



BIG DATA  
DEVELOPMENT

**ACADGILD**

# Session 14: Hive Introduction

---

## Assignment 1

## *Session 14: Hive Introduction*

### *Assignment –Hive*

#### **Table of Contents**

1. Introduction .....	3
2. Objective .....	3
3. Prerequisites .....	3
4. Associated Data Files .....	3
5. Problem Statement.....	3
7. Approximate Time to Complete Task .....	4

## Big Data and Hadoop Development

### **1. Introduction**

In this assignment you need to write answers for questions on Hive.

### **2. Objective**

This assignment will help you to understand Hive concepts in depth before diving deep into Hive programming.

### **3. Prerequisites**

Acadgild's VM, or Linux operating system with Hadoop installed in it.

### **4. Associated Data Files**

N/A

### **5. Problem Statement**

**Give brief answers to the below questions:**

- 1. When Hive is best suited and when is it not?**
- 2. When should one use Hive over MapReduce?**
- 3. What is Hive metastore?**
- 4. How can Hive improve performance with orc file format tables?**
- 5. What is thrift server and client, jdbc and odbc driver importance in hive?**
- 6. What is the importance of partition in hive?**
- 7. What is the use of bucketing in hive?**
- 8. What is the difference between static partitioning and dynamic partitioning in hive?**

**6. Expected output**

N/A

**7. Approximate Time to Complete Task**

120 min