

One reducer —> Everything is sorted..

10 reducers —> Output in each reducer is sorted but its not globally

Total Order Partitioner...

if  $k_1 > k_2$  then reducer for key 1 > reducer for key 2

reducer	reducer
a	
b	d
c	

Dude, how do i write this code

Dude, how do i bundle this code...

## How to create the code:

### OPTION 1 USING Eclipse:

- Take the code from LMS
- Unzip it
- Take the wordcount folder and put in Workspace
- Open Eclipse
- New Java Project

- Include libraries from  
/usr/lib/hadoop, /usr/lib/hadoop/lib, /usr/lib/hadoop/hadoop-mapreduce, /usr/lib/hadoop-yarn
  - Export the code as JAR file
  - Run the JAR
- Bang you are done

## **OPTION 2 USING Build Tool called ANT:**

- Install ANT: `sudo yum install ant`
- Copy wordcount folder from the LMS code files to say a new location of your choice, we did it in america
- `cd wordcount` and delete existing jar
- Run `$>ant`
- Bang your JAR is ready..
- Run your code...

If you want more details each program has a read me file...

At this moment, we are clear with:

- HDFS
- How code gets submitted
- How code gets bundled from Eclipse as well as from command line using ANT

Last question to be answered today is... HOW TO  
CODE 😊 MR...

Lets jump on to it..

Before that lets take up any questions...

