Workshop – R Bootcamp: For Newcomers to R

Details

- Course Time: Sunday, June 16th, 2019
    - Schedule:
        - Workshop starts at 2:00pm
        - End of the Workshop: 4:30pm

Overview

This class will be an introduction to the statistical programming language R. This afternoon workshop launches your tenure as a user of R, the well-known open-source platform for data science and machine learning. The workshop stands alone as the perfect way to get started with R, or may serve to prepare for the more advanced full-day hands-on workshop, "R for Predictive Modeling".

Goals

This is a "short course" of an afternoon. Class will be given in a lab setting, with student exercises mixed with lectures. Students ***must*** bring a laptop to class. Due to the focused nature of this course, there will be no individual class projects but the instructors will be available to help students who are applying R to their own work outside of class.

Syllabus

Unit 1: Data Science and R Intro

- BIG DATA
- Data Science
- Roles in Data Science
- Use Cases
- Data'isms
- Class Format overview
- R Background
- R Intro
- R Studio
- R intro exercises

INTRODUCTION TO DATA SCIENCE OVERVIEW

Data science has become the central approach to tackling data-heavy problems in both business and academia. In this course, students learn how data science is done in the wild, with a focus on data acquisition, cleaning, and aggregation, exploratory data analysis and visualization, feature engineering, and model creation and validation. Students use the R statistical programming language to work through real-world examples that illustrate these concepts. Concurrently, students learn some of the statistical and mathematical foundations that power the data-scientific approach to problem solving.

WHO IS THIS COURSE FOR?

The R Intro for newcomers is for anyone with a basic understanding of data analysis techniques and those interested in improving their ability to tackle problems involving multi-dimensional data in a systematic, principled way. A familiarity with the R programming language is helpful, but unnecessary.

PREREQUISITES

Students may have some experience with programming and have some familiarity with basic statistical and linear algebraic concepts such as mean, median, mode, standard deviation, correlation, and the difference between a vector and a matrix.

OUTCOME

Upon completing the course, students have:

An understanding using R and an ability to use for further data science use cases. Familiarity with the R data science ecosystem, strategizing and the various tools to continue developing as a data scientist

--

STIPULATIONS

The above agenda, course structure and content is subject to change given time restrictions allotted for the course. Outcomes are not guaranteed. Students should download all course content during the class time as the content may not be available at a future date.