Импорт таблицы в Hadoop с использованием SQOOP в Hive-таблицу:

```
taya@MacBook-Air-Taisia ~ % ssh student4_1@manager.novalocal  -i /Users/taya/private_key_hadoop
Last login: Mon Jun 29 19:09:10 2020 from 46.160.232.254
[student4_1@manager ~]$ ssh node3.novalocal
Last login: Mon Jun 29 19:09:19 2020 from manager.novalocal
[student4_1@node3 ~]$ sqoop import --connect jdbc:postgresql://node3.novalocal/pg_db --username
exporter -P --table character --hive-import --hive-database student4_1 --hive-table character4_1
Warning: /opt/cloudera/parcels/CDH-5.16.2-1.cdh5.16.2.p0.8/bin/../lib/sqoop/../accumulo does not
exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
20/06/29 19:42:24 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6-cdh5.16.2
Enter password:
20/06/29 19:42:37 INFO tool.BaseSqoopTool: Using Hive-specific delimiters for output. You can
override
20/06/29 19:42:37 INFO tool.BaseSqoopTool: delimiters with --fields-terminated-by, etc.
20/06/29 19:42:37 INFO manager.SqlManager: Using default fetchSize of 1000
20/06/29 19:42:37 INFO tool.CodeGenTool: Beginning code generation
20/06/29 19:42:38 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "character"
AS t LIMIT 1
20/06/29 19:42:38 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /opt/cloudera/parcels/CDH/
lib/hadoop-mapreduce
Note: /tmp/sqoop-student4_1/compile/62a664f2fda2864fa62f9fbc98d72561/character.java uses or
overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
20/06/29 19:42:46 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-student4_1/compile/
62a664f2fda2864fa62f9fbc98d72561/character.jar
20/06/29 19:42:46 WARN manager.PostgresqlManager: It looks like you are importing from
postgresql.
20/06/29 19:42:46 WARN manager.PostgresqlManager: This transfer can be faster! Use the --direct
20/06/29 19:42:46 WARN manager.PostgresqlManager: option to exercise a postgresql-specific fast
path.
20/06/29 19:42:46 INFO mapreduce.ImportJobBase: Beginning import of character
20/06/29 19:42:47 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use
mapreduce.job.jar
20/06/29 19:42:50 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use
mapreduce.job.maps
20/06/29 19:42:50 INFO client.RMProxy: Connecting to ResourceManager at manager.novalocal/
89.208.221.132:8032
20/06/29 19:43:21 INFO db.DBInputFormat: Using read commited transaction isolation
20/06/29 19:43:21 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN("charid"),
MAX("charid") FROM "character"
20/06/29 19:43:21 WARN db.TextSplitter: Generating splits for a textual index column.
20/06/29 19:43:21 WARN db.TextSplitter: If your database sorts in a case-insensitive order, this
may result in a partial import or duplicate records.
20/06/29 19:43:21 WARN db.TextSplitter: You are strongly encouraged to choose an integral split
column.
20/06/29 19:43:24 INFO mapreduce.JobSubmitter: number of splits:6
20/06/29 19:43:25 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1592839005008_0060
20/06/29 19:43:27 INFO impl.YarnClientImpl: Submitted application application_1592839005008_0060
20/06/29 19:43:27 INFO mapreduce.Job: The url to track the job: http://manager.novalocal:8088/
proxy/application_1592839005008_0060/
20/06/29 19:43:27 INFO mapreduce.Job: Running job: job_1592839005008_0060
20/06/29 19:43:45 INFO mapreduce.Job: Job job_1592839005008_0060 running in uber mode : false
20/06/29 19:43:45 INFO mapreduce.Job:  map 0% reduce 0%
20/06/29 19:44:24 INFO mapreduce.Job:  map 17% reduce 0%
20/06/29 19:44:29 INFO mapreduce.Job:  map 33% reduce 0%
20/06/29 19:44:36 INFO mapreduce.Job:  map 50% reduce 0%
20/06/29 19:44:41 INFO mapreduce.Job:  map 67% reduce 0%
20/06/29 19:44:47 INFO mapreduce.Job:  map 83% reduce 0%
20/06/29 19:44:52 INFO mapreduce.Job:  map 100% reduce 0%
20/06/29 19:44:55 INFO mapreduce.Job: Job job_1592839005008_0060 completed successfully
20/06/29 19:44:55 INFO mapreduce.Job: Counters: 30
        File System Counters
                FILE: Number of bytes read=0
                FILE: Number of bytes written=1053846
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=803
                HDFS: Number of bytes written=92512
                HDFS: Number of read operations=24
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=12
        Job Counters
                Launched map tasks=6
                Other local map tasks=6
                Total time spent by all maps in occupied slots (ms)=121356
                Total time spent by all reduces in occupied slots (ms)=0
```

```
        Total time spent by all map tasks (ms)=30339
        Total vcore-milliseconds taken by all map tasks=30339
        Total megabyte-milliseconds taken by all map tasks=31067136
    Map-Reduce Framework
        Map input records=1954
        Map output records=1954
        Input split bytes=803
        Spilled Records=0
        Failed Shuffles=0
        Merged Map outputs=0
        GC time elapsed (ms)=914
        CPU time spent (ms)=5310
        Physical memory (bytes) snapshot=1321046016
        Virtual memory (bytes) snapshot=16763236352
        Total committed heap usage (bytes)=1160249344
    File Input Format Counters
        Bytes Read=0
    File Output Format Counters
        Bytes Written=92512
20/06/29 19:44:55 INFO mapreduce.ImportJobBase: Transferred 90,3438 KB in 125,4708 seconds
(737,3188 bytes/sec)
20/06/29 19:44:55 INFO mapreduce.ImportJobBase: Retrieved 1954 records.
20/06/29 19:44:55 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "character"
AS t LIMIT 1
20/06/29 19:44:56 INFO hive.HiveImport: Loading uploaded data into Hive

Logging initialized using configuration in jar:file:/opt/cloudera/parcels/
CDH-5.16.2-1.cdh5.16.2.p0.8/jars/hive-common-1.1.0-cdh5.16.2.jar!/hive-log4j.properties
OK
Time taken: 10.99 seconds
Loading data to table student4_1.character4_1
Table student4_1.character4_1 stats: [numFiles=6, totalSize=92512]
OK
Time taken: 2.632 secondstaya@MacBook-Air-Taisia ~ % ssh student4_1@manager.novalocal  -i /
Users/taya/private_key_hadoop
Last login: Mon Jun 29 19:09:10 2020 from 46.160.232.254
[student4_1@manager ~]$ ssh node3.novalocal
Last login: Mon Jun 29 19:09:19 2020 from manager.novalocal
[student4_1@node3 ~]$ sqoop import --connect jdbc:postgresql://node3.novalocal/pg_db --username
exporter -P --table character --hive-import --hive-database student4_1 --hive-table character4_1
Warning: /opt/cloudera/parcels/CDH-5.16.2-1.cdh5.16.2.p0.8/bin/../lib/sqoop/../accumulo does not
exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
20/06/29 19:42:24 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6-cdh5.16.2
Enter password:
20/06/29 19:42:37 INFO tool.BaseSqoopTool: Using Hive-specific delimiters for output. You can
override
20/06/29 19:42:37 INFO tool.BaseSqoopTool: delimiters with --fields-terminated-by, etc.
20/06/29 19:42:37 INFO manager.SqlManager: Using default fetchSize of 1000
20/06/29 19:42:37 INFO tool.CodeGenTool: Beginning code generation
20/06/29 19:42:38 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "character"
AS t LIMIT 1
20/06/29 19:42:38 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /opt/cloudera/parcels/CDH/
lib/hadoop-mapreduce
Note: /tmp/sqoop-student4_1/compile/62a664f2fda2864fa62f9fbc98d72561/character.java uses or
overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
20/06/29 19:42:46 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-student4_1/compile/
62a664f2fda2864fa62f9fbc98d72561/character.jar
20/06/29 19:42:46 WARN manager.PostgresqlManager: It looks like you are importing from
postgresql.
20/06/29 19:42:46 WARN manager.PostgresqlManager: This transfer can be faster! Use the --direct
20/06/29 19:42:46 WARN manager.PostgresqlManager: option to exercise a postgresql-specific fast
path.
20/06/29 19:42:46 INFO mapreduce.ImportJobBase: Beginning import of character
20/06/29 19:42:47 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use
mapreduce.job.jar
20/06/29 19:42:50 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use
mapreduce.job.maps
20/06/29 19:42:50 INFO client.RMProxy: Connecting to ResourceManager at manager.novalocal/
89.208.221.132:8032
20/06/29 19:43:21 INFO db.DBInputFormat: Using read commited transaction isolation
20/06/29 19:43:21 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN("charid"),
MAX("charid") FROM "character"
20/06/29 19:43:21 WARN db.TextSplitter: Generating splits for a textual index column.
20/06/29 19:43:21 WARN db.TextSplitter: If your database sorts in a case-insensitive order, this
may result in a partial import or duplicate records.
20/06/29 19:43:21 WARN db.TextSplitter: You are strongly encouraged to choose an integral split
column.
20/06/29 19:43:24 INFO mapreduce.JobSubmitter: number of splits:6
```

```
20/06/29 19:43:25 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1592839005008_0060
20/06/29 19:43:27 INFO impl.YarnClientImpl: Submitted application application_1592839005008_0060
20/06/29 19:43:27 INFO mapreduce.Job: The url to track the job: http://manager.novalocal:8088/
proxy/application_1592839005008_0060/
20/06/29 19:43:27 INFO mapreduce.Job: Running job: job_1592839005008_0060
20/06/29 19:43:45 INFO mapreduce.Job: Job job_1592839005008_0060 running in uber mode : false
20/06/29 19:43:45 INFO mapreduce.Job:  map 0% reduce 0%
20/06/29 19:44:24 INFO mapreduce.Job:  map 17% reduce 0%
20/06/29 19:44:29 INFO mapreduce.Job:  map 33% reduce 0%
20/06/29 19:44:36 INFO mapreduce.Job:  map 50% reduce 0%
20/06/29 19:44:41 INFO mapreduce.Job:  map 67% reduce 0%
20/06/29 19:44:47 INFO mapreduce.Job:  map 83% reduce 0%
20/06/29 19:44:52 INFO mapreduce.Job:  map 100% reduce 0%
20/06/29 19:44:55 INFO mapreduce.Job: Job job_1592839005008_0060 completed successfully
20/06/29 19:44:55 INFO mapreduce.Job: Counters: 30
        File System Counters
                FILE: Number of bytes read=0
                FILE: Number of bytes written=1053846
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=803
                HDFS: Number of bytes written=92512
                HDFS: Number of read operations=24
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=12
        Job Counters
                Launched map tasks=6
                Other local map tasks=6
                Total time spent by all maps in occupied slots (ms)=121356
                Total time spent by all reduces in occupied slots (ms)=0
                Total time spent by all map tasks (ms)=30339
                Total vcore-milliseconds taken by all map tasks=30339
                Total megabyte-milliseconds taken by all map tasks=31067136
        Map-Reduce Framework
                Map input records=1954
                Map output records=1954
                Input split bytes=803
                Spilled Records=0
                Failed Shuffles=0
                Merged Map outputs=0
                GC time elapsed (ms)=914
                CPU time spent (ms)=5310
                Physical memory (bytes) snapshot=1321046016
                Virtual memory (bytes) snapshot=16763236352
                Total committed heap usage (bytes)=1160249344
        File Input Format Counters
                Bytes Read=0
        File Output Format Counters
                Bytes Written=92512
20/06/29 19:44:55 INFO mapreduce.ImportJobBase: Transferred 90,3438 KB in 125,4708 seconds
(737,3188 bytes/sec)
20/06/29 19:44:55 INFO mapreduce.ImportJobBase: Retrieved 1954 records.
20/06/29 19:44:55 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "character"
AS t LIMIT 1
20/06/29 19:44:56 INFO hive.HiveImport: Loading uploaded data into Hive

Logging initialized using configuration in jar:file:/opt/cloudera/parcels/
CDH-5.16.2-1.cdh5.16.2.p0.8/jars/hive-common-1.1.0-cdh5.16.2.jar!/hive-log4j.properties
OK
Time taken: 10.99 seconds
Loading data to table student4_1.character4_1
Table student4_1.character4_1 stats: [numFiles=6, totalSize=92512]
OK
Time taken: 2.632 seconds
```

```
taya@MacBook-Air-Taisia ~ % ssh student4_1@manager.novalocal  -i /Users/taya/private_key_hadoop
Last login: Mon Jun 29 18:51:38 2020 from 46.160.232.254
[student4_1@manager ~]$ ssh node3.novalocal
Last login: Mon Jun 29 18:51:50 2020 from manager.novalocal
[student4_1@node3 ~]$ sqoop import --connect jdbc:postgresql://node3.novalocal/pg_db --username
exporter --password ●●●●●● --table character --hive-import --hive-database student4_1 --hive-
table my_character --target-dir hdfs://manager.novalocal:8020/user/student4_1/character_tmp
Warning: /opt/cloudera/parcels/CDH-5.16.2-1.cdh5.16.2.p0.8/bin/../lib/sqoop/../accumulo does not
exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
20/06/29 19:09:54 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6-cdh5.16.2
20/06/29 19:09:54 WARN tool.BaseSqoopTool: Setting your password on the command-line is
insecure. Consider using -P instead.
20/06/29 19:09:54 INFO tool.BaseSqoopTool: Using Hive-specific delimiters for output. You can
override
20/06/29 19:09:54 INFO tool.BaseSqoopTool: delimiters with --fields-terminated-by, etc.
20/06/29 19:09:55 INFO manager.SqlManager: Using default fetchSize of 1000
20/06/29 19:09:55 INFO tool.CodeGenTool: Beginning code generation
20/06/29 19:09:55 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "character"
AS t LIMIT 1
20/06/29 19:09:56 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /opt/cloudera/parcels/CDH/
lib/hadoop-mapreduce
Note: /tmp/sqoop-student4_1/compile/7c07cfea60039e2078a36056ac2ca4d1/character.java uses or
overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
20/06/29 19:10:02 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-student4_1/compile/
7c07cfea60039e2078a36056ac2ca4d1/character.jar
20/06/29 19:10:02 WARN manager.PostgresqlManager: It looks like you are importing from
postgresql.
20/06/29 19:10:02 WARN manager.PostgresqlManager: This transfer can be faster! Use the --direct
20/06/29 19:10:02 WARN manager.PostgresqlManager: option to exercise a postgresql-specific fast
path.
20/06/29 19:10:02 INFO mapreduce.ImportJobBase: Beginning import of character
20/06/29 19:10:03 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use
mapreduce.job.jar
20/06/29 19:10:05 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use
mapreduce.job.maps
20/06/29 19:10:06 INFO client.RMProxy: Connecting to ResourceManager at manager.novalocal/
89.208.221.132:8032
20/06/29 19:10:30 INFO db.DBInputFormat: Using read commited transaction isolation
20/06/29 19:10:30 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN("charid"),
MAX("charid") FROM "character"
20/06/29 19:10:30 WARN db.TextSplitter: Generating splits for a textual index column.
20/06/29 19:10:30 WARN db.TextSplitter: If your database sorts in a case-insensitive order, this
may result in a partial import or duplicate records.
20/06/29 19:10:30 WARN db.TextSplitter: You are strongly encouraged to choose an integral split
column.
20/06/29 19:10:32 INFO mapreduce.JobSubmitter: number of splits:6
20/06/29 19:10:32 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1592839005008_0059
20/06/29 19:10:33 INFO impl.YarnClientImpl: Submitted application application_1592839005008_0059
20/06/29 19:10:33 INFO mapreduce.Job: The url to track the job: http://manager.novalocal:8088/
proxy/application_1592839005008_0059/
20/06/29 19:10:33 INFO mapreduce.Job: Running job: job_1592839005008_0059
20/06/29 19:10:55 INFO mapreduce.Job: Job job_1592839005008_0059 running in uber mode : false
20/06/29 19:10:55 INFO mapreduce.Job:  map 0% reduce 0%
20/06/29 19:11:15 INFO mapreduce.Job:  map 17% reduce 0%
20/06/29 19:11:16 INFO mapreduce.Job:  map 33% reduce 0%
20/06/29 19:11:22 INFO mapreduce.Job:  map 50% reduce 0%
20/06/29 19:11:23 INFO mapreduce.Job:  map 67% reduce 0%
20/06/29 19:11:27 INFO mapreduce.Job:  map 83% reduce 0%
20/06/29 19:11:30 INFO mapreduce.Job:  map 100% reduce 0%
20/06/29 19:11:30 INFO mapreduce.Job: Job job_1592839005008_0059 completed successfully
20/06/29 19:11:30 INFO mapreduce.Job: Counters: 30
        File System Counters
                FILE: Number of bytes read=0
                FILE: Number of bytes written=1055130
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=803
                HDFS: Number of bytes written=92512
                HDFS: Number of read operations=24
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=12
        Job Counters
                Launched map tasks=6
                Other local map tasks=6
                Total time spent by all maps in occupied slots (ms)=134432
                Total time spent by all reduces in occupied slots (ms)=0
                Total time spent by all map tasks (ms)=33608
```

```
                    Total vcore-milliseconds taken by all map tasks=33608
                    Total megabyte-milliseconds taken by all map tasks=34414592
            Map-Reduce Framework
                    Map input records=1954
                    Map output records=1954
                    Input split bytes=803
                    Spilled Records=0
                    Failed Shuffles=0
                    Merged Map outputs=0
                    GC time elapsed (ms)=840
                    CPU time spent (ms)=4830
                    Physical memory (bytes) snapshot=1329287168
                    Virtual memory (bytes) snapshot=16799383552
                    Total committed heap usage (bytes)=1131413504
            File Input Format Counters
                    Bytes Read=0
            File Output Format Counters
                    Bytes Written=92512
20/06/29 19:11:30 INFO mapreduce.ImportJobBase: Transferred 90,3438 KB in 84,972 seconds (1,0632
KB/sec)
20/06/29 19:11:30 INFO mapreduce.ImportJobBase: Retrieved 1954 records.
20/06/29 19:11:31 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "character"
AS t LIMIT 1
20/06/29 19:11:31 INFO hive.HiveImport: Loading uploaded data into Hive

Logging initialized using configuration in jar:file:/opt/cloudera/parcels/
CDH-5.16.2-1.cdh5.16.2.p0.8/jars/hive-common-1.1.0-cdh5.16.2.jar!/hive-log4j.properties
OK
Time taken: 9.371 seconds
Loading data to table student4_1.my_character
Table student4_1.my_character stats: [numFiles=6, totalSize=92512]
OK
Time taken: 1.284 seconds
```

Импорт таблицы в HDFS с использованием SQOOP в формате avro:

```
[student4_1@node3 ~]$ sqoop import --connect jdbc:postgresql://node3.novalocal/pg_db --username
exporter -P --table character --target-dir /user/student4_1/character_avro --compress --
compression-codec org.apache.hadoop.io.compress.SnappyCodec --as-avrodatafile
Warning: /opt/cloudera/parcels/CDH-5.16.2-1.cdh5.16.2.p0.8/bin/../lib/sqoop/../accumulo does not
exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
20/06/30 11:35:49 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6-cdh5.16.2
Enter password:
20/06/30 11:36:10 INFO manager.SqlManager: Using default fetchSize of 1000
20/06/30 11:36:10 INFO tool.CodeGenTool: Beginning code generation
20/06/30 11:36:10 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "character"
AS t LIMIT 1
20/06/30 11:36:10 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /opt/cloudera/parcels/CDH/
lib/hadoop-mapreduce
Note: /tmp/sqoop-student4_1/compile/cf793b8869ad477916ae09f7d8c58d11/character.java uses or
overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
20/06/30 11:36:12 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-student4_1/compile/
cf793b8869ad477916ae09f7d8c58d11/character.jar
20/06/30 11:36:12 WARN manager.PostgresqlManager: It looks like you are importing from
postgresql.
20/06/30 11:36:12 WARN manager.PostgresqlManager: This transfer can be faster! Use the --direct
20/06/30 11:36:12 WARN manager.PostgresqlManager: option to exercise a postgresql-specific fast
path.
20/06/30 11:36:12 INFO mapreduce.ImportJobBase: Beginning import of character
20/06/30 11:36:13 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use
mapreduce.job.jar
20/06/30 11:36:14 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "character"
AS t LIMIT 1
```

```
20/06/30 11:36:14 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM "character"
AS t LIMIT 1
20/06/30 11:36:14 INFO mapreduce.DataDrivenImportJob: Writing Avro schema file: /tmp/sqoop-
student4_1/compile/cf793b8869ad477916ae09f7d8c58d11/character.avsc
20/06/30 11:36:14 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use
mapreduce.job.maps
20/06/30 11:36:15 INFO client.RMProxy: Connecting to ResourceManager at manager.novalocal/
89.208.221.132:8032
20/06/30 11:36:19 INFO db.DBInputFormat: Using read commited transaction isolation
20/06/30 11:36:19 INFO db.DataDrivenDBInputFormat: BoundingValsQuery: SELECT MIN("charid"),
MAX("charid") FROM "character"
20/06/30 11:36:19 WARN db.TextSplitter: Generating splits for a textual index column.
20/06/30 11:36:19 WARN db.TextSplitter: If your database sorts in a case-insensitive order, this
may result in a partial import or duplicate records.
20/06/30 11:36:19 WARN db.TextSplitter: You are strongly encouraged to choose an integral split
column.
20/06/30 11:36:19 INFO mapreduce.JobSubmitter: number of splits:6
20/06/30 11:36:19 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1592839005008_0064
20/06/30 11:36:19 INFO impl.YarnClientImpl: Submitted application application_1592839005008_0064
20/06/30 11:36:20 INFO mapreduce.Job: The url to track the job: http://manager.novalocal:8088/
proxy/application_1592839005008_0064/
20/06/30 11:36:20 INFO mapreduce.Job: Running job: job_1592839005008_0064
20/06/30 11:36:27 INFO mapreduce.Job: Job job_1592839005008_0064 running in uber mode : false
20/06/30 11:36:27 INFO mapreduce.Job:  map 0% reduce 0%
20/06/30 11:36:33 INFO mapreduce.Job:  map 17% reduce 0%
20/06/30 11:36:34 INFO mapreduce.Job:  map 50% reduce 0%
20/06/30 11:36:35 INFO mapreduce.Job:  map 83% reduce 0%
20/06/30 11:36:36 INFO mapreduce.Job:  map 100% reduce 0%
20/06/30 11:36:38 INFO mapreduce.Job: Job job_1592839005008_0064 completed successfully
20/06/30 11:36:38 INFO mapreduce.Job: Counters: 30
        File System Counters
                FILE: Number of bytes read=0
                FILE: Number of bytes written=1057326
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=803
                HDFS: Number of bytes written=57794
                HDFS: Number of read operations=24
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=12
        Job Counters
                Launched map tasks=6
                Other local map tasks=6
                Total time spent by all maps in occupied slots (ms)=131000
                Total time spent by all reduces in occupied slots (ms)=0
                Total time spent by all map tasks (ms)=32750
                Total vcore-milliseconds taken by all map tasks=32750
                Total megabyte-milliseconds taken by all map tasks=33536000
        Map-Reduce Framework
                Map input records=1954
                Map output records=1954
                Input split bytes=803
                Spilled Records=0
                Failed Shuffles=0
                Merged Map outputs=0
                GC time elapsed (ms)=1027
                CPU time spent (ms)=12640
                Physical memory (bytes) snapshot=1698119680
                Virtual memory (bytes) snapshot=16922517504
                Total committed heap usage (bytes)=1148715008
        File Input Format Counters
                Bytes Read=0
        File Output Format Counters
                Bytes Written=57794
```