An Introduction to BIOF 439

Abhijit Dasgupta, PhD

Objectives of this course

- Understand principles of good data visualization
- Know what might make a visualization poor or ineffective
- Get you going using R **for visualization**
    - Various packages
- Creating static and dynamic visualizations using R
- Using the web as a presentation medium

Course resources

Website: Canvas

Slack: https://biof439.slack.com

You will be sent an invitation link to the Slack channel on Canvas. Please sign in, since that will be the primary means of communication between us, and between your fellow students.

General class practices

- Each week's module will be released by 10am each Monday.
- Each week will include
  - Lectures/Screencasts
  - Readings
  - Assigned tutorials that you will do on your own
  - Formative assessments
  - One quiz
  - Homework

Quizzes

- Each week will have a quiz to assess your understanding of that week's material
- Quizzes will typically be multiple choice
- Quizzes will typically be time-limited

Homework

Rules

- Homework assignments are due each Saturday at 11:59pm
- Homework may be submitted up to 24 hours after the due date, i.e. by Sunday 11:59 pm, with a 20% penalty
- Homework may not be submitted after Sunday 11:59 pm since solutions will be provided on Monday

Homework

Format

- Unless otherwise specified, homework must be submitted as a RMarkdown file and the corresponding HTML file.
    - **Both files are required**
    - Homework for Week 1 will not require RMarkdown.
- Homework files will be submitted on Canvas using the Canvas assignment submission system
    - See screencast for how this works

Collaboration

You are encouraged to collaborate through Slack, especially since there will be a fair amount of variability of R expertise in this class. However, your submitted work should be your own.

Final Project

- A R Markdown dashboard using the `flexdashboard` package. Submit both RMarkdown and HTML files
- Use your own data
- Use R package(s) to visualize your data sets in at least 3 ways, to show what your data looks like and what your analytic results look like.
- Each student will be randomly assigned to 2 peers
    - Critiques based on quality and effectiveness of visualizations

- All final projects will be shared through a dedicated website, so we can learn from each other (unless you let me know first, for example, if the data is private and embargoed)
- If you use Shiny as part of your dashboard, please create an free account on shinyapps.io, deploy your dashboard, and provide a link as part of your submission
- I fully expect some of you to blow me away!!

Class participation

- Ask questions
- Comment on the strengths and weaknesses of visualizations when we work on them each week
- I will start a discussion topic every Wednesday on Canvas. Participate in the discussion, which will remain open until Saturday
- Discuss topics on the Slack channel

Multiple sections

We may have multiple sections for a session depending on demand

- We will all share the same Slack environment
- Each section will have a different Canvas environment
- All the teaching material will be identical in each section
- I will reserve the right to make quizzes, formative assessments, assignments and discussions different, though I don't expect to do that often.

Grading

Grade components

Component	Percentage
Quizzes	20%
Assignments	30%
Participation	25%
Final Project	25%

Contact info

Email: adasgupta+biof439@araastat.com (don't use my NED email)

Slack

Software requirements

For this class you will need R and RStudio.

1. You can install R on your machine from https://cran.r-project.org. There are versions for Windows, Mac OS X and Linux. On windows machines, R can be installed without administrative privileges.

2. You can install RStudio from RStudio. If you have a Windows machine and do not have administrative privileges, you should download the version from the Zip/Tarballs section, open the zip file and just run RStudio.exe from there.

3. You can access R and RStudio online without any installations using RStudio Cloud. You can get an account for free.

    - RStudio Cloud is excellent as a teaching tool, but is not usually sufficient for most real analyses

Software requirements

R is a modular software system with **packages** providing various functionality. This class will use several packages the provide visualization and reporting functionalities.

On the website, there will be a file *install.R*. This will have the R script to install these packages. This is described in a screencast.

Thr following are a selection of the packages we will use in this class.

1. `tidyverse` (a meta-package containing 8 other packages)
2. `flexdashboard` (the main reporting package)
3. `RColorBrewer` and `viridis` (color palettes)
4. `shiny` for some interactivity
5. `learnr` for interactive tutorials
6. `htmlwidgets`, `plotly`,`DataTables`, `highcharter` for dynamic plotting and reporting

Other potentially useful software

R can provide wonderful visualizations, but these may require a tweak or two to make them the way you want for publication or dissemination. Each item provides a commercially available software and an open-source free alternative.

1. Adobe Illustrator / Inkscape (all OS)
2. Adobe Photoshop / GIMP (all OS)
3. Adobe Reader Pro / Preview (Mac OSX)
4. ImageMagick (accessible via the R package `magick`) (all OS)