://forms.gle/vxp8vxdN2Ck2j7uY6) - <https://forms.gle/vxp8vxdN2Ck2j7uY6>
- [Pulse Check](https://forms.gle/JAsLbyRDtCtxwgFG9) - <https://forms.gle/JAsLbyRDtCtxwgFG9>

## Summary

Image by [Gerd Altmann](#) from [Pixabay](#)