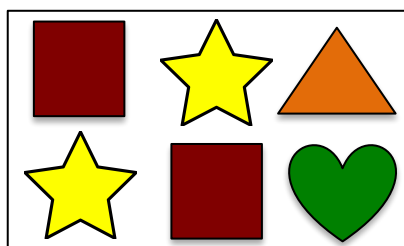
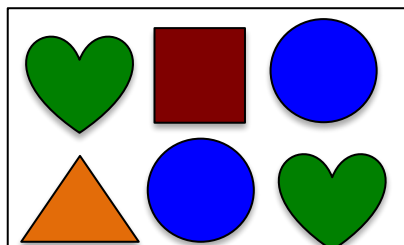
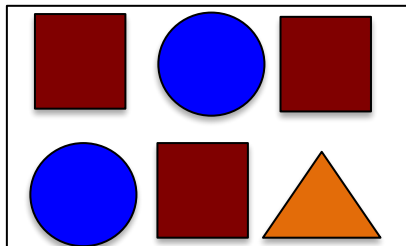
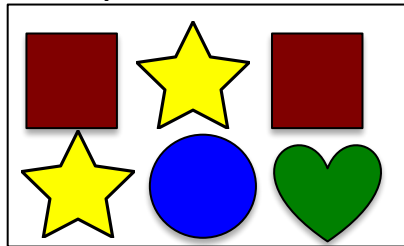














STEP 0 – STORE TO HDFS







Assume 4 data
partitions.









1 - MAP




( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)









( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)








( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)

( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)






2 – SHUFFLE and SORT

( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)

( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)

( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)
( , 1)

3 - REDUCE

( , 8)
( , 4)
( , 5)
( , 4)
( , 3)