

```
ssh student3_10@37.139.32.56
```

```
hdfs dfs -ls
```

```
hdfs dfs -ls lesson
```

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
```

```
Файл Правка Вид Терминал Вкладки Справка
```

```
igor@igor-MS-7808:~$ ssh student3_10@37.139.32.56
Last login: Mon Apr 19 06:12:37 2021 from 109-252-26-235.nat.spd-mgts.ru
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
Found 3 items
drwx-----  - student3_10 student3_10          0 2021-04-17 18:00 .staging
drwxr-xr-x   - student3_10 student3_10          0 2021-04-15 19:19 lesson
drwxrwxrwx   - student3_10 student3_10          0 2021-04-18 19:32 test
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls lesson
Found 1 items
-rw-r--r--   2 student3_10 student3_10  257081603 2021-04-15 19:07 lesson/urbandict.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```

```
head -n 2 urbandict.csv
```

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
```

```
Файл Правка Вид Терминал Вкладки Справка
```

```
igor@igor-MS-7808:~$ ssh student3_10@37.139.32.56
Last login: Mon Apr 19 06:12:37 2021 from 109-252-26-235.nat.spd-mgts.ru
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls
Found 3 items
drwx-----  - student3_10 student3_10          0 2021-04-17 18:00 .staging
drwxr-xr-x   - student3_10 student3_10          0 2021-04-15 19:19 lesson
drwxrwxrwx   - student3_10 student3_10          0 2021-04-18 19:32 test
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls lesson
Found 1 items
-rw-r--r--   2 student3_10 student3_10  257081603 2021-04-15 19:07 lesson/urbandict.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ head -n 2 urbandict.csv
word_id,word,up_votes,down_votes,author,definition
0000007,Janky,296,255,dc397b2f,"Undesirable; less-than optimum."
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$
```

```
use student3_10;
```

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
```

```
Файл Правка Вид Терминал Вкладки Справка
```

```
Connecting to jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/19 08:00:33 [main]: INFO jdbc.HiveConnection: Connected to bigdataanalytics2-head-shdpt-v31-1-0.novalocal:1000
Connected to: Apache Hive (version 3.1.0.3.1.4.0-315)
Driver: Hive JDBC (version 3.1.0.3.1.4.0-315)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 3.1.0.3.1.4.0-315 by Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdpt> use student3_10;
INFO : Compiling command(queryId=hive_20210419080127_ac2fc80-a2a5-4acb-95d3-a76d4224d9eb): use student3_10
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210419080127_ac2fc80-a2a5-4acb-95d3-a76d4224d9eb); Time taken: 0.029 seconds
INFO : Executing command(queryId=hive_20210419080127_ac2fc80-a2a5-4acb-95d3-a76d4224d9eb): use student3_10
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210419080127_ac2fc80-a2a5-4acb-95d3-a76d4224d9eb); Time taken: 0.011 seconds
INFO : OK
No rows affected (0,209 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt>
```

```
show tables;
```

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

Файл Правка Вид Терминал Вкладки Справка

```

INFO : Completed executing command(queryId=hive_20210419080127_ac2fca80-a2a5-4acb-95d3-a76d4224d9eb); Time taken: 0.011 seconds
INFO : OK
No rows affected (0,209 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> show tables;
INFO : Compiling command(queryId=hive_20210419080742_980bd165-b285-41ea-8bab-bd149bfa2f0a): show tables
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:tab_name, type:string, comment:from deserializer)], properties:null)
INFO : Completed compiling command(queryId=hive_20210419080742_980bd165-b285-41ea-8bab-bd149bfa2f0a); Time taken: 0.01 seconds
INFO : Executing command(queryId=hive_20210419080742_980bd165-b285-41ea-8bab-bd149bfa2f0a): show tables
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210419080742_980bd165-b285-41ea-8bab-bd149bfa2f0a); Time taken: 0.011 seconds
INFO : OK
+-----+
| tab_name |
+-----+
| vehicles |
+-----+
1 row selected (0,138 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> 
```

create table urbandict (

~/BigData. Введение в экосистему Hadoop/4. Hive & HUE/string • - Sublime Text (UNREGISTERED)

File Edit Selection Find View Goto Tools Project Preferences Help

```

string          create external table vehicles ( 
1  create external table urbandict (
2    word_id string,
3    word string,
4    up_votes int,
5    down_votes int,
6    author string,
7    definition
8  )
9  partitioned by ( p_date string )
10 ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
11 WITH SERDEPROPERTIES ('field.delim' = ',')
12 location "/user/student3_10/lesson"
13 tblproperties ("skip.header.line.count"="1")
14 ;
15 | 
```

use student3\_10;

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

Файл Правка Вид Терминал Вкладки Справка

```

1-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.
novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/19 08:26:43 [main]: INFO jdbc.HiveConnection: Connected to bigdataanalytics2-head-shdpt-v31-1-0.novalocal:10000
Connected to: Apache Hive (version 3.1.0.3.1.4.0-315)
Driver: Hive JDBC (version 3.1.0.3.1.4.0-315)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 3.1.0.3.1.4.0-315 by Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdp> use student3_10;
INFO : Compiling command(queryId=hive_20210419082704_e130b08c-2ae1-41fa-b303-81105b984352): use student3_10
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210419082704_e130b08c-2ae1-41fa-b303-81105b984352); Time taken: 0.005 seconds
INFO : Executing command(queryId=hive_20210419082704_e130b08c-2ae1-41fa-b303-81105b984352): use student3_10
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210419082704_e130b08c-2ae1-41fa-b303-81105b984352); Time taken: 0.005 seconds
INFO : OK
No rows affected (0,157 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> 
```

```
hdfs dfs -chmod 666 lesson/urbandict.csv
```

```
hdfs dfs -chmod 777 lesson/urbandict.csv
```

```
Файл    Правка    Вид    Терминал    Вкладки    Справка
INFO  : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO  : Completed compiling command(queryId=hive_20210419092320_2107180d-3727-45a1-bfb1-2d3d17c6e995); Time taken: 0.0
3 seconds
INFO  : Executing command(queryId=hive_20210419092320_2107180d-3727-45a1-bfb1-2d3d17c6e995): create external table urb
andict (
word_id string,
word string,
up_votes int,
down_votes int,
author string,
definition string
)
ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
WITH SERDEPROPERTIES ('field.delim' = ',')
location "/user/student3_10/lesson"
tblproperties ("skip.header.line.count"="1")
INFO  : Starting task [Stage-0:DDL] in serial mode
INFO  : Completed executing command(queryId=hive_20210419092320_2107180d-3727-45a1-bfb1-2d3d17c6e995); Time taken: 0.0
59 seconds
INFO  : OK
No rows affected (0,15 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> █
```

```
select * from urbandict;
```

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
Файл Правка Вид Терминал Вкладки Справка
en in attendance. Ideally expressed as such: 1:2 and abbreviated as BTH." |
| 3604213 | Ghandi's Revenge | 8 | 32 | 76db3e41
| "EXTREMELY spicy curry. Most non-Indians who eat this will die of massive brain trauma or heart attack" |
| 3604216 | Layli | 11 | 3 | e1f4e3a3
| "a supa dupa cool name from a persian love story. ;; a term for a supa good looking person." |
| 3604218 | fauxking | 4 | 2 | efda79c0
| "The act of using an electronic cigarette to imitate smoking a real one. Usually used in areas where smoking is prohibited." |
| 3604230 | Boomskoot | 5 | 0 | 53f9899f
| "A word to represent anything you would like as long as it's within the proper context. It can be a noun" |
| 3604234 | man-jamas | 3 | 3 | d6a93036
| "Pajamas wore by men. A phrase used to add a more manly tone to wearing pajama pants." |
| 3604235 | vagina mouth | 11 | 3 | 4c2aba10
| "the smell of vagina on someones breath; someones mouth that looks like a vagina" |
| 3604242 | fruit neck | 6 | 0 | 6a6daf75
| "someone so fat they have a fat neck" |
| 3604243 | DC-licious | 12 | 4 | 2dd25d86
| "Describes an effin hot pair of sneakers made by DC Shoes." |
+-----+
1 290 460 rows selected (52,878 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp>

```

select \* from urbandict limit 10;

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
Файл Правка Вид Терминал Вкладки Справка
| 0000008 | slumpin' | 16 | 37 | dc397b2f | "low down and funky"
| 0000009 | yayeyay | 19 | 27 | dc397b2f | "affirmation; suggestion of encouragement"
| 0000012 | hard-core | 162 | 96 | d1610749 | "anything out of our league that can be good or bad."
| 0000013 | brutal | 12 | 45 | 40ece1ef | "anything that makes you sweat"
| 0000014 | skanky | 9 | 48 | 485e4db7 | "Anything off pertaining to a $10
| 0000015 | ho-bag | 26 | 27 | b37fba05 | "a term of endearment"
| 0000016 | massive | 36 | 45 | b9dcf126 | "really really good. excellently good."
| 0000017 | wtf | 183 | 99 | a6c97ba3 | "what the fuck? ;; use it in place of expletives. a more polite alternative."
| 0000019 | Hazy | 272 | 184 | 49bc960d | "A guys state of mind after he sees the girl of his dreams...He just can't believe it."
+-----+
10 rows selected (0,216 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp>

```

select author, sum(up\_votes) from urbandict group by author order by author;

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
Файл Правка Вид Терминал Вкладки Справка
| fffedc32 | 3 |
| ffffef936 | 3 |
| fffff098f | 6 |
| fffff107e | 248 |
| fffff181d | 28 |
| fffff211f | 4 |
| fffff283b | 72 |
| fffff2e74 | 1 |
| fffff334b | 22 |
| fffff4599 | 13 |
| fffff49b9 | 184 |
| fffff7d28 | 19 |
| fffff80a4 | 0 |
| fffff9d9b | 63 |
| fffffa986 | 11 |
| fffffd1af | 9 |
| fffffeb25 | 37 |
| ffffff773 | 20 |
| ffffffacc | 3 |
+-----+
825 035 rows selected (40,781 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp>

```

select author, sum(down\_votes) from urbandict group by author order by author;

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

```

Файл Правка Вид Терминал Вкладки Справка
| fffedc32 | 9
| ffffef936 | 12
| fffff098f | 3
| fffff107e | 182
| fffff181d | 31
| fffff211f | 4
| fffff283b | 93
| fffff2e74 | 0
| fffff334b | 6
| fffff4599 | 47
| fffff49b9 | 325
| fffff7d28 | 14
| fffff80a4 | 7
| fffff9d9b | 24
| fffffa986 | 10
| fffffd1af | 14
| fffffeb25 | 16
| ffffff773 | 45
| ffffffacc | 10
+-----+
825 035 rows selected (27,515 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp>

```

ALTER TABLE urbandict change up\_votes up\_votes string;  
ALTER TABLE urbandict change down\_votes down\_votes string;

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

```

Файл Правка Вид Терминал Вкладки Справка
INFO : Executing command(queryId=hive_20210419101152_9228c42a-b5d2-4214-8bb7-555779c06c65): ALTER TABLE urbandict change up_votes up_votes string
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210419101152_9228c42a-b5d2-4214-8bb7-555779c06c65); Time taken: 0.166 seconds
INFO : OK
No rows affected (0,227 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ALTER TABLE urbandict change down_votes down_votes string;
INFO : Compiling command(queryId=hive_20210419101226_03eb31c6-718c-4993-9fe1-d95e0d81e7b4): ALTER TABLE urbandict change down_votes down_votes string
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210419101226_03eb31c6-718c-4993-9fe1-d95e0d81e7b4); Time taken: 0.052 seconds
INFO : Executing command(queryId=hive_20210419101226_03eb31c6-718c-4993-9fe1-d95e0d81e7b4): ALTER TABLE urbandict change down_votes down_votes string
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210419101226_03eb31c6-718c-4993-9fe1-d95e0d81e7b4); Time taken: 0.071 seconds
INFO : OK
No rows affected (0,159 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp>

```

select \* from urbandict where author == 'fffffacc';

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

```

Файл Правка Вид Терминал Вкладки Справка
g, comment:null), FieldSchema(name:urbandict.definition, type:string, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20210419101626_343d9fe7-7fab-4b3a-945b-7d5b68510ed2); Time taken: 0.3 seconds
INFO : Executing command(queryId=hive_20210419101626_343d9fe7-7fab-4b3a-945b-7d5b68510ed2): select * from urbandict where author == 'fffffacc'
INFO : Completed executing command(queryId=hive_20210419101626_343d9fe7-7fab-4b3a-945b-7d5b68510ed2); Time taken: 0.01 seconds
INFO : OK
+-----+-----+-----+-----+-----+
| urbandict.word_id | urbandict.word | urbandict.up_votes | urbandict.down_votes | urbandict.author | urbandict.definition |
+-----+-----+-----+-----+-----+
| 2740811 | emo | 3 | 10 | fffffacc | "Emo is not people who cut themselves or hate the world. Emo's do not hate themsevles either. Emo is not another word for gothic. A person who cuts themself is a cutter not an emo. Emo people are realist. Just because you are emo that doesn't mean you don't believe in God. People misunderstand emo people so much. I just wanted to inform people what emo is not." |
+-----+-----+-----+-----+-----+
1 row selected (2,67 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp>

```

select author, sum(down\_votes) from urbandict group by author order by author;

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~ - □ ×
Файл Правка Вид Терминал Вкладки Справка
| ffffedc32 | 9.0 |
| fffffef936 | 12.0 |
| fffff098f | 3.0 |
| fffff107e | 182.0 |
| fffff181d | 31.0 |
| fffff211f | 4.0 |
| fffff283b | 93.0 |
| fffff2e74 | 0.0 |
| fffff334b | 6.0 |
| fffff4599 | 47.0 |
| fffff49b9 | 325.0 |
| fffff7d28 | 14.0 |
| fffff80a4 | 7.0 |
| fffff9d9b | 24.0 |
| fffffa986 | 10.0 |
| fffffd1af | 14.0 |
| fffffeb25 | 16.0 |
| ffffff773 | 45.0 |
| ffffffacc | 10.0 |
+-----+
825 035 rows selected (37,91 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■
```

select author, sum(up\_votes) from urbandict group by author order by author;

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~ - □ ×
Файл Правка Вид Терминал Вкладки Справка
| ffffedc32 | 3.0 |
| fffffef936 | 3.0 |
| fffff098f | 6.0 |
| fffff107e | 248.0 |
| fffff181d | 28.0 |
| fffff211f | 4.0 |
| fffff283b | 72.0 |
| fffff2e74 | 1.0 |
| fffff334b | 22.0 |
| fffff4599 | 13.0 |
| fffff49b9 | 184.0 |
| fffff7d28 | 19.0 |
| fffff80a4 | 0.0 |
| fffff9d9b | 63.0 |
| fffffa986 | 11.0 |
| fffffd1af | 9.0 |
| fffffeb25 | 37.0 |
| ffffff773 | 20.0 |
| ffffffacc | 3.0 |
+-----+
825 035 rows selected (28,317 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■
```

set hive.execution.engine=spark;

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~ - □ ×
Файл Правка Вид Терминал Вкладки Справка
| fffff098f | 6.0 |
| fffff107e | 248.0 |
| fffff181d | 28.0 |
| fffff211f | 4.0 |
| fffff283b | 72.0 |
| fffff2e74 | 1.0 |
| fffff334b | 22.0 |
| fffff4599 | 13.0 |
| fffff49b9 | 184.0 |
| fffff7d28 | 19.0 |
| fffff80a4 | 0.0 |
| fffff9d9b | 63.0 |
| fffffa986 | 11.0 |
| fffffd1af | 9.0 |
| fffffeb25 | 37.0 |
| ffffff773 | 20.0 |
| ffffffacc | 3.0 |
+-----+
825 035 rows selected (28,317 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> set hive.execution.engine=spark;
No rows affected (0.028 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■
```

ALTER TABLE urbandict change down\_votes down\_votes string;

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~ - □ ×
Файл Правка Вид Терминал Вкладки Справка
| ffffffeb25 | 37.0 |
| ffffff773 | 20.0 |
| ffffffacc | 3.0 |
+-----+
825 035 rows selected (28,317 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> set hive.execution.engine=spark;
No rows affected (0,028 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ALTER TABLE urbandict change down_votes down_votes string;
INFO : Compiling command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc): ALTER TABLE urbandict change down_votes down_votes string
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc); Time taken: 0.0
42 seconds
INFO : Executing command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc): ALTER TABLE urbandict change down_votes down_votes string
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc); Time taken: 0.0
47 seconds
INFO : OK
No rows affected (0,126 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp>
```

set hive.execution.engine=mr;

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~ - □ ×
Файл Правка Вид Терминал Вкладки Справка
| ffffffacc | 3.0 |
+-----+
825 035 rows selected (28,317 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> set hive.execution.engine=spark;
No rows affected (0,028 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ALTER TABLE urbandict change down_votes down_votes string;
INFO : Compiling command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc): ALTER TABLE urbandict change down_votes down_votes string
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc); Time taken: 0.0
42 seconds
INFO : Executing command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc): ALTER TABLE urbandict change down_votes down_votes string
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc); Time taken: 0.0
47 seconds
INFO : OK
No rows affected (0,126 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> set hive.execution.engine=mr;
Error: Error while processing statement: hive execution engine mr is not supported. (state=42000,code=1)
0: jdbc:hive2://bigdataanalytics2-worker-shdp>
```

set hive.execution.engine=tez;

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~ - □ ×
Файл Правка Вид Терминал Вкладки Справка
825 035 rows selected (28,317 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> set hive.execution.engine=spark;
No rows affected (0,028 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ALTER TABLE urbandict change down_votes down_votes string;
INFO : Compiling command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc): ALTER TABLE urbandict change down_votes down_votes string
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc); Time taken: 0.0
42 seconds
INFO : Executing command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc): ALTER TABLE urbandict change down_votes down_votes string
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210419102957_6d66a8d6-f9ef-44a7-88aa-1ae63fdaecc); Time taken: 0.0
47 seconds
INFO : OK
No rows affected (0,126 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> set hive.execution.engine=mr;
Error: Error while processing statement: hive execution engine mr is not supported. (state=42000,code=1)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> set hive.execution.engine=tez;
No rows affected (0,01 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp>
```

select author, sum(up\_votes) from urbandict group by author order by author;

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
Файл Правка Вид Терминал Вкладки Справка
| fffedc32 | 3.0 |
| ffffef936 | 3.0 |
| fffff098f | 6.0 |
| fffff107e | 248.0 |
| fffff181d | 28.0 |
| fffff211f | 4.0 |
| fffff283b | 72.0 |
| fffff2e74 | 1.0 |
| fffff334b | 22.0 |
| fffff4599 | 13.0 |
| fffff49b9 | 184.0 |
| fffff7d28 | 19.0 |
| fffff80a4 | 0.0 |
| fffff9d9b | 63.0 |
| fffffa986 | 11.0 |
| fffffd1af | 9.0 |
| fffffeb25 | 37.0 |
| ffffff773 | 20.0 |
| ffffffacc | 3.0 |
+-----+
825 035 rows selected (134,756 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■

```

---

cat urbandict.csv | python3.7 mapper.py

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
Файл Правка Вид Терминал Вкладки Справка
319ff6a2
97b18a12
2cff5051
c3cdel2e
ff17879
a1901627
c795a9ca
a6c11229
684aab14
999a333b
b8256651
595b54e1
45a1e4b9
c795a9ca
4dda8004
76db3e41
elf4e3a3
efda79c0
53f9899f
d6a93036
4c2aba10
6a6daf75
2dd25d86
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ cat urbandict.csv | python3.7 mapper.py

```

hdfs dfs -rm r result

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
Файл Правка Вид Терминал Вкладки Справка
-rw-rw-r-- 1 student3_10 student3_10 525 арп 16 21:27 mapper.py
-rw-rw-r-- 1 student3_10 student3_10 529 арп 16 19:39 Mapper.py
drwxrwxr-x 2 student3_10 student3_10 4,0K арп 3 18:35 .oracle_jre_usage
-rw-rw-r-- 1 student3_10 student3_10 1,2K арп 16 21:33 reducer.py
-rw-rw-r-- 1 student3_10 student3_10 1,2K арп 16 19:55 Reducer.py
drwx----- 2 student3_10 student3_10 4,0K арп 3 18:40 .ssh
-rw-rw-r-- 1 student3_10 student3_10 246M арп 15 18:53 urbandict.csv
-rw-rw-r-- 1 student3_10 student3_10 331M арп 13 17:56 vehicles.csv
-rw----- 1 student3_10 student3_10 3,7K арп 15 21:36 .viminfo
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ cat urbandict.csv | python3.7 mapper.py | sort | python3.7 Reducer.py
^[[A^[[ATraceback (most recent call last):
  File "Reducer.py", line 40, in <module>
    print(current_word+'/'+str(current_count))
TypeError: unsupported operand type(s) for +: 'NoneType' and 'str'
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ cat urbandict.csv | python3.7 mapper.py | sort | python3.7 Reducer.py
Traceback (most recent call last):
  File "Reducer.py", line 40, in <module>
    print(current_word+'/'+str(current_count))
TypeError: unsupported operand type(s) for +: 'NoneType' and 'str'
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ cat urbandict.csv | python3.7 mapper.py | sort | python3.7 reducer.py
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -rm r result
rm: 'r': No such file or directory
rm: 'result': Is a directory
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ ■

```

yarn jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-streaming.jar -input lesson -output result -mapper "python3.7 mapper.py" -reducer "python3.7 reducer.py" -file mapper.py -file reducer.py

---

head -n 2 vehicles.csv  
show databases;  
use student3\_10;

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

Файл Правка Вид Терминал Вкладки Справка

```

| student21_19
| student21_2
| student21_24
| student21_24_test
| student21_25
| student21_26
| student21_3
| student21_5
| student21_50
| student21_6
| student21_7
| student2_3
| student2_4
| student2_5
| student3_1
| student3_10
| student3_12
| student3_13
| student3_23
| student3_39
| student3_4
| sys
| testdb
| titanic
| trending_youtube_videos
| vaccin_db
| valerias_hive_db
| vorozhbitov
| vyacheslavk2
| wine
| zzz
+-----+
86 rows selected (0,414 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> use student3_10;
INFO : Compiling command(queryId=hive_20210418134923_93661e47-8d85-4e26-8150-859607c314aa): use student3_10
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210418134923_93661e47-8d85-4e26-8150-859607c314aa); Time taken: 0.075 seconds
INFO : Executing command(queryId=hive_20210418134923_93661e47-8d85-4e26-8150-859607c314aa): use student3_10
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210418134923_93661e47-8d85-4e26-8150-859607c314aa); Time taken: 0.027 seconds
INFO : OK
No rows affected (0,159 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■

```

show tables;

File Edit Selection Find View Goto Tools Project Preferences Help

string rid string, \* - Sublime Text (UNREGISTERED)

```

1   rid string,
2   id int,
3   url string,
4   region string,
5   region_url string,
6   price int,
7   year string,
8   manufacturer string,
9   model string,
10  condition string,
11  cylinders string,
12  fuel string,
13  odometer double,
14  title_status string,
15  transmission string,
16  VIN string,
17  drive string,
18  size string,
19  type string,
20  paint_color string,
21  image_url string,
22  description string,
23  state string,
24  lat string,
25  long string,
26  posting_date string
27  )
28 ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
29 WITH SERDEPROPERTIES ('field.delim' = ',')
30 location "/user/student3_10/test"
31 tblproperties ("skip.header.line.count"="1")
32 ;

```

File Edit Selection Find View Goto Tools Project Preferences Help

string create table vehicles (\* - Sublime Text (UNREGISTERED))

```

1   create table vehicles (
2     rid string,
3     id int,
4     url string,
5     region string,
6     region_url string,
7     price int,
8     year string,
9     manufacturer string,
10    model string,
11    condition string,
12    cylinders string,
13    fuel string,
14    odometer double,
15    title_status string,
16    transmission string,
17    VIN string,
18    drive string,
19    size string,
20    type string,
21    paint_color string,
22    image_url string,
23    description string,
24    state string,
25    lat string,
26    long string,
27    posting_date string
28  )
29 ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
30 WITH SERDEPROPERTIES ('field.delim' = ',')
31 location "/user/student3_10/test"
32tblproperties ("skip.header.line.count"="1")
33 ;

```

Line 31, Column 33 Tab Size: 4 Plain Text

Файл Правка Вид Терминал Вкладки Справка

```

ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
WITH SERDEPROPERTIES ('field.delim' = ',')
location "/user/student3_10/test"
tblproperties ("skip.header.line.count"="1")
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210418161711_f86ce341-e855-4c61-8a0c-14678e0e04bd); Time taken: 0.046 seconds
INFO : Executing command(queryId=hive_20210418161711_f86ce341-e855-4c61-8a0c-14678e0e04bd): create table vehicles (
rid string,
id int,
url string,
region string,
region_url string,
price int,
year string,
manufacturer string,
model string,
condition string,
cylinders string,
fuel string,
odometer double,
title_status string,
transmission string,
VIN string,
drive string,
size string,
type string,
paint_color string,
image_url string,
description string,
state string,
lat string,
long string,
posting_date string
)
ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
WITH SERDEPROPERTIES ('field.delim' = ',')
location "/user/student3_10/test"
tblproperties ("skip.header.line.count"="1")
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210418161711_f86ce341-e855-4c61-8a0c-14678e0e04bd); Time taken: 0.171 seconds
INFO : OK
No rows affected (0,285 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■

```

select \* from vehicles;

Файл Правка Вид Терминал Вкладки Справка

```

)
ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
WITH SERDEPROPERTIES ('field.delim' = ',')
location "/user/student3_10/test"
tblproperties ("skip.header.line.count"="1")
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210418161711_f86ce341-e855-4c61-8a0c-14678e0e04bd); Time taken: 0.171 seconds
INFO : OK
No rows affected (0,285 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> select * from vehicles;
INFO : Compiling command(queryId=hive_20210418161909_ef95d9d5-7157-4d4c-8765-fd9ca68c0021): select * from vehicles
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:vehicles.rid, type:string, comment:null), FieldSchema(name:vehicles.id, type:int, comment:null), FieldSchema(name:vehicles.url, type:string, comment:null), FieldSchema(name:vehicles.region, type:string, comment:null), FieldSchema(name:vehicles.region_url, type:string, comment:null), FieldSchema(name:vehicles.price, type:int, comment:null), FieldSchema(name:vehicles.year, type:string, comment:null), FieldSchema(name:vehicles.manufacturer, type:string, comment:null), FieldSchema(name:vehicles.model, type:string, comment:null), FieldSchema(name:vehicles.condition, type:string, comment:null), FieldSchema(name:vehicles.cylinders, type:string, comment:null), FieldSchema(name:vehicles.fuel, type:string, comment:null), FieldSchema(name:vehicles.odometer, type:double, comment:null), FieldSchema(name:vehicles.title_status, type:string, comment:null), FieldSchema(name:vehicles.transmission, type:string, comment:null), FieldSchema(name:vehicles.vin, type:string, comment:null), FieldSchema(name:vehicles.drive, type:string, comment:null), FieldSchema(name:vehicles.size, type:string, comment:null), FieldSchema(name:vehicles.type, type:string, comment:null), FieldSchema(name:vehicles.paint_color, type:string, comment:null), FieldSchema(name:vehicles.image_url, type:string, comment:null), FieldSchema(name:vehicles.description, type:string, comment:null), FieldSchema(name:vehicles.state, type:string, comment:null), FieldSchema(name:vehicles.lat, type:string, comment:null), FieldSchema(name:vehicles.long, type:string, comment:null), FieldSchema(name:vehicles.posting_date, type:string, comment:null)], properties:null)
INFO : Completed compiling command(queryId=hive_20210418161909_ef95d9d5-7157-4d4c-8765-fd9ca68c0021); Time taken: 0.194 seconds
INFO : Executing command(queryId=hive_20210418161909_ef95d9d5-7157-4d4c-8765-fd9ca68c0021): select * from vehicles
INFO : Completed executing command(queryId=hive_20210418161909_ef95d9d5-7157-4d4c-8765-fd9ca68c0021); Time taken: 0.005 seconds
INFO : OK
Error: java.io.IOException: java.lang.RuntimeException: ORC split generation failed with exception: org.apache.orc.FileFormatException: Malformed ORC file hdfs://bigdataanalytics2-head-shdp-v31-1-0.novalocal:8020/user/student3_10/test/vehicles.csv. Invalid postscript. (state=,code=0)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> select * from vehicles;

```

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

```

Файл Правка Вид Терминал Вкладки Справка
at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.checkAccess(ClientNamenodeProtocolServerSideTranslatorPB.java:1659)
at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:524)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1025)
at org.apache.hadoop.ipc.Server$RpcCall.run(Server.java:876)
at org.apache.hadoop.ipc.Servers$RpcCall.run(Server.java:822)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1730)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2682)
)
INFO : Completed executing command(queryId=hive_20210418162308_d5882107-5b1e-41e0-9bb1-bfffd269ad036); Time taken: 0.058 seconds
Error: Error while processing statement: FAILED: Execution Error, return code 1 from org.apache.hive.ql.exec.DDLTask. MetaException(message:java.security.AccessControlException: Permission denied: user=hive, access=WRITE, inode="/user/student3_10/test";student3_10:student3_10:drwxr-xr-x
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:399)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:261)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:193)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1857)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1841)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPathAccess(FSDirectory.java:1791)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkAccess(FSNamesystem.java:7804)
at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.checkAccess(NameNodeRpcServer.java:2217)
at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.checkAccess(ClientNamenodeProtocolServerSideTranslatorPB.java:1659)
at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
at org.apache.hadoop.ipc.ProtobufRpcEngines$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:524)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1025)
at org.apache.hadoop.ipc.Servers$RpcCall.run(Server.java:876)
at org.apache.hadoop.ipc.Servers$RpcCall.run(Server.java:822)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1730)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2682)
) (state=08501,code=1)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt:> 
```

hdfs dfs -ls test

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

```

Файл Правка Вид Терминал Вкладки Справка
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:399)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:261)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:193)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1857)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1841)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPathAccess(FSDirectory.java:1791)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkAccess(FSNamesystem.java:7804)
at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.checkAccess(NameNodeRpcServer.java:2217)
at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.checkAccess(ClientNamenodeProtocolServerSideTranslatorPB.java:1659)
at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
at org.apache.hadoop.ipc.ProtobufRpcEngines$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:524)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1025)
at org.apache.hadoop.ipc.Servers$RpcCall.run(Server.java:876)
at org.apache.hadoop.ipc.Servers$RpcCall.run(Server.java:822)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1730)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2682)
)
(state=08501,code=1)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt:> hdfs dfs -ls test
.
.
.
Error: Error while compiling statement: FAILED: ParseException line 1:0 cannot recognize input near 'hdfs' 'dfs' '-' (state=42000,code=40000)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt: Closing: 0: jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls test
Found 1 items
-rw-r--r-- 2 student3_10 student3_10 346353924 2021-04-15 17:36 test/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ 
```

hdfs dfs -chmod 666 test

hive

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

```

Файл Правка Вид Терминал Вкладки Справка
-rw-r--r-- 2 student3_10 student3_10 346353924 2021-04-15 17:36 test/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -chmod 666 test
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/Log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Connecting to jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/18 16:29:57 [main]: INFO jdbc.HiveConnection: Connected to bigdataanalytics2-head-shdpt-v31-1-0.novalocal:10000
Connected to: Apache Hive (version 3.1.0.3.1.4.0-315)
Driver: Hive JDBC (version 3.1.0.3.1.4.0-315)
Transaction isolation: TRANSACTION_REPEATABLE_READ
BeeLine version 3.1.0.3.1.4.0-315 by Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdpt:> select * from vehicles;
Error: Error while compiling statement: FAILED: SemanticException [Error 10001]: Line 1:14 Table not found 'vehicles' (state=42502,code=10001)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt:> use student3_10;
INFO : Compiling command(queryId=hive_20210418163255_3200d6a4-62e0-437c-b8e8-2e6602996225): use student3_10
INFO : Semantic Analysis Completed (retry= false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210418163255_3200d6a4-62e0-437c-b8e8-2e6602996225); Time taken: 0.013 seconds
INFO : Executing command(queryId=hive_20210418163255_3200d6a4-62e0-437c-b8e8-2e6602996225): use student3_10
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210418163255_3200d6a4-62e0-437c-b8e8-2e6602996225); Time taken: 0.011 seconds
INFO : OK
No rows affected (0.099 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt:> 
```

OK

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0~  
Файл Правка Вид Терминал Вкладки Справка  
t_color, type:string, comment:null), FieldSchema(name:vehicles.image_url, type:string, comment:null), FieldSchema(name:vehicles.description, type:string, comment:null), FieldSchema(name:vehicles.state, type:string, comment:null), FieldSchema(name:vehicles.lat, type:string, comment:null), FieldSchema(name:vehicles.long, type:string, comment:null), FieldSchema(name:vehicles.posting_date, type:string, comment:null), properties:null)  
INFO : Completed compiling command(queryId=hive_20210418163341_a4ccaa2c-76eb-4a3b-a854-9b089328a878); Time taken: 0.151 seconds  
INFO : Executing command(queryId=hive_20210418163341_a4ccaa2c-76eb-4a3b-a854-9b089328a878); select * from vehicles  
INFO : Completed executing command(queryId=hive_20210418163341_a4ccaa2c-76eb-4a3b-a854-9b089328a878); Time taken: 0.003 seconds  
INFO : OK  
Error: java.io.IOException: org.apache.hadoop.security.AccessControlException: Permission denied: user=hive, access=READ_EXECUTE, inode="/user/student3_10/test":student3_10:student3_10:drw-rw-rw-  
    at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:399)  
    at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:261)  
    at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:193)  
    at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1857)  
    at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1841)  
    at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPathAccess(FSDirectory.java:1791)  
    at org.apache.hadoop.hdfs.server.namenode.FSDirStatAndListingOp.getListingInt(FSDirStatAndListingOp.java:79)  
    at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getListing(FSNamesystem.java:3788)  
    at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getListing(NameNodeRpcServer.java:1141)  
    at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getListing(ClientNamenodeProtocolServerSideTranslatorPB.java:719)  
    at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)  
    at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:524)  
    at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1025)  
    at org.apache.hadoop.ipc.Servers$RpcCall.run(Server.java:876)  
    at org.apache.hadoop.ipc.Servers$RpcCall.run(Server.java:822)  
    at java.security.AccessController.doPrivileged(Native Method)  
    at javax.security.auth.Subject.doAs(Subject.java:422)  
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1730)  
    at org.apache.hadoop.ipc.Servers$Handler.run(Server.java:2682) (state=,code=0)  
0: jdbc:hive2://bigdataanalytics2-worker-shdpt> ■
```

hdfs dfs -chmod 777 test

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0~  
Файл Правка Вид Терминал Вкладки Справка  
INFO : Executing command(queryId=hive_20210418163341_a4ccaa2c-76eb-4a3b-a854-9b089328a878); select * from vehicles  
INFO : Completed executing command(queryId=hive_20210418163341_a4ccaa2c-76eb-4a3b-a854-9b089328a878); Time taken: 0.003 seconds  
INFO : OK  
Error: java.io.IOException: org.apache.hadoop.security.AccessControlException: Permission denied: user=hive, access=READ_EXECUTE, inode="/user/student3_10/test":student3_10:student3_10:drw-rw-rw-  
    at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:399)  
    at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:261)  
    at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:193)  
    at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1857)  
    at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1841)  
    at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPathAccess(FSDirectory.java:1791)  
    at org.apache.hadoop.hdfs.server.namenode.FSDirStatAndListingOp.getListingInt(FSDirStatAndListingOp.java:79)  
    at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getListing(FSNamesystem.java:3788)  
    at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getListing(NameNodeRpcServer.java:1141)  
    at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getListing(ClientNamenodeProtocolServerSideTranslatorPB.java:719)  
    at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)  
    at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:524)  
    at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1025)  
    at org.apache.hadoop.ipc.Servers$RpcCall.run(Server.java:876)  
    at org.apache.hadoop.ipc.Servers$RpcCall.run(Server.java:822)  
    at java.security.AccessController.doPrivileged(Native Method)  
    at javax.security.auth.Subject.doAs(Subject.java:422)  
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1730)  
    at org.apache.hadoop.ipc.Servers$Handler.run(Server.java:2682) (state=,code=0)  
0: jdbc:hive2://bigdataanalytics2-worker-shdpt> Closing: 0: jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-6.novalocal:2181,serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2  
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -chmod 777 test  
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ ■
```

OK

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0~  
Файл Правка Вид Терминал Вкладки Справка  
INFO : Compiling command(queryId=hive_20210418163606_347f014c-4a9d-4381-b537-2eacf21c117); use student3_10  
INFO : Semantic Analysis Completed (retrial = false)  
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)  
INFO : Completed compiling command(queryId=hive_20210418163606_347f014c-4a9d-4381-b537-2eacf21c117); Time taken: 0.005 seconds  
INFO : Executing command(queryId=hive_20210418163606_347f014c-4a9d-4381-b537-2eacf21c117); use student3_10  
INFO : Starting task [Stage-0:DDL] in serial mode  
INFO : Completed executing command(queryId=hive_20210418163606_347f014c-4a9d-4381-b537-2eacf21c117); Time taken: 0.011 seconds  
INFO : OK  
No rows affected (0,154 seconds)  
0: jdbc:hive2://bigdataanalytics2-worker-shdpt> select * from vehicles;  
INFO : Compiling command(queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7); select * from vehicles  
INFO : Semantic Analysis Completed (retrial = false)  
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:vehicles.rid, type:string, comment:null), FieldSchema(name:vehicles.id, type:int, comment:null), FieldSchema(name:vehicles.url, type:string, comment:null), FieldSchema(name:vehicles.region, type:string, comment:null), FieldSchema(name:vehicles.region_url, type:string, comment:null), FieldSchema(name:vehicles.price, type:int, comment:null), FieldSchema(name:vehicles.year, type:string, comment:null), FieldSchema(name:vehicles.manufacturer, type:string, comment:null), FieldSchema(name:vehicles.model, type:string, comment:null), FieldSchema(name:vehicles.condition, type:string, comment:null), FieldSchema(name:vehicles.cylinders, type:string, comment:null), FieldSchema(name:vehicles.fuel, type:string, comment:null), FieldSchema(name:vehicles.odometer, type:double, comment:null), FieldSchema(name:vehicles.title_status, type:string, comment:null), FieldSchema(name:vehicles.transmission, type:string, comment:null), FieldSchema(name:vehicles.vin, type:string, comment:null), FieldSchema(name:vehicles.drive_type, type:string, comment:null), FieldSchema(name:vehicles.size, type:string, comment:null), FieldSchema(name:vehicles.type, type:string, comment:null), FieldSchema(name:vehicles.paint_color, type:string, comment:null), FieldSchema(name:vehicles.state, type:string, comment:null), FieldSchema(name:vehicles.lat, type:string, comment:null), FieldSchema(name:vehicles.long, type:string, comment:null), FieldSchema(name:vehicles.posting_date, type:string, comment:null)], properties:null)  
INFO : Completed compiling command(queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7); Time taken: 0.119 seconds  
INFO : Executing command(queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7); select * from vehicles  
INFO : Completed executing command(queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7); Time taken: 0.001 seconds  
INFO : OK  
Error: java.io.IOException: java.lang.RuntimeException: ORC split generation failed with exception: org.apache.orc.FileFormatException: Malformed ORC file hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_10/test/vehicles.csv. Invalid postscript. (state=,code=0)  
0: jdbc:hive2://bigdataanalytics2-worker-shdpt> ■
```

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0~

Файл Правка Вид Терминал Вкладки Справка
INFO : Compiling command[queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7]: select * from vehicles
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:vehicles.id, type:int, comment:null), FieldSchema(name:vehicles.url, type:string, comment:null), FieldSchema(name:vehicles.region, type:string, comment:null), FieldSchema(name:vehicles.region_url, type:string, comment:null), FieldSchema(name:vehicles.price, type:int, comment:null), FieldSchema(name:vehicles.year, type:string, comment:null), FieldSchema(name:vehicles.manufacturer, type:string, comment:null), FieldSchema(name:vehicles.model, type:string, comment:null), FieldSchema(name:vehicles.condition, type:string, comment:null), FieldSchema(name:vehicles.fuel_type, type:string, comment:null), FieldSchema(name:vehicles.odometer, type:double, comment:null), FieldSchema(name:vehicles.title_status, type:string, comment:null), FieldSchema(name:vehicles.fuel, type:string, comment:null), FieldSchema(name:vehicles.transmission, type:string, comment:null), FieldSchema(name:vehicles.vin, type:string, comment:null), FieldSchema(name:vehicles.description, type:string, comment:null), FieldSchema(name:vehicles.drive_type, type:string, comment:null), FieldSchema(name:vehicles.size, type:string, comment:null), FieldSchema(name:vehicles.type, type:string, comment:null), FieldSchema(name:vehicles.paint_color, type:string, comment:null), FieldSchema(name:vehicles.lat, type:string, comment:null), FieldSchema(name:vehicles.long, type:string, comment:null), FieldSchema(name:vehicles.posting_date, type:string, comment:null)], properties:null)
INFO : Completed compiling command[queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7]: Time taken: 0.119 seconds
INFO : Executing command[queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7]: select * from vehicles
INFO : Completed executing command[queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7]: Time taken: 0.001 seconds
INFO : OK
Error: java.io.IOException: java.lang.RuntimeException: ORC split generation failed with exception: org.apache.orc.FileFormatException: Malformed ORC file hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_10/test/vehicles.csv. Invalid postscript. (state=-code=0)
0: jdbc:hive2://bigdataanalytics2-worker-shdps> drop table vehicles;
INFO : Compiling command[queryId=hive_20210418163724_d50bddbc-4f43-4fe1-9eb4-0a673b50be19]: drop table vehicles
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command[queryId=hive_20210418163724_d50bddbc-4f43-4fe1-9eb4-0a673b50be19]: Time taken: 0.034 seconds
INFO : Executing command[queryId=hive_20210418163724_d50bddbc-4f43-4fe1-9eb4-0a673b50be19]: drop table vehicles
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command[queryId=hive_20210418163724_d50bddbc-4f43-4fe1-9eb4-0a673b50be19]: Time taken: 0.28 seconds
INFO : OK
No rows affected (0,354 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdps>

```

hdfs dfs -chmod 666 test/vehicles.csv

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0~

Файл Правка Вид Терминал Вкладки Справка
schema(name:vehicles.vin, type:string, comment:null), FieldSchema(name:vehicles.drive_type, type:string, comment:null), FieldSchema(name:vehicles.size, type:string, comment:null), FieldSchema(name:vehicles.type, type:string, comment:null), FieldSchema(name:vehicles.paint_color, type:string, comment:null), FieldSchema(name:vehicles.image_url, type:string, comment:null), FieldSchema(name:vehicles.description, type:string, comment:null), FieldSchema(name:vehicles.state, type:string, comment:null), FieldSchema(name:vehicles.lat, type:string, comment:null), FieldSchema(name:vehicles.long, type:string, comment:null), FieldSchema(name:vehicles.posting_date, type:string, comment:null), properties:null)
INFO : Completed compiling command[queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7]: Time taken: 0.119 seconds
INFO : Executing command[queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7]: select * from vehicles
INFO : Completed executing command[queryId=hive_20210418163620_f92dd711-7388-48a9-8c54-cb6748fac2d7]: Time taken: 0.001 seconds
INFO : OK
Error: java.io.IOException: java.lang.RuntimeException: ORC split generation failed with exception: org.apache.orc.FileFormatException: Malformed ORC file hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_10/test/vehicles.csv. Invalid postscript. (state=-code=0)
0: jdbc:hive2://bigdataanalytics2-worker-shdps> drop table vehicles;
INFO : Compiling command[queryId=hive_20210418163724_d50bddbc-4f43-4fe1-9eb4-0a673b50be19]: drop table vehicles
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command[queryId=hive_20210418163724_d50bddbc-4f43-4fe1-9eb4-0a673b50be19]: Time taken: 0.034 seconds
INFO : Executing command[queryId=hive_20210418163724_d50bddbc-4f43-4fe1-9eb4-0a673b50be19]: drop table vehicles
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command[queryId=hive_20210418163724_d50bddbc-4f43-4fe1-9eb4-0a673b50be19]: Time taken: 0.28 seconds
INFO : OK
No rows affected (0,354 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdps> Closing: 0: jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls test
Found 1 items
-rw-r--r-- 2 student3_10 student3_10 346353924 2021-04-15 17:36 test/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -chmod 666 test/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ 

```

hive

select \* from vehicles;

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0~

Файл Правка Вид Терминал Вкладки Справка
-rw-r--r-- 2 student3_10 student3_10 346353924 2021-04-15 17:36 test/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -chmod 666 test/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Connecting to jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/18 16:40:27 [main]: INFO jdbc.HiveConnection: Connected to bigdataanalytics2-head-shdpt-v31-1-0.novalocal:10000
Connected to: Apache Hive (version 3.1.0.3.1.4.0-315)
Driver: Hive JDBC (version 3.1.0.3.1.4.0-315)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 3.1.0.3.1.4.0-315 by Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdps> use student3_10;
INFO : Compiling command[queryId=hive_20210418164630_d4a4e268-c3f5-404f-b5e1-82e1551f4a09]: use student3_10
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command[queryId=hive_20210418164630_d4a4e268-c3f5-404f-b5e1-82e1551f4a09]: Time taken: 0.027 seconds
INFO : Executing command[queryId=hive_20210418164630_d4a4e268-c3f5-404f-b5e1-82e1551f4a09]: use student3_10
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command[queryId=hive_20210418164630_d4a4e268-c3f5-404f-b5e1-82e1551f4a09]: Time taken: 0.011 seconds
INFO : OK
No rows affected (0,282 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdps> select * from vehicles;
Error: Error while compiling statement: FAILED: SemanticException [Error 10001]: Line 1:14 Table not found 'vehicles' (state=42502, code=10001)
0: jdbc:hive2://bigdataanalytics2-worker-shdps>

```

ok

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~

Файл Правка Вид Терминал Вкладки Справка

manufacturer string,
model string,
condition string,
cylinders string,
fuel string,
odometer double,
title_status string,
transmission string,
VIN string,
drive string,
size string,
paint_color string,
image_url string,
description string,
state string,
lat string,
long string,
posting_date string
)
ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
WITH SERDEPROPERTIES ('field.delim' = ',')
location "/user/student3_10/test"
tblproperties ("skip.header.line.count"="1")
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210418164822_87d03ea1-2111-4cab-be1d-dcef2b909596); Time taken: 0.143 seconds
INFO : OK
No rows affected (0,23 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■

```

drop table vehicles;

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~

Файл Правка Вид Терминал Вкладки Справка

INFO : Semantic Analysis Completed (retryal = false)
INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:vehicles.rid, type:string, comment:null), FieldSchema(name:vehicles.id, type:int, comment:null), FieldSchema(name:vehicles.url, type:string, comment:null), FieldSchema(name:vehicles.region, type:string, comment:null), FieldSchema(name:vehicles.region_url, type:string, comment:null), FieldSchema(name:vehicles.price, type:int, comment:null), FieldSchema(name:vehicles.year, type:string, comment:null), FieldSchema(name:vehicles.condition, type:string, comment:null), FieldSchema(name:vehicles.model, type:string, comment:null), FieldSchema(name:vehicles.cylinder, type:string, comment:null), FieldSchema(name:vehicles.fuel, type:string, comment:null), FieldSchema(name:vehicles.odometer, type:double, comment:null), FieldSchema(name:vehicles.title_status, type:string, comment:null), FieldSchema(name:vehicles.transmission, type:string, comment:null), FieldSchema(name:vehicles.vin, type:string, comment:null), FieldSchema(name:vehicles.drive, type:string, comment:null), FieldSchema(name:vehicles.size, type:string, comment:null), FieldSchema(name:vehicles.image_url, type:string, comment:null), FieldSchema(name:vehicles.description, type:string, comment:null), FieldSchema(name:vehicles.state, type:string, comment:null), FieldSchema(name:vehicles.lat, type:string, comment:null), FieldSchema(name:vehicles.long, type:string, comment:null), FieldSchema(name:vehicles.posting_date, type:string, comment:null)], properties:{null})
INFO : Completed compiling command(queryId=hive_20210418164958_71c5c151-35d5-44f7-9001-c5735d9e4e40); Time taken: 0.157 seconds
INFO : Executing command(queryId=hive_20210418164958_71c5c151-35d5-44f7-9001-c5735d9e4e40): select * from vehicles
INFO : Completed executing command(queryId=hive_20210418164958_71c5c151-35d5-44f7-9001-c5735d9e4e40); Time taken: 0.001 seconds
INFO : OK
Error: java.io.IOException: java.lang.RuntimeException: ORC split generation failed with exception: org.apache.orc.FileFormatException: Malformed ORC file hdfs://bigdataanalytics2-head-shdpt-v31-1-0.novalocal:8020/user/student3_10/test/vehicles.csv. Invalid postscript. (state=,code=0)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> drop table vehicles;
INFO : Compiling command(queryId=hive_20210418165047_c0dea426-9d84-48ba-a135-66c88306a000): drop table vehicles
INFO : Semantic Analysis Completed (retryal = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210418165047_c0dea426-9d84-48ba-a135-66c88306a000); Time taken: 0.025 seconds
INFO : Executing command(queryId=hive_20210418165047_c0dea426-9d84-48ba-a135-66c88306a000): drop table vehicles
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210418165047_c0dea426-9d84-48ba-a135-66c88306a000); Time taken: 0.164 seconds
INFO : OK
No rows affected (0,229 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> drop table vehicles; ■

```

use student3\_10;

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~

Файл Правка Вид Терминал Вкладки Справка

ocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls test
Found 1 items
-rw-rw-rw- 2 student3_10 student3_10 346353924 2021-04-15 17:36 test/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.log4j.Log4jLoggerFactory]
Connecting to jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/18 16:54:28 [main]: INFO jdbc.HiveConnection: Connected to bigdataanalytics2-head-shdpt-v31-1-0.novalocal:10000
Connected to: Apache Hive (version 3.1.0.3.1.4.0-315)
Driver: Hive JDBC (version 3.1.0.3.1.4.0-315)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 3.1.0.3.1.4.0-315 by Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdp> use student3_10;
INFO : Compiling command(queryId=hive_20210418165529_66af3d74-37b8-4810-a097-ecaefaa47489): use student3_10
INFO : Semantic Analysis Completed (retryal = false)
INFO : Returning Hive schema: Schema(fieldSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210418165529_66af3d74-37b8-4810-a097-ecaefaa47489); Time taken: 0.011 seconds
INFO : Executing command(queryId=hive_20210418165529_66af3d74-37b8-4810-a097-ecaefaa47489): use student3_10
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210418165529_66af3d74-37b8-4810-a097-ecaefaa47489); Time taken: 0.009 seconds
INFO : OK
No rows affected (0,187 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■

```

select \* from vehicles;

```
select manufacturer, sum(price) from vehicles group by manufacturer order by manufacturer;
```

```
Файл Правка Вид Терминал Вкладки Справка

student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~ student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~

infiniti | 13550397 |
jaguar | 6137143 |
jeep | 69589094 |
kia | 19559407 |
land rover | 94400 |
lexus | 30034843 |
lincoln | 7796232 |
mazda | 14377882 |
mercedes-benz | 65436510 |
mercury | 1162558 |
mini | 7507965 |
mitsubishi | 3672210 |
morgan | 74977 |
nissan | 67049248 |
pontiac | 3417008 |
porsche | 17043208 |
ram | 92764848 |
rover | 9843502 |
saturn | 885346 |
subaru | 30836333 |
tesla | 10943535 |
toyota | 134519824 |
volkswagen | 26791675 |
volvo | 8615208 |

+-----+
135 rows selected (21,539 seconds)
0: idbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-0:~
```

```
select * from vehicles where manufacturer == '323ci';
```

```
hdfs dfs -cat result/part* | grep 323ci
```

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

```

Файл Правка Вид Терминал Вкладки Справка
student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~ student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
+-----+-----+-----+-----+-----+-----+-----+-----+
| 528xi 528xi S4 X3 X5 VW GOLF JETTA PASSAT RABBIT CC GTI VOLVO S40 S60 S80 | NULL | OUTBACK IMPREZA ALLROAD AVANT | 323i | 325i | 528i | NULL | 5
+-----+-----+-----+-----+-----+-----+-----+-----+
| NULL | 330i | 323ci | 325ci | 328ci | 525i | 528i | NULL | 5
| 525xi | 330xi | 525it | 528it | 540it | audi | a4 | s4
| a6 | vw | gti | passat | | | | |
+-----+-----+-----+-----+-----+-----+-----+-----+
| z3 z4 x5 x3 525i 528i 530i 540i 335i 535i 550i 635Csi 640i 645Ci 740i 740L 750i 750Li 745i 745Li 750i 760i 760Li M3 M5 M6 325xi 330xi 525it 528it 540it 325xit 330
+-----+-----+-----+-----+-----+-----+-----+-----+
| 528xi 528xi S4 X3 X5 VW GOLF JETTA PASSAT RABBIT CC GTI VOLVO S40 S60 S80 | NULL | OUTBACK IMPREZA ALLROAD AVANT | 323i | 325i | 528i | NULL | 5
+-----+-----+-----+-----+-----+-----+-----+-----+
| NULL | 330i | 323ci | 325ci | 328ci | 525i | 528i | NULL | 5
| 525xi | 330xi | 525it | 528it | 540it | audi | a4 | s4
| a6 | vw | gti | passat | | | | |
+-----+-----+-----+-----+-----+-----+-----+-----+
| z3 z4 x5 x3 525i 528i 530i 540i 335i 535i 550i 635Csi 640i 645Ci 740i 740L 750i 750Li 745i 745Li 750i 760i 760Li M3 M5 M6 325xi 330xi 525it 528it 540it 325xit 330
+-----+-----+-----+-----+-----+-----+-----+-----+
| 528xi 528xi S4 X3 X5 VW GOLF JETTA PASSAT RABBIT CC GTI VOLVO S40 S60 S80 | NULL | OUTBACK IMPREZA ALLROAD AVANT | 323i | 325i | 528i | NULL | 5
+-----+-----+-----+-----+-----+-----+-----+-----+
| NULL | 330i | 323ci | 325ci | 328ci | 525i | 528i | NULL | 5
| 525xi | 330xi | 525it | 528it | 540it | audi | a4 | s4
| a6 | vw | gti | passat | | | | |
+-----+-----+-----+-----+-----+-----+-----+-----+
5 rows selected (3,387 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt: Closing: 0: jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -cat result/part* | grep 323ci
cat: result/part*: No such file or directory
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ ■

```

create table vehicles (\* - Sublime Text (UNREGISTERED))

```

File Edit Selection Find View Goto Tools Project Preferences Help
string create table vehicles (
16 transmission string,
17 VIN string,
18 drive string,
19 size string,
20 type string,
21 paint_color string,
22 image_url string,
23 description string,
24 state string,
25 lat string,
26 long string,
27 posting_date string
28 )
29 ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
30 WITH SERDEPROPERTIES ('field.delim' = ',')
31 location "/user/student3_10/test"
32 tblproperties ("skip.header.line.count"="1")
33 ;
34
35 ALTER TABLE vechicles change id id string;
36
37 ALTER TABLE vechicles change price price string;
38
39 ALTER TABLE vechicles change odometer odometer string;

```

use student3\_10;

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

```

Файл Правка Вид Терминал Вкладки Справка
student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~ student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~
+-----+-----+-----+-----+-----+-----+-----+
0: jdbc:hive2://bigdataanalytics2-worker-shdpt: Closing: 0: jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -cat result/part* | grep 323ci
cat: result/part*: No such file or directory
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Connecting to jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/18 17:50:57 [main]: INFO jdbc.HiveConnection: Connected to bigdataanalytics2-head-shdpt-v31-1-0.novalocal:10000
Connected to: Apache Hive (version 3.1.0.3.1.4.0-315)
Driver: Hive JDBC (version 3.1.0.3.1.4.0-315)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 3.1.0.3.1.4.0-315 by Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdpt: use student3_10;
INFO : Compiling command(queryId=hive_20210418175214_d9104189-0176-4c75-be26-62dbb6e13f5e): use student3_10
INFO : Semantic Analysis Completed (retryal = false)
INFO : Returning Hive schema: Schema(fieldschemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210418175214_d9104189-0176-4c75-be26-62dbb6e13f5e); Time taken: 0.03 seconds
INFO : Executing command(queryId=hive_20210418175214_d9104189-0176-4c75-be26-62dbb6e13f5e): use student3_10
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210418175214_d9104189-0176-4c75-be26-62dbb6e13f5e); Time taken: 0.013 seconds
INFO : OK
No rows affected (0,261 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt> ■

```

ALTER TABLE vechicles change id id string;

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:-
Файл Правка Вид Терминал Вкладки Справка
ocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -cat result/part* | grep 323ci
cat: 'result/part*': No such file or directory
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.1.4.0-315/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.log4j.Log4jLoggerFactory]
Connecting to jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
21/04/18 17:50:57 [main]: INFO jdbc.HiveConnection: Connected to bigdataanalytics2-head-shdpt-v31-1-0.novalocal:10000
Connected to: Apache Hive (version 3.1.0.3.1.4.0-315)
Driver: Hive JDBC (version 3.1.0.3.1.4.0-315)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 3.1.0.3.1.4.0-315 by Apache Hive
0: jdbc:hive2://bigdataanalytics2-worker-shdpt-> use student3_10;
INFO : Compiling command(queryId=hive_20210418175214_d9104189-0176-4c75-be26-62dbb6e13f5e): use student3_10
INFO : Semantic Analysis Completed (retrial = false)
INFO : Returning Hive schema: Schema(fieldsSchemas:null, properties:null)
INFO : Completed compiling command(queryId=hive_20210418175214_d9104189-0176-4c75-be26-62dbb6e13f5e); Time taken: 0.03 seconds
INFO : Executing command(queryId=hive_20210418175214_d9104189-0176-4c75-be26-62dbb6e13f5e): use student3_10
INFO : Starting task [Stage-0:DDL] in serial mode
INFO : Completed executing command(queryId=hive_20210418175214_d9104189-0176-4c75-be26-62dbb6e13f5e); Time taken: 0.013 seconds
INFO : OK
No rows affected (0,261 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt-> ALTER TABLE vechicles change id id string;
Error: Error while compiling statement: FAILED: SemanticException [Error 10001]: Table not found vechicles (state=42502, code=10001)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt-> ■

```

hdfs dfs -cat result/part\* | grep tesla

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:-
Файл Правка Вид Терминал Вкладки Справка
| kia | 19559407 |
| land rover | 94400 |
| lexus | 30034843 |
| lincoln | 7796232 |
| mazda | 14377882 |
| mercedes-benz | 65436510 |
| mercury | 1162558 |
| mini | 7507965 |
| mitsubishi | 3672210 |
| morgan | 74977 |
| nissan | 67049248 |
| pontiac | 3417008 |
| porsche | 17043208 |
| ram | 92764848 |
| rover | 9843502 |
| saturn | 885346 |
| subaru | 30836333 |
| tesla | 10943535 |
| toyota | 134519824 |
| volkswagen | 26791675 |
| volvo | 8615208 |
+-----+
135 rows selected (18,444 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt-> Closing: 0: jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -cat result/part* | grep tesla^C
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -cat result/part* | grep tesla
cat: 'result/part*': No such file or directory
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ ■

```

hdfs dfs -ls test/p\_date=1204

```

Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:-
Файл Правка Вид Терминал Вкладки Справка
INFO : Completed executing command(queryId=hive_20210418191545_51332722-f862-4563-a8f5-644284f024bb); Time taken: 0.003 seconds
INFO : OK
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| vehicles.rid | vehicles.id | vehicles.url | vehicles.region | vehicles.region_url | vehicles.price | vehicles.year | vehicles.manufacturer | vehicles.model | |
| vehicles.condition | vehicles.cylinders | vehicles.fuel | vehicles.odometer | vehicles.title_status | vehicles.transmission | vehicles.vin | vehicles.drive |
| vehicles.size | vehicles.type | vehicles.paint_color | vehicles.image_url | vehicles.description | vehicles.state | vehicles.lat | vehicles.long | vehicles.posting_date | vehicles.p_date |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
No rows selected (0,373 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdpt-> Closing: 0: jdbc:hive2://bigdataanalytics2-worker-shdpt-v31-1-5.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-0.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-4.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-3.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-2.novalocal:2181,bigdataanalytics2-worker-shdpt-v31-1-1.novalocal:2181/default;password=hive;serviceDiscoveryMode=zooKeeper;user=hive;zooKeeperNamespace=hiveserver2
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -mkdir test/p_date=1204
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -mv test/vehicles.csv test/p_date=1204.
mv: 'test/vehicles.csv': No such file or directory
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -mv test/vehicles.csv test/p_date=1204.
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ hdfs dfs -ls test/p_date=1204
Found 1 items
-rw-rw-rw- 2 student3_10 student3_10 346353924 2021-04-15 17:36 test/p_date=1204/vehicles.csv
[student3_10@bigdataanalytics2-head-shdpt-v31-1-0 ~]$ ■

```

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~  
Файл Правка Вид Терминал Вкладки Справка  
manufacturer string,  
model string,  
condition string,  
cylinders string,  
fuel string,  
odometer double,  
title_status string,  
transmission string,  
VIN string,  
drive string,  
size string,  
type string,  
paint_color string,  
image_url string,  
description string,  
state string,  
lat string,  
long string,  
posting_date string  
)  
partitioned by ( p_date string )  
ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'  
WITH SERDEPROPERTIES ('field.delim' = ',')  
location "/user/student3_10/test"  
tblproperties ("skip.header.line.count"="1")  
INFO : Starting task [Stage-0:DDL] in serial mode  
INFO : Completed executing command(queryId=hive_20210418192259_f0915666-e637-4024-9a39-425d4ce318cf); Time taken: 0.086 seconds  
INFO : OK  
No rows affected (0,372 seconds)  
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■
```

msck REPAIR TABLE vehicles;

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~  
Файл Правка Вид Терминал Вкладки Справка  
size string,  
type string,  
paint_color string,  
image_url string,  
description string,  
state string,  
lat string,  
long string,  
posting_date string  
)  
partitioned by ( p_date string )  
ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'  
WITH SERDEPROPERTIES ('field.delim' = ',')  
location "/user/student3_10/test"  
tblproperties ("skip.header.line.count"="1")  
INFO : Starting task [Stage-0:DDL] in serial mode  
INFO : Completed executing command(queryId=hive_20210418192259_f0915666-e637-4024-9a39-425d4ce318cf); Time taken: 0.086 seconds  
INFO : OK  
No rows affected (0,372 seconds)  
0: jdbc:hive2://bigdataanalytics2-worker-shdp> msck REPAIR TABLE vehicles;  
INFO : Compiling command(queryId=hive_20210418192635_cbad9b10-f121-4e15-bd8a-60c4128b56bb): msck REPAIR TABLE vehicles  
INFO : Semantic Analysis Completed (retrial = false)  
INFO : Returning Hive schema: Schema(fieldschemas:null, properties:null)  
INFO : Completed compiling command(queryId=hive_20210418192635_cbad9b10-f121-4e15-bd8a-60c4128b56bb); Time taken: 0.05 seconds  
INFO : Executing command(queryId=hive_20210418192635_cbad9b10-f121-4e15-bd8a-60c4128b56bb): msck REPAIR TABLE vehicles  
INFO : Starting task [Stage-0:DDL] in serial mode  
INFO : Completed executing command(queryId=hive_20210418192635_cbad9b10-f121-4e15-bd8a-60c4128b56bb); Time taken: 0.501 seconds  
INFO : OK  
No rows affected (0,59 seconds)  
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■
```

select \* from vehicles;

```
Терминал - student3_10@bigdataanalytics2-head-shdpt-v31-1-0:~  
Файл Правка Вид Терминал Вкладки Справка  
size string,  
type string,  
paint_color string,  
image_url string,  
description string,  
state string,  
lat string,  
long string,  
posting_date string  
)  
partitioned by ( p_date string )  
ROW FORMAT serde 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'  
WITH SERDEPROPERTIES ('field.delim' = ',')  
location "/user/student3_10/test"  
tblproperties ("skip.header.line.count"="1")  
INFO : Starting task [Stage-0:DDL] in serial mode  
INFO : Completed executing command(queryId=hive_20210418192259_f0915666-e637-4024-9a39-425d4ce318cf); Time taken: 0.086 seconds  
INFO : OK  
No rows affected (0,372 seconds)  
0: jdbc:hive2://bigdataanalytics2-worker-shdp> msck REPAIR TABLE vehicles;  
INFO : Compiling command(queryId=hive_20210418192635_cbad9b10-f121-4e15-bd8a-60c4128b56bb): msck REPAIR TABLE vehicles  
INFO : Semantic Analysis Completed (retrial = false)  
INFO : Returning Hive schema: Schema(fieldschemas:null, properties:null)  
INFO : Completed compiling command(queryId=hive_20210418192635_cbad9b10-f121-4e15-bd8a-60c4128b56bb); Time taken: 0.05 seconds  
INFO : Executing command(queryId=hive_20210418192635_cbad9b10-f121-4e15-bd8a-60c4128b56bb): msck REPAIR TABLE vehicles  
INFO : Starting task [Stage-0:DDL] in serial mode  
INFO : Completed executing command(queryId=hive_20210418192635_cbad9b10-f121-4e15-bd8a-60c4128b56bb); Time taken: 0.501 seconds  
INFO : OK  
No rows affected (0,59 seconds)  
0: jdbc:hive2://bigdataanalytics2-worker-shdp> select * from vehicles; ■
```

select manufacturer, sum(price) from vehicles group by manufacturer order by manufacturer;

Терминал - student3\_10@bigdataanalytics2-head-shdpt-v31-1-0:~

```
Файл Правка Вид Терминал Вкладки Справка
INFO : MERGE_PHASE_TIME: 145
INFO : NUM_DISK_TO_DISK_MERGES: 0
INFO : NUM FAILED SHUFFLE INPUTS: 0
INFO : NUM MEM TO DISK MERGES: 0
INFO : NUM SHUFFLED INPUTS: 0
INFO : NUM_SKIPPED_INPUTS: 2
INFO : REDUCE_INPUT_GROUPS: 0
INFO : REDUCE_INPUT_RECORDS: 0
INFO : SHUFFLE_BYTES: 0
INFO : SHUFFLE_BYTES_DECOMPRESSED: 0
INFO : SHUFFLE_BYTES_DISK_DIRECT: 0
INFO : SHUFFLE_BYTES_TO_DISK: 0
INFO : SHUFFLE_BYTES_TO_MEM: 0
INFO : SHUFFLE_PHASE_TIME: 140
INFO : SPILLED_RECORDS: 0
INFO : TaskCounter Reducer_3_OUTPUT_out_Reducer_3:
INFO : OUTPUT_RECORDS: 0
INFO : org.apache.hadoop.hive.ql.exec.tez.HiveInputCounters:
INFO : GROUPED_INPUT_SPLITS Map_1: 0
INFO : INPUT_DIRECTORIES_Map_1: 0
INFO : INPUT_FILES_Map_1: 0
INFO : RAW_INPUT_SPLITS_Map_1: 0
INFO : Completed executing command(queryId=hive_20210418193705_90967940-dc82-4329-903b-a6439cb73748); Time taken: 10.835 seconds
INFO : OK
+-----+-----+
| manufacturer | _c1 |
+-----+-----+
+-----+-----+
No rows selected (11,24 seconds)
0: jdbc:hive2://bigdataanalytics2-worker-shdp> ■
```