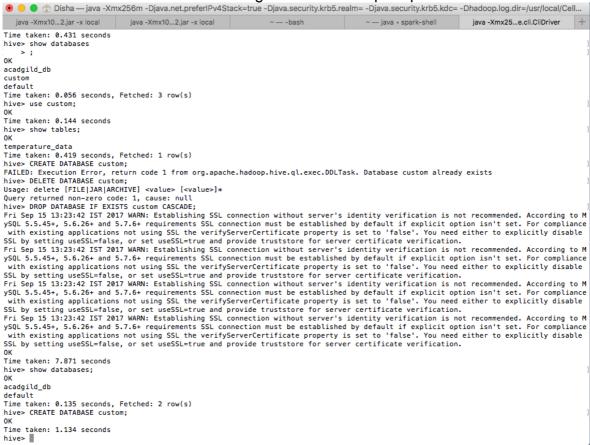
## step 1:

create database named custom using hive command prompt:



2. copy dataset.txt to hadoop using put command and then check the data in the hadoop path using cat command:

```
java -Xmx10...2.jar -x local java -Xmx10...2.jar -x local
                                                                                                                                                 ~ — -bash
                                                                                                                                                                                                  ~ — java • spark-shell java -Xmx25...e.cli.CliDriver
                   [-touchz <path> ...]
[-truncate [-w] <length> <path> ...]
[-usage [cmd ...]]
Generic options supported are

-conf <configuration file> specify an application configuration file
-D -D -cpreperty=value> use value for given property
-fs <file:///|hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.
-jt <local|resourcemanager:port> specify a ResourceManager
-files <comma separated list of files> specify comma separated files to be copied to the map reduce cluster
-libjars <comma separated list of jars> specify comma separated jar files to include in the classpath.
-archives <comma separated list of archives> specify comma separated archives to be unarchived on the compute machines.
 The general command line syntax is command [genericOptions] [commandOptions]
 Chaitanyas-MacBook-Pro:~ Disha$ hadoop fs -cat /hive_data/temperature_data.txt 17/09/15 13:31:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where
 17/09/15 13:31:57 WAI applicable 10-01-1990,123112,10 14-02-1991,283901,11 10-03-1990,381920,15 10-01-1991,302918,22 12-02-1990,384902,9 10-01-1991,123112,11
 14-02-1990,283901,12
10-03-1991,381920,16
10-01-1990,302918,23
 12-02-1991,384902,10
10-01-1993,123112,11
14-02-1994,283901,12
 10-03-1993.381920.16
 10-01-1994,302918,23
12-02-1991,384902,10
10-01-1991,123112,11
 14-02-1990,283901,12
10-03-1991,381920,16
 10-01-1990,302918,23
 12-02-1991.384902.10
```

- 3. Create a table named temperature\_data inside custom having below fields:
- 1. date (mm-dd-yyyy) format
- 2. zip code
- 3. temperature

The table will be loaded from comma-delimited file.

commad for table creation:

```
create external table temperature_data (
dt string,
zipcode BIGINT,
temperature int
)
row format delimited
fields terminated by ','
location '/hive_data/';
```