Spark the Definitive Guide 2nd Edition
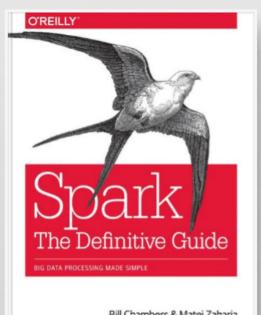
Chapter 04

Structured API Overview

Structured API Overview

Text Book

Objectives and Outcomes

- ▶ Introduced to Spark's Structured APIs, DataSets, DataFrames, and SQL Views
- ▶ Learn how Spark transforms into a physical execution plan on a cluster

Review

So far:

- ▶ We learned about Spark's programming model
- ▶ We learned how to run production code
- ▶ We were introduced to type-safe data structures in Spark
- ▶ We were introduced to Structured Streaming on Spark
- ▶ We were introduced to Machine Learning on Spark
- ▶ We were introduced to 3rd party Spark packages

Spark Overview

Conclusion

- We learned Spark's programming model
- We learned how to run production code
- We were introduced to type-safe data structures in Spark
- We were introduced to Structured Streaming on Spark
- We were introduced to Machine Learning on Spark
- We were introduced to 3rd party Spark packages