

Data Querying Digital + Syllabus

Faculty

Mentor	Dr. Ernesto Lee	
Moderator	Heera Lal	helal@visa.com
Learner's Experience Manager	Sofia Vergara	sofia.vergara@visa.com
Hands on Lab Support	Nuvepro	support@nuvepro.freshdesk.com

General Information

Description

This is the first of two advanced courses in Data Querying with Hive. Students will learn how to go deeper than simple queries as we focus on Hive queries. Learners will have the opportunity to improve queries formations to reduce stability and performance issues.

Prerequisites:

- Experience with SQL, including grouping, simple and correlated subqueries, functions, and triggers
- Basic understanding of Hadoop HDFS and YARN

Upon completion of the Data Querying with Hive 101 Digital + course, learners will be able to:

- Feel comfortable working with Data Warehousing
- Use the Hive Query Language
- Use different Storage Patterns
- Be familiar with Partitioning and Bucketing
- Use the Metastore and HCatalog
- Compare the use of Hive and Spark

Expectations and Goals

This is a learner-centric instead of instructor-centric course, with heavy student's responsibility in the learning process.

There are six modules in this course, each module includes video lectures, hands on labs, community channel discussions and, a live mentor session with Hive experts to discuss Visa's use cases related to the module subjects and to answer questions.

You should complete the module's self-paced assignments before attending the 1-hour mentor live session.

Self-paced resources

Below are the links to access the self-paced resources

- Video Lectures

Visa University Learning Path: <https://degreed.com/pathway/79xx73jw9k?orgsso=visa>

- Hands on Labs

- VM Access: <https://cloudlabs.nuvepro.com/company/VISAInc/home>

Login-id: your visa email

Password:

- Community Channel

MS Teams Course Channel: [Data Querying D+ Oct 11th_2022 Cohort](#)

Data Querying Digital + Group Oct 11th_2022 Cohort

5f615bee.visainc.onmicrosoft.com@amer.teams.ms

Additional Materials

Included @ general/additional materials in [Data Querying D+](#)

Frequent Asked Questions FAQ Data Querying with Hive 101 Digital +.docx

Syllabus Data Querying with Hive 101 Digital +

Course Schedule/Assignments

All assignments and labs should be completed before each live session

Due Date	Topic	Live Session	Assignments	Hands on Labs	Reading
Week 1 Monday	Course Launch Module 1	Kick off session (1h 30 m)	Pre-assessment (10m)		Programming Hive Chapter 1 Introduction to Apache Hive
Week 1 Tuesday	Big Data and SQL (1h 40m total)	Mentor session (1h)	Module 1 Video lectures – Hive Architecture (2h)	Configuring Hive and getting access (1m)	
Week 1 Thursday	Module 2 Data warehousing and Big Data (3h 43m total)	Mentor session (1h)	Module 2 Video lectures – Hive Architecture and SQL vs Hive (1h 43m)	Labs Module 2 (1h)	N/A
Week 2 Tuesday	Module 3 Data Storage (4h 19m total)	Mentor session (1h)	Quiz – Week 1 (10m) Module 3 Video lectures – Hive Functions (1h 9m)	Labs Module 3 (2h)	N/A
Week 2 Thursday	Module 4 File Formats Partitioning and Bucketing (3h 31m total)	Mentor session (1h)	Module 4 Video lectures – Views in Hive (31m)	Labs Module 4 (2h)	N/A