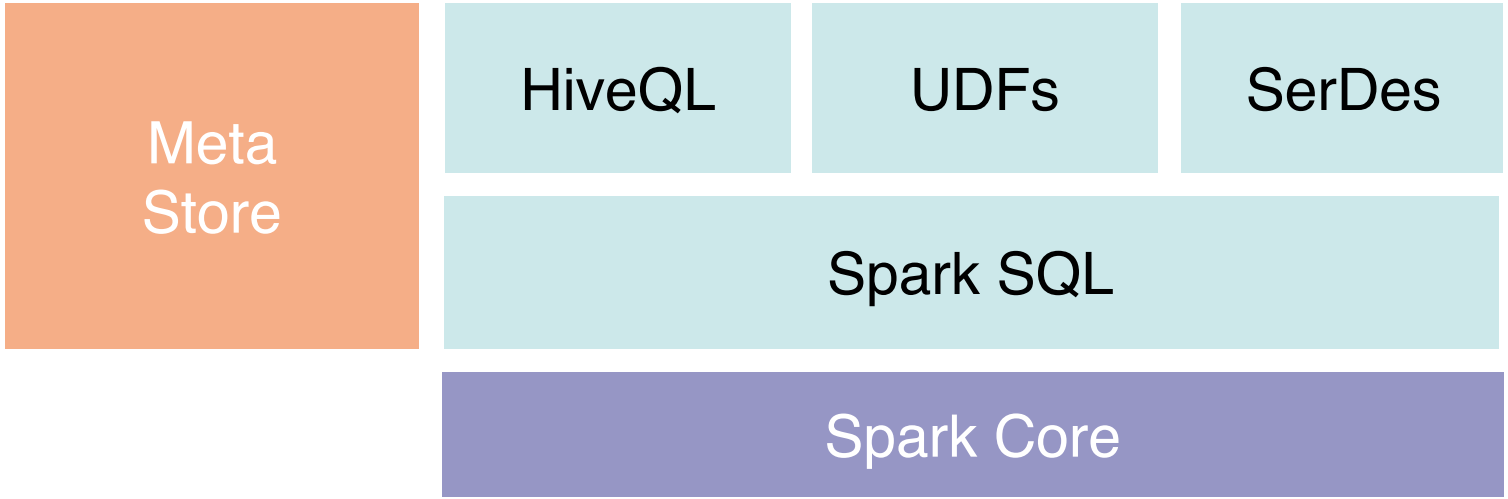
livelessons
video instruction from technology experts

Lesson 2: Spark Programming APIs

2.10 Spark SQL: Integration into Existing Workflows

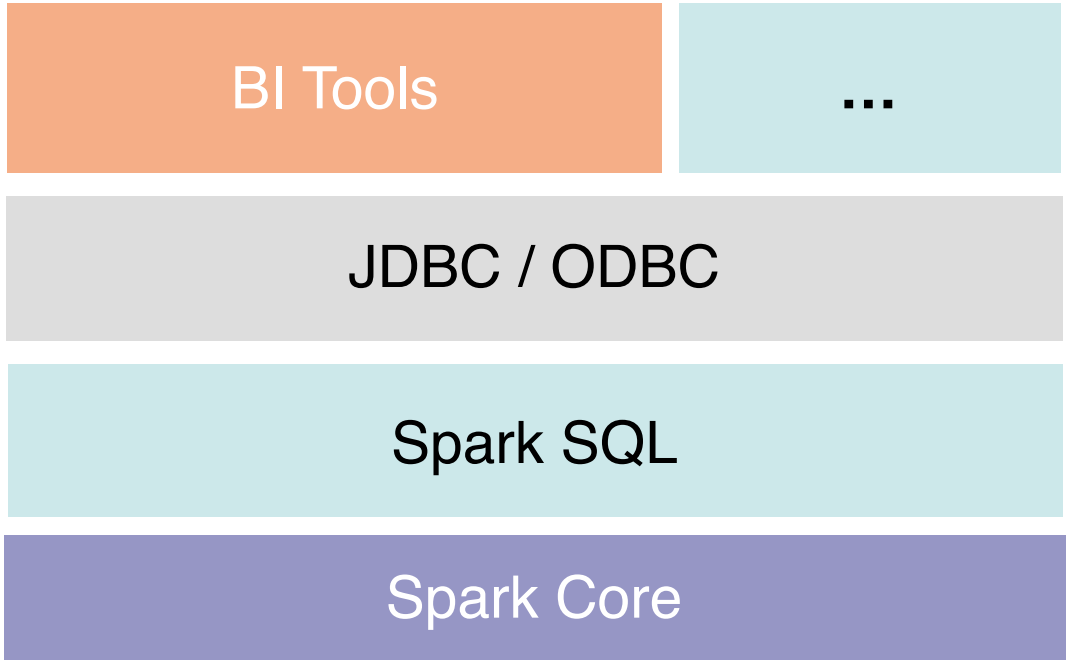
livelessons
©2016 Pearson, Inc.

Hive Integration



livelessons
©2016 Pearson, Inc.

Standard RDBMS Integration



livelessons
©2016 Pearson, Inc.

Review

- Interoperate SQL with other APIs (Python, R, etc.)
- Some data is best represented in tabular form
- Declarative syntax of SQL more natural (and accessible) for complex filters/aggregates
- Utilize performance gains of a query optimizer
- Leverage existing technologies and skill sets

livelessons
©2016 Pearson, Inc.

Next Up: Exploratory Data Analysis

livelessons
©2016 Pearson, Inc.