



# **Final Exam Review**

# Final Exam Details

- o Final exam covers lectures 8, 9, 10 and 11
- o Exam Day – Thursday (Dec 22)
- o Exam time – 2.5 hrs
- o Total points - 60
- o Total questions – 6
- o 15 MCQs
- o Closed book/Closed notes

# Lesson-wise Points Distribution

- o Apache HBase - 16 points
- o Apache Hive - 13 points
- o Apache Spark - 12 points
- o Apache Sqoop - 10 points
- o Apache Flume - 9 points

# Important Topics

## o Apache HBase

- o Architecture
- o Features and Concepts
- o Supported Operations
- o Major Components in details
- o Read/Write Paths
- o Compactions

# Important Topics

## o Apache Hive

- o Internal and External tables
  - o Create table statements
  - o Data loading statements
- o Partitioning
- o Warehouse and MetaStore
- o No complicated SerDe

# Important Topics

## o Apache Sqoop

- o Important commands and their use
- o Important import/export control arguments
- o Behind the scenes – MapReduce

## o Apache Flume

- o Different components
- o Transactions in Flume
- o Flume configuration file

# Important Topics

## ◦ Apache Spark

- Transformations and Actions on RDDs
- RDD Persistence & Lineage
- Spark deploy modes in Yarn
- Spark runtime components



All the Best!

And do TM before coming  
for the Final Exam