Spark in Scala

First Program In Spark (02-Spark-First-Program)

We saw in our walkthrough that we count the number of stock quotes for AAPL using code that looks like this:

```scala
val rdd = sc.textFile("hdfs:///data/stocks-flat/input")
val aapl = rdd.filter( line => line.contains("AAPL") )
aapl.count
```

Try using that pattern to count the number of lines in the file /data/shakespeare/input that contain the word "king". You'll want to use the toLowerCase function before the contains function to ensure all the text is lower case.

One solution looks like this:

```scala
val rdd = sc.textFile("hdfs:///data/shakespeare/input")
val kings = rdd.filter( line => line.toLowerCase().contains("king"))
kings.count
```

You should have gotten 4773 lines that reference a king in all of Shakespeare. That's almost 3% of the 175,376 lines in all of Shakepeare's plays and poems.