**CCA Certification Prep**

<https://www.cloudera.com/more/training/certification/cca-data-analyst.html>

|  |  |  |
| --- | --- | --- |
| Week | Topics | Description |
| Apr25 – May1 | Unix, HDFS and Sqoop | Hands-on Exercises |
| May1 – May 8 | Hive/Impala Basics | Create/Alter, Managed/External Tables |
| May8 – May 15 | Hive/Impala Advanced | Complex Queries,  Aggregations, Partitioned Tables. |

**Week 1 :**

* **Learn UNIX Basics**
  + Follow YoBi Onboard Program for UNIX
  + Practice the exercises attached.
* **HDFS Commands**
  + Practice all basic HDFS operations through command line
* **Sqoop –**
  + Sqoop Import
    - Import table as TEXT file with '\001' delimiter
    - Import table as AVRO/Parquet in HDFS
    - Import table as Hive table with ‘^’ delimiter
    - Import table as Hive Parquet table
  + Sqoop Export
    - Export a TEXT file from HDFS to MySQL
    - Export an AVRO file from HDFS to MySQL
    - Export a Hive table to MySQL
* **Bash Scripts**
  + Create a Bash script which does the following
    - Accepts MySQL database name as first argument
    - Comma separated MySQL table names or “ALL” as second argument
    - If second argument is “ALL”, then
      * Use sqoop-import-all-tables tool to import whole database to Hive
    - If second argument is “tbl1,tbl2,tbl3”, then
      * Import the specified 3 tables to Hive
      * Use bash loops to import tables one by one
  + Schedule this job from your CRONTAB scheduler.