

PrjWk2



default ▼

```
val libraryEvtText = sc.textFile("/home/sravan/Desktop/CapeStone/aarhus_libraryEvents_dl.csv")
```

```
libraryEvtText: org.apache.spark.rdd.RDD[String] = /home/scarface/Desktop/sem-3/Cape/aarhus_libraryEvents_d1.csv
MapPartitionsRDD[61] at textFile at <console>:27
```

Took 1 sec. Last updated by anonymous at January 29 2017, 7:39:33 AM. (outdated)

```
val df = sqlContext.read.format("com.databricks.spark.csv").option("header", "true").option("delimiter", ";").load("/home/sravan/Desktop/CapeStone/aarhus libraryEvents edit.csv").cache()
```

```
df: org.apache.spark.sql.Dataset[org.apache.spark.sql.Row] = [lid: string, city: string ... 21 more fields]
```

Took 1 sec. Last updated by anonymous at January 29 2017, 8:34:54 AM.

```
case class LibEvt(libraryId:String, city:String, endtime : String, title : String, url : String, price: String,
imageurl : String, teaser : String , street : String ,status : String ,longitude : String ,starttime : String ,l
```

```
defined class LibEvt
```

Took 0 sec. Last updated by anonymous at January 29 2017, 7:44:52 AM.

```
val libEvt = libraryEvtText.map(s=>s.split('|')).filter(s=>s(0)!="lid").map(  
s=>LibEvtnt(s(0), s(1).replaceAll(","," "), s(2), s(3).replaceAll("\\",""), s(4).replaceAll(","," "), s(5)  
s(9).replaceAll("\\","").replaceAll(","," "), s(10), s(11).replaceAll("\\","").replaceAll(","," "), s(12).  
s(14), s(15), s(16), s(17), s(18))  
)
```

```
libEvt: org.apache.spark.rdd.RDD[LibEvt] = MapPartitionsRDD[64] at map at <console>:31
```

Took 1 sec. Last updated by anonymous at January 29 2017, 7:39:35 AM.

```
df.toDF().registerTempTable("library")
```

warning: there was one deprecation warning; re-run with -deprecation for details

Took 1 sec. Last updated by anonymous at January 29 2017, 8:34:58 AM.

```
%sql
select * from library limit 5
```



lid

1

<p><img title="" src=""https://www.aakb.dk/files/content_images/bogmaerkertilweb.jpg"" alt="" width=""267



PrjWk2



◀

2/2