EmbRace R

Summary: Duration – 2 days / 8 hours (2x4 hours), Level - 200 (MS Scale), Delivery method - Virtual Instructor-led class or In-class

R is the most popular environment and language for statistical analyses, data mining, and machine learning. Managed and scalable version of R runs in SQL Server and Azure ML. Learn how to program in R to squeeze the information from your data. This course is both, an independent seminar, and a complement to the Python for SQL Server Specialists seminar, meaning that the two courses overlap only partially; in each course, some different algorithms and techniques are introduced.

As being an open-source development, R is the most popular analytical engine and programming language for data scientists worldwide. The number of libraries with new analytical functions is enormous and continuously growing. However, there are also some drawbacks. R is a programming language, so you have to learn it to use it. Open-source development also means less control over code. Finally, the free R engine is not scalable.

Microsoft added support for R code in SQL Server 2016 and, Azure Machine Learning, or Azure ML, and in Power BI. A parallelized highly scalable execution engine is used to execute the R scripts. In addition, not every library is allowed in these two environments.

Attendees of this seminar learn to program with R from the scratch. Basic R code is introduced using the free R engine and RStudio IDE. Then the seminar shows some more advanced data manipulations, matrix calculations and statistical analysis together with graphing options. The mathematics behind is briefly explained as well. Then the seminar switches more advanced data mining and machine learning analyses. Attendees also learn how to use the R code in SQL Server.

AUDIENCE:

The target audience is everybody that wants to start analyzing data with R. Database developers that deal with SQL Server and code in T-SQL and want to move more to advanced analytics can get the most of this course.

AFTER THE TRAINING ATTENDEES WILL BE ABLE TO:

- Start analyzing their data with statistical and machine learning methods immediately
- They can also prepare the data accordingly for the target analysis and deploy the solution in SQL Server
- Besides practical skills, the delegates also learn the basics of the mathematics behind the algorithms

TOPICS:

Module 1. Introduction to R
Module 2. Data overview and manipulation
Module 3. Basic and advanced visualizations
Module 4. Data mining and machine learning methods
Module 5. Scalable R in SQL Server