**TUGAS BIG PRAKTIK BIG DATA UAS**

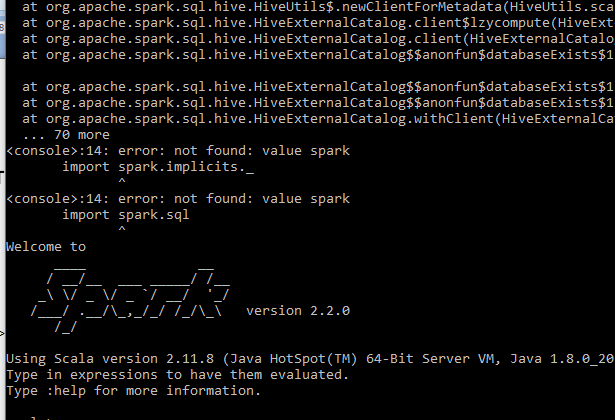
**(Apache Spark intro examples)**

NAMA : HERIBERTUS ALYSIUS NELE

NIM : 155610066

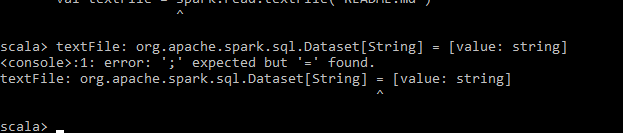
PRODI : SISTEM INFORMASI/S1

Yang pertama kita buka cmd lalu ketikan perintah > spark-shell maka akan tampil seperti berikut ini:



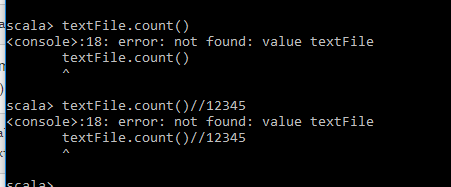
Kemudian ketikan perintah berikut > **val** textFile **=** spark.read.textFile("README.md")

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

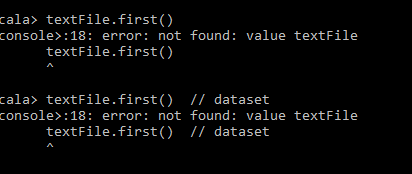


Tampilan di bawah ini masi eror.

Anda bisa mendapatkan nilai dari Dataset secara langsung, dengan memanggil beberapa tindakan, atau mengubah Dataset untuk mendapatkan yang baru



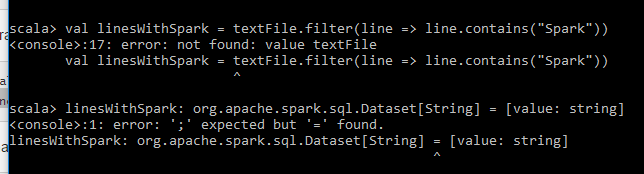
Di atas ini kita menggunakan perintah textfile.count() tapi hasilnya masih erorr kemudian kita melanjutkan perintah > scala> textFile.first()



Hasilnya masih eror

Sekarang mari kita ubah Dataset ini menjadi yang baru. Kami memanggil filter untuk mengembalikan Dataset baru dengan subkumpulan item dalam file. Yang pertama kita menggunakan perintah scala> val linesWithSpark = textFile.filter(line => line.contains("Spark"))

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



Sekarang Kita bisa menggabungkan transformasi dan tindakan bersama dengan menggunakan perintah

Sebagai berikut scala> textFile.filter(line **=>** line.contains("Spark")).count() *// How many lines contain "Spark"?*

res3**:** Long = 15

