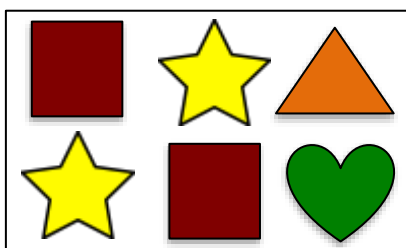
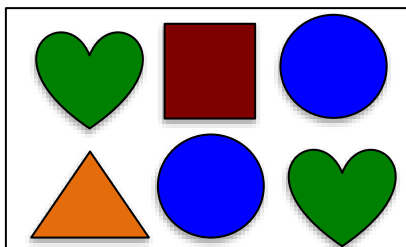
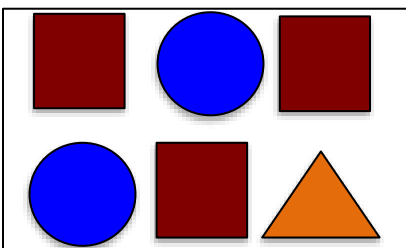
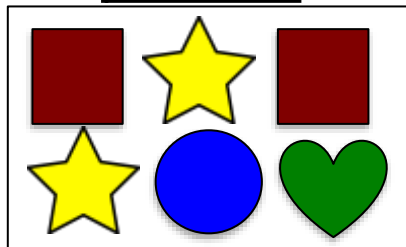


STEP 0 – STORE TO HDFS

Assume 4 data
partitions.



1 - MAP

2 square
2 star
1 circle
1 heart

3 square
2 circle
1 triangle

2 heart
1 square
2 circle
1 triangle

2 square
2 star
1 triangle
1 heart

2 – SHUFFLE and SORT

2 square
3 square
1 square
2 square

2 star
2 star
1 circle
2 circle
2 circle


1 heart
2 heart
1 heart
1 triangle
1 triangle
1 triangle

3 - REDUCE

( , 8)

( , 4)

( , 5)

( , 4)

( , 3)