R for Beginners: Data Science for Workforce and Economic Development Research

October 20 - 29, 2020

Online – T, R – 2:30pm – 5:30pm ET

Week 1

Session

Setup on RStudio Cloud

Key Takeaway: Audience will ensure they can login to RStudio cloud using login established prior to the session. After login, instructors will provide link to an R session with all workshop materials.

Introduction to RStudio setup and the workflow

Key Takeaway: Learn main components of R and RStudio: R script, console, environment, directory, history, and workflow fundamentals

Introduction to "tidyverse" (https://www.tidyverse.org/) & Data Science Workflow

- Data import for various file types
- Basic introduction to data classes in r (vectors and data frames)
- Introduction to tidy datasets – is data ready to analyze?
- Data tidying

Key Takeaway: Learn the key steps typically required for proper analysis and visualization and how R and the tidyverse packages help accomplish these.

Introduction to "tidyverse" (continued)

- Data wrangling
  - o Filtering/querying data
  - o Selecting columns
- Data export to various file types (.xls, .cvs, .txt)

Key Takeaway: Learn key functions for wrangling data to desired format and save your changes in various file formats.

Introduction to data visualization with ggplot2 package

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

Week 2

Session

Continue to data visualization from day 1 and take a deep dive

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

Key Takeaway: Make your graphs look more appealing and lay the foundation for important model diagnostic visuals.

Exporting graphics

Key Takeaway: Save your graphs and share it with others.

Interactive visualizations

Key Takeaway: Learn how to make your ggplots interactive.

Introduction to modeling

- Univariate and Multivariate Linear models

Key Takeaway: Soft introduction to foundational modeling techniques. Modeling helps to explore relevant relationships between the variables of the dataset and determine which variables matter most for our outcome of interest.

Modeling and Prediction

Key Takeaway: After variables are selected, learn the final steps of building a predictive model and evaluating accuracy

---

Basics of reporting using RMarkdown

Key Takeaway: Learn how to make repeatable reports and share your R work with others

---

Basics of reporting using RMarkdown (continued)

Wrap-up and notes on RStudio local use

Key Takeaway: Learn the basics of using RStudio locally (vs. in the cloud environment).