EmbRaceR: Introducing R

Dejan Sarka

Instructor Bio

- ## Dejan Sarka
  - ~35 years of experience
  - Data Platform MVP, MCT
  - 18 books
  - 20+ courses
  - Focus:
    - Data science
    - Data quality
    - Data modeling

Course Introduction (1)

- R is the most popular environment and language for statistical analyses, data mining, and machine learning. 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.

Course Introduction (2)

- Microsoft added support for R code in SQL Server 2016, and continues to support it in later versions. A parallelized highly scalable execution engine is used to execute the R scripts. In addition, not every library is allowed in these two environments.

Course Introduction (3)

- You will learn the R language through a collection of courses under the common name EmbRaceR. In the first course, attendees learn to program with R from the scratch. Basic R code is introduced using the free R engine and RStudio IDE. A lifecycle of a data science project is explained in details. You will also learn how to use RStudio IDE, the most popular development environment for R.

Modules

- Starting with R

End of module 00: Course Introduction