# Table of
# Contents

# About Us

www.palin.co.in

Palin Analytics is a leading edtech and consulting firm specializing in Azure data engineering and cloud-based data solutions. We train individuals and professionals to master Microsoft Azure tools and build scalable data infrastructure for the modern world.

Our Azure Data Engineering Certification Program focuses on Azure Data Factory, Azure Databricks, Azure Synapse Analytics, and Azure SQL Database, meeting industry standards. With hands-on, project-based learning, you'll build real-world Azure data pipelines and systems.

Our expert trainers, with experience in finance, healthcare, and retail, guide you through live projects and case studies to ensure job-ready skills.

Flexible Learning Modes

Industry Experienced Trainers

Certifications & Recognition

Hands-on Learning Approach

Personalized Mentorship

Comprehensive Curriculum

Live Projects & Case Studies

Placement Assistance

**LEARNING OBJECTIVE**

### Build Azure Data Pipelines

Design ETL/ELT workflows with Azure Data Factory and Synapse Pipelines.

### Master Azure SQL

Manage and query data with Azure SQL Database and Cosmos DB.

### Leverage Azure Databricks

Process big data using Spark in Azure Databricks.

### Integrate Azure Storage

Use Azure Blob Storage and Data Lake for scalable data management.

### Process Big Data on Azure

Handle large datasets with Azure Databricks and HDInsight.

### Ensure Data Quality

Implement validation and governance in Azure Synapse.

### Automate Data Workflows

Orchestrate pipelines with Azure Data Factory for efficiency.

Designed for beginners and pros aiming to excel in Azure data engineering.

# PREREQUISITES

### Basic Computer Skills

**Familiarity with files and software.**

### Database Basics

**Knowledge of tables and queries.**

### Logical Thinking

**Ability to design workflows.**

### Basic Programming

**Python or SQL knowledge helps.**

### No Azure Experience Needed

**We teach Azure from scratch.**

### Interest in Cloud Data

**Passion for scalable Azure solutions.**

# ELIGIBILITY

## Students & Fresh Graduates

B.Tech, B.Sc, B.Com grads entering Azure data engineering.

## Working Professionals

IT, database, or cloud pros upskilling in Azure.

## Database Administrators

Pros scaling skills with Azure tools.

## Tech Enthusiasts & Entrepreneurs

Those leveraging Azure for business data solutions.

# KEY FEATURES OF AZURE DATA ENGINEERING

## Comprehensive Curriculum

Comprehensive curriculum designed to equip you with in-demand skills and real-world expertise.

## Industry-Expert Trainers

Learn from industry expert trainers with hands-on experience and deep domain knowledge.

## Live Projects & Capstone Assignments

Work on industry-relevant projects to build a strong portfolio.

## Interactive Sessions

Engage in interactive sessions for practical learning and real-time problem-solving. 40 instructor-led sessions (2 hours each) with live demonstrations and Q&A.

## Flexible Learning Modes

Online, offline, and hybrid training options available.

# KEY FEATURES OF AZURE DATA ENGINEERING

## Hands-on Learning Approach

Practical, project-based training with real-world case studies and projects.

## Placement Assistance

Resume building, interview preparation, and job referrals through our hiring network.

## Certification

Earn an industry-recognized certificate upon course completion.

## Personalized Mentorship

One-on-one doubt-clearing sessions and career guidance.

## Affordable & Value-Packed

High-quality training at competitive pricing.

# JOB ROLES WE CAN TARGET IN DATA ENGINEERING

Azure Data Engineer

ETL Developer

Azure Databricks Engineer

Data Pipeline Developer

Azure Database Administrator

Cloud Data Engineer

Data Operations Engineer

Azure Data Architect

Streaming Data Engineer

Data Integration Specialist

Azure Synapse Specialist

# Our Trainers

## ANURAG PANDEY

**Data Science | Business Intelligence | Machine Learning**

Senior Leader having close to half decade of expertise in leading large, globally distributed, cross-functional teams to strategize and deliver, Product Engineering led large scale SaaS and Digital Transformation programs, leveraging new-age technologies in Big Data , AWS, Azure, Power BI, MSBI ,complete SQL and TQL development.

### PROFESSIONAL EXPERIENCE

- I have worked on a complete setup of Data Warehouse and Data Lake based on AWS, Azure Cloud and On-Premises solutions.
- I have total of nearly 8 + years of IT experience and which was completely dedicated to Big Data, Database development and MSBI and have extensive experience in AWS, Glue, ETL, EMR, EC2, Hive, Hadoop, Spark, ADF, Data Lake Gen2, Databricks, MSBI Technologies like SQL Server, SSIS, SSAS,and POWER BI.

## SAHIL ARORA

**Generative AI | Python | PySpark | RStudio | Azure ML | SAS | SQL**

- I am working as a Senior Data Engineer in Tiger Analytics with 7+ years of experience.
- I have experience in functions including data extraction, transformation, loading and integration in support of enterprise data warehouses using Azure services in Retail and Healthcare domains.

### PROFESSIONAL EXPERIENCE

- ETL: Azure Data Factory, Delta Live Tables, DBT for Transformation, Metallion, Talend DI, DataStage
- Scripting Language: Python, SQL, Pyspark
- Big Data Solution: Azure Databricks, Azure Synapse Analytics Studio, MS Fabric
- Storage: Azure Data Lake Storage, Delta Lake with Unity Catalog
- Data Warehouse: Azure Dedicated SQL pool, Redshift, Teradata
- Database: Azure SQL DB, PostgreSQL
- AI: Harnessing Generative AI on Databricks
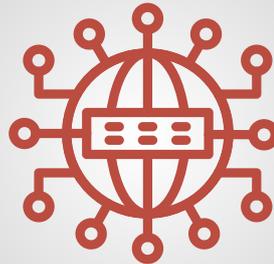- Version Control: Azure Devops

Add On: Elastic Search,Logic Apps, Azure Functions, Fast API, LUIS, Docker,Timbr.ai

# Program Innovations :

Our Azure Live Project Lab provides hands-on experience with Azure tools, solving real-world data challenges under expert guidance.

**1** Domains



**2** Placement Support

# 1 Domains

Specialize in Azure data engineering for specific industries:

### - Finance & Banking

Transaction pipelines, fraud detection flows.

### - Healthcare & Pharma

Patient data integration, trial processing.

### - E-Commerce & Retail

Inventory sync, customer data pipelines.

### - Supply Chain & Logistics

Demand forecasting, route optimization.

### - Telecom & Energy

Usage tracking, smart grid pipelines.

### HR & Workforce Analytics

Employee data pipelines, payroll integration.

### Telecom & Network Analytics

Network performance streams, churn pipelines.

### Energy & Utilities Analytics

Power usage flows, equipment data systems.

Each domain includes Azure projects and datasets.

# 2 Placement Support

Palin Analytics, we are committed to ensuring that every learner successfully transitions into a rewarding career in data engineering. Our placement support provides continuous assistance for up to one year, offering multiple interview opportunities until you secure a job.

## Mock Interviews

Palin Analytics, we prepare our learners for real-world job interviews by conducting Mock Interviews with industry professionals. These sessions help candidates gain confidence, improve technical skills, and refine communication to ace their job interviews.

## ATS Friendly Resume

Palin Analytics, we ensure that every learner has an ATS-friendly resume to increase their chances of getting shortlisted for job interviews. Applicant Tracking Systems (ATS) are used by companies to filter resumes before they reach recruiters. Our expert-designed resumes help you bypass these filters and land more interview calls!

## Interview Calls

Palin Analytics, we ensure that our learners get maximum job opportunities through regular interview calls from top companies across multiple industries. Our dedicated Placement Assistance Program helps candidates secure jobs in data engineer and related roles.

## Phase 1: Foundation & Azure Setup

**Class 1 – Intro to Data Engineering + Azure Essentials**
- Data Engineering lifecycle
- Azure overview: ADLS, Databricks, ADF
- Data Lake vs Data Warehouse vs Lakehouse
- Practical: Azure Free Tier setup + workspace overview

**Class 2 – ADLS Gen2 + Mount in Databricks**
- ADLS containers, permissions (RBAC, SPN)
- Mount ADLS in Databricks (secure way)
- Practical: Upload sample data, mount ADLS, list files

## Phase 2: Spark & Delta Lake Core

**Data modeling + good-level Spark coding covered here**

**Class 3 – DataFrames: Load, Transform, Save**
- Read CSV/JSON/Parquet, schema inference
- Filtering, withColumn, casts, writing back
- Practical: Load raw data, clean and write to ADLS

**Class 4 – Data Warehousing Concepts (DW Intro + Modeling)**
- OLTP vs OLAP
- Star vs Snowflake Schema
- Fact and Dimension tables
- Dimensional Modeling + Date Dimension + SCD Types
- Practical: Design schema for retail dataset (FactSales, DimCustomer, etc.)

### Class 5 – Joins, Aggregations & Window Functions
- Inner, Left, Semi joins
- Aggregations, rollups
- Window functions (row_number, lag/lead)
- Practical: Customer-level aggregations, top-N products

### Class 6 – Delta Lake Deep Dive
- Delta format, ACID, Versioning, Time Travel
- Upserts with MERGE INTO (SCD Type 1 logic)
- Practical: Convert parquet to Delta, perform upserts

### Class 7 – Slowly Changing Dimensions (SCD Type 2)
- Concept of SCD2 (DW concept)
- Implement SCD2 with Spark using Merge
- Practical: Track changes to DimCustomer or DimProduct over time

### Class 8 – Performance Optimization
- Partitioning, bucketing
- Broadcast joins, caching
- Spark UI, job optimization tips
- Practical: Optimize a slow query and compare timings

### Class 9 – SQL Views & Table Management
- Spark SQL, global/temp views, managed vs external tables
- Lakehouse Architecture: Bronze, Silver, Gold
- Practical: Create views on curated Delta tables + query

### Class 10 – Structured Streaming (Optional Real-Time Bonus)
- Micro-batch vs continuous
- Trigger options, watermarking
- Practical: Simulate stream → Delta (or skip if not needed)

## ◆ Phase 3: Orchestration + Integration

### Class 11 – ADF Basics + Ingestion
- ADF Pipelines, Linked Services, Datasets
- Copy raw data to ADLS
- Practical: Use ADF to land raw data into Bronze zone

### Class 12 – ADF Triggering Databricks
- Use Databricks notebook in ADF
- Parameterize + monitor pipeline
- Practical: Full ADF pipeline (Ingest → Process via notebook)

**Class 13 – Job Scheduling, Logging, Alerts**
- Databricks Job UI, task dependencies
- Email alerts, logging with dbutils.notebook.exit()
- Practical: Create production-ready job with alerting

## Phase 4: End-to-End DW Project

Retail Sales Lakehouse Project (Fact + Dim, ADF, Spark, Power BI)

**Class 14 – Ingestion + Bronze Layer**
- Load raw CSVs via ADF → ADLS
- Read in Spark → Raw Delta (Bronze)
- Practical: ADF pipeline + Raw zone setup

**Class 15 – Transform to Curated Tables (DW Layer)**
- Build FactSales, DimCustomer, DimProduct (Star Schema)
- Handle SCD logic (SCD1 or SCD2)
- Practical: Write Silver/Gold layer Delta tables

**Class 16 – Reporting + Project Review**
- Power BI → connect to Delta tables via Direct Lake or Synapse
- Present KPIs: Sales by Product, Top Customers
- Recap + Interview prep + Best practices
- Practical: Build report dashboard + review entire flow

## Optional Add-ons

- GitHub CI/CD integration (Databricks Repos)
- Interview-focused class
- Real-time + alerting extension (Kafka or Event Hub)