

TEKLRN

HADOOP LEVEL 10

1. Advanced MapReduce Techniques
 - Simple, Advanced and in-Between
 - Joins
 - Map-side versus Reduce-side Joins
 - Matching Account and Sales information
2. Reduce side joins using Multiple Inputs
 - DataJoinMapper and TaggedMapperOutput
3. Implementing Map Side Joins
 - Using Distributed Cache
 - Pruning data to fit in the cache
 - Using a data representation instead of Raw Data
 - Using multiple mappers