# Implementing Predictive Analytics with Spark in Azure HDInsight

**Modality: On Demand**

**Duration: 24 Hours**

***This course prepares you for the 70-475 Exam leading to 70-475 Certification. This course does not include the*** <span style="color:green">***Official Exam Voucher***</span>***, however, you can request to purchase the Official Exam Voucher separately.***

## About this course:

Are you ready for big data science? In this course, learn how to implement predictive analytics solutions for big data using Apache Spark in Microsoft Azure HDInsight. See how to work with Scala or Python to cleanse and transform data and build machine learning models with Spark ML (the machine learning library in Spark).Note: To complete the hands-on elements in this course, you will require an Azure subscription and a Windows client computer. You can sign up for a free Azure trial subscription (a valid credit card is required for verification, but you will not be charged for Azure services). Note that the free trial is not available in all regions. This course also helps in the preparation for the Microsoft 70-475: Designing and Implementing Big Data Analytics Solutions certification exam.

The average salary for Data Analyst with Apache Spark skills is <span style="color:green">**$63,882**</span> per year.

## Course Objective:

After completing this course, students will be able to:

- Using Spark to explore data and prepare for modeling
- Build supervised machine learning models
- Evaluate and optimize models
- Build recommenders and unsupervised machine learning models

## Audience:

This course is intended for

- Big Data Professionals

## Prerequisites:

- Familiarity with Azure HDInsight.
- Familiarity with databases and SQL.
- Some programming experience.
- A willingness to learn actively in a self-paced manner.

# Suggested prerequisites courses:

- [Microsoft Azure Fundamentals - MOC On Demand (MS-10979)](#)
- [Programming for Absolute Beginners](#)
- [SQL Database for Beginners](#)

# Career path:

- [Cloud Computing Career Path…](#)

# Course Outline:

### Introduction to Data Science with Spark

- Getting Started with Spark
- Exploring Data with Spark
- Further Reading
- Lab
- Review

### Getting Started with Machine Learning

- Introduction to Machine Learning in Spark
- Pipelines and Text Analysis
- Further Reading
- Lab
- Review

### Evaluating and Optimizing Machine Learning Models

- Evaluating Machine Learning Models
- Optimizing Model Parameters
- Further Reading
- Lab
- Review

### Recommenders and Unsupervised Models

- Recommenders
- Clustering
- Further Reading
- Lab
- Review

### Final Exam

- Final Exam?

https://www.quickstart.com/ *Contact Us: (866) 991-3924*