INTRODUCTION TO THE R PROGRAMMING LANGUAGE &
STATISTICAL METHODS FOR ENVIRONMENTAL DATA ANALYSIS
March 09-13, 2020, Brasilia, Brazil
Course outline, subject to revision prior to start of course

MONDAY

8:30 am - 10:00 am – Introduction - navigating RStudio, packages, functions, help

10:00 am - 10:30 am – ***Break***

10:30 am - 12:00 pm – Get - basic operators, classes/data structures, read data

12:00 pm - 2:00 pm – ***Lunch***

2:00 pm - 3:30 pm – Clean - subsetting, dplyr/tidyverse functions

3:30 pm - 4:00 pm – ***Break***

4:00 pm - 5:30 pm – Clean, continued - pipes and joining data

5:30 pm - 6:00 pm – Free time - revisit previous sections, ask questions, work with our class examples or individual data.

TUESDAY

8:30 am - 10:30 am – Explore - summary statistics, dplyr functions, basic plots

10:30 am - 11:00 am – ***Break***

11:00 am - 12:00 pm – Analyze
Introduction to linear models, correlation, plotting models (preview of Wednesday)

12:00 pm - 2:00 pm – ***Lunch***

2:00 pm - 3:30 pm – Graphs with ggplot2

3:30 pm - 4:00 pm – ***Break***

4:00 pm - 5:30 pm – Repeat - loops, if-else, functions, R Markdown

5:30 pm - 6:00 pm – Free time - revisit previous sections, ask questions, work with our class examples or individual data.

WEDNESDAY

8:30 am - 9:30 am - Activity - Continuous water-quality monitoring problem (Part 1) - Exploratory data analysis and plotting

9:30 am - 10:30 am – Lecture - Simple Linear Regression
- Definitions, model fitting
- Hypothesis tests
- Assumptions, residual plots

10:30 am - 11:00 am – ***Break***

11:00 am - 12:00 pm – Lecture - Simple Linear Regression
- Transformation of explanatory and response variables
- What is a good regression model?

12:00 pm - 2:00 pm – ***Lunch***

2:00 pm - 2:15 pm – Activity - Continuous water-quality monitoring problem (Part 2a) - Simple linear regression

2:15 pm - 3:15 pm – Lecture - Multiple Regression
- Measures of a good model
- Linear vs. nonlinear model
- Model diagnostics

3:15 pm - 3:30 pm – Activity: Continuous water-quality monitoring problem (Part 2b) - Multiple linear regression

3:30 pm - 4:00 pm – ***Break***

4:15 pm - 4:30 pm – Activity: Continuous water-quality monitoring problem (Part 2b) - Multiple linear regression

4:30 pm - 5:00 pm – Lecture: Multiple Regression
- Multi-collinearity

Model selection (all subsets)

5:15 pm - 5:30 pm – Activity: Continuous water-quality monitoring problem (Part 3) - Model diagnostics and comparison

5:30 pm - 5:45 pm – Lecture: Multiple Regression
Analysis of Covariance (indicator regression)

5:45 pm - 6:00 pm – Activity: Continuous water-quality monitoring problem (Part 4) - Indicator regression

THURSDAY

8:30 am - 10:30 am – R Package Overview
What is a package?
Basic Package Mechanics
Documentation
Maintenance

10:30 am - 11:00 am – ***Break***

11:00 am - 12:00 pm – Data Retrieval US - What is in this package?

12:00 pm - 2:00 pm – ***Lunch***

2:00 pm - 3:30 pm – Introduction to GitHub

3:30 pm - 4:00 pm – ***Break***

4:00 - 6:00 pm – Setting up on GitHub

FRIDAY

8:30 am - 10:30 am – Data Retrieval Brazil
What would this look like?
How will we collaborate throughout the year?

10:30 am - 11:00 am – ***Break***

11:00 am - 12:00 pm – Concluding discussion
Feedback & Future Planning

12:00 pm - 2:00 pm – ***Lunch***

2:00 pm - USGS Departure (likely to visit Embassy)