Beginner R Training: Data Science for Workforce and Economic Development Research

February 8, 2022 – February 24, 2022
Online – Tuesdays + Thursdays, 4:30pm – 6:30pm ET

Week 1

Session 1: 2/8

RStudio Setup Essentials & Workflow Intro

- Learn basics of R, R Packages, and RStudio IDE
- Get set up with project
- Introduction and initial walkthrough of workflow
- Basic introduction to data classes in r (vectors and data frames)

Key Takeaway: Audience will ensure they can open RStudio and are able to install necessary packages. Learn main components of R and RStudio: RMarkdown, R script, console, environment, directory, history, and workflow fundamentals

Office Hours: 2/9; 12 – 1pmET

Session 2: 2/10

Introduction to "tidyverse" (https://www.tidyverse.org/)

- Continue with basic introduction to data classes in R (vectors and data frames)
- Importing and exporting data with R
- Gain familiarity with tidyverse packages, syntax, and practices
- Intro to data munging and joining with dplyr

Key Takeaway: Start to use key functions for data tidying and munging with examples and hands on tutorials to start coding on your own

Week 2

Session 3: 2/15

"Tidyverse" deep dive

- Pick back up with dplyr and data munging
- Reshaping data with tidyr
- Working with strings (stringr)

Key Takeaway: Learn how to do modern day data wrangling and transformations. These are skills necessary to work with your data and prepare it for visualizations and modeling.

Office Hours: 2/16; 1 – 2pmET

Session 4: 2/17

Introduction to data visualization with ggplot2 package

- Aesthetic styling and theme elements
- Key graphs for model diagnostics and exploratory analysis

Key Takeaways: Generate most commonly used plots and understand the basics of the key R visualization package ggplot2.

Session 5: 2/22

Data visualization with ggplot2 package continued & Case Study Intro

- Go further with ggplot2 - exploring different plot types
- Adjust axes, aesthetics, and theme with ease
- Introduce Case Study for Final day

Key Takeaways: Make your graphs look more appealing and lay the foundation for important model diagnostic visuals. Start to get hands on with final project

---

Office Hours: 2/23; 2 – 3pmET

Session 6: 2/24

Example Case study - Real world applicable dataset

- Follow up on useful packages for BLS and census dataset.
- Utilize all the skills learned throughout the course on real publicly available data.

Key takeaway: Practice makes perfect. The best way to improve one's data science skills is to code and work with real data. Here we will give a glimpse on what that type of workflow would be like.
