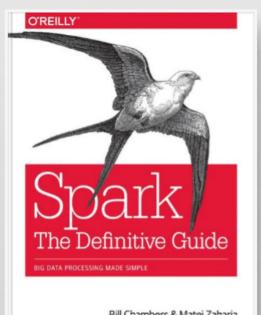
Spark the Definitive Guide 2nd Edition

Chapter 05

Basic Structured Operations

Basic Structured Operations

Text Book



Bill Chambers & Matei Zaharia

Objectives and Outcomes

- Introduce the tools we will use to manipulate DataFrames
- ► Focus on fundamental DataFrame operations

Review

So far:

- ▶ We were introduced to Spark's Structured APIs, Datasets, DataFrames, and SQL Views
- We learned how Spark transform a logical plan into a physical execution plan on a cluster

Conclusion

► Conclusion here

Questions

- ► Any questions?
- ▶ Read Chapter 06 and do any exercises in the book.