



# ExampleNote

## Introduction

### This notebook supports multiple language backend

- default: scala with SparkContext available as sc
- %md: markdown
- %sql: hive on spark by default. hiveContext object is also available by default. If you specify zeppelin.spark.useHiveContext as false in interpreter settings then this will use spark sql and sqlContext object becomes available
- %pyspark: pyspark
- %sh: shell

Interpreters are loaded lazily and hence takes some time the first time.

Note uses a websocket for communication. So changes propagate to all open instances of this notebook instantaneously

User can configure parameters with which sc object is created by going to interpreters.

User can create as many sc interpreters as they want with different settings

Qubole Notes supports auto assist for code completion. You can view the suggestions using the key combination **Ctrl + .(period)**

## Note that

- If you use sc.stop in the notebook then spark context will stop and you have to go to interpreter page and click on restart to restart that interpreter again.

## Run Spark Command

```
res0: String = 1.3.0
```

## Scala SparkPi Example

```
import scala.math.random
import org.apache.spark._
slices: Int = 6
n: Int = 600000
count: Int = 471305
Pi is roughly 3.1420333333333335
```

## Top 10 Words In Dataset In S3

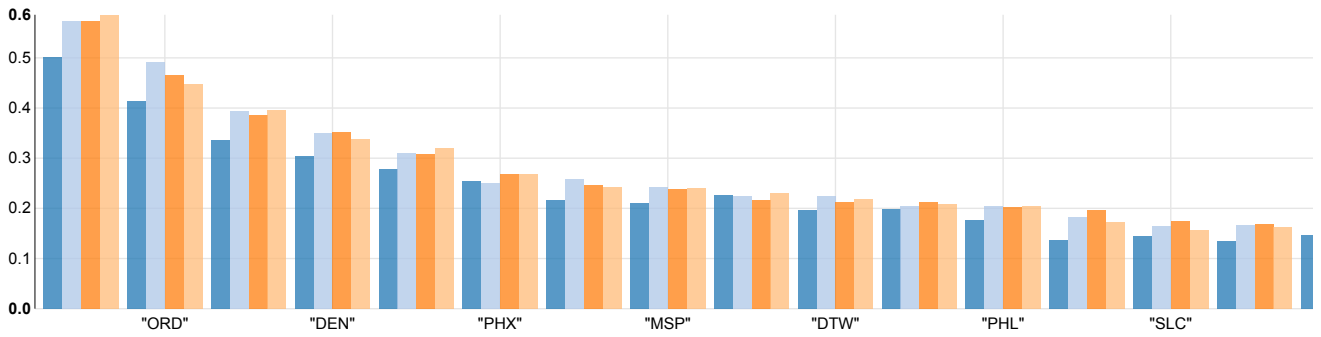
```
file: org.apache.spark.rdd.RDD[String] = s3://paid-qubole/default-datasets/gutenberg/pg20417.txt MapPartitionsRDD[3] at textFile at <console>:28
counts: Array[(String, Int)] = Array(("",14637), (the,7906), (of,5425), (and,2759), (a,2422), (to,2168), (is,2068), (in,2048), (that,1273), (are,921))
(,14637)
(the,7906)
(of,5425)
(and,2759)
(a,2422)
(to,2168)
(is,2068)
(in,2048)
(that,1273)
(are,921)
```

## Spark Sql

### sql example

- while using hive on spark you can fetch list of hive tables in qubole and query those tables

## Visualizations On Airline Data For Four Quarters



| site  | ts                  | phr  | Inks                             |
|---|---------------------|--|----------------------------------|
| <a href="http://eighteezbaby.com/2008/12/30/you-know-you-are-new-money-whenwill-post-funny-pics-later">http://eighteezbaby.com/2008/12/30/you-know-you-are-new-money-whenwill-post-funny-pics-later</a> | 2009-01-01 00:00:50 | []   | []                               |
| <a href="http://thelondonreviewer.com/2008/12/31/uk-sunday-times-round-up-28-dec-08">http://thelondonreviewer.com/2008/12/31/uk-sunday-times-round-up-28-dec-08</a>                                     | 2009-01-01 00:01:10 | ["not only do they patients get to lose their love handles or chubby belly but they get to take part in saving the earth"] | ["http://entertainment.timeso"]  |
| <a href="http://thelondonreviewer.com/2008/12/28/film-review-twilight">http://thelondonreviewer.com/2008/12/28/film-review-twilight</a>   | 2009-               | []   | ["http://imdb.com/title/tt1099"] |

### Pyspark Pi Calculation Example

```
u'1.3.0'
'2.6.9 (unknown, Sep 13 2014, 00:25:11) \n[GCC 4.8.2 20140120 (Red Hat 4.8.2-16)]'
Pi is roughly 3.143740
```

### Temp Table Example: Prepare Data

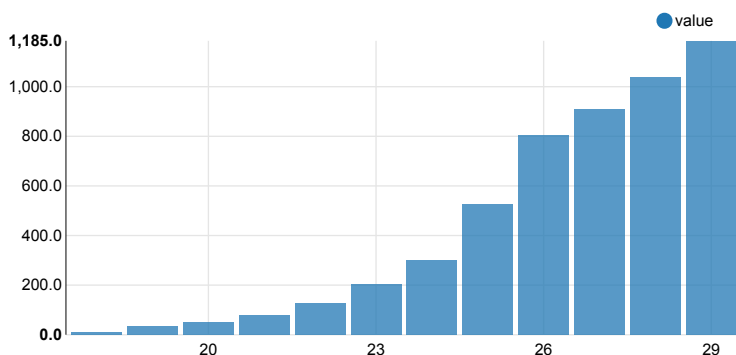
```
Archive: bank.zip
  inflating: /tmp/data/bank-full.csv
  inflating: /tmp/data/bank-names.txt
  inflating: /tmp/data/bank.csv
bank.csv
bank-full.csv
bank-names.txt
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
```

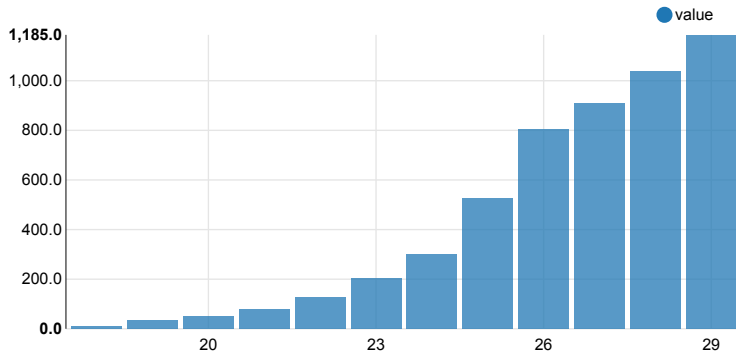
```
15/04/08 18:38:07 INFO Configuration.deprecation: fs.default.name is deprecated. Instead, use fs.defaultFS
15/04/08 18:38:09 INFO Configuration.deprecation: fs.default.name is deprecated. Instead, use fs.defaultFS
```

### Temp Table Example: Load Data In Table

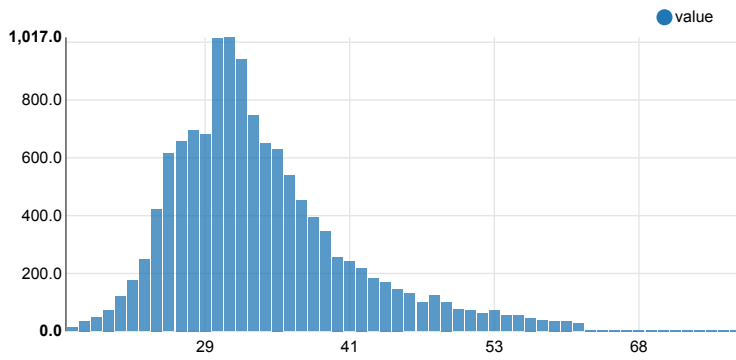
```
bankText: org.apache.spark.rdd.RDD[String] = /tmp/bank-full.csv MapPartitionsRDD[49] at textFile at <console>:27
defined class Bank
bank: org.apache.spark.sql.DataFrame = [age: int, job: string, marital: string, education: string, balance: int]
```

maxAge





marital



Other Features

simple printing data

- printIn can be used
- data can be printed in tabular form using %table
- input forms are also supported

```
age      numbers
2        10
3        20
8        70
12       50
20       10
```

| age | numbers |
|-----|---------|
| 2   | 10      |
| 3   | 20      |
| 8   | 70      |
| 12  | 50      |
| 20  | 10      |

maxAge

| <b>age</b> | <b>numbers</b> |
|------------|----------------|
| 2          | 10             |
| 3          | 20             |
| 12         | 70             |
| 12         | 50             |
| 20         | 10             |