



DATA ENGINEERING

Professional

Delivery mode: self-paced, onsite, live online

Course length: 2 days

Target Audience: Data Engineer, Business Analyst, Data Scientist, Domain Expert

Prerequisite courses: none



Data Engineer



Business Analyst



Domain Expert



Data Scientist



Executive



Administrator

Practice building data preparation processes using common techniques in RapidMiner.

Overview

This course provides the opportunity to directly work on creating and modifying RapidMiner Processes to prepare data for machine learning. It covers the most commonly used operations in RapidMiner with detail on how to implement and configure them. It provides a solid foundation for building RapidMiner Processes and preparing data.

Consider taking *Applications & Use Cases Professional* before this course

Course Objectives

- Access Data
- Complete Common Data Preparation tasks
- Perform Joins & Set Operations
- Transform Data with Pivots & Aggregations
- Be Able to Leverage Process Control
- Understand How to Handle Text



Data Engineer



Business Analyst



Domain Expert



Data Scientist



Executive



Administrator



Data Engineering Professional Topic Guide:

Learn how to access data from RapidMiner.

Practice importing data.

Learn when Wizards can be helpful.

Go through the process of bringing data in, using Turbo Prep and setting up Auto Model.

Practice Basic Data Transformations with process design and with Turbo Prep.

Perform exercises with the most common data preparation steps including: rename attributes, Filtering, Attribute Selection, Type Conversion, and Value Replacements

Practice working with Multiple Data Sets with process design and Turbo Prep.

Learn about combining data sets with: Joins, appending, and Set Operators.

Learn how and when to Pivot and Aggregate data with process design and Turbo Prep.

Work on reshaping data with: Aggregation, Pivots, De-Pivot, and Transpositions.

Understand and be able to use processes effectively.

Learn what Processes are and how to use them.

Learn how to organize operators in Subprocesses.

Learn how to Execute a process from another process.

Learn how to create, use, and manage Building Blocks.

Know the basics of basic Text Processing in RapidMiner.

Inspect the process of turning text into structured data.

Know how and when to use Turbo Prep and Process Design.



DATA ENGINEERING

Professional



Delivery mode: self-paced, onsite, live online
Course length: 2 days



Data Engineer



Business Analyst



Domain Expert



Data Scientist



Executive



Administrator

We designed our program with the most common user personas and the required areas of expertise for applied data science in mind:

