Exploring and Analyzing Data with Splunk

This 9-hour course is for users who want to attain operational intelligence level 4, (business insights) and covers exploratory data analysis by using statistical tools and custom visualizations.

Course Topics

- Analytics Framework
- Exploring and visualizing data
- Cleaning and Preprocessing Data
- Numerical and String based clustering
- Data Correlation
- Meta Transactions
- Detecting Anomalies
- Forecasting

Prerequisite Knowledge

To be successful, students should have a solid understanding of the following courses:

- Intro to Splunk
- Using Fields
- Scheduling Reports and Alerts
- Visualizations
- Working with Time
- Statistical Processing
- Comparing Values
- Result Modification
- Leveraging Lookups and Sub-searches
- Correlation Analysis
- Search Under the Hood
- Intro to Knowledge Objects
- Creating Field Extractions
- Search Optimization

Course Format

Instructor-led lecture with labs, delivered via virtual classroom or at your site.

Course Objectives

Topic 1 – What is Data Science

- Define terms related to analytics and data science
- Describe the analytics workflow
- Describe Artificial Intelligence and Machine Learning
- Examine common Machine Learning myths
- Describe Splunk's Machine Learning tools

Topic 2 – Exploratory Data Analysis

- Use bin and makecontinuous to restructure and visualize data
- Examine field statistics with fieldsummary
- Transform fields with eval and fillnull
- Clean text with the rex and cleantext commands
- Solve Anscombe's Quartet
- Apply boxplots and 3d scatterplots to visualize data

Topic 3 – Event Clustering

- Take a behavioral based approach to cluster data
- Cluster numerical fields using the kmeans command
- Cluster based of string similarity with the cluster command
- Find patterns in clusters

Topic 4– Correlations and Transactions

- Define correlation and co-occurrence
- Use SPL correlation commands
- Use the statistical tests from the Machine Learning Toolkit to correlate fields
- Use streamstats and chart commands to correlate data

Topic 5– Anomaly Detection

- Define Statistical Outliers
- Use Add-hoc methods of numerical anomaly detection
- Find numerical or categorical anomalies with the AnomalyDetection command

Topic 6 – Forecasting

- Define forecasting use cases
- Use the predict command to forecast future timeseries

About Splunk Education

Splunk classes are designed for specific roles such as Splunk Administrator, Developer, User, Knowledge Manager, or Architect. To contact us, email Education_AMER@splunk.com

Certification Tracks

Our certification tracks provide comprehensive education for Splunk customer and partner personnel according to their areas of responsibility.
To view all Splunk Education's course offerings, or to register for a course, go to http://www.splunk.com/education

Splunk, Inc.
270 Brannan St. San Francisco, CA 94107
+1 866.GET.SPLUNK (1 866.438.7758)
Contact sales