**Recommendations/Observations:**

* Use Cube technology (whether it be Kylin, Kyvos, Atscale) to access data in Hadoop as a general practice for ad-hoc reports
  + Runs queries in sub-second speeds
  + Should be implemented ASAP in my opinion
* Cube should be used for generating static reports with fixed or minimal changes in the data-warehouse/starr schema
* Kylin seems to be more compatible with SQL clients (Squirrel SQL) than BI applications like Pentaho or Qlikview
  + Kylin still connects to most BI applications, however, it is hard read all the data into reports/charts because the BI app itself does not have enough memory AKA it is not scalable; this happened with Qlikview
* Don’t use Spark SQL for just one query – it makes more sense to use for multiple queries, possible transformation or for pre-processing data
* Use Hive for a one-time query view – data queried in Hive is not actually stored anywhere
  + Good for quick snapshot of what a query output from Hive tables will look like