

Workshop abstracts

Workshop 1 – Data Visualization

Dr. Xijin Ge – South Dakota State University

This hands-on workshop assumes that the attendees have basic knowledge of R. The goal is to introduce beginners and intermediate level R users to some packages that can easily generate high-quality, interactive, or web-based graphics. Please make sure you have a recent version of R and Rstudio installed, and also install these packages below (plus devtools and tidyverse) ahead of time.

Main topics:

1. Some basic principles of effective data visualization
2. ggplot2, a quick introduction.
3. Simplifying ggplot2 with ggfortify
4. Interactive plots with plot_ly and ggplotly
5. Streamline EDA with DataExplorer and GGally
6. Heat maps with complexHeatmap visualizes data matrices (Install via www.bioconductor.org)
7. VennDiagrams and UpSetR visualizes overlapping sets
8. Introducing animations to your plots with ganimate *
9. Shiny and Dashboards*
10. Interactive network visualization with visNetwork*
11. Other tricks and cools plots I learned from my students*

* Time permitting.

Workshkop 2 – Web scraping

Peter Claussen - Gylling Data Management, Inc and SDSU

Open source tools for web scrapping

Web scraping or web mining involves interacting with distributed files and information systems through abstract interfaces, where the analyst has little direct control over the computer hardware or services. Programming practices that support web scraping include

- Language-independent file transfer protocols (i.e. HTTP)
- Self-documenting document structuring languages (HTML, XML, JSON)
- Abstract programming interfaces (API) through which data providers allow systematic queries to data repositories
- Text mining via pattern matching (regular expressions)

This workshop will cover open-source tools available to assist with these practices, with an emphasis on libraries that can be interface via either Python or R

Workshop 3 – R library creation fundamentals

Md Riaz Khan

The goal of the workshop is to learn about the fundamentals of creating an R package using R Markdown. Main topics include:

- R package- what and how
- Preparing the system for package building
- Package creating and package metadata
- Writing functions
- Package documentation using roxygen2
- User defined class and S3 methods
- Adding example dataset to the package
- Vignette using R Markdown
- Package testing and release

Prerequisite

- Basic knowledge of R programming
- Most recent versions of R and R Markdown installed