

Data Science without a Ph.D. Using IBM SPSS Modeler (v18.1.1) SPVC

Code: 0A018G-SPVC
URL: [View Online](#)

This IBM Self-Paced Virtual Class (SPVC) includes:

- PDF course guide available to attendee during and after course
- Lab environment where students can work through demonstrations and exercises at their own pace

Contains PDF course guide, as well as a lab environment where students can work through demonstrations and exercises at their own pace.

This course focuses on reviewing concepts of data science, where participants will learn the stages of a data science project. Topics include using automated tools to prepare data for analysis, build models, evaluate models, and deploy models. To learn about these data science concepts and topics, participants will use IBM SPSS Modeler as a tool.

If you are enrolling in a Self Paced Virtual Classroom or Web Based Training course, before you enroll, please review the Self-Paced Virtual Classes and Web-Based Training Classes on our Terms and Conditions page, as well as the system requirements, to ensure that your system meets the minimum requirements for this course.

Skills Gained

Please refer to course overview

Who Can Benefit

- Business Analysts
- Data Scientists
- Participants who want to get started with data science

Prerequisites

- It is recommended that you have an understanding of your business data

Course Details

Course Outline

- 1: Introduction to data science and IBM SPSS Modeler
 - Explain the stages in a data-science project, using the CRISP-DM methodology
 - Create IBM SPSS Modeler streams
 - Build and apply a machine learning model
 - 2: Setting measurement levels
 - Explain the concept of "field measurement level"
 - Explain the consequences of incorrect measurement levels
 - Modify a field's measurement level
 - 3: Exploring the data
 - Audit the data
 - Check for invalid values
 - Take action for invalid values
 - Impute missing values
 - Replace outliers and extremes
 - 4: Using automated data preparation
 - Automatically exclude low quality fields
 - Automatically replace missing values
 - Automatically replace outliers and extremes
 - 5: Partitioning the data
 - Explain the rationale for partitioning the data
 - Partition the data into a training set and testing set
 - 6: Selecting predictors
 - Automatically select important predictors (features) to predict a target
 - Explain the limitations of automatically selecting features
 - 7: Using automated modeling
 - Find the best model for categorical targets
 - Find the best model for continuous targets
 - Explain what an ensemble model is
 - 8: Evaluating models
 - Evaluate models for categorical targets
 - Evaluate models for continuous targets
 - 9: Deploying models
 - List two ways to deploy models
 - Export scored data
-

Download Whitepaper: Accelerate Your Modernization Efforts with a Cloud-Native Strategy

Get Your Free Copy Now