

Data Engineering Curriculum (32 weeks ~ 8 months)

Module	Topics Covered	Learning Objectives	Duration
Introduction to Data Engineering	Overview of Data Engineering, Importance, and Career Paths	Understand the role and significance of a Data Engineer	1 Week
Programming Fundamentals	Python Programming, Data Structures, and Algorithms	Develop programming skills essential for data engineering	2 Weeks
SQL and Relational Databases	SQL Queries, Database Design, Normalization, Indexing, and Joins	Efficiently query and manage relational databases	2 Weeks
NoSQL Databases	Key-Value Stores, Document Stores, Graph Databases, Use Cases	Understand and use NoSQL databases for unstructured data	2 Weeks

Data Warehousing	Data Modeling, ETL Concepts, Star and Snowflake Schemas, Redshift	Design and implement data warehouses	2 Weeks
ETL and Data Pipelines	Extract, Transform, Load Processes, Apache Airflow, Data Integration	Build and automate data pipelines for seamless data flow	3 Weeks
Big Data Ecosystems	Hadoop, HDFS, MapReduce, Apache Spark Basics	Process large datasets using distributed computing systems	3 Weeks
Cloud Data Engineering	AWS, Azure, or GCP for Data Engineering, Cloud Storage Solutions	Deploy data engineering solutions on cloud platforms	3 Weeks
Data Lake and Lakehouse	Concepts of Data Lakes, Delta Lake, and Lakehouse Architecture	Differentiate and apply data storage techniques	2 Weeks

Streaming Data and Real-Time Processing	Apache Kafka, Apache Flink, Streaming vs Batch Processing	Implement streaming solutions for real-time data analysis	3 Weeks
Data Quality and Validation	Data Profiling, Validation Techniques, Data Cleaning Tools	Ensure the quality and integrity of data	2 Weeks
Data Security and Governance	GDPR, Data Encryption, Role-Based Access, Data Masking	Apply best practices for securing and governing data	2 Weeks
Data Engineering on Spark	RDDs, DataFrames, SparkSQL, and Optimizations	Build scalable data processing workflows using Apache Spark	3 Weeks
Machine Learning for Data Engineers	ML Pipelines, Data Preprocessing for ML, Model Deployment	Prepare data for machine learning models	2 Weeks



**R-107, First Floor, Metro Tower, Near
Mangal City, Vijay Nagar, Indore (M.P.)
INDIA Pin code: 452010**

Capstone Project	End-to-End Data Engineering Project	Apply learned skills in a real-world project scenario	4 Week
-------------------------	-------------------------------------	---	--------