```
Administrator: Command Prompt - spark-shell

C:\Spark\spark-2.4.8-bin-hadoop2.6\cd bin

C:\Spark\spark-2.4.8-bin-hadoop2.6\bin\spark-shell

21/96/15 10:27:09 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN":

To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).

Spark context Web UI available at http://vic:4040
Spark context available as 'sc' (master = local1*1, app id = local-1623733042535).

Spark session available as 'spark'.

Welcome to

Using Scala version 2.11.12 (Java HotSpot(TM) Client UM, Java 1.8.0_241)

Type in expressions to have them evaluated.

Type :help for more information.
```

1 Loading Data

```
Layout
Sam
                                                                                                                                                                                                 Ċ
                                                                                                                                                                   _ 🗆
        C:5.
                                                        Administrator: Command Prompt - spark-shell
       A is 10
                                                                                                                                                                                               laı
       scala> car cards = sc.textFile("G:\Sampledata\largedeck.txt")
<console>:1: error: invalid escape character
car cards = sc.textFile("G:\Ṣampledata\largedeck.txt")
                                                                                                                                                                                               Te
       <console>:1: error: invalid escape character
car cards = sc.textFile("G:\Sampledata\largedeck.txt")
       scala> car cards = sc.textFile("G:/Sampledata/largedeck.txt")
<console>:26: error: not found: value car
val $ires7 = car.cards
                                                                                                                                                                                               Dat
        <console>:24: error: not found: value car
    car cards = sc.textFile<"G:/Sampledata/largedeck.txt">
                                                                                                                                                                                               Siz
       scala> var cards = sc.textFile("G:/Sampledata/largedeck.txt")
21/06/15 10:41:20 WARN SizeEstimator: Failed to check whether UseCompressedOops
is set; assuming yes
cards: org.apache.spark.rdd.RDD[String] = G:/Sampledata/largedeck.txt MapPartiti
onsRDD[1] at textFile at <console>:24
        scala>
```

2 Filter Black Cards

3 Counting Black Cards

```
_ _ _
                          Administrator: Command Prompt - spark-shell
C:4.
Kconsole>:24: error: not found: value cls
scala> var blackcards =cards.filter(value => value.contains("BLACK">>
blackcards: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[3] at filter at
{console>:25
(0
(1
(4
(7
(8
(12
(14
(16
(17
(19
(20
                                                                                 + 4)
+ 4)
+ 4)
+ 4)
+ 5)
+ 4)
+ 3)
+ 2)
+ 1)
       1:------>
1:----->
1:----->
1:----->
1:----->
1:----->
[Stage
[Stage
[Stage
[Stage
res6: Long = 27262976
scala≻
```

4 Filter Black Queens from Cards

```
Administrator: Command Prompt - spark-shell

| Console>:25: error: type mismatch; found : String("q") required: Int var blackqueens =cards.filter(value => value.contains("BLACK")).filter(value => value("q"))
| Console>:25: error: type mismatch; found : String("q") required: Int var blackqueens =cards.filter(value => value.contains("BLACK")).filter(value => value("q")) required: Int var blackqueens =cards.filter(value => value.contains("BLACK")).filter(value => value("q"))
| Console>:25: error: type mismatch; found : String("q") required: Int var blackqueens =cards.filter(value => value.contains("BLACK")).filter(value => value.contains("BLACK"))
```

5 Count Black Queens from Cards

```
_ 🗆
C:4.
                    Administrator: Command Prompt - spark-shell
scala> var blackqueens =cards.filter(value => value.contains("BLACK">>).filter(va
lue => value.contains("Q">)
blackqueens: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[5] at filter at
<console>:25
(0
(4
(8
(9
(11
(12
(13
(16
(17
(19
(20
(21
                                                                      22]
22]
22]
22]
22]
22]
22]
22]
                                                                 4) 4) 4) 3) 2) 1)
                                                                      22 1
22 1
22 1
22 1
22 1
      2:---->
[Stage
[Stage 2:----->
res7: Long = 2097152
scala>
```