**INFS 774 Big Data Analytics**

**Guidelines for Week 5**

Relevant Learning Objectives:

1. Gain an understanding of Apache Hive
2. Gain an understanding of Apache Pig (optional)
3. Try to learn how to import database tables into HDFS using Sqoop

Required activities:

1. Watch the LinkedIn Learning course “Analyzing Big Data with Hive” by Ben Sullins.

To go to the course, you need to go through the following steps:

1. Go to https://library.dsu.edu/, scroll down, in the box “Featured Databases” find LinkedIn Learning (Lynda.com) and click it.
2. Click “sign in” in the upper right corner and sign in with your university email address. Then you will see the message “Your organization requires you to sign in with Single Sign-On”. For DSU student, please use your DSU network credential to log in. For SDSU Students: Please use your DSU VM account credential that has been assigned to you (please see the document “Access Hadoop VM”) to log in.
3. In the search box, type “apache hive” and you can find the course.
4. Read the Hive chapter in the book: Hadoop: the definitive guide
5. Finish Quiz -Hive (posted on 6/17, due 6/24, with no grace period). Almost all quiz questions are from the LinkedIn Learning course.

Optional activities:

1. Read Cloudera’s slides on Apache Pig
2. Read the chapter on Pig in the book: Hadoop: the definitive guide

Top of Form