

## ORDER BY vs SORT BY

- Let's take a small sample of first 50 employees from the table -
  - **create table emp\_50 as select \* from emp limit 50;**
- Order records in ascending order of their salary using ORDER BY clause-
  - **select info.gender, job.salary from emp\_50 order by job.salary;**
- Use the SORT BY clause to write the above query -
  - **select info.gender, job.salary from emp\_50 sort by salary;**
- Increase the number of reducers. We can increase the number of reducers with the following command -
  - **set mapreduce.job.reduces;**
  - Set to two reducers -
    - **set mapreduce.job.reduces=2;**
- Run the ORDER BY command -
  - **select info.gender, job.salary from emp\_50 order by job.salary;**
  - The output is in order over here.
  - Notice that ORDER BY is using only 1 of the 2 reducers here.
- Run the SORT BY command -
  - **select info.gender, job.salary from emp\_50 sort by salary;**
  - Notice that the output is not in order. This is because the records here were distributed to two reducers. And both the reducers ordered the records locally.
  - Notice SORT BY is using both the reducers here.