

MATLAB for Data Analysis and Predictive Modeling

Course Duration: 5 days

Course Objectives

1	Master data science workflow and analyze datasets effectively.
2	Import and manipulate diverse data types in MATLAB.
3	Visualize data and customize figures for exploration.
4	Perform calculations and generate summary statistics.
5	Explore data distributions using plots and statistical measures.
6	Organize and prepare data for analysis efficiently.
7	Clean and handle messy data, address missing values and outliers.
8	Engineer features for predictive modeling and analysis.

Course Modules

Module 01: Introduction to the Data Science Workflow

In this module you'll learn about the key steps in a data science workflow and begin exploring a data set using a script provided for you. As you work with the file, take note of the different elements in the script. As you progress through the course, you'll create a similar script yourself.

Module 02: Importing Data

In this module you'll import data into MATLAB, customize the import options, and generate code to automate the process. You'll also work with different types of data, such as numeric, dates, and text.

Module 03: Visualizing and Filtering Data

In this module you'll create visualizations and learn how to customize figures. You'll also filter your data to select only what is needed for your analysis. You'll create new tables and save them to use in the future or share with others outside of MATLAB.

Module 04: Performing Calculations

In this module you'll write small pieces of code to extend your analysis. You'll calculate summary statistics on groups of data and determine if variables are correlated. You'll extend your ability to filter data to defining conditions across multiple variables. You'll also modify categorical data to remove, combine, or create new categories to use for defining groups.

Module 05: Exploring the Data

In this module you'll apply the skills gained in Exploratory Data Analysis with MATLAB on a new dataset. You'll explore different types of distributions and calculate quantities like the skewness and interquartile range. You'll also learn about more types of plots for visualizing multi-dimensional data.

Module 06: Organizing the Data

In this module you'll learn to prepare data for analysis. Often data is not recorded as required. You'll learn to manipulate string variables to extract key information. You'll create a single datetime variable from date and time information spread across multiple columns in a table. You'll efficiently load and combine data from multiple files to create a final table for analysis.

Module 07: Cleaning the Data

In this module you'll clean messy data. Missing data, outliers, and variables with very different scales can obscure trends in the data. You'll find and address missing data and outliers in a data set. You'll compare variables with different scales by normalizing variables.

Module 08: Feature Engineering

In this module you'll create new features to better understand your data. You'll evaluate features to determine if a feature is potentially useful for making predictions.

Module 09: Creating Regression Models

In this module you'll apply the skills gained from the first two courses in the specialization on a new dataset. You'll be introduced to the Supervised Machine Learning Workflow and learn key terms. You'll end the module by creating and evaluating regression machine learning models.

Module 10: Creating Classification Models

In this module you'll learn the basics of classification models. You'll train several types of classification models and evaluate the results.

Module 11: Applying the Supervised Machine Learning Workflow

In this module you'll apply the complete supervised machine learning workflow. You'll use validation data inform model creation. You'll apply different feature selection techniques to reduce model complexity. You'll create ensemble models and optimize hyperparameters.