FIAP – FACULDADE DE INFORMÁTICA E ADMINISTRAÇÃO PAULISTA

BRUNO BIANCCHI – RM 84351

LUIS HENRIQUE CALDAS ALTERO – RM 88670

PEDRO GUILHERME POLLONI BARRETO - RM 88964

VITOR LAMPRECHT – RM 86691

Sprint 4 – Arquitetura de BI & Big Data

São Paulo

2021

Comando: cd Downloads

```
[oracle@bigdatalite ~]$ cd Downloads
[oracle@bigdatalite Downloads]$ █
```

Comando: pig -x local sentiment_analysis_challenge.pig

```
[oracle@bigdatalite Downloads]$ pig -x local sentiment_analysis_challenge.pig
log4j:WARN No appenders could be found for logger (org.apache.hadoop.util.Shell)
.
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more in
fo.
2021-09-20 11:09:47,229 [main] INFO  org.apache.pig.Main - Apache Pig version 0.
12.0-cdh5.7.0 (rexported) compiled Mar 23 2016, 11:34:31
2021-09-20 11:09:47,229 [main] INFO  org.apache.pig.Main - Logging error message
s to: /home/oracle/Downloads/pig_1632150587213.log
2021-09-20 11:09:47,243 [main] INFO  org.apache.hadoop.conf.Configuration.deprec
ation - user.name is deprecated. Instead, use mapreduce.job.user.name
2021-09-20 11:09:47,387 [main] INFO  org.apache.pig.impl.util.Utils - Default bo
otup file /home/oracle/.pigbootup not found
2021-09-20 11:09:47,438 [main] INFO  org.apache.hadoop.conf.Configuration.deprec
ation - fs.default.name is deprecated. Instead, use fs.defaultFS
2021-09-20 11:09:47,439 [main] INFO  org.apache.hadoop.conf.Configuration.deprec
ation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.addr
ess
2021-09-20 11:09:47,440 [main] INFO  org.apache.pig.backend.hadoop.executionengi
ne.HExecutionEngine - Connecting to hadoop file system at: file:///
```

```
Success!

Job Stats (time in seconds):
JobId   Alias   Feature Outputs
job_local1588913684_0001        clean_tokens,dictionary,english_tweets,token_sen
timent,tokenized,tweets HASH_JOIN       file:///home/oracle/Downloads/sentiment_
analysis_teste,

Input(s):
Successfully read records from: "file:///home/oracle/Downloads/united_airlines_t
weets.tsv"
Successfully read records from: "file:///home/oracle/Downloads/dictionary.tsv"

Output(s):
Successfully stored records in: "file:///home/oracle/Downloads/sentiment_analysi
s_teste"

Job DAG:
job_local1588913684_0001


2021-09-20 11:10:50,032 [main] INFO  org.apache.pig.backend.hadoop.executionengi
ne.mapReduceLayer.MapReduceLauncher - Success!
[oracle@bigdatalite Downloads]$ █
```

Comando: hadoop fs -copyFromLocal sentiment_analysis_teste/part-r-00000 analise_sentimento

```
[oracle@bigdatalite Downloads]$ hadoop fs -copyFromLocal sentiment_analysis_test
e/part-r-00000 analise_sentimento
```

Comando: hadoop fs -ls

```
[oracle@bigdatalite Downloads]$ hadoop fs -ls
Found 11 items
drwxr-xr-x   - oracle oracle          0 2016-05-23 20:42 .Trash
drwxr-xr-x   - oracle oracle          0 2021-09-14 08:26 .sparkStaging
drwx------   - oracle oracle          0 2021-08-31 09:06 .staging
-rw-r--r--   1 oracle oracle      12085 2021-09-20 09:52 analise_sentimento
drwxr-xr-x   - oracle oracle          0 2021-09-14 08:11 dados
drwxr-xr-x   - oracle oracle          0 2016-06-01 18:59 mediademo
drwxr-xr-x   - oracle oracle          0 2016-05-15 12:02 moviedemo
drwxr-xr-x   - oracle oracle          0 2016-05-15 12:03 moviework
drwxr-xr-x   - oracle oracle          0 2016-05-15 12:03 oggdemo
drwxr-xr-x   - oracle oracle          0 2016-05-15 12:03 oozie-oozi
-rw-r--r--   1 oracle oracle      10563 2021-09-14 08:14 population.csv
[oracle@bigdatalite Downloads]$
```

Comando: hadoop fs -chmod 664 analise_sentimento

```
[oracle@bigdatalite Downloads]$ hadoop fs -chmod 664 analise_sentimento
[oracle@bigdatalite Downloads]$
```

Comando: spark-shell

```
[oracle@bigdatalite Downloads]$ spark-shell
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel).
Welcome to
      ____              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /___/ .__/\_,_/_/ /_/\_\   version 1.6.0
      /_/

Using Scala version 2.10.5 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_91)
Type in expressions to have them evaluated.
Type :help for more information.
21/10/10 13:04:08 WARN util.Utils: Your hostname, bigdatalite.localdomain resolv
es to a loopback address: 127.0.0.1; using 10.0.2.15 instead (on interface eth1)
21/10/10 13:04:08 WARN util.Utils: Set SPARK_LOCAL_IP if you need to bind to ano
ther address
Spark context available as sc (master = yarn-client, app id = application_163162
2112260_0011).
SQL context available as sqlContext.
```

Comando: val analise = sc.textFile("analise_sentimento")

```
scala> val analise = sc.textFile("analise_sentimento")
analise: org.apache.spark.rdd.RDD[String] = analise_sentimento MapPartitionsRDD[
4] at textFile at <console>:27
```

Comando: analise.first()

```
scala> analise.first()
res3: String = 475641126129369088      wow      wow      1
```

Comando: analise.filter(line => line.contains("thanks")).count

```
scala> analise.first()
res3: String = 475641126129369088      wow      wow      1

scala> analise.filter(line => line.contains("thanks")).count
res4: Long = 36
```

Comando: analise.filter(line => line.contains("luxurytravel")).count

```
scala> analise.filter(line => line.contains("luxurytravel")).count
res7: Long = 28
```

Comando: analise.filter(line => line.contains("great")).count

```
scala> analise.filter(line => line.contains("great")).count
res8: Long = 9
```