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

# June 15 – 26, 2020

# TENTATIVE AGENDA

|  |  |
| --- | --- |
| **Day** | **Session** |
| **June 15** | **Session 1, Instructor: Ezgi Karaesmen, TA: Abbas Rizvi** |
|  | 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 a link to a R project with all workshop materials. |
| Introduction to RStudio setup and the workflow**Key Takeaway:** Learn main components of R and RStudio: RMarkdown, R script, console, environment, directory, history, and workflow fundamentals |
| Introduction to “tidyverse” ( <https://www.tidyverse.org/>) & Data Science Workflow* Basic introduction to data classes in r (vectors and data frames)
 |
| **June 17** | **Session 2, Instructor: Abbas Rizvi, TA: Ezgi Karaesmen** |
|  | Introduction to “tidyverse” (continued) * Data import (readr)
* Data wrangling (dplyr)
	+ Filtering/querying data
	+ 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 |
| **June 19** | **Session 3, Instructor: Abbas Rizvi, TA: Ezgi Karaesmen** |
|  | Data Wrangling Part 2* Tidying data and pivot functions (tidyr)
* Working with dates (lubridate)
* Working with strings (stringr)
* Joining and merging (dplyr)
* Iterations (purrr)
* Writing R functions

**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.  |
| **June 22** | **Session 4, Instructor: Ezgi Karaesmen, TA: Abbas Rizvi** |
| 2:45 | 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. Make your graphs look more appealing and lay the foundation for important model diagnostic visuals.  |
| **June 24** | **Session 5, Instructor Ezgi Karaesmen, TA: Abbas Rizvi** |
|  | Introduction to modeling and prediction with tidymodels* 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.After variables are selected, learn the final steps of building a predictive model and evaluating accuracy |
| **June 26** | **Session 6, Instructor Abbas Rizvi, TA: Ezgi Karaesmen**  |
|  | Case study - Real world applicable dataset* Break out into 2 to 4 person teams.
* Utilize all the skills learned throughout the course on real publicly available data.
* Prepare data for analysis, create visualizations, and conduct basic modeling. All reported in RMarkdown.
* Present your analysis for 5 minutes with your group and discuss what your analysis and results.

**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.  |
|  | **Adjourn -- course complete** |