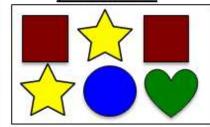
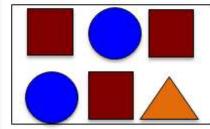
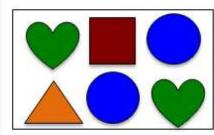


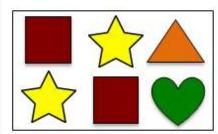
## STEP 0 -STORE TO **HDFS**

Assume 4 data partitions.









## 1 - MAP



( , 1 )( , 1 ) $(\blacksquare,1)$ (0,1)(0,1) $(\triangle,1)$ 

(\*,1)(0,1)(0,1) $(\triangle,1)$ ( > , 1 )

(**1**,1)  $(\blacksquare,1)$ (<del>\(\alpha\)</del>,1)  $(^{\uparrow}_{\uparrow},1)$ (<u>A</u>,1) (\*,1)

## 2 - SHUFFLE 3 - REDUCE and SORT

(**1**,1)  $(\blacksquare,1)$  $( \blacksquare, 1)$  $(\blacksquare,1)$  $(\blacksquare,1)$  $(\blacksquare,1)$  $(\blacksquare,1)$  $(\blacksquare,1)$ 

 $(\diamondsuit,1)$  $(\updownarrow,1)$  $(^{+}_{1}, 1)$  $(^{1}_{2}, 1)$  $(\mathbf{2},1)$  $(\mathbf{v},1)$  $(\mathbf{v},1)$ (\*,1)

 $(\triangle,1)$  $(\triangle,1)$  $(\triangle,1)$ (0,1)(0,1)(0,1)

( , 3 )