

```
scala> hive
<console>:24: error: not found: value hive
hive
^

scala> [hadoop@ip-172-31-77-33 ~]$ hive

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j2.properties Async: true
hive> CREATE TABLE IF NOT EXISTS delay_flights (
  > id int,
  > Year int,
  > Month int,
  > DayOfMonth int,
  > DayOfWeek int,
  > DepTime string,
  > CRSDepTime string,
  > ArrTime string,
  > CRSArrTime string,
  > UniqueCarrier string,
  > FlightNum int,
  > TailNum string,
  > ActualElapsedTime string,
  > CRSElapsedTime string,
  > AirTime string,
  > ArrDelay string,
  > DepDelay string,
  > Origin string,
  > Dest string,
  > Distance int,
  > TaxiIn int,
  > TaxiOut int,
  > Cancelled int,
  > CancellationCode string,
  > Diverted int,
  > CarrierDelay string,
  > WeatherDelay string,
  > NASDelay string,
  > SecurityDelay string,
  > LateAircraftDelay string)
  > COMMENT 'Delay flights Table'
  > ROW FORMAT DELIMITED
  > FIELDS TERMINATED BY ',';

OK
Time taken: 0.68 seconds
```

```
hive> CREATE TABLE IF NOT EXISTS delay_flights (
  > CRSElapsedTime string,
  > AirTime string,
  > ArrDelay string,
  > DepDelay string,
  > Origin string,
  > Dest string,
  > Distance int,
  > TaxiIn int,
  > TaxiOut int,
  > Cancelled int,
  > CancellationCode string,
  > Diverted int,
  > CarrierDelay string,
  > WeatherDelay string,
  > NASDelay string,
  > SecurityDelay string,
  > LateAircraftDelay string)
  > COMMENT 'Delay flights Table'
  > ROW FORMAT DELIMITED
  > FIELDS TERMINATED BY ',';

OK
Time taken: 0.68 seconds
hive> LOAD DATA INPATH 's3://bigdataassignm/DelayedFlights-updated.csv' INTO TABLE delay_flights;
Loading data to table default.delay_flights
OK
Time taken: 8.825 seconds
hive> set hive.cli.print.header=true;
hive> SELECT Year, avg((CarrierDelay /ArrDelay)*100) from delay_flights WHERE Year=2003 AND Year<=2010 GROUP BY Year ORDER BY Year;
Query ID = hadoop_20230305154131_c834ec5e-6a51-4ad2-a226-3f84c1f5af43
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1678026367764_0005)

-----
VERTICES    MODE        STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... container  SUCCEEDED   1         1         0         0         0         0
Reducer 2 ..... container  SUCCEEDED   2         2         0         0         0         0
Reducer 3 ..... container  SUCCEEDED   1         1         0         0         0         0
-----
VERTICES: 03/03  [=====] 100% ELAPSED TIME: 7.25 s
-----

OK
year      cl
```

```

PuTTY (inactive)
Query ID = hadoop_20230305154208_a184263d-9c75-4410-997f-d4551b1f2f42
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1678026367764_0005)

-----
VERTICES    MODE      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... container  SUCCEEDED   1         1         0         0         0         0
Reducer 2 ..... container  SUCCEEDED   2         2         0         0         0         0
Reducer 3 ..... container  SUCCEEDED   1         1         0         0         0         0
-----
VERTICES: 03/03 [=====>>] 100% ELAPSED TIME: 0.57 s
-----
OK
year      _cl
2003      24.55754975557537
2004      43.64459443230064
2005      28.019776372022875
2006      30.453296261292582
2007      19.85000701797128
2008      28.88346981456985
2009      28.33058554238574
2010      21.893102460159868
Time taken: 1.546 seconds, Fetched: 0 row(s)
hive> SELECT Year, avg((CarrierDelay /ArrDelay)*100) from delay_flights WHERE Year>=2003 AND Year<=2010 GROUP BY Year ORDER BY Year;
Query ID = hadoop_20230305154212_dedc1ce9-494a-454f-a3a4-8a8b4ff28b9e
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1678026367764_0005)

-----
VERTICES    MODE      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... container  SUCCEEDED   1         1         0         0         0         0
Reducer 2 ..... container  SUCCEEDED   2         2         0         0         0         0
Reducer 3 ..... container  SUCCEEDED   1         1         0         0         0         0
-----
VERTICES: 03/03 [=====>>] 100% ELAPSED TIME: 0.38 s
-----
OK
year      _cl
2003      24.55754975557537
2004      43.64459443230064

```

```
Putty (inactive)
VERTICES: 03/03 [=====]>>] 100% ELAPSED TIME: 0.46 s
-----
OK
year      cl
2003      24.55754975557537
2004      43.64459443230064
2005      28.019776372022875
2006      30.453296261292582
2007      19.85000701797128
2008      28.88346981456985
2009      28.33058554239574
2010      21.893102460159568
Time taken: 1.355 seconds, Fetched: 8 row(s)
hive> SELECT Year, avg((CarrierDelay /ArrDelay)*100) from delay_flights WHERE Year>=2003 AND Year<=2010 GROUP BY Year ORDER BY Year;
Query ID = hadoop_20230305154150_a99e5a0e-8c17-4ecc-8fd8-df1ed8e600b9
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1678026367764_0005)

-----
VERTICES      MODE      STATUS      TOTAL      COMPLETED      RUNNING      PENDING      FAILED      KILLED
-----
Map 1 ..... container      SUCCEEDED      1              1              0              0              0              0
Reducer 2 ..... container      SUCCEEDED      2              2              0              0              0              0
Reducer 3 ..... container      SUCCEEDED      1              1              0              0              0              0
-----
VERTICES: 03/03 [=====]>>] 100% ELAPSED TIME: 0.43 s
-----
OK
year      cl
2003      24.55754975557537
2004      43.64459443230064
2005      28.019776372022875
2006      30.453296261292582
2007      19.85000701797128
2008      28.88346981456985
2009      28.33058554239574
2010      21.893102460159568
Time taken: 1.295 seconds, Fetched: 8 row(s)
hive> SELECT Year, avg((CarrierDelay /ArrDelay)*100) from delay_flights WHERE Year>=2003 AND Year<=2010 GROUP BY Year ORDER BY Year;
Query ID = hadoop_20230305154159_635e61f1-655b-40df-ba95-9b8194ce4770
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1678026367764_0005)
```

```
Putty (inactive)
hive> LOAD DATA INPATH 's3://bigdataassign/DelayedFlights-updated.csv' INTO TABLE delay_flights;
Loading data to table default.delay_flights
OK
Time taken: 0.825 seconds
hive> set hive.cli.print.header=true;
hive> SELECT Year, avg((CarrierDelay /ArrDelay)*100) from delay_flights WHERE Year>=2003 AND Year<=2010 GROUP BY Year ORDER BY Year;
Query ID = hadoop_20230305154131_c834ec5e-6a51-4ad2-a226-3f84c1f5af43
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1678026367764_0005)

-----
VERTICES      MODE      STATUS      TOTAL      COMPLETED      RUNNING      PENDING      FAILED      KILLED
-----
Map 1 ..... container      SUCCEEDED      1              1              0              0              0              0
Reducer 2 ..... container      SUCCEEDED      2              2              0              0              0              0
Reducer 3 ..... container      SUCCEEDED      1              1              0              0              0              0
-----
VERTICES: 03/03 [=====]>>] 100% ELAPSED TIME: 7.25 s
-----
OK
year      cl
2003      24.55754975557537
2004      43.64459443230064
2005      28.019776372022875
2006      30.453296261292582
2007      19.85000701797128
2008      28.88346981456985
2009      28.33058554239574
2010      21.893102460159568
Time taken: 12.777 seconds, Fetched: 8 row(s)
hive> SELECT Year, avg((CarrierDelay /ArrDelay)*100) from delay_flights WHERE Year>=2003 AND Year<=2010 GROUP BY Year ORDER BY Year;
Query ID = hadoop_20230305154146_d5dcf990-6766-4da7-a3b7-1162c9cf25b4
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1678026367764_0005)

-----
VERTICES      MODE      STATUS      TOTAL      COMPLETED      RUNNING      PENDING      FAILED      KILLED
-----
Map 1 ..... container      SUCCEEDED      1              1              0              0              0              0
Reducer 2 ..... container      SUCCEEDED      2              2              0              0              0              0
Reducer 3 ..... container      SUCCEEDED      1              1              0              0              0              0
-----
```