



## TRAINING OFFERING

---

**You can reach us at:**

Arrow ECS, Nidderdale House, Beckwith Knowle, Harrogate, HG3 1SA

Email: [education.ecs.baltic@arrow.com](mailto:education.ecs.baltic@arrow.com)

Phone: 0870 251 1000



# Splunk for Analytics and Data Science

<b>CODE:</b>	<b>LENGTH:</b>	<b>PRICE:</b>
SPL_SFADS	16 Hours (2 days)	Request Price

## Description

This 13.5-hour course is for users who want to attain operational intelligence level 4, (business insights) and covers implementing analytics and data science projects using Splunk's statistics, machine learning, built-in and custom visualization capabilities.

## Objectives

### Topic 1 – Analytics Workflow

- Define terms related to analytics and data science
- Describe the analytics workflow
- Describe common usage scenarios
- Navigate Splunk Machine Learning Toolkit

### Topic 2 – Training and Testing Models

- Split data for testing and training using the sample command
- Describe the fit and apply commands
- Use the score command to evaluate models

### Topic 3 – Regression: Predict Numerical Values

- Differentiate predictions from estimates
- Identify prediction algorithms and assumptions
- Model numeric predictions in the MLTK and Splunk Enterprise

### Topic 4 – Clean and Preprocess the Data

- Define preprocessing and describe its purpose

- Describe algorithms that preprocess data for use in models
- Use FieldSelector to choose relevant fields
- Use PCA and ICA to reduce dimensionality
- Normalize data with StandardScaler and RobustScaler
- Preprocess text using Imputer, NPR, TF-IDF, and HashingVectorizer

### **Topic 5 – Clustering**

- Define Clustering
- Identify clustering methods, algorithms, and use cases
- Use Smart Clustering Assistant to cluster data
- Evaluate clusters using silhouette score
- Validate cluster coherence
- Describe clustering best practices

### **Topic 6 – Forecasting Fields**

- Differentiate predictions from forecasts
- Use the Smart Forecasting Assistant
- Use the StateSpaceForecast algorithm
- Forecast multivariate data
- Account for periodicity in each time series

### **Topic 7 – Detect Anomalies**

- Define anomaly detection and outliers

- Identify anomaly detection use cases
- Use Splunk Machine Learning Toolkit Smart Outlier Assistant
- Detect anomalies using the Density Function algorithm
- View results with the Distribution Plot visualization

## **Topic 8 – Classify: Predict Categorical Values**

- Define key classification terms
- Identify when to use different classification algorithms
- Evaluate classifier tradeoffs
- Evaluate results of multiple algorithms

## **Audience**

Splunk classes are designed for specific roles such as Splunk Administrator, Developer, User, Knowledge Manager, or Architect.

## **Prerequisites**

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

- Fundamentals 1, 2, & 3
- Advanced Searching & Reporting Or the following single-subject courses:
- What is Splunk?
- 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

- Introduction to Knowledge Objects
- Creating Field Extractions
- Search Optimization

## **Programme**

- Analytics Framework
- Regression for Prediction
- Cleaning and Preprocessing Data
- Algorithms, Preprocessing and Feature Extraction
- Clustering Data
- Detecting Anomalies
- Forecasting
- Classification

## **Test and Certification**

Our certification tracks provide comprehensive education for Splunk customer and partner personnel according to their areas of

## **Further Information**

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

## **Session Dates**

On request. Please [Contact Us](#)

## **Additional Information**

[This training is also available as onsite training. Please contact us to find out more.](#)