

## Data Mining with SQL Server 2012

### Course Summary

#### Description

In this course, attendees learn how to use data mining to find advanced patterns in their data, and perform predictions based on the patterns found using SQL Server 2012 Analysis Services. This is a high-end course, enabling students to go beyond simple usage of wizards. In-depth understanding of the data mining algorithms, how they are built and how they work, is provided.

#### Objectives

At the end of this course, students will be able to:

- Describe what data mining is and what business questions it can answer
- Explain the process of a data mining project
- Explore and understand your data using descriptive statistics, OLAP cubes, reports and other tools
- Prepare the data to make better models
- Understand the data mining algorithms and when to use them
- Create data mining models and browse them
- Evaluate models to find the one that gives best results Use SQL Server 2012 Integration Services data mining tasks
- Do text mining with Integration Services and with Full-Text and Semantic Search
- Understand and use the Data Mining Extensions (DMX) language
- Deploy data mining models in production using custom application, OLAP cubes or reports developed with SQL Server 2008 Reporting Services.
- 

#### Topics

- Introduction to Data Mining
- Understanding the Data
- Data Preparation
- Data Mining Algorithms, Part 1
- Data Mining Algorithms, Part 2
- Data Mining Algorithms, Part 3
- Data Mining Extensions (DMX) Language
- Integration with SQL Server BI Suite
- Data Mining with Excel
- Analyzing Texts
- Developing Data Mining Applications
- Administering Data Mining Models

#### Audience

This course is intended for business intelligence application developers, advanced database administrators, and advanced analysts.

#### Prerequisites

Before attending this course, it is recommended that students have the following skills:

- At least moderate experience with data warehousing, reporting and On-Line Analytical Processing
- Familiarity with the Transact-SQL language
- Knowledge of a .NET language like C# or VB.NET is welcome as well.

The following courses are recommended as a learning path to this course:

- Transact-SQL Fundamentals
- BI Boot Camp 2008 R2 or BI Immersion 2012.

#### Duration

Three days

Due to the nature of this material, this document refers to numerous hardware and software products by their trade names. References to other companies and their products are for informational purposes only, and all trademarks are the properties of their respective companies. It is not the intent of ProTech Professional Technical Services, Inc. to use any of these names generically

## Data Mining with SQL Server 2012

### Course Outline

- I. Introduction to Data Mining**
  - A. Introduction to Data Mining
  - B. Business Questions
  - C. Data Mining Process
  - D. SQL Server Data Mining Tools
  - E. *Lab 01: Using SSMS and SSDT to Explore Sample Databases and a Sample Project*
- II. Understanding the Data**
  - A. Cases, Variables, and Measures
  - B. Descriptive Statistics with T-SQL and CLR
  - C. Data Profiling with SSIS, SSAS and SSRS
  - D. Measuring Information
  - E. Lab 02: Using an SSAS Multidimensional Cube and Excel to Get a Data Overview
- III. Data Preparation**
  - A. Missing Values and Outliers
  - B. Derived Variables
  - C. Time Series
  - D. Sampling and Testing the Sampling
  - E. Lab 03: Using SSIS to Split the Data into Training and Test Set and Checking the Split with Decision Trees
- IV. Data Mining Algorithms, Part 1**
  - A. Naïve Bayes
  - B. Decision Trees
  - C. Linear Regression and Regression Trees
  - D. Lab 04: Using the Naïve Bayes and Decision Trees Algorithms
  - E. Day 2
- V. Data Mining Algorithms, Part 2**
  - A. Neural Network and Logistic Regression
  - B. Predictive Models Evaluation
  - C. Time Series
  - D. Lab 05: Using the Neural Network and Time Series Algorithms, and Evaluating Predictive Models
- VI. Data Mining Algorithms, Part 3**
  - A. Association Rules
  - B. Clustering
  - C. Sequence Clustering
- D. Lab 06: Using the Association Rules, Clustering, and Sequence Clustering Algorithms**
- VII. Data Mining Extensions (DMX) Language**
  - A. XMLA and Data Mining Objects
  - B. DDL Statements
  - C. DML Statements
  - D. DMX Select Statement
  - E. Lab 07: Using the DMX Language
- VIII. Integration with SQL Server BI Suite**
  - A. Integration with SSIS
  - B. Integration with SSRS
  - C. Integration with SSAS
  - D. Lab 08: Integrating Data Mining with SSIS, SSRS, and SSAS
- IX. Data Mining with Excel**
  - A. Excel Data Preparation and Data Mining
  - B. Table Analysis with Excel
  - C. Combining Data Mining with PowerPivot
  - D. Lab 09: Using the Data Mining Client for Excel for Data Preparation and Analysis
- X. Analyzing Texts**
  - A. Text Mining with SSIS
  - B. Using Full-Text Search (FTS)
  - C. Using Semantic Search (SS)
  - D. Lab 10: Text Mining with SSIS
- XI. Developing Data Mining Applications**
  - A. Advanced DMX
  - B. Developing Data Mining Models with AMO
  - C. Using ADOMD.NET Client
  - D. Server Procedures and ADOMD.NET Server Components
  - E. Lab 11: Using Advanced DMX
- XII. Administering Data Mining Models**
  - A. Managing SSAS Databases
  - B. Monitoring SSAS
  - C. Data Mining Security
  - D. Lab 12: Securing, Backing Up, and Processing Data Mining Models