# CLOUD AND BIG DATA ENGINEER
## MASTER PROGRAM

Become a Data Engineer with In-demand Bootcamp and Land highpaying jobs at top tech Companies

INDEX


01 Our Mission & Vision
02 About the Program
03 Roadmap
04 Tools Covered
05 Data Engineer Curriculum
06 Projects
07 Placement Support
08 Know your mentor
09 Contact us


## 01 Our Mission & Vision

### Our Mission

Our mission is simple — to help you upgrade your knowledge and skills to become a high-paying tech job. We believe in a world where everyone deserves to have the opportunity to thrive, survive and grow. With our program, you'll build the right skillsets to sustain and take your career to the next level. Whether you want to pivot, take the leap, or level up.

### Our Vision

Our vision is to UPSKILL Professionals and prepare them for real-world industry needs. We want to build a strong, skilled workforce that not only meets the demands of the industry but contributes to the growth of the country. Our goal is to become a leader in this space in a practical, practical training that makes a real difference in people's lives.

## 02 About the Program

The Data Engineer Job Oriented Program is a comprehensive 6-month bootcamp that will give you the skills and confidence you need to land a job as a data engineer.

You'll learn how to build the Microsoft Azure, Google Cloud, and Amazon Web Services data platforms, Data Pipelines and Synapse Analytics – all hands-on, real-world, and industry-ready.

We don't just teach — we guide. You'll get:

1. 1-on-1 mentorship from experienced Data Engineers
2. Support with your resume, LinkedIn profile and interview preparation
3. Experience working on live, cloud-based data projects
4. Skills for roles like Azure Data Engineer, Data Integration Specialist, Cloud Data Engineer Data Architect, and more

This program is your step-by-step path to a high-growth career in data. Want to learn more or get started? Book a free call today!

## 03 Roadmap

## 04 Tools Covered

## 05 Data Engineer Curriculum

**Module 1** – Introduction to Big Data

1.1-What is Data
1.2-What is database
1.3 Types of databases
1.4-What is Big Data
1.5- 5 V's of Big Data
1.6- Types of data (structure, semi-structure, unstructured)
1.7- Store and process big data

**Module 2** – Azure Cloud Fundamentals

2.1- What is cloud computing?
2.2- Categories of cloud services:
2.3- Features of cloud:
2.4- Different cloud platforms, Azure, AWS, GCP
2.5- Steps to create an Azure account
2.6- Azure Subscription
2.7- Accessing the Azure account:
2.8- Create Azure Resource Group Manager
2.9- Overview of Azure Resources / Services
- Data Factory
- Azure Data bricks
- BLOB Storage, Data Lake Storage Gen1 and Gen2
- Azure SQL Server, SQL Database
- Logic Apps
- Synapse Analytics

**Module 3** – Data Storage on Azure Cloud

3.1- Creating a storage account in Azure
3.2- Create Blob Storage and a container, upload files
3.3- Create Data lake Gen2 and upload files
3.4- Difference between blob storage and ADLS GEN2
3.5- Access tiers
3.6- Data Replication Policies
3.7- Set up an Azure storage account with a different tier
3.8- Maintain a version of files in the storage account whenever a file is modified.
3.9- exclusive access to the file for one Application
3.10 Version control of files in Blob and ADLS

**Module 4** – SQL and Azure SQL

4.1-SQL Server Download
4.2- CLUSTER by, GROUP BY, DCL, and the Difference Between DELETE and TRUNCATE
4.3 -Operators (AND, OR, NULL, LIKE), WHERE and HAVING
4.4- Stored Procedures, Functions, Table Dependencies, and Data Folder
4.5-Constraints, UNION, UNION ALL, String Functions, and Date Functions
4.6-Aggregation and GROUP BY, PARTITION, Case and WHEN Statements and of Subqueries
4.7-Joins in SQL
4.8-Date Functions, Window Functions, CASE and IF Statements
4.9-CTE, Ranking Functions, Higher Order by Department, CONCAT, ISNULL and COALESCE
- Correlated Subquery, Views, Temp Tables, Variables, Correlated Subqueries, and Finding a Particular Column
4.10-Order of SQL Execution, LEAD, LAG, PIVOT, and Finding Relationships Between Primary Key and Foreign Key
4.11-SQL Optimization and Normalization

**Module 5** – Python

5.1-What is python
5.2-Why python for Data Engineers
5.3-Advantage of Python
5.4-Installation of Python
5.5-Variables
5.6-Collections (List, Tuple, Set, Dictionary)
5.7- If-Else statements
5.8- Loops , Functions , Operators

**Module 6** – Introduction to Hadoop

6.1- What is Hadoop?
6.2- How Hadoop overcomes big data challenges
6.3-Hadoop Architecture
6.4-Hadoop Daemons
6.5-HDFS
6.6-YARN
6.7-MapReduce
6.8- Hadoop Ecosystem Components

**Module 7** – Introduction to Spark

7.1- What is Spark
7.2-What is pyspark
7.3-Challenges of Hadoop
7.4-Why spark needed
7.5-A Spark Interrupts
7.6-Spark Architecture & Internals
7.7- Spark Data Frame and datasets

**Module 8** – Azure Data Factory (ADF)

8.1- Introduction to Azure and Big Data
8.2-Big Data, Loading Data from Rub to SQL, Redundancy, and Blob Storage
8.3-Difference Between Blob Storage and ADLS Gen2, Loading Excel to Blob, Blob to Snol, Adding Columns, Range, Delimiter Types, and Blob Storage
8.4-Loading Data From Rub to Sand Using Dynamic Parameters, Security Mechanisms, and Changing Data Types Using Set Metadata and Forlook Activities
8.5-Loading Data From Sand to SQL as Single and Multiple Tables Using Authentication Types (SAS and Vaults)
8.6-Key Vaults, Degree of Parallelism, Loading Data Files from the Last 3-3 Days to Azure Blob Storage and Azure SQL Database
8.7- Lookup Stored Procedure, Stoorage Explorer Connection, Triggers, and Delete Activity (Normal and Event Triggers)
8.8- Copying and Filtering Multiple CSV and TXT Files into Single and Multiple Tables in the Sink Table Using Wildcards, Assigned Managed Identity Authentication
8.9 -Exporting and Importing Pipelines, Running Pipelines from One Data Factory to Another Data Factory
8.10 - Loading All Tables and Specific Tables From Azure SQL Database to Azure Blob Storage, Self-Hosted Integration Runtime, and ER
8.11 - Nested Looping (Execute Pipeline) and Merging Multiple Files into a Single Folder
8.12 - Concept of Synapse Analytics
8.13 - Synapse Analytics (Part and Dimension Table), Star and SnowFlake Schema, Fact Table
8.14- Connecting Azure Databricks from Azure Data Factory, running notebooks, and optimization
8.15- Synapse Analytics Transformations (Narrow and Wide Tables), Using Notebooks to Read and Write Data
8.16 - Data Flow - Part 1
8.17 - Data Flow - Part 2
8.18 - Data Flow (SCD Type 0, Type 1, and SCD Type 2)
8.19 - Debugging and Monitoring Data Pipelines, Audit Logs, Pipeline Troubleshooting, Performance Optimization, and Cost Optimization
8.20 - Agile Methodology, Git, CI/CD, and Jira
8.21 - Incremental Load for a Single Table
8.22 - Incremental Load for Multiple Tables

**Module 9** – Databricks New Features

9.1-Introduction to Data Lake Tables
9.2-SQL Warehouse in Databricks
9.3-Embedding Dashboards notebooks
9.4-Git Integration in Databricks
9.5-DevOps for Databricks
9.6-Databricks cost optimization strategies
9.7-End-to-end integration of Databricks with Power BI

**Module 10** – Spark Streaming

9.1-Introduction to Spark Streaming
9.2-Streaming data using Files
9.3- End-to-end practical implementation of Spark Streaming

**Module 11** – Unity Catalog in Databricks

9.1-Unity catalog introduction
9.2-Key features of Unity Catalog
9.3-Unity Catalog object model
9.3-Managed vs. External tables
9.3-End-To-End Implementation of Unity Catalog

**PROJECT**

- Data Lakehouse & Medallion Architecture Introduction
- Data Ingestion using Azure Data Factory (ADF) from Web URL
- Connect ADLS with Azure Databricks
- Advanced Transformations with Databricks
- Connect Power Bi Data to Silver Layer
- Aggregations and Business Logic for Gold Layer
- Data Modeling using Synapse Serverless SQL Pool
- Build Dashboards with Power BI

## What kind of mentorship you get?

End-to-End Mentorship | I guide you through every stage of the project—from data ingestion to transformation, modeling, and deployment—ensuring you gain real-world hands-on experience.

Real World Integrations | Tackle other programs with multiple tutors for different topics, I teach every tool and technique myself. Why? Because in real jobs, a Data Engineer handles the entire stack, and want you to be job-ready with the same mindset.

Career-Focused Learning | My goal is simple—transform your career with relevant, highpaying education and give you the confidence to enter the data engineering world. Let's build your future

Website - Wisdomsail.com

## CONTACT US


Email ID : hello@wisdomsail.com
Mobile : +91 81248 38828
Website : wisdomsail.com


Wisdom Sail Data Science Institute

## THANKS!

Thank You for Your Time! We Appreciate Your Attention!

Wisdom Sail Data Science Institute