

**Defining and Scoping Data Science Projects**

**Course Number:** DATA-132  
**Duration:** 2 days

**Overview**

Data science projects can be complex and time-consuming. By planning the project early and establishing parameters, you can ensure that the project is aligned with business goals and that resources are appropriately allocated.

This Data Science training course teaches attendees how to define and scope data science projects from start to finish. Participants learn how to identify the business problem, define project goals, estimate resources, identify risks, set project boundaries, define success criteria and metrics, and communicate results to stakeholders. Students also learn how to gather data from various sources, clean and prepare data for analysis, and conduct exploratory data analysis to ensure their data science projects succeed.

**Prerequisites**

* General familiarity with data and data science pipelines
* Experience working on data science and machine learning projects

**Materials**

All Data Science training students will receive comprehensive courseware.

**Software Needed on Each Student PC**

A modern web browser and an Internet connection.

**Objectives**

* Understand the importance of defining and scoping data science projects
* Identify and understand the business problem
* Estimate resources and identify risks
* Set project boundaries and scope
* Define success criteria and metrics

**Outline**

* Introduction to Data Science Projects
  + Data Science and Its Impact on Business
  + Understanding the Data Science Project Lifecycle
  + The Importance of Defining and Scoping in Data Science Projects
* Defining the Data Science Project
  + Identifying and Understanding the Business Problem
  + Translating Business Problems into Data Science Tasks
  + Creating Effective Problem Statements
  + Understanding Stakeholder Expectations
* Scoping the Data Science Project
  + Estimating the Resources: Data, Time, Skills, and Tools
  + Identifying Risks and Assumptions
  + Setting Project Boundaries and Scope
  + Defining Success Criteria and Metrics
* Data Requirements and Preparation
  + Understanding Data Sources and Collection
  + Data Cleaning and Preparation
  + Exploratory Data Analysis
  + Feature Engineering and Selection
* Creating and Executing a Project Plan
  + Drafting the Project Plan
  + Model Building, Validation, and Selection
  + Deployment and Maintenance
  + Communicating Results and Findings
* Final Project
  + Students will scope and plan a data science project
  + Presentation
* Conclusion