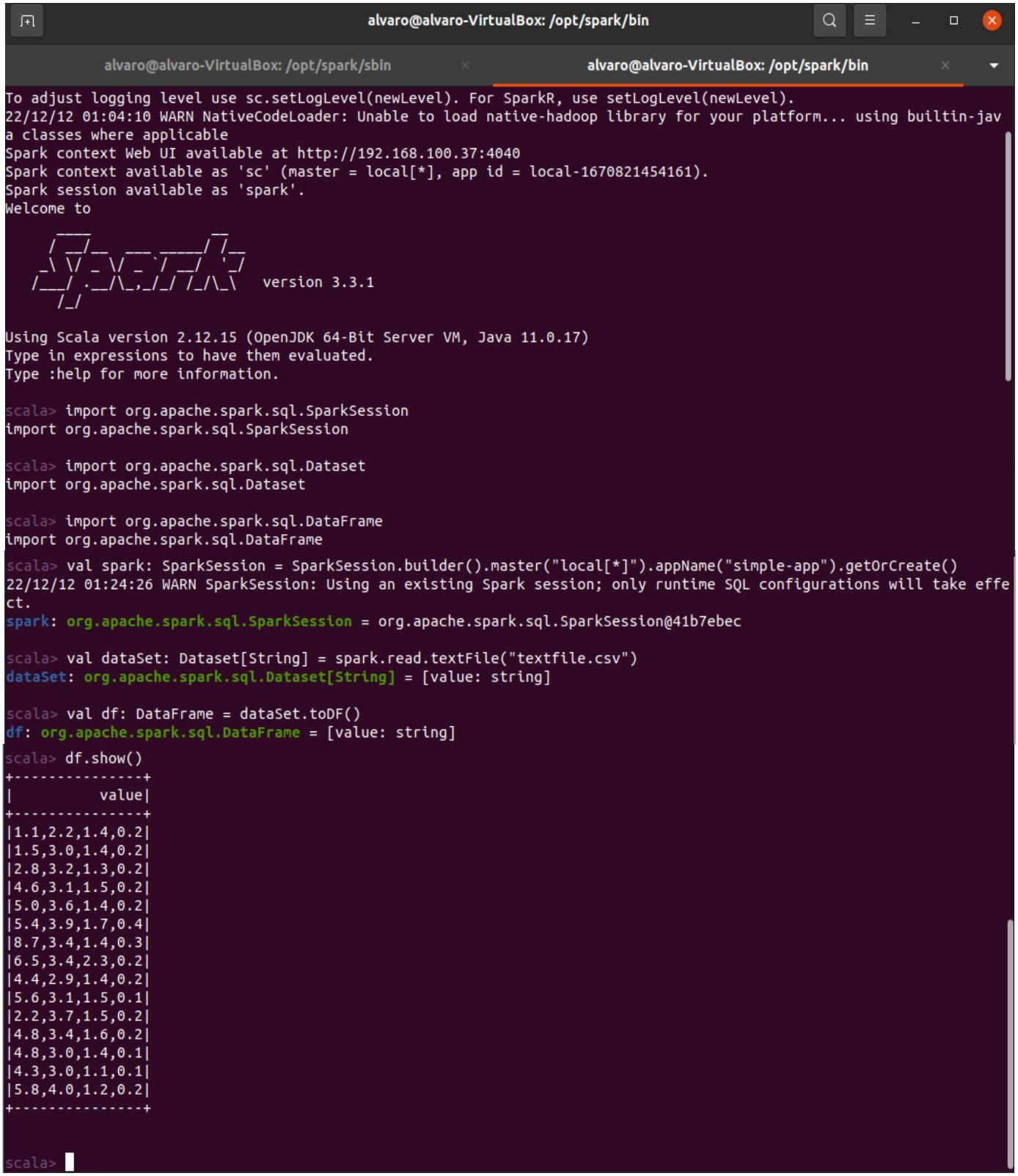
2. Realice el siguiente código, documente su funcionamiento en apache spark

Sesiones

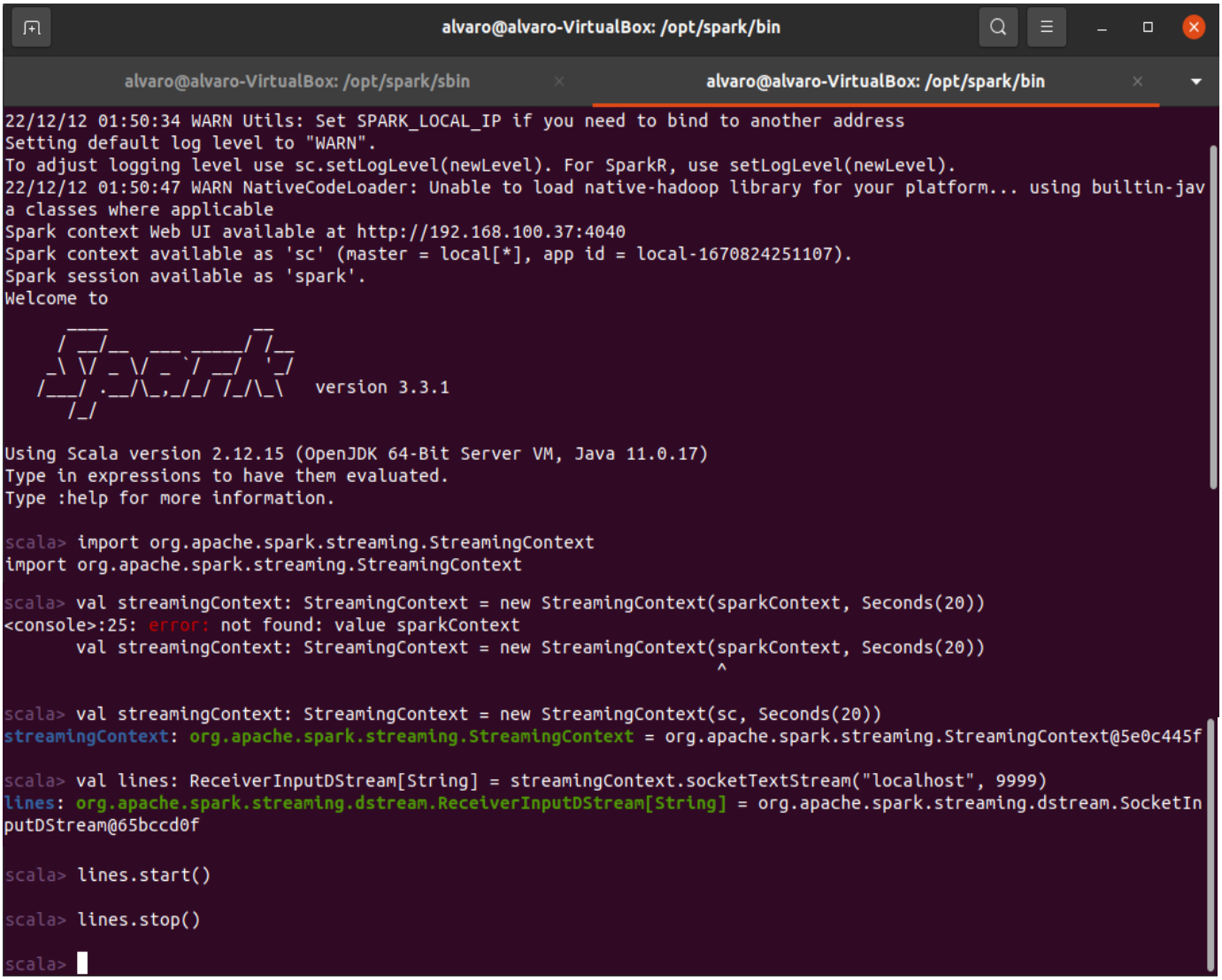
```
val spark: SparkSession = SparkSession.builder()
   .master("local[*]")
   .appName("simple-app")
   .getOrCreate()

val dataSet: Dataset[String] = spark.read.textFile("textfile.csv")
val df: DataFrame = dataSet.toDF()
```



Streaming

```
val streamingContext: StreamingContext = new StreamingContext(sparkContext, Seconds(20))
val lines: ReceiverInputDStream[String] = streamingContext.socketTextStream("localhost", 9999)
```



RDD

```
val cadenas = Array(“Docentes”, “inteligenciaArtificial”, “quefinal”)
val cadenasRDD = sc . parallelize (cadenas)
cadenasRDD.collect()
file.collect()
val filtro = cadenasRDD.filter(line => line.contains(“quefinal”))
val fileNotFound = sc.textFile(“/7añljdlsjd/alkls/”, 6)
fileNotFound.collect()
```