R Class Descriptions (Advanced) with Learning Objectives

Doug Joubert

7/10/23

Table of contents

Data Visualization in R: Customization in ggplot

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. Students will need to install R, RStudio, and the tidyverse package, before the class so that they can follow along with the instructor.

Learning Objectives

By the end of this class, students should be able to:

1. Describe options for time series data
2. Create a line plot in ggplot
3. Learn how to facet a plot
4. Demonstrate options for customizing the title and axis
5. Apply different ggplot themes
6. Arrange multiple plots (handout)

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. First, it assumed that you have already installed R and RStudio. Second, is that you have experience with R. If not, here are some resources for getting started:

- NIH Library Introduction to R and RStudio

- A (very) short introduction to R

- Quick-R

- Getting up to speed with R

Third, is experience working in RStudio and creating scripts and/or markdown files.

Data Visualization in R: 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 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. Students are encouraged to install R, RStudio, and the tidyverse package, before the class so that they can follow along with the instructor.

Learning Objectives

By the end of this class, students should be able to:

1. Discuss the connection between data, aesthetics, & the grammar of graphics
2. Describe how ggplot works
3. Define geoms and distinguish between individual geoms and collective geoms

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. First, it assumed that you have already installed R and RStudio. Second, is that you have experience with R. If not, here are some resources for getting started:

- NIH Library Introduction to R and RStudio

- A (very) short introduction to R

- Quick-R

- Getting up to speed with R

Third, is experience working in RStudio and creating scripts and/or markdown files.

Data Visualization in ggplot: Visualizing Relationships and Linear Regression

Description

This class provides a basic 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. Participants are expected to have taken the Introduction to Data Visualization in R: ggplot class. Participants must also install R, RStudio, and the tidyverse package, before the class so that they can follow along with the instructor.

Learning Objectives

Upon completion of this class, students should be able to:

- Define bivariate data

- Define scatterplot

- Create a scatterplot using ggplot

- Define linear regression

- Demonstrate how to perform a simple linear regression in R

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. First, it assumed that you have already installed R and RStudio. Second, is that you have experience with R. If not, here are some resources for getting started:

- NIH Library Introduction to R and RStudio class

- A (very) short introduction to R

- Quick-R

- Getting up to speed with R

Third, is experience working in RStudio and creating scripts and/or markdown files.

Version Control and GitHub (New)

Description

This class provides an introduction to version control and using GitHub for project versioning. Students will have a better understanding of version control and its advantages for collaboration and version control.

Learning Objectives

Upon completion of this class students should be able to:

- Recognize why version control is useful
- Discuss the difference between Git and GitHub
- List the options for authenticating to GitHub
- Create a personal access token (PAT)

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio:

1. You have already installed R and RStudio.
2. You have experience with R. If not, here are some resources for getting started:
- NIH Library Introduction to R and RStudio class
- A (very) short introduction to R
- Quick-R
- Getting up to speed with R
3. You have already created a GitHub account

Git in RStudio (Modified)

Description

This class focuses on using Git and GitHub, with RStudio. Using integrated RStudio tools learners will have a chance to experiment with this integration and understand its advantages for collaboration and version control. Students need to install R and the current version of RStudio before the class so that they can follow along with the instructor.

Learning Objectives

Upon completion of this class students should be able to:

- Discuss the difference between Git and GitHub
- List the options for authenticating to GitHub
- Create a personal access token (PAT)
- Create a new R project using a GitHub repository

- Distinguish between pulling and pushing data from a repository

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio:

1. You have already installed R and RStudio.
2. You have experience with R. If not, here are some resources for getting started:
- NIH Library Introduction to R and RStudio class
- A (very) short introduction to R
- Quick-R
- Getting up to speed with R
3. You have already created a GitHub account
4. You have downloaded, installed, and verified Git on your computer

Data Visualization in ggplot: Visualizing Relationships and Linear Regression

Description

This class provides a basic 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. Participants are expected to have taken the Introduction to Data Visualization in R: ggplot class. Participants must also install R, RStudio, and the tidyverse package, before the class so that they can follow along with the instructor.

Learning Objectives

Upon completion of this class, students should be able to:

- Define bivariate data

- Define scatterplot

- Create a scatterplot using ggplot

- Define linear regression

- Demonstrate how to perform a simple linear regression in R

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. First, it assumed that you have already installed R and RStudio. Second, is that you have experience with R. If not, here are some resources for getting started:

- NIH Library Introduction to R and RStudio class

- A (very) short introduction to R

- Quick-R

- Getting up to speed with R

Introduction to Bioconductor

Description

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

Learning Objectives

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

- Ensure that they are using the correct version of R to reproduce exactly the contents of this lesson

- Describe how to join and communicate with the Bioconductor community

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. First, it assumed that you have already installed R and RStudio. Second, is that you have experience with R. If not, here are some resources for getting started:

- NIH Library Introduction to R and RStudio class

- A (very) short introduction to R

- Quick-R

- Getting up to speed with R

Third, is experience working in RStudio and creating scripts and/or markdown files.

Reproducibility in RStudio: Basic Markdown Manual

Description

This class is designed for those who want to extend the basics of R Markdown and apply those skills in Quarto. 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.

In this intermediate-level class you will learn about the similarities and differences between R-markdown and Quarto. You will also learn how to use Quarto to render documents in multiple formats, with a focus on scholarly publishing. Some familiarity or experience in R and RStudio is recommended but not required. Students are encouraged to install the most current versions of R and RStudio before the class so that they can follow along with the instructor.

Learning Objectives

Upon completion of this class students should be able to:

- Distinguish between R-markdown and Quarto
- Identify publishing workflows using markdown
- Demonstrated the difference between the visual and source editors
- Create basic markdown elements
- Learn how to create and run code-blocks
- Insert images into a markdown document and modify attributes
- Render a markdown document

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. First, it assumed that you have already installed R and RStudio. Second, is that you have experience with R. If not, here are some resources for getting started:

- NIH Library Introduction to R and RStudio

- A (very) short introduction to R

- Quick-R

- Getting up to speed with R

Third, is experience working in RStudio and creating scripts and/or markdown files.

Project Management in RStudio Manual

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 intermediate-level course is designed to be relevant to students from different disciplines.

Learning Objectives

Upon completion of this class students 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
- Discuss best practices for organizing data in an RStudio project

Reproducibility in RStudio: Advanced Markdown

Description

This class is designed for those who want to extend the basics covered in the Reproducibility in RStudio: Basic Markdown class. This class uses Quarto to render format citations and bibliographies using Zotero. Zotero is a free, easy-to-use tool to help you collect, organize, annotate, cite, and share research. Some familiarity or experience in R and RStudio is recommended but not required. Students are encouraged to install R and RStudio before the class so that they can follow along with the instructor.

Learning Objectives

Upon completion of this class students should be able to:

- Download and install Zotero
- 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
- Download and link a CSL file which specifies the formatting to use when generating the citations and bibliography

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. First, it assumed that you have already installed R and RStudio. Second, is that you have experience with R. If not, here are some resources for getting started:

- NIH Library Introduction to R and RStudio

- A (very) short introduction to R

- Quick-R

- Getting up to speed with R

Third, is experience working in RStudio and 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.

Learning Objectives

Upon completion of this class participants should be able to:

- Locate resources on S4Vector classes

- Review standard R datatypes

- List the 6 basic Bioconductor classes

- Discuss methods for working with biological sequences

- Locate information on Bioconductor, Bioconductor packages, and Bioconductor Courses

Assumptions for This Class

This class makes a few assumptions about your understanding of R and RStudio. First, it assumed that you have already installed R and RStudio. Second, is that you have experience with R. If not, here are some resources for getting started:

· NIH Library Introduction to R and RStudio class

· A (very) short introduction to R

· Quick-R

· Getting up to speed with R

Third, is experience working in RStudio and creating scripts and/or markdown files.