Day 1:

1. Scala and SQL coding challenges
2. HDFS Review
   1. review all HDFS commands (i.e [HDFS basic commands](https://hadoop.apache.org/docs/r2.4.1/hadoop-project-dist/hadoop-common/FileSystemShell.html))
   2. local install of HDFS ([Hadoop Environment Setup](https://www.tutorialspoint.com/hadoop/hadoop_enviornment_setup.htm)) those having trouble with HDP Sandbox
3. Hive Review
   1. Review Hive Query Language ([Hive Docs](https://cwiki.apache.org/confluence/display/Hive/))
      1. utilize joins, subqueries, and loading to and from json, txt, csv, and orc files
      2. managed vs. external tables
      3. bucketing and partitioning
   2. Use HiveContext in Spark to write and run queries in Spark with Scala (or Python) already tested on the command line using Hive
4. Associates review all SQL topics in curriculum in small groups
   1. write SQL examples as a group that encompass examples from all of curriculum covered
   2. do not forget to implement atleast one stored procedure or trigger

Day 2:

1. Scala and SQL coding challenges
2. Spark Review ([Spark-examples](https://sparkbyexamples.com/))
   1. RDDs
      1. loading and saving
      2. transformations (i.e. map, filter, groupBy) and actions (i.e reduce, collect)
   2. Dataframes and Datasets
      1. loading from json, csv, txt, parquet, orc and saving as json, txt, csv, parquet, orc
      2. load these files to and from HDFS
      3. also implement bucketting and partitioning ([good overview](https://selectfrom.dev/apache-spark-partitioning-bucketing-3fd350816911))
   3. For all examples above implement wordcounts with a few variations and create and run some queries
   4. spark-submit
      1. local
      2. and on AWS EMR
3. Associates review all Scala topics in curriculum in small groups
   1. write Scala examples as a group that encompass examples from all of curriculum covered
   2. make sure to use as many different collections as possible

Day 3:

1. Python and SQL coding challenge
2. Kafka Review ([Kafka Docs](https://kafka.apache.org/documentation/))
   1. Create Producer and Consumer
   2. Publish to topics and subscribe from topics
3. Final Project Review
   1. implement final project for each individual and in small groups
      1. in small groups ensure each person can do a full run and walkthrough of the basic processes for the project demo
4. Associates review all Python topics in small groups
   1. write Python examples as a group that encompass examples from all of curriculum covered

Day 4:

1. Python and SQL coding challenge
2. Practice interviews
   1. give some interview tips
      1. be confident
      2. avoid “i don’t know” if blank on topic could describe something similar
      3. try to lead a conversation as if explaining to a colleague
      4. do not forget to study what you know, including basic code reviewed
   2. in small groups, review possible list of interview questions
   3. in small groups, interview each member of the group for 20-30 minutes
   4. rotate and continue interviewing until each associate has 3-4 interviews