Cloudera Data Science Workbench Training (DATASWB-ON)

Overview of CDSW

- Introduction to CDSW
- Who Can Use CDSW
- How to Access CDSW
- Navigating around CDSW
- User Settings
- Hadoop Authentication

Projects in CDSW

- Creating a New Project
- Navigating around a Project
- Project Settings

The CDSW Workbench Interface

- Using the Workbench
- Using the Sidebar
- Using the Code Editor
- Engines and Sessions

Running Python and R Code in CDSW

- Running Code
- Using the Session Prompt
- Using the Terminal
- Installing Packages
- Using Markdown in Comments

Using Apache Spark 2 in CDSW

- Scenario and Dataset
- Copying Files to HDFS
- Interfaces to Apache Spark 2
- Connecting to Spark
- Reading Data
- Inspecting Data

Data Science and Machine Learning in CDSW

- Transforming Data
- Using SQL Queries
- Visualizing Data from Spark
- Machine Learning with MLlib

- Session History

Experiments and Models in CDSW

- Machine Learning Workflow
- Running Experiments
- Using Packages in Experiments
- Deploying Models
- Calling Models
- Using Packages in Models

Teams and Collaboration in CDSW

- Collaboration in CDSW
- Teams in CDSW
- Using Git for Collaboration
- Conclusion