Getting started

INTRODUCTION TO SPARK WITH SPARKLYR IN R

Richie Cotton
Course Creator at DataCamp

datacamp

APACHE
Spark
TM

sparklyr

Connect-work-disconnect

1. Connect using `spark_connect()`
2. Work
3. Disconnect using `spark_disconnect()`

dplyr verbs

1. Select columns using `select()`
2. Filter rows using `filter()`
3. Arrange rows using `arrange()`
4. Change or add columns using `mutate()`
5. Calculate summary stats using `summarize()`

Let's practice!

INTRODUCTION TO SPARK WITH SPARKLYR IN R