

Screenshot of query results:

```
br33@dsc650-breannaparker: x + v
> tblproperties("skip.header.line.count"="1");
OK
Time taken: 1.975 seconds
hive> select * from grades;
OK
Time taken: 2.603 seconds
hive> LOAD DATA INPATH '/grades.csv' INTO TABLE grades;
Loading data to table default.grades
OK
Time taken: 0.417 seconds
hive> select * from grades;
OK
Alfalfa Aloysius      123-45-6789    40.0    90      100.0    83.0    49.0    D-
Alfred University    123-12-1234    41.0    97      96.0    97.0    48.0    D+
Gerty Gramma        567-89-0123    41.0    80      60.0    40.0    44.0    C
Android Electric     087-65-4321    42.0    23      36.0    45.0    47.0    B-
Bumpkin Fred        456-78-9012    43.0    78      88.0    77.0    45.0    A-
Rubble Betty        234-56-7890    44.0    90      80.0    90.0    46.0    C-
Noshov Cecil        345-67-8901    45.0    11      -1.0    4.0     43.0    F
Buff Bif            632-79-9939    46.0    20      30.0    40.0    50.0    B+
Airlump Andrew      223-45-6789    49.0    1       90.0    100.0   83.0    A
Backus Jim          143-12-1234    48.0    1       97.0    96.0    97.0    A+
Carnivore Art        565-89-0123    44.0    1       80.0    60.0    40.0    D+
Dandy Jim           087-75-4321    47.0    1       23.0    36.0    45.0    C+
Elephant Ima        456-71-9012    45.0    1       78.0    88.0    77.0    B-
Franklin Benny       234-56-2890    50.0    1       90.0    80.0    90.0    B-
George Boy          345-67-3901    40.0    1       11.0    -1.0    4.0     B
Heffalump Harvey    632-79-9439    30.0    1       20.0    30.0    40.0    C
Time taken: 0.162 seconds, Fetched: 16 row(s)
hive>
```

## Query 1:

```

br33@dsc650-breannaparker: X + v
hive> select * from grades order by Grade ASC;
2023-12-12 21:45:30,138 INFO [c03a0337-1e58-4ffa-a03c-9c4ecb21bb8d main] reducesink.VectorReduceSinkInfo@166ddfb7
constructor vectorReduceSinkInfo org.apache.hadoop.hive.ql.plan.VectorReduceSinkInfo@166ddfb7
Query ID = root_20231212214529_d3fc152a-2baa-4cd9-8f80-dd20dc387433
Total jobs = 1
Launching Job 1 out of 1
Tez session was closed. Reopening...
2023-12-12 21:45:30,813 INFO [c03a0337-1e58-4ffa-a03c-9c4ecb21bb8d main] client.RMPProxy: Connecting to ResourceManager at dsc650-breannaparker:8032
Session re-established.
Session re-established.
Status: Running (Executing on YARN cluster with App id application_1702416591548_0004)

-----
VERTICES      MODE      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... container  SUCCEEDED    1         1         0         0         0         0
Reducer 2 ..... container  SUCCEEDED    1         1         0         0         0         0
-----
VERTICES: 02/02 [=====] 100% ELAPSED TIME: 5.48 s
-----
OK
Airpump Andrew 223-45-6789 49.0 1 90.0 100.0 83.0 A
Backus Jim 143-12-1234 48.0 1 97.0 96.0 97.0 A+
Bumpkin Fred 456-78-9012 43.0 78 88.0 77.0 45.0 A-
George Boy 345-67-3901 40.0 1 11.0 -1.0 4.0 B
Buff Bif 632-79-9939 46.0 20 30.0 40.0 50.0 B+
Android Electric 087-65-4321 42.0 23 36.0 45.0 47.0 B-
Franklin Benny 234-56-2890 50.0 1 90.0 80.0 90.0 B-
Elephant Ima 456-71-9012 45.0 1 78.0 88.0 77.0 B-
Heffalump Harvey 632-79-9439 30.0 1 20.0 30.0 40.0 C
Gerty Gramma 567-89-0123 41.0 80 60.0 40.0 44.0 C
Dandy Jim 087-75-4321 47.0 1 23.0 36.0 45.0 C+
Rubble Betty 234-56-7890 44.0 90 80.0 90.0 46.0 C-
Carnivore Art 565-89-0123 44.0 1 80.0 60.0 40.0 D+
Alfred University 123-12-1234 41.0 97 96.0 97.0 48.0 D+
Alfalfa Aloysius 123-45-6789 40.0 90 100.0 83.0 49.0 D-
Noshow Cecil 345-67-8901 45.0 11 -1.0 4.0 43.0 F
Time taken: 13.419 seconds, Fetched: 16 row(s)
hive>

```

## Query 2:

```

br33@dsc650-breannaparker: X + v
hive> select * from grades where grade = 'A';
OK
Airpump Andrew 223-45-6789 49.0 1 90.0 100.0 83.0 A
Time taken: 0.24 seconds, Fetched: 1 row(s)
hive>

```

### Query 3:

```
br33@dsc650-breannaparker: X + v
hive> select count (grade) from grades;
2023-12-12 21:58:42,488 INFO [c03a0337-1e58-4ffa-a03c-9c4ecb21bb8d main] reducesink.VectorReduceSinkEmptyKeyOperator: V
ectorReduceSinkEmptyKeyOperator constructor vectorReduceSinkInfo org.apache.hadoop.hive.ql.plan.VectorReduceSinkInfo@20a
18cc9
Query ID = root_20231212215842_9938739f-c450-48cd-aac5-f50b3f77cfd
Total jobs = 1
Launching Job 1 out of 1
Tez session was closed. Reopening...
2023-12-12 21:58:42,667 INFO [c03a0337-1e58-4ffa-a03c-9c4ecb21bb8d main] client.RMPProxy: Connecting to ResourceManager
at master/172.28.1.1:8032
Session re-established.
Session re-established.
Status: Running (Executing on YARN cluster with App id application_1702416591548_0005)

-----
VERTICES      MODE      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... container  SUCCEEDED    1         1         0         0         0         0
Reducer 2 ..... container  SUCCEEDED    1         1         0         0         0         0
-----
VERTICES: 02/02 [=====>>] 100% ELAPSED TIME: 4.84 s
-----
OK
16
Time taken: 11.44 seconds, Fetched: 1 row(s)
hive>
```

From my own dataset. Found a list of 100 organizations and their data:

Query 1:

```
hive> select name from orgs where num_employees > 6000;  
OK  
Mayer Group  
Hansen-Everett  
Carr Inc  
Gaines Inc  
Kidd Group  
Crane-Clarke  
Glover-Pope  
Hodge-Ayers  
Branch-Mann  
Ayala LLC  
Burch-Ewing  
Cuevas-Moss  
Arroyo Inc  
Prince PLC  
Brock-Blackwell  
Gonzales Ltd  
Cherry PLC  
Melton-Nichols  
Potter-Walsh  
Soto Group  
Riley Ltd  
Duffy-Levy  
Wagner LLC  
Hicks LLC  
Cummings-Rojas  
Crawford-Rivera  
Montes-Hensley  
Eaton-Morales  
Pineda-Morton  
Huang and Sons  
Rowe and Sons  
Time taken: 0.156 seconds, Fetched: 31 row(s)  
hive>
```

## Query 2:

```
br33@dsc650-breannaparker: X + v
hive> select count(founded), founded from orgs group by founded order by count(founded) desc;
2023-12-12 22:30:20,689 INFO [428b21e3-6527-43ee-8cf3-5c9c0ed114a0 main] reducesink.VectorReduceSinkObjectHashOperator:
VectorReduceSinkObjectHashOperator constructor vectorReduceSinkInfo org.apache.hadoop.hive ql.plan.VectorReduceSinkInfo
@c8ed958
Query ID = root_20231212223020_ac308a8f-8700-4ee6-9dd5-af90714de25a
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1702416591548_0008)

-----
VERTICES      MODE      STATUS  TOTAL  COMPLETED  RUNNING  PENDING  FAILED  KILLED
-----
Map 1 ..... container  SUCCEEDED    1         1         0         0         0         0
Reducer 2 ..... container  SUCCEEDED    1         1         0         0         0         0
Reducer 3 ..... container  SUCCEEDED    1         1         0         0         0         0
-----
VERTICES: 03/03 [=====>>>] 100% ELAPSED TIME: 4.80 s
-----
OK
4      2012
4      1991
3      2021
3      1970
3      1972
3      1990
3      1971
3      2009
3      1986
2      1976
2      1998
2      1999
2      2001
2      2002
2      1987
2      2006
2      1988
```

## Query 3:

```
br33@dsc650-breannaparker: X + v
hive> select index,name from orgs where country = 'Sweden' or industry = 'Ceramics';
OK
15      Pacheco-Spears
44      Prince PLC
Time taken: 0.133 seconds, Fetched: 2 row(s)
hive>
```