

Файл Действия Правка Вид Справка

georgiy@georgiy-extensa2510g: ~

```
georgiy@georgiy-extensa2510g:~$ sudo docker cp Загрузки/dataset.csv namenode:/dataset.csv
```

```
[sudo] пароль для georgiy:
```

```
Successfully copied 1.15MB to namenode:/dataset.csv
```

```
georgiy@georgiy-extensa2510g:~$ sudo docker exec -it namenode bash
```

```
root@namenode:/# ls
```

AB_NYC_2019.csv	dataset.csv	employees.java	hadoop	lib	metastore_db	proc	run.sh	sys	var
bin	derby.log	entrypoint.sh	hadoop-data	lib64	mnt	root	sbin	tmp	
boot	dev	etc	home	media	opt	run	srv	usr	

```
root@namenode:/# hdfs dfs -ls /
```

```
Found 2 items
```

drwxrwxr-x	-	root	supergroup	0	2024-02-01 18:19	/tmp
drwxr-xr-x	-	root	supergroup	0	2024-02-01 18:19	/user

```
root@namenode:/# hdfs dfs -put /dataset.csv /
```

```
root@namenode:/# hdfs dfs -ls /
```

```
Found 3 items
```

-rw-r--r--	3	root	supergroup	1152966	2024-02-10 06:19	/dataset.csv
drwxrwxr-x	-	root	supergroup	0	2024-02-01 18:19	/tmp
drwxr-xr-x	-	root	supergroup	0	2024-02-01 18:19	/user

```
root@namenode:/#
```

```
Connected to: Apache Hive (version 2.3.2)
Driver: Hive JDBC (version 2.3.2)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 2.3.2 by Apache Hive
0: jdbc:hive2://hive-server:10000/> show tables;
+-----+
| tab_name |
+-----+
+-----+
No rows selected (1.787 seconds)
0: jdbc:hive2://hive-server:10000/> CREATE TABLE airbnb (price INT) row format delimited fields terminated by ',' stored
as textfile;
No rows affected (0.379 seconds)
0: jdbc:hive2://hive-server:10000/> show tables;
+-----+
| tab_name |
+-----+
| airbnb   |
+-----+
1 row selected (0.177 seconds)
0: jdbc:hive2://hive-server:10000/> describe airbnb;
+-----+-----+-----+
| col_name | data_type | comment |
+-----+-----+-----+
| price    | int       |         |
+-----+-----+-----+
1 row selected (0.083 seconds)
0: jdbc:hive2://hive-server:10000/> █
```

```
0: jdbc:hive2://hive-server:10000/> show tables;
+-----+
| tab_name |
+-----+
| airbnb   |
+-----+
1 row selected (0.177 seconds)
0: jdbc:hive2://hive-server:10000/> describe airbnb;
+-----+-----+-----+
| col_name | data_type | comment |
+-----+-----+-----+
| price    | int       |         |
+-----+-----+-----+
1 row selected (0.083 seconds)
0: jdbc:hive2://hive-server:10000/> LOAD DATA INPATH '/dataset.csv' INTO TABLE airbnb;
No rows affected (6.968 seconds)
0: jdbc:hive2://hive-server:10000/> select * from airbnb limit 5;
+-----+
| airbnb.price |
+-----+
| NULL         |
| 149          |
| 225          |
| 150          |
| 89           |
+-----+
5 rows selected (0.788 seconds)
0: jdbc:hive2://hive-server:10000/> █
```


Файл Действия Правка Вид Справка

georgiy@georgiy-extensa2510g: ~ x georgiy@georgiy-extensa2510g: ~ x

```
Driver: Hive JDBC (version 2.3.2)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 2.3.2 by Apache Hive
0: jdbc:hive2://hive-server:10000/> show tables;
+-----+
| tab_name |
+-----+
| airbnb   |
+-----+
1 row selected (2.667 seconds)
0: jdbc:hive2://hive-server:10000/> select * from airbnb limit 5;
+-----+
| airbnb.price |
+-----+
| NULL         |
| 149          |
| 225          |
| 150          |
| 89           |
+-----+
5 rows selected (6.755 seconds)
0: jdbc:hive2://hive-server:10000/> select avg(price) from airbnb;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
Unknown HS2 problem when communicating with Thrift server.
Error: org.apache.thrift.transport.TTransportException: java.net.SocketException: Broken pipe (Write failed) (state=08S01,code=0)
0: jdbc:hive2://hive-server:10000/> georgiy@georgiy-extensa2510g:~$
```