## Bytewise Itd Fellow ship

# Data Engineering

### Technologies

- SQL (Structured Query Language)
- Formula 1 Project
- Hacker Rank
- Data Engineering
- Data Lakes
- Delta Lakes
- Data Warehouse
- ETL





### SQL (Structured Query Language)



### **SQL Basics**

In Basics I learn about the Basic syntax of the SQL In which ho to extract the record ,understand the use of Where Clause,



#### **Functions**

I understand about the concepts of Functions how we use the built-in functions and create our own functions



#### Joins

Understand the terms of joins, in join I understand how e use joins and when we use joins,



#### Procedures

Understand the concepts of procedures how we will use procedure and how to create the procedure



### Group By & SUM OVER

Understand the concepts of Functions, Group By "Having Clause and Sum Over Clause

### Data Engineering



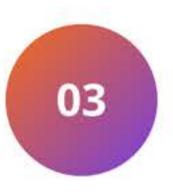
### **Data Engineering Basic**

Understand the Basic Concepts of Data Engineering like they are responsible for managing, monitoring the data and storage and analysis the data



### **Roles of Data Engineer**

Understand the core concept of Generalist, Pipeline-centric and database centric



#### **End to End Architecture**

Understand the whole concept of the architecture

### **Data Lake**

Centralized repository designed to Store, Process Secure large amounts of Structured, Semi structured Unstructured data. It can store data in its native format and process any variety of it, ignoring size limits.



### **Delta Lake**

Delta Lake is an open source storage layer that brings reliability to data lakes. Delta Lake provides ACID transactions, scalable metadata handling, and unifies streaming and batch data processing. Delta Lake runs on top of your existing data lake and is fully compatible with Apache Spark APIs



### Data Warehouse

A data warehouse is a central repository of information that can be analyzed to make more informed decisions. Data flows into a data warehouse from transactional systems, relational databases, and other sources, typically on a regular cadence.



### Extract, Transform & Load



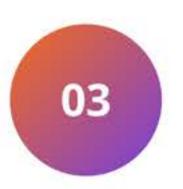
#### Extract

In extraction we extract the data from any source like files csv files json files etc



#### Transform

In Transformation, we transform the extracted data do some Transformation on it



#### Load

Load the Extracted and transformation data into data warehousing



### Session



#### **Databricks Fundamentals**

in this session we understand the basics of Fundamentals of databricks



#### **Azure Fundamental**

Azure fundamentals is a six-part series that teaches you basic cloud concepts, provides a streamlined overview of many Azure services, and guides you with hands-on exercises to deploy your very first services for free.



#### **Azure Data Fundamentals**

Students will learn about core data concepts such as relational, non-relational, big data, and analytics, and build their foundational knowledge of cloud data services within Microsoft Azure



### **Azure Databricks**



### Umar Ali

has successfully completed the requirements to obtain the following Databricks Academy Accreditation:

#### Databricks Accredited Lakehouse Fundamentals

Date of Issue:

September 27, 2022

Date of Expiration:

September 27, 2023



Jody Speiro de Faria

Sr. Director, Curriculum & User Success

https://credentials.databricks.com/fdf7a9f6-6dcb-4e18-9a01-4419924c294a

### HackerRank SQL (Basic)



### HackerRank SQL(Intermediate)





### Formula 1

- Setup a system
- Mounting Data lake Container
- Upload Raw data
- · Ingested the raw data
  - CSV Files
  - Json Files
  - Multi Files
- Raw data -> Processed Data
  - Filters and Join transformation
  - Aggregations
  - Spark SQLs
- Presentation Data
  - Analysis the data
- Incremental Load



