

TEKLRN

HADOOP LEVEL 9

1. Analyzing a large Dataset
 - Getting the UFO Sighting dataset
 - Getting a feel for the Dataset
2. Summarizing the UFO dataset
 - Examining the UFO Shapes
3. Summarizing the shape Data
4. Correlating the sighting duration to UFO Shape
 - Using Streaming Scripts outside Hadoop
5. Perform Shape time analysis from outside the command line
 - Java Shape and location Analysis
6. Using Chain/Mapper for field validation/analysis
 - Too many abbreviations
 - Using the distributed Cache
7. Using the distributed Cache to improve location Output
8. Counter, Status and other output
9. Creating counters, Task states, and writing log output