



## Fundamentals of Data Science and Visualization

**Virtual Training**  
**November 9 – 19, 2020**

Classes: Nov 9, 10, 16, 17, 19 from 2:00-4:00 PM (EST)  
Office Hours: Nov 9, 10, 16, 17, 19 from 4:00-5:00 PM (EST)

A picture is worth a thousand words. Use data to state your case using easy-to-understand data visualization tools. Give your audience the freedom to adapt your data in new ways in interactive dashboards that answer immediate questions and uncovers new insights. Data visualization tools can help you communicate better both internally and with your partners.

With a deeper understanding of data visualization software packages, your organization can produce more intuitive data visualizations in less time and identify the best software solutions to optimize your team's workflows. In this course, we will review best practices in data visualization design and use cases for Excel, Tableau, and R (programming language).

Learn how to clean and format data in Excel, create interactive dashboards in Tableau, and clean and visualize data in R. This course will help participants identify use-cases for each software package that maximize impact with minimal effort, expanding participants' toolbox as an analyst. Join us to learn about how your organization can better leverage data visualization software!

### **Meet Your Instructor:**

Lee Winkler joined the Center for Regional Economic Competitiveness (CREC) in 2018 after graduating with a Master's in Public Policy from the George Washington University. He currently supports projects analyzing state-level certification and license attainment and the prevalence of educational and workforce credentials. Lee regularly uses Tableau to clean data, mine insights and create interactive visualizations and is excited to help the class find how Tableau can add value to their workflow.



## Training Prerequisites:

None! This course is designed as an introduction these software packages and data visualization methods. All participants need is to be ready to learn, ask questions, and overcome challenges.

**Tools:** Laptop, wired/wireless mouse, Microsoft Excel, Tableau Desktop (personal, professional, or public version), and R Studio Cloud. Please install all applications and do any set-up before class. We will provide instructions for setting up R Studio Cloud before the first session.

- Microsoft Excel Online (free) is available at: [https://www.microsoft.com/en-us/microsoft-365/buy/compare-all-microsoft-365-products?tab=1&irgwc=1&OCID=AID2000142\\_aff\\_7593\\_10078&tduid=\(ir\\_ynbhesti90kftxvckk0sohznxf2xi2a36hwec3qn00\)\(7593\)\(10078\)\(techadvisor.co.uk\)\(803X112721X967a3d93a3d5d08fdcf4c1f4e036e21d\)&irclickid=ynbhesti90kftxvckk0sohznxf2xi2a36hwec3qn00&rtc=1](https://www.microsoft.com/en-us/microsoft-365/buy/compare-all-microsoft-365-products?tab=1&irgwc=1&OCID=AID2000142_aff_7593_10078&tduid=(ir_ynbhesti90kftxvckk0sohznxf2xi2a36hwec3qn00)(7593)(10078)(techadvisor.co.uk)(803X112721X967a3d93a3d5d08fdcf4c1f4e036e21d)&irclickid=ynbhesti90kftxvckk0sohznxf2xi2a36hwec3qn00&rtc=1). Create an account/log in for access.
- Public version of the Tableau desktop is available at: <https://public.tableau.com/s/download>
- R Studio Cloud <https://rstudio.cloud/>

## Agenda

(all times in Eastern)

### Introduction to Data Science and Visualization in Excel

November 9<sup>th</sup> 2:00-4:00 PM

- **Covering:** Tall vs Wide Data; Pivoting Data; Database Design Elements; Cleaning and Preparing Data; Calculations and Formulas; Chart types and pivot tables; Using Vlookup to merge data; Useful plugins; Introduction to Microsoft Access; Creating a database; Using data from IPEDS on educational programs and completers to create a state-level dashboard for key occupations.
- **Homework:** Finish your in-class assignment and customize your Excel Dashboard.
- **Key Takeaways:** Participants will learn how to clean data in Excel and create a dashboard that visualizes IPEDS completer data by CIP code, matched with SOC data.

### Introduction to Tableau

November 10<sup>th</sup> 2:00-4:00 PM

- **Covering:** Tableau basics and workspace; Connecting to a file and connecting multiple files; Creating a text table; Duplicating a worksheet; Sorting and filtering; Creating charts and maps; Mark Labels, Tooltips, Detail, Coloring; Using State Certifications and Licenses data to create an interactive, regional map of Certification and License Attainment and Earnings
- **Homework:** Finish your in-class assignment and continue to customize your Tableau Map.
- **Key Takeaways:** Participants will learn how to upload data to Tableau and use parameters to create an interactive data visualization.



## Dashboard Design and Interactive Visualizations in Tableau

November 16<sup>th</sup> 2:00-4:00 PM

- **Covering:** ‘Case, When’ and ‘If, then’ functions; Creating Groups; Creating hierarchies; Creating Sets; Sets and Filters – What’s the Difference; Creating Sets – Exercise; What are Parameters; Use Cases for Parameters; Using Parameters in Filters; Using QCEW data to visualize earnings and employment for filtered industries for your selected geography.
- **Homework:** Finish your in-class assignment and continue to customize your Tableau Dashboard.
- **Key Takeaways:** Participants will learn how to create an interactive dashboard that allows users to visualize data by selected industry.

## Introduction to R and R Basics

November 18<sup>th</sup> 2:00-4:00 PM

- **Covering:** Introduction to GGplot and R; Assigning Objects; Object Names and Syntax; Data Types; Installing R Packages; Importing Data with readxl; Pipe Operator; Logical and Boolean Operators; Introduction to Tidyverse; Summarizing Data; Removing Duplicates; Sampling; Selecting Columns with Select and Grepl; Mutate data with dplyr; Pivot data with tidyr
- **Homework:** Complete R Studio [Primer 1](#) and [Primer 2](#)
- **Key Takeaways:** Participants will learn how to perform basic operations in R, including importing data into R and cleaning datasets. This will prepare students to visualize data in R with ggplot.

## Visualizing your data in R with ggplot

November 19<sup>th</sup> 2:00-4:00 PM

- **Covering:** Loading ggplot; AES to define aesthetics; Changing axis labels and titles; faceting data; Installing the blscrapeR package; importing data; cleaning data; bar charts; histogram; box plots; month to month line plots; plotting maps; Users will use LAUS data to visualize unemployment in the United States, their state, and their group of counties.
- **Homework:** Complete R Studio [Primer 3](#) and [Primer 4](#) and explore the remaining [R primers](#)!
- **Key Takeaways:** Participants will learn how to visualize and map LAUS data for their LDD.