

IBM Interactive Experience

# Apache Spark

**Hands on Introduction to Apache Spark  
for Data Engineers, Data Scientist and  
Developers**

March 15<sup>th</sup> 2016

IBM McLean Tech Center  
8401 Greensboro Drive, 1st Floor, McLean, VA 22102

**Instructors:**

Carlo Appugliese  
Braden Callahan  
Mokhtar Kandil  
Brandon Mackenzie  
Jacques Roy  
Daniel Kikuchi

**Handcrafted for life**



## Summary

IBM is offering a free proof of technology for clients and practitioners on Apache Spark. This is a full day of education on Spark with hands on exercises instructed in person by Spark experts. The POT will provide a detail overview of Apache Spark. The exercises will be performed on Jupyter notebooks with publicly available datasets. Participants will use IBM's fully managed free Cloud platform available for educational purposes.

### Title:

"Hands on Introduction to Apache Spark for Data Engineers, Data Scientist and Developers"

### Who should go:

"Anyone interested in learning more about Apache Spark."

### Prerequisite:

"A working knowledge of Coding (Preferred Python and/or Scala), understand distributed computing, Spark and SQL."

### What to expect:

"Expect to spend a full day of lecture and hands on exercises attacking real-world data challenges using Apache Spark. In 8 hours you will learn the basic essentials of Apache Spark and why it's important to your organization. This workshop will focus on data wrangling and machine learning."

### Full Day Agenda:

8:30am - 9am	Breakfast, Socialize
9:00am – 10:00am	Kickoff, Apache Spark Overview
10:00am – 11:00am	Lab 1, Hello Spark - Hand on exercise
11:00 am – 12:00pm	Apache Spark SQL Overview
12:00 pm – 1pm	Lunch
1pm – 2pm	Lab 2, Spark SQL - Hands on exercises
2pm – 3pm	Overview of Data Science & Machine Learning w/ Apache Spark
3pm – 4pm	Lab 3, Machine Learning w/ Spark – Hands on exercises
4pm – 4:30pm	Wrap up – Feedback from attendees

