**Week 9 Study Guide**

**Course Learning Objectives**

1. Explain the contemporary big data computing paradigm.

2. Compare NoSQL database systems with traditional database systems.

3. Perform system administration on contemporary big data platforms

**Module Learning Objectives**

1. Perform the regression task in data analysis using Spark
2. Monitor and track Spark tasks

Module Learning Objective 1 contribute to Course Learning Objective 1

Module Learning Objective 2 contribute to Course Learning Objective 3

In this module we will continue to learn how to use Spark to perform the regression task. We first start with discussions on important concepts in regression, then we move on to hands-on with PySpark on two models, namely Linear Regression and Gradient Boosting Tree.

Please go through all the provided materials and lectures.