

Intro to R

Introduction

Welcome!

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



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

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 Web: <https://avahoffman.com>



About Us

Carrie Wright (she/her)

Senior Staff Scientist, Fred Hutchinson Cancer Center

Associate, Department of Biostatistics, JHSPH

PhD in Biomedical Sciences

Email: cwright2@fredhutch.org Web: <https://carriewright11.github.io>



The Learning Curve

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

We recommend fully diving in and making lots of mistakes through trial and error.

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](#) (aka, free!)



- Powerful and flexible - especially for data wrangling and visualization
- Extensive add-on software (packages)
- Strong community – <https://rladies.org/>

[source: <http://www.r-project.org/>, [https://en.wikipedia.org/wiki/R_\(programming_language\)](https://en.wikipedia.org/wiki/R_(programming_language))]

Workshop Website

https://hutchdatascience.org/SeattleStatSummer_R/



Learning Objectives

- Understanding basic programming syntax
- Reading data into R
- Summarizing and grouping data
- Filtering data
- Recoding data
- Making plots with your data

Installing R

- Install the [latest R version](#)
- [Install RStudio](#)

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!

Getting files from downloads

This course will involve moving files around on your computer and downloading files.

If you are new to this - check out these videos.

If you have a PC: <https://youtu.be/we6vwB7DsNU>

If you have a Mac: <https://www.youtube.com/watch?v=Ao9e0cDzMrE>

You can find these on the resource page of the website.

Useful (+ mostly Free) Resources

Found on our website under the Resources tab:

https://hutchdatascience.org/SeattleStatSummer_R/resources.html

- videos from our Intro to R Course
- cheatsheets from that course

Useful (+ mostly Free) Resources

Want more?

- Tidyverse Skills for Data Science Book: <https://jhudatascience.org/tidyversecourse/> (more about the tidyverse, some modeling, and machine learning)
- Tidyverse Skills for Data Science Course: <https://www.coursera.org/specializations/tidyverse-data-science-r>
(same content with quizzes, can get certificate with \$)
- R for Data Science: <http://r4ds.had.co.nz/>
(great general information)
- R basics by Rafael A. Irizarry: <https://rafalab.github.io/dsbook/r-basics.html> (great general information)
- Open Case Studies: <https://www.opencasestudies.org/>
(resource for specific public health cases with statistical implementation and interpretation)
- Dataquest: <https://www.dataquest.io/>
(general interactive resource)

Useful (+ mostly Free) Resources

Need help?

- Various “Cheat Sheets”: <https://www.rstudio.com/resources/cheatsheets/>
- R reference card: <http://cran.r-project.org/doc/contrib/Short-refcard.pdf>
- R jargon: <https://link.springer.com/content/pdf/bbm%3A978-1-4419-1318-0%2F1.pdf>
- R vs Stata: <https://link.springer.com/content/pdf/bbm%3A978-1-4419-1318-0%2F1.pdf>
- R terminology: <https://cran.r-project.org/doc/manuals/r-release/R-lang.pdf>

Summary

- R is a powerful data visualization and analysis software language.
- Lots of **resources** can be found on the website.

[Workshop Website](#)