Introduction to R For Data Science

R is the gold standard programming language today for data scientist. R provides a wealth of libraries for data manipulation, visualization, and data modeling. R packages allows a beginner data scientist to build simple to sophisticated models quickly and easily in a few lines of code.

The course aims to quickly bring up to speed a programmer or business analyst who already know how to program in other language or have done advanced Excel MACROS to begin using R as a data science tool.

In the course, we will define data science and explore the first two things a data scientist must do – cleaning and visualizing data. We will learn and use R's dplyr, ggplot and ggvis packages for these tasks. We will move on to further steps in the Data Science Workflow - training models and testing them. We will do this by applying machine learning models to various industry-relevant data science problems. Our tool of choice will be the caret package.

At the end, you should have a working knowledge of how to solve data science problems with R.

DATE
12 - 13 April 2018

DURATION
2 Days; 9.00am to 5.30pm daily

VENUE
National University of Singapore

FEE
S$1,605 (includes GST)

Pre-Requisites
Familiarity with a programming language such as Java, C/ C++ or Python but NOT essential, and statistics 101 at a pre-university level.

Who Should Attend
Software Engineers, Programmers or Data Analysts