R Class Descriptions (Advanced) with Learning Objectives

Doug Joubert

2023-10-27

Table of contents

[Project Management in RStudio 2](#_Toc150841425)

[Description 2](#_Toc150841426)

[Assumptions for This Class (send out to registrants via email) 2](#_Toc150841427)

[Version Control and GitHub 2](#_Toc150841428)

[Description 2](#_Toc150841429)

[Assumptions for This Class (send out to registrants via email) 3](#_Toc150841430)

[Git in RStudio 3](#_Toc150841431)

[Description 3](#_Toc150841432)

[Assumptions for This Class (send out to registrants via email) 3](#_Toc150841433)

[Data Visualization in ggplot 4](#_Toc150841434)

[Description 4](#_Toc150841435)

[Assumptions for This Class (send out to registrants via email) 4](#_Toc150841436)

[Data Visualization in ggplot: Customization 4](#_Toc150841437)

[Description 4](#_Toc150841438)

[Assumptions for This Class 4](#_Toc150841439)

[Introduction to Quarto for Scholarly Publishing 5](#_Toc150841440)

[Description 5](#_Toc150841441)

[Assumptions for This Class 5](#_Toc150841442)

[Quarto for Scholarly Publishing: Advanced Formatting 5](#_Toc150841443)

[Description 5](#_Toc150841444)

[Assumptions for This Class 6](#_Toc150841445)

[Quarto for Scholarly Publishing: Working with Citations 6](#_Toc150841446)

[Description 6](#_Toc150841447)

[Assumptions for This Class (send out to registrants via email) 6](#_Toc150841448)

[Data Visualization in ggplot: Visualizing Relationships and Linear Regression 7](#_Toc150841449)

[Description 7](#_Toc150841450)

[Assumptions for This Class 7](#_Toc150841451)

[Introduction to Bioconductor 7](#_Toc150841452)

[Description 7](#_Toc150841453)

[Assumptions for This Class 7](#_Toc150841454)

[Working with Data in Bioconductor 8](#_Toc150841455)

[Description 8](#_Toc150841456)

[Assumptions for This Class 8](#_Toc150841457)

# Project Management in RStudio

## Description

This class focuses on data and project management using R and RStudio. RStudio makes it possible to work on a complete research project in a more efficient, integrated, and organized manner. This course is designed to be relevant to participants from different disciplines.

Upon completion of this class participants should be able to define project management from a data science perspective, list the advantages of using RStudio projects, apply best practices for setting up RStudio for projects, create a new RStudio Project, and discuss best practices for organizing data in an RStudio project.

## Assumptions for This Class (send out to registrants via email)

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Experience creating scripts and/or markdown files.

# Version Control and GitHub

## Description

This class introduces version control and how to use [GitHub](https://github.com/) for project versioning. Participants will have a better understanding of version control, GitHub, and their advantages for managing projects.

Upon completion of this class participants should be able to recognize why version control is useful, discuss the difference between Git and GitHub, list the options for authenticating to GitHub, and list the options for creating a personal access token (PAT).

## Assumptions for This Class (send out to registrants via email)

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Experience creating scripts and/or markdown files
4. You have already created a GitHub account

# Git in RStudio

## Description

This class focuses on using Git and [GitHub](https://github.com/), with RStudio. Using integrated RStudio tools, participants will have a chance to experiment with this integration and understand its advantages for collaboration and managing projects. You must have taken Version Control and GitHub class to be successful in this class.

Upon completion of this class participants should be able to discuss the difference between Git and GitHub, list the options for authenticating to GitHub, create a new R project using a GitHub repository, and distinguish between pulling and pushing data from a repository.

## Assumptions for This Class (send out to registrants via email)

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Taken the Version Control and GitHub class
4. Created a GitHub account
5. Downloaded, installed, and verified Git on your computer

# Data Visualization in ggplot

## Description

This class provides a basic overview of creating plots using ggplot. ggplot is a part of the Tidyverse, a [collection of R packages](https://www.tidyverse.org/packages) designed for data science. This class will focus on identifying the appropriate packages for plotting, defining plot aesthetics, and demonstrating how to add layers to ggplot graphs.  **You must have taken Introduction to R and RStudio class to be successful in this class**.

By the end of this class, participants should be able to discuss the connection between data, aesthetics, & the grammar of graphics, describe how ggplot works, define geoms, and distinguish between individual geoms and collective geoms.

## Assumptions for This Class (send out to registrants via email)

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Experience creating scripts and/or markdown files.
4. Taken Introduction to R and RStudio class to be successful in this class

# Data Visualization in ggplot: Customization

## Description

This class provides an overview of options for customizing a ggplot graph. This class will focus on methods for creating small multiples, options for customizing a graph, and how to apply ggplot themes. **You must have taken Data Visualization in R: ggplot class to be successful in this class**.

By the end of this class, participants should be able to describe options for time series data, create a line plot in ggplot, learn how to facet a plot, demonstrate options for customizing the title and axis, and apply different ggplot themes.

## Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Experience creating scripts and/or markdown files.
4. Taken Data Visualization in R: ggplot class to be successful in this class

# Introduction to Quarto for Scholarly Publishing

## Description

This class is designed for those who want to extend the basics of R Markdown and apply those skills in [Quarto](https://quarto.org/). Quarto is an open-source scientific and technical publishing system that offers multilingual programming language support to create dynamic and static documents, books, presentations, blogs, and other online resources.

Upon completion of this class participants will be able to distinguish between R-markdown and Quarto, identify publishing workflows using markdown, demonstrate the differences between the visual and source editors, create basic markdown elements, learn how to create and run code-blocks, and render a markdown document in multiple formats.

## Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Experience creating scripts and/or markdown files.

# Quarto for Scholarly Publishing: Advanced Formatting

## Description

This class is designed for those who want to extend the basics covered in the Introduction to Quarto for Scholarly Publishing class. Upon completion of this class participants should be able to create tables, customize code-blocks, create LaTeX equations, and insert images into a markdown document. **You must have taken Introduction to Quarto for Scholarly Publishing to be successful in this class**.

## Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Taken the Introduction to Quarto for Scholarly Publishing class

# Quarto for Scholarly Publishing: Working with Citations

## Description

This class is designed for those who want to extend the basics covered in the Introduction to Quarto for Scholarly Publishing to Formatting class. This class uses [Quarto](https://quarto.org/) to render formatted citations and bibliographies included in a journal article, report, or presentation.  This class also discusses the Zotero API, which is supported in RStudio. [Zotero](https://www.zotero.org/) is a free, easy-to-use tool to help you collect, organize, annotate, cite, and share research. This class also covers exporting citations from Endnote in a format supported by Quarto. EndNote is a software package which is designed to help you to organize citations and create a bibliography. The current version of EndNote [available from the NIH Library](http://www.nihlibrary.nih.gov/resources/subject-guides/endnote/endnote-download) is Endnote 21.

**You must have taken Introduction to Quarto for Scholarly Publishing to be successful in this class**.

Upon completion of this class participants should be able to link RStudio to Zotero, create a bibliography and link it to a markdown document, insert citations using RStudio Visual Interface, and via the command line, and download and link a CSL file which specifies the formatting to use when generating the citations and bibliography.

## Assumptions for This Class (send out to registrants via email)

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Taken the Introduction to Quarto for Scholarly Publishing class

# Data Visualization in ggplot: Visualizing Relationships and Linear Regression

## Description

This class provides an overview of the methods used to visualize the association among two or more quantitative variables. This class will focus on scatterplots, scatterplot matrix, and visualizing paired data. Upon completion of this class, participants should be able to define bivariate data, create a scatterplot using ggplot, define linear regression, and demonstrate how to perform a simple linear regression in R. **You must have taken Data Visualization in R: ggplot class to be successful in this class**.

## Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Experience creating scripts and/or markdown files
4. Taken the Visualization in R: ggplot class

# Introduction to Bioconductor

## Description

This class will provide an overview of Bioconductor. We will learn how to identify Bioconductor packages that are appropriate for a project, explore package documentation, and demonstrate how to download and install R packages from Bioconductor.

Upon completion of this class participants should be able to describe what the Bioconductor project comprises, navigate the Bioconductor website to find packages for a particular task, install and update Bioconductor package, open a package vignette and practice running through the examples that they contain, and ensure that they are using the correct version of R to work with packages.

## Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Experience creating scripts and/or markdown files.

# Working with Data in Bioconductor

## Description

This class will provide an overview of common Bioconductor datatypes and explore options for working with biological sequence data.  Specifically, this class will focus on the object types for storing and manipulating genomic features and sequences.

Upon completion of this class participants should be able to locate resources on S4Vector classes, understand the standard R datatypes, list the 6 basic Bioconductor classes, and discuss methods for working with biological sequences.

## Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. Specifically, you have:

1. Installed [R](https://cran.r-project.org/) and [RStudio](https://www.rstudio.com/products/RStudio/#Desktop)
2. Experience with R and RStudio. If not, here are some resources for getting started:
   1. [Introduction to R and RStudio](https://carpentries-incubator.github.io/R-ecology-lesson-alternative/introduction-r-rstudio.html)
   2. [A Few Notes to Get Started with R](https://rcompanion.org/rcompanion/a_06.html)
   3. [Getting started with R and RStudio](http://rafalab.dfci.harvard.edu/dsbook/getting-started.html)
3. Experience creating scripts and/or markdown files.