

2018 USEPA R User Group Workshop Agenda

On-line participation details:

There will be a remote participation option for the presentations and workshops. Details will be posted here closer to workshop date.

On-site details:

The workshop is being held at the US EPA campus in Research Triangle Park, NC.

Workshop summaries:

- **Introduction to R (Full Day):** Participants in this workshop will be introduced to R through the lense of the “Tidyverse” as well as learn basics of the R Language for statistical computing. In particular we will: 1) learn the fundamentals of R including syntax, functions, and packages; 2) spend quality time with the core R data structure, the data frame; 3) learn how to manipulate our data with the `dplyr` and `tidyr` packages; and 4) gain experience creating data visualizations with `ggplot2`. Our goal is to set learners up with a foundation for future exploration of R.
- **Working with Spatial Data in R (Half Day):** This workshop will focus on the fundamentals of working with spatial data in R – reading and writing spatial data, dealing with map projections, packages for handling both vector and raster data, and visualizing and mapping spatial data in R. After covering fundamentals, we’ll move through example exercise demonstrating spatial operations such as spatial joins, spatial subsetting, spatial aggregation, cropping raster data, and extracting raster data for points and we’ll make use of packages such as `sf`, `sp`, `raster`, `rgdal`, `maptools`, and several others. A basic working knowledge of R will be assumed.
- **Using LaTeX and Markdown for Reproducible Research (Half Day):** Taking a modular approach and a stepwise fashion, participants will be introduced to the basics of LaTeX, Markdown, and the R package `knitr`. Participants will then get to use these modules synchronously to facilitate reproducible reports and research with R.
- **Network Analysis/Graph Analysis techniques (Half Day):** Networks and graphs are a flexible framework for organizing complex interactions between large number of entities. In this workshop, using the `igraph` package in R, we will start from scratch by building a small network. We will then import a larger more complicated network and go through the process of encoding node and edge attributes. We will discuss graph layout options and complete some standard analyses. In the end, participants will gain a familiarity with the how to build, import, visualize and analyze networks in R.
- **An Introduction to Shiny, R Markdown & htmlwidgets with Applications in Research (Half Day):** RStudio will be presenting an overview of creating interactive visualizations, dashboards and Shiny Apps. This short course will provide an introduction to flexible and powerful tools for statistical analysis, reproducible research and interactive visualizations. The hands-on course will include an overview of how to build Shiny apps and R Markdown documents.
- **Version Control, GitHub, and Bitbucket Workshop (Half Day):** The basics of using git and GitHub for version control and reproducible research at EPA, with an introduction to EPA’s new Bitbucket services. An overview presentation will be followed by a hands on session.

Mon, Sep. 10

Webinar Details

| Time | Room | Title | Speaker |
|-----------------------|------|--------------|---------|
| 11:00 AM - 1:00 PM | | Registration | |

| Time | Room | Title | Speaker |
|---------------------------|---|---|---|
| 1:00 PM - 1:10 PM | C111 A,B,C | Introduction | Jeff Hollister, Ann Vega and Richard Allen, US EPA, Office of Research and Development |
| 1:10 PM - 1:30 PM | C111 A,B,C | Welcome | Vaughn Noga, Principal Deputy Assistant Administrator and Deputy Chief Information Officer, US EPA, OEI |
| Keynote Address | | | |
| 1:30 PM - 2:30 PM | C111 A,B,C | Data Science for Public Goods | Noam Ross, EcoHealth Alliance |
| 2:30 PM - 3:00 PM | | Break | |
| Talks | | | |
| 3:00 PM - 3:20 PM | C111 A,B,C | Packaging your research problem: the case of CityWaterBalance for R | Laura Erban, US EPA, ORD/NHEERL |
| 3:25 PM - 3:45 PM | C111 A,B,C | Package Management for Reproducible R Code | Cole Arendt, RStudio |
| 3:50 PM - 4:10 PM | C111 A,B,C | Using R and MCMC for Amphibian Exposure Model Selection and Parameterization | Tom Purucker, US EPA, ORD/NERL |
| 4:15 PM - 4:35 PM | C111 A,B,C | Evaluation of a numerical air quality model: Data wrangling large NetCDF files with R | Kristen Foley, US EPA, ORD/NERL |
| 4:40 PM - 5:00 PM | C111 A,B,C | Using R and Machine Learning to Predict Inspection Outcomes | Stephanie Wilson, US EPA, Region 2 |
| 5:05 PM - 5:25 PM | C111 A,B,C | Ambient Air Monitoring Network Assessment Tools using R shiny | Benjamin Wells, US EPA, OAQPS |
| Evening Activities | | | |
| 5:30 PM - 6:30 PM | Hilton Garden Inn Durham South- point, 7007 Fayetteville Road, Durham, NC | Informal Social | All! |

Tues, Sep. 11

| Time | Room | Title | Speaker |
|-----------------------------|-------|--|--|
| Morning Work-shops | | | |
| 8:00 AM - 11:30 AM | C111A | Introduction to R - Part 1 | Jeff Hollister, US EPA, ORD/NHEERL, Clinton McCrowey, US EPA, Region 3 and Noam Ross, EcoHealth Alliance |
| 8:00 AM - 11:30 AM | C112 | Working with Spatial Data in R | Marc Weber, US EPA, ORD/NHEERL |
| 8:00 AM - 11:30 AM | C111C | Using LaTeX and Markdown for Reproducible Research | Eric Beck, US EPA, Region 1 and Emily Li, US EPA, ORD/NRMRL |
| Lunch | | | |
| 11:30 AM - 1:30 PM | | On your own | |
| Afternoon Work-shops | | | |
| 1:30 PM - 5:00 PM | C111A | Introduction to R - Part 2 | Jeff Hollister, US EPA, ORD/NHEERL, Clinton McCrowey, US EPA, Region 3 and Noam Ross, EcoHealth Alliance |
| 1:30 PM - 5:00 PM | C114 | Network Analysis/Graph Analysis techniques | Nate Pollesch, US EPA, ORD/NHEERL |
| 1:30 PM - 5:00 PM | C111C | An Introduction to Shiny, R Markdown & htmlwidgets with Applications in Research | Phil Bowsher, RStudio |

Wed, Sep. 12

| Time | Room | Title | Speaker |
|--------------------|-------|---|---|
| 8:00 AM - 8:15 AM | | Welcome Day 2 and Welcome to DDES Folks | |
| Workshop | | | |
| 8:15 AM - 11:50 AM | C111 | Version Control, GitHub, and Bitbucket | Jameel Alsalam, Terry Brown, Sharon Kenny |
| 11:50 AM - Noon | A,B,C | Workshop Wrap-up | Ann Vega |

Webinar Links

| Date | Time | Title | Link | Call-in |
|---------------|--------------------|----------------------------|--------------------|--|
| Mon, Sep 10 | 1:00 PM - 5:25 PM | Keynote and Presentations | Skype Webinar Link | Call-in number: 1-202-991-0477, Access code/Conference ID: 8603858 |
| Tues, Sep. 11 | 8:00 AM - 11:30 AM | Introduction to R - Part 1 | Skype Webinar Link | Call-in number: 1-202-991-0477, Access code/Conference ID: 7404609 |

| Date | Time | Title | Link | Call-in |
|---------------|--------------------|--|--------------------|--|
| Tues, Sep. 11 | 8:00 AM - 11:30 AM | Working with Spatial Data in R | Skype Webinar Link | Call-in number: 1-202-991-0477, Access code/Conference ID: 3802797 |
| Tues, Sep. 11 | 8:00 AM - 11:30 AM | Using LaTeX and Markdown for Reproducible Research | Skype Webinar Link | Call-in number: 1-202-991-0477, Access code/Conference ID: 8603858 |
| Tues, Sep. 11 | 1:30 PM - 5:00 PM | Introduction to R - Part 2 | Skype Webinar Link | Call-in number: 1-202-991-0477, Access code/Conference ID: 7404609 |
| Tues, Sep. 11 | 1:30 PM - 5:00 PM | Network Analysis/Graph Analysis techniques | Skype Webinar Link | Call-in number: 1-202-991-0477, Access code/Conference ID: 3802797 |
| Tues, Sep. 11 | 1:30 PM - 5:00 PM | An Introduction to Shiny, R Markdown & htmlwidgets with Applications in Research | Skype Webinar Link | Call-in number: 1-202-991-0477, Access code/Conference ID: 8603858 |
| Wed, Sep. 12 | 8:15 AM - 11:50 AM | Version Control, GitHub, and BitBucket Workshop | Skype Webinar Link | Call-in number: 1-202-991-0477, Access code/Conference ID: 8603858 |