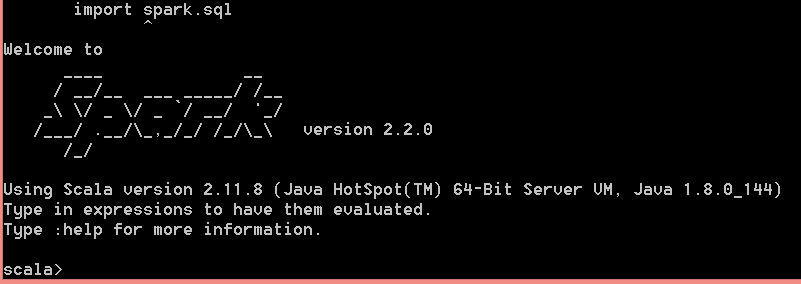
Tugas big data Apache spark into example

Perintah untuk memluai apche spark

./bin/spark-shell



* perintah dibawah merupakan dataset dengan nama README

scala> **val** textFile **=** spark.read.textFile("README.md")

textFile**:** org.apache.spark.sql.Dataset[String] **=** [value: string]

hasil:



* disini kita bisa mengambil nilai dari dataset secara langsung.dengan mengambil beberapa perintah,atau mengubah dataset untuk mendapatkan yang baru.

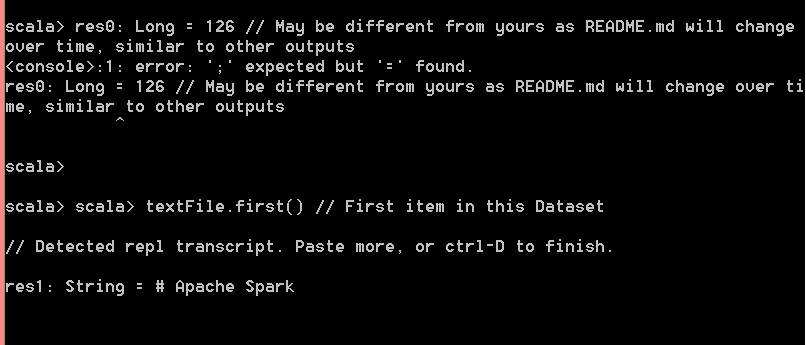
scala> textFile.count() *// Number of items in this Dataset*

res0**:** Long = 126 *// May be different from yours as README.md will change over time, similar to other outputs*

scala> textFile.first() *// First item in this Dataset*

res1**:** String = **#** **Apache** **Spark**

hasil:



* perintah ini digunakan mengubah dataset yang lama menjadi yang baru dengan mengumpulkan item di dalam file

scala> **val** linesWithSpark **=** textFile.filter(line **=>** line.contains("Spark"))

linesWithSpark**:** org.apache.spark.sql.Dataset[String] **=** [value: string]

hasil:

