**TUGAS BIG PRAKTIK BIG DATA UAS**

**(Apache Spark SQL)**

NAMA : HERIBERTUS ALYSIUS NELE

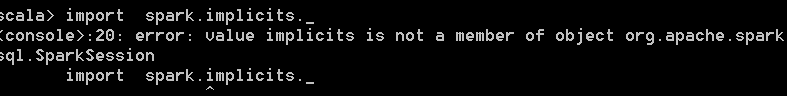
NIM : 155610066

PRODI : SISTEM INFORMASI/S1

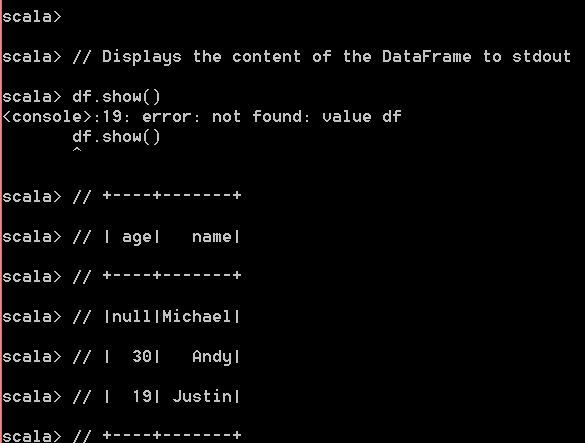
* Untuk masuk ke semua fungsi spark ada di sparksession class,dengan perintah sparksession.builder().



Hasil:

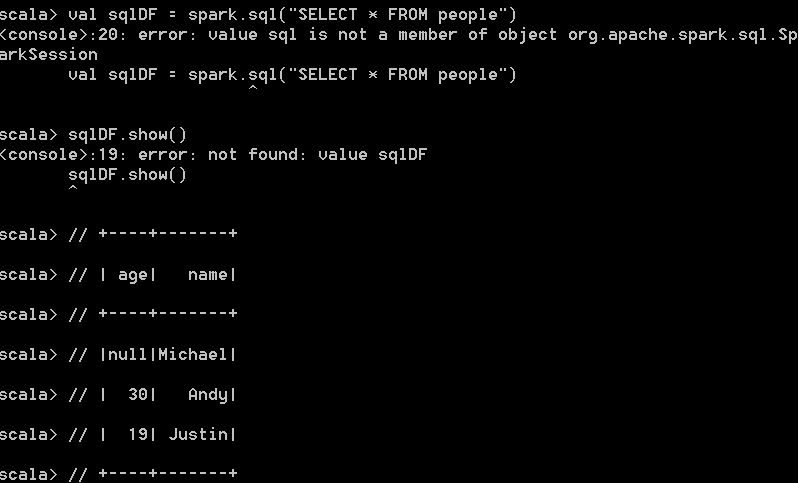


* Membuat data frame ,dengan sparksession aplikasi dapat membuat dataframe dari adaRDDD, table hive, atau dari sumber data spark



* Menjalankan query sql secara terprogram

Sql pada fungsi spark session memungkinkan aplikasi untuk menjalankan query sql pemograman dan mengembalikan hasilny sebagai dataframe



*// Note: Case classes in Scala 2.10 can support only up to 22 fields. To work around this limit,*

*// you can use custom classes that implement the Product interface*

**case** **class** **Person**(name**:** String, age**:** Long)

*// Encoders are created for case classes*

**val** caseClassDS **=** **Seq**(**Person**("Andy", 32)).toDS()

caseClassDS.show()

*// +----+---+*

*// |name|age|*

*// +----+---+*

*// |Andy| 32|*

*// +----+---+*

*// Encoders for most common types are automatically provided by importing spark.implicits.\_*

**val** primitiveDS **=** **Seq**(1, 2, 3).toDS()

primitiveDS.map(**\_** + 1).collect() *// Returns: Array(2, 3, 4)*

*// DataFrames can be converted to a Dataset by providing a class. Mapping will be done by name*

**val** path **=** "examples/src/main/resources/people.json"

**val** peopleDS **=** spark.read.json(path).as[Person]

peopleDS.show()

*// +----+-------+*

*// | age| name|*

*// +----+-------+*

*// |null|Michael|*

*// | 30| Andy|*

*// | 19| Justin|*

*// +----+-------+*

