Sachiye's Data Boot Camp

One-Day Boot Camp Geared Toward a Career in Data Analytics in R!

CURRICULUM

Course: This course is a virtual one-day boot camp meant to be very comprehensive and practical for a career in data analytics. The instructor was a senior consultant in Washington, D.C., working at various well-known firms and organizations doing data analytics. Her goal is gear this course towards a successful career in data analysis in R, a tool that can be used at any company or organization.

Section	Description
About R	R is a programming language and software environment for statistical computing and graphics. R is widely used by statisticians, data scientists, and data analysts for statistical software. R uses a command line interface and many of its standard functions are written in R itself. The course will go over why R is popular for data analytics and why it is so important.
About the Data	The course uses public data from the United States that tracks fatalities from motor vehicle accidents across the states and the R package Ggplot2's data too. With these data sets, the class will have practical examples on how to format data, join data, and create tables, graphs, and maps to explore the many capabilities of R.
Get Started	R software operates with a working directory and library. The class teaches how to sets up working directories and libraries in R to save files and download packages. Also the class goes over the use of packages for functions, how to download packages for functions, and how to attach packages for functions in R.
Load Data	The class teaches the three ways to load data: uploading data from files on your computer, pulling data from databases, and hard coding data in vectors, matrices, or "lists" in R. This allows the beginning of analysis work.
Calculations and Conditional Statements	The class equips you with the knowledge of operators, conditional statements, and the know-how to create a new variable in your data set. The course gives practical examples that will be typical in the work place or research environment with operators and conditional statements in creating new variables for data sets and analysis work.
Transform Data with SQL	SQL is the most common language to pull and manage data held in a relational database that is taught in this course along with R language to use for data transformations in R. Data transformations include joins, pulling data, filtering data, and sorting data. There will be practical examples of these actions with the public data being used in the course to show data transformations often done in the "prep" work before analysis.
Formats and Strings	The class equips you with knowing how to format variables and edit strings. This is often needed as data sets are sometimes not formatted in the right way. Strings also often have to be edited to do analysis. This is also done often in the "prep" work before analysis, depending on the type of data being used.
Reports	The class teaches R's easy capability to create reports through the package R Markdown. A R Markdown document is written in an easy to write plain text format and contains chunks of

	embedded R code to create reports. It is a simple formatting syntax for authoring HTML, PDF, and Word documents. There are also a host of templates that can be used with R Markdown that will be reviewed.
Visualizations and Tables	The class teaches the R package, Ggplot2, that is dedicated to data visualizations that improves the quality and aesthetics of the graphics with efficient functions. The course will cover how to create tables, line graph, scatter plot, bar graph, and maps in R. These are the most common analysis visualizations in R for analysis work.
Statistics	The course covers descriptive statistics to get you started with any analysis. The class teaches finding the min, max, mean, median, and quartiles. Also it covers finding percentiles to understand the distribution of the data more in-depth and allows students to explore standard deviations, correlations, and regressions.
Automate Reports	The class teaches how to automate any repetitive series of reports you might have in the work place by going over inline code, parameters, loops, and the render function. Students will learn high-level concepts of coding through these automation steps and be able to learn skills to help bring efficacy to their work for analytical reporting.
Customize Reports	The class teaches that R has the capability through its package R Markdown to customize reports in an easy way by referencing a template document. A template document is often used and available at many organizations and company. A template document in R can be referenced to use its text style, margins, logos, and tables without specifying all the details in your code. It is a very quick, fast, and practical way to customize reports in a work place environment.
Type of Reports	The class teaches that R has the capability with R Markdown a suite of other packages to give templates for resumes, CVs, posters, presentations, etc. that can be easily generated and learned. The course will try one of these packages, Pagedown, to make a resume that will be posted on GitHub as an example. The class will also cover presentations as that is most commonly used.
Write Functions	The course teaches you to define and make functions. This allows the reuse of a chunk of code without endlessly copy and pasting by putting it as a function. This allows more efficiency when coding and allows sharing of your code for easy use.
Create Packages	The class teaches packages allow you to save, reproduce, and share functions that are created for more efficiency and repetitive tasks. The course will go through all the steps of creating a package with an example functions to try out and make into a package.
Best Practices	The course teaches best practices in saving, sharing, editing code, and project management in R. This is very practical and useful for the work place. The class will be specifically covering GitHub, the most widely used version control system.

Sachiye Day is the President and instructor of Sachiye's Data Boot Camp in Hemet, CA, which teaches virtual one-day boot camps in data analytics. She has worked at companies, such as the State Department, the Executive Office of the President, Fannie Mae, and the International Monetary Fund. Sachiye has extensive experience to equip you with the right skills to be successful.

Tel 240.423.6048 Email: Sachiyeday@gmail.com