Getting Started with R for Mass Spectrometry Data Analysis

2-Day Short Course • 2021 ASMS Annual Meeting

Ryan Benz
Seer, Inc.

Jeffrey Jones
SoCal Bioinformatics

N. Heath Patterson
Vanderbilt School of Medicine
Learn the fundamentals of R and add new data analysis skills to your scientific tool belt!

Course Overview
This course will cover the fundamentals of the R programming language and how it can be used to perform data analysis tasks, with a focus on the practical essentials to get you started on your way to becoming an effective R user.

Target Audience
Anyone new to R (and perhaps coding) who wants to learn the practical fundamentals for data analysis work.

Prerequisites
- Desire to learn one of the best data analysis tools around
- Willing to try something new and challenging
- No prior coding experience needed
- Bring a laptop for the example exercises and working sessions
The course will focus on 3 main topics

- R Fundamentals
- Data Visualization with ggplot2
- Data Analysis with the tidyverse
Get started with R and RStudio, learn best practices for data analysis with R

- RStudio provides an integrated environment for coding
  - code editor
  - interactive terminal
  - variables, plots, documentation
  - version control
  - and much more

- Streamlines all aspects of the data analysis process

- Learn best practices for organizing analyses into projects and make your work reproducible
Start making plots and visualizations from your own data

- ggplot2 is a powerful plotting package in R that allows you to easily make plots from your data.
- Provides the fundamental pieces to make all kinds of visualizations, from simple to complex.
- Quickly make plots for exploratory data analysis.
- Fine tune the look and style of your plots, make them ready for presentation and publication.

Learn about tidy data and how the tidyverse collection of packages can streamline your analyses

- Learn how tidy data and the tidyverse can make your data analysis code easier to read, write and understand

- Learn about dplyr, an R package that allows you to create streamlined data analysis pipelines

- The tidyverse covers all steps of the analysis process, from data import & cleaning to visualization and final reporting
Additional course info…

• This course is designed to help participants get over the initial hurdles new R users often face

• Mass spectrometry and related data sets / examples will be used during the course to help put the material into context

• Information on any required pre-course set-up (i.e. software to install) will be provided in advance

• If you’ve been thinking about trying to learn R, this course is a great way to get started