Data Science and Data Engineering Bootcamp
Unleash the Data Scientist in You

data science dojo

Target Audience
Software engineers, data analysts, business analysts, technical program managers, architects, database administrators, and researchers, with an interest in data science and big data engineering.

Prerequisites
Knowledge of at least one programming language. Interest in data science and data engineering.

Format
- 50% Lectures + 50% Labs, Exercises, and Demos
- 8 hours of preparatory work through webinars and online training materials
- 40+ hours of in-class training
- 10 hours of mentored project participation

Duration
5 days of in-person training + 1 day of project work

Course Outline

Preparatory Material (via webinars)
- Introduction to Big Data, Data Science, and Predictive Analytics
- Introduction to Azure ML Studio
- Fundamentals of Data Mining
- Introduction to R Programming

Fundamentals of Data Science
- Data Exploration, Visualization, and Feature Engineering
- Hands-On Lab: Data Exploration, Visualization, and Feature Engineering
- Machine Learning Fundamentals

Classification Algorithms
- Introduction to Predictive Modeling
- Decision Tree Learning
- Logistic Regression
- Hands-On Labs: Decision Tree Learning and Logistic Regression

Operationalizing Machine Learning Models
- Hands-On Lab: Building a Classification Model in Azure ML Studio
- Hands-On Lab: Deploying a Predictive Model as a Service

Regression Algorithms
- Linear Regression
- Regularized Regression Models
- Metrics and Methods for Evaluating Classification and Regression Models
- Hands-On Lab: Building a Regression Model

Unsupervised Learning
- K-Means Clustering

- Hands-On Lab: Using K-Means Clustering

Recommender Systems
- Content-Based Filtering
- Collaborative Filtering
- Hands-On Lab: Deploying a Recommender System as a Service

Ensemble Methods
- Bootstrapping, Bagging and Boosting
- AdaBoost
- Random Forests
- Hands-On Lab: Building a Random Forest Classifier

Fundamentals of Big Data Engineering
- Introduction to Large-Scale Online Systems
- Hive Tutorial
- Hands-On Labs: Creating a Hadoop Cluster and Writing Hive Queries

Handling Real-Time and Streaming Data
- Message Queues and Real-time Analytics
- Hands-On Lab: Creating a Streaming Analytics Pipeline

Distributed Databases and Data Warehousing
- Hands-On Lab: Setting Up Relational Databases in the Cloud
- NoSQL Databases and HBase
- Hands-On Lab: Twitter and HBase
- ETL and Data Warehousing
- Hands-On Lab: Setting Up a Data Warehouse in the Cloud

Data Science Essentials
- Introduction to Online Experimentation and A/B Testing

Hack Day Project
Spend one full day building an 'Internet of Things' solution using Azure EventHubs, Azure Stream Analytics, Power BI, and Azure ML studio. You will be using a Texas Instruments sensor tag to obtain temperature, humidity, pressure, and accelerometer readings and observe various analytics in a cloud BI dashboard in real-time.

About Data Science Dojo
Data Science Dojo has trained more than 600 professionals from over 40 companies around the world through a variety of workshops and bootcamps. Unlike other offerings in the industry, our trainings are short, hands-on, and personal. We enable you to solve real-world business problems in the shortest duration possible. Visit our website, www.datasciencedojo.com, to learn more.