1. Create a table to store show and channel details:

```
create table show_to_channel(show String,channel String)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

STORED AS TEXTFILE;
```

2. Insert records from a file:

```
LOAD DATA LOCAL INPATH 'c:\dataset\join2_genchanA.txt' INTO TABLE show_to_channel;

LOAD DATA LOCAL INPATH 'c:\dataset\join2_genchanB.txt' INTO TABLE show_to_channel;

LOAD DATA LOCAL INPATH 'c:\dataset\join2_genchanB.txt' INTO TABLE show_to_channel;
```

3. Create a table to store show and number of viewers details:

```
create table show_to_viewers(show String,no_of_viewers int)

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'

STORED AS TEXTFILE;
```

4. Insert records from a file:

LOAD DATA LOCAL INPATH 'c:\dataset\join2_gennumA.txt' INTO TABLE show_to_viewers;

LOAD DATA LOCAL INPATH 'c:\dataset\join2_gennumB.txt' INTO TABLE show_to_viewers;

LOAD DATA LOCAL INPATH 'c:\dataset\join2_gennumC.txt' INTO TABLE show_to_viewers;

5. What is the total number of viewers for shows on ABC?

select sum(no_of_viewers) from show_to_viewers where show in (select distinct(show) from show_to_channel where channel='ABC');

```
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
2019-03-27 16:54:06,643 main WARN Unable to instantiate org.fusesource.jansi.WindowsAnsiOutputStream
2019-03-27 16:54:06,651 main WARN Unable to instantiate org.fusesource.jansi.WindowsAnsiOutputStream
2019-03-27 16:54:07     Starting to launch local task to process map join;     maximum memory = 477626368
2019-03-27 16:54:08     Dump the side-table for tag: 1 with group count: 20 into file: file:/C:/Users/jigim/AppDa
1312499-1/-local-10006/HashTable-Stage-4/MapJoin-mapfile01--.hashtable
2019-03-27 16:54:08     Uploaded 1 File to: file:/C:/Users/jigim/AppData/Local/Temp/jigim/d18817bb-2848-4c74-98e
pJoin-mapfile01--.hashtable (851 bytes)
2019-03-27 16:54:08     End of local task; Time Taken: 1.012 sec.
1115974
1 row selected (75.573 seconds)
hive>
```

===

6. What is the number of viewers for the BAT channel?

select sum(no_of_viewers) from show_to_viewers where show in (select distinct(show) from show_to_channel where channel='BAT');

```
MapReduce Total cumulative CPU time: 3 seconds 513 msec
Ended Job = job_1553670772018_0006
MapReduce Jobs Launched:
Stage-Stage-3: Map: 1  Reduce: 1   Cumulative CPU: 4.793 sec   HDFS Read: 18573 HDFS Write: 1627 SUCCESS
Stage-Stage-4: Map: 1   Cumulative CPU: 2.717 sec   HDFS Read: 109041 HDFS Write: 117 SUCCESS
Stage-Stage-2: Map: 1  Reduce: 1   Cumulative CPU: 3.513 sec   HDFS Read: 5113 HDFS Write: 107 SUCCESS
Total MapReduce CPU Time Spent: 11 seconds 23 msec
OK
3031762
```

7. What is the most viewed show on ABC channel?

select show,sum(no_of_viewers) s from show_to_viewers where show in (select distinct(show) from show_to_channel where channel='ABC') group by show order by s desc limit 1;

```
=========================================
select show,sum(no_of_viewers) s from show_to_viewers where show in (select distinct(show) from show_to_channel where channel='ABC') group by show order b;
desc limit 1;

OK
Hourly_Talking 108163
1 row selected (95.656 seconds)
===========================
```

8. What are the aired shows on ZOO,NOX, ABC channels ?

select distinct show from show_to_channel where channel='ZOO';

```
hive> select distinct show  from show_to_channel where channel='ZOO';
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.
Query ID = jigim_20190327003922_2d30d251-f4ac-42db-8ef5-002759c7819b
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1553624240821_0008, Tracking URL = http://DESKTOP-
FHNBR9V:8088/proxy/application_1553624240821_0008/
Kill Command = C:\hadoop-2.7.6\bin\hadoop.cmd job  -kill job_1553624240821_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2019-03-27 00:39:37,949 Stage-1 map = 0%,   reduce = 0%
2019-03-27 00:39:44,247 Stage-1 map = 100%,   reduce = 0%, Cumulative CPU 2.498 sec
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.
2019-03-27 00:39:51,568 Stage-1 map = 100%,   reduce = 100%, Cumulative CPU 5.059 sec
MapReduce Total cumulative CPU time: 5 seconds 59 msec
Ended Job = job_1553624240821_0008
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 5.059 sec   HDFS Read: 18892 HDFS Write: 87 SUCCESS
Total MapReduce CPU Time Spent: 5 seconds 59 msec
OK
No rows selected (29.82 seconds)
```

select distinct show from show_to_channel where channel='NOX';

```
Almost_Games
Almost_News
Almost_Show
Almost_Sports
Baked_Cooking
Baked_Games
Baked_News
Baked_Show
Baked_Talking
Cold_Cooking
Cold_News
Cold_Sports
Cold_Talking
Dumb_Cooking
Dumb_Games
Dumb_Show
Dumb_Talking
Hot_Cooking
Hot_Games
Hot_News
Hot_Show
Hot_Talking
Hourly_Games
Hourly_News
Hourly_Show
Hourly_Sports
Hourly_Talking
Loud_Cooking
Loud_Sports
Loud_Talking
PostModern_Cooking
PostModern_Show
PostModern_Sports
PostModern_Talking
Surreal_Cooking
Surreal_Show
Surreal_Talking
37 rows selected (25.619 seconds)
=======================================
```

select distinct show from show_to_channel where channel='ABC';

```
Launching Job 1 out of 1
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1553624240821_0010, Tracking URL = http://DESKTOP-FHNBR9V:8088/proxy/applicati
Kill Command = C:\hadoop-2.7.6\bin\hadoop.cmd job  -kill job_1553624240821_0010
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2019-03-27 00:44:10,727 Stage-1 map = 0%,  reduce = 0%
2019-03-27 00:44:17,028 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 2.279 sec
2019-03-27 00:44:24,374 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 4.901 sec
MapReduce Total cumulative CPU time: 4 seconds 901 msec
Ended Job = job_1553624240821_0010
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 4.901 sec   HDFS Read: 18892 HDFS Write: 575 S
Total MapReduce CPU Time Spent: 4 seconds 901 msec
OK
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Cons
using Hive 1.X releases.
Almost_Games
Almost_News
Almost_Show
Baked_Games
Baked_News
Cold_News
Cold_Sports
Dumb_Show
Dumb_Talking
Hot_Games
Hot_Show
Hourly_Cooking
Hourly_Show
Hourly_Talking
Loud_Games
Loud_Show
PostModern_Games
PostModern_News
Surreal_News
Surreal_Sports
20 rows selected (27.183 seconds)
```

****