

COURSE: Data Science with R and SQL Server

Introducing the language, statistics, data mining, and machine learning with R and using R in SQL Server and Microsoft BI stack.

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.

Summary

<i>Duration:</i>	<i>4 days/ 32 hours</i>
<i>Level:</i>	<i>300*</i>
<i>Delivery method:</i>	<i>In class</i>
<i>Language:</i>	<i>English</i>
<i>Price:</i>	<i>1050 EUR excl VAT</i>

* The difficulty level is consistent with the widely accepted scale of technical difficulty of training on Microsoft Corp

AUDIENCE:

Attendees should have basic understanding of data analysis and basic familiarity with SQL Server tools.

AFTER THE TRAINING ATTENDEES WILL BE ABLE TO:

Attendees of this course 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. The attendees learn how to perform the data overview and do the most tedious task in a project, the data preparation task. After data overview and preparation, the analytical part begins with intermediate statistics in order to analyze associations between pairs of variables. Then the course introduces more advanced methods for researching linear dependencies.

Too many variables in a model can make its own problem. The course shows how to do feature selection, starting with the basics of matrix calculations. Then the course switches more advanced data mining and machine learning analyses, including supervised and unsupervised learning. The course also introduces the currently modern topics, including forecasting, text mining, and reinforcement learning. Finally, the attendees also learn how to use the R code in SQL Server, Azure ML, and Power BI.

TOPICS:

- Module 1. **Introducing data science and R** (What are statistics, data mining, machine learning, Data science projects and their lifetime; Introducing R; R tools; R data structures)
- Module 2. **Data overview** (Datasets, cases and variables; Types of variables; Introductory statistics for discrete variables; Descriptive statistics for continuous variables; Basic graphs; Sampling, confidence level, confidence interval)
- Module 3. **Data preparation** (Derived variables; Missing values and outliers; Smoothing and normalization; Time series; Training and test sets)
- Module 4. **Associations between two variables and visualizations of associations** (Covariance and correlation; Contingency tables and chi-squared test; T-test and analysis of variance; Bayesian inference; Linear models)
- Module 5. **Feature selection and matrix operations** (Feature selection in linear models; Basic matrix algebra; Principal component analysis; Exploratory factor analysis; Lab 5)
- Module 6. **Unsupervised learning** (Hierarchical clustering; K-means clustering; Association rules;)
- Module 7. **Supervised learning** (Neural Networks; Logistic Regression; Decision and regression trees; Random forests; Gradient boosting trees; K-nearest neighbors)
- Module 8. **Modern topics** (Support vector machines; Time series; Text mining; Deep learning; Reinforcement learning)
- Module 9. **R in SQL Server and MS BI**

FUNDAMENTALS

SQL Server Fundamentals courses give you an excellent base for successful career as administrator or data base developer. Their practical purpose, use of scenarios of working practice will help both the beginners, as well as the more experienced students. The courses will assist the students to develop skills to protect in a better way the data of their organizations, to react quicker and adequate to all issues, which are data base related, and to avoid following some harmful myths and wrong concepts in administration and data base development.

- ❖ **MS Windows & AD Administration for DBA's**
- ❖ **Querying SQL Server Fundamentals**
- ❖ **SQL Server Administration Fundamentals**
- ❖ **SQL Server Development Fundamentals**

MASTERING

The trainings of SQL Server Mastering series consist of intensive modules which are comparatively complicated regarding the technical level and which are designed and intended for professionals in the field of SQL Server. Trainings include solid theoretical material, part of which comes directly from the program team, who develops SQL Server in Microsoft also numerous demonstrations from the "workshop" of SQL Server, unpublished tuning techniques for increase of SQL Server performance.

- ❖ **Mastering SQL Server for Database Admins**
- ❖ **Mastering SQL Server for Database Developers**
- ❖ **SQL Server Performance Tuning and Optimization**

DEEP DIVE

Deep Dive Series consists two and three day trainings with technical level 300+ (level 100 – beginners, level 400 – masters). Those courses are focused on specific components of SQL Server. They consist of a significant number of practical advices, a lot of demonstrations, huge theoretical part and also examples of real infrastructural solutions, discussions of real situations which students have in their infrastructures and want to discuss.

- ❖ **SQL Server Enterprise Infrastructure Optimization and Deployment**
- ❖ **SQL Server Upgrade**
- ❖ **SQL Server Replication Deep Dive**
- ❖ **SQL Server High Availability and Disaster Recovery Solutions**

BI

SQL Master Academy works in close partnership with SolidQ in the field of Business Intelligence trainings, due to which fact we have in our portfolio a unique programme of professional trainings in the field of Data Quality, Multidimensional Design, SSAS, ETL and SSIS. The courses train BI developers and designers of BI solutions, including back end EIM, as well as front end and reporting.

- ❖ **Personal Business Intelligence with MS Power BI**
- ❖ **Introduction to MS Business Intelligence**
- ❖ **BI Fundamentals**
- ❖ **Accelerated SQL Server 2012-2016 Integration Services**
- ❖ **Multidimensional Database Design with SSAS 2012-2016**
- ❖ **Data Quality and Master Data Management with SQL Server 2012-2016**
- ❖ **Data Mining Algorithms in SSAS, Excel, R, and Azure ML**

DATABASE DESIGN

The design of relational data base is one important and inseparable part of the solution development. Those are knowledge and skills which every DB Dev or DB Designer should have. A huge part of the problems in production sphere are due to a not good base model or to such, which does not scale well. The courses for design of relational data base we offer in SQL Master Academy, are oriented not only to mastering of the relational model and the good design of the subject filed, but also to the mastering of good practices in modelling of scalable data base.

- ❖ **Relational Database Design Fundamentals**
- ❖ **Relational Database Design for Performance and Scalability**

SQL SERVER 2016

Explore the new version of the fastest performing database system in the market today – SQL Server 2016! Be prepared for InMemory OLTP, the new columnstore extensions and the brand new operational analytics possibilities, which can give your workload 30x-100x performance improvement. See how the Query Store, temporals, the new built-in security features and AlwaysOn groups works. Get all that you need to put your skills and your databases on track.

- ❖ **SQL Server 2016 Get Ready for DBA and DB Devs**
- ❖ **SQL Server 2016 for Business Intelligence Experts**
- ❖ **SQL Server Cloud based Solutions**