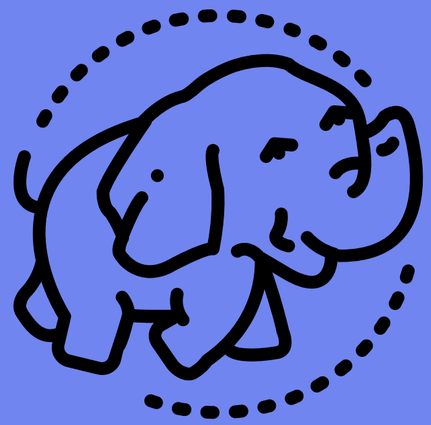
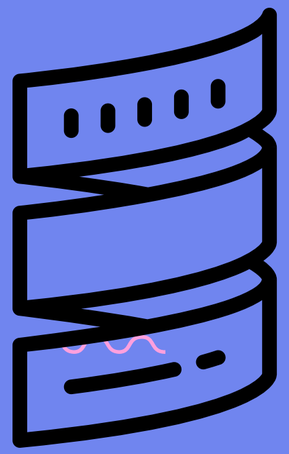
BIG DATA HANDLING

SCALA DEMO

BY: VANSHIKA AGGARWAL

OBJECTIVE





01 UNDERSTANDING RDD

02 FUNCTIONS IN RDD

03 DIRECTLY UPLOADING CSV TO SPARK SHELL

04 CONVERTING DF TO TEMP TABLE

05 DIFFERENT VARIABLES

06 TYPE INFERENCE

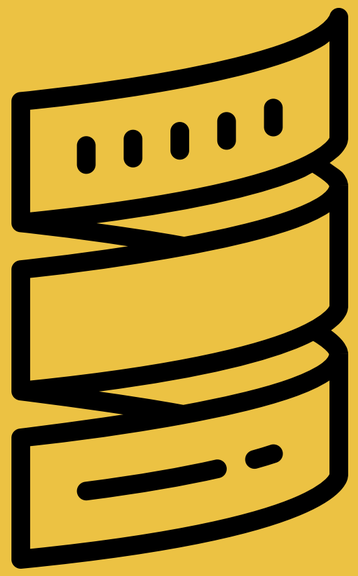
07 STRING OPERATIONS

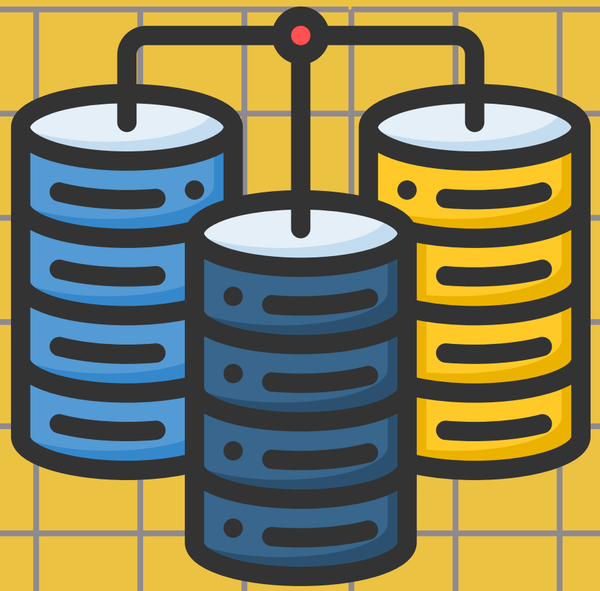
08 STRING SPLITTING

09 CONDITIONAL STATEMENTS

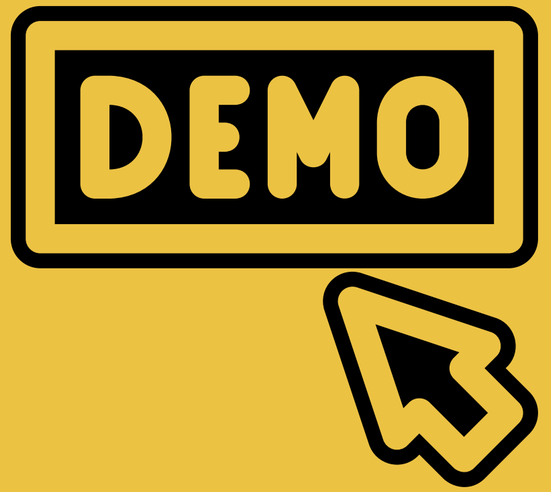
10 PATTERN MATCHING

DEMO





APACHE Spark™



ADVANTAGES OF SCALA

Scala combines object-oriented and functional programming in one concise, high-level language. Scala's static types help avoid bugs in complex applications, and its JVM and JavaScript runtimes let you build high-performance systems with easy access to huge ecosystems of libraries.

LESS CODE

Scala has an exact syntax, eliminating boilerplate code. Programs written in Scala require less code than similar programs written in Java

WIDE RANGE OF LIBRARIES

Besides accessing to Java's vast ecosystem, Scala has a wide verity of native libraries for scientific computing and big data projects

TYPE INFERENCE

Scala has a built-in type inference that allows to omit certain type annotations

THANK YOU!

THAT WAS MY TIME.
NOW IT'S YOUR TIME TO ASK QUESTIONS!!!