

**Driving Excellence:
Your BI and AI Journey Starts Here!**



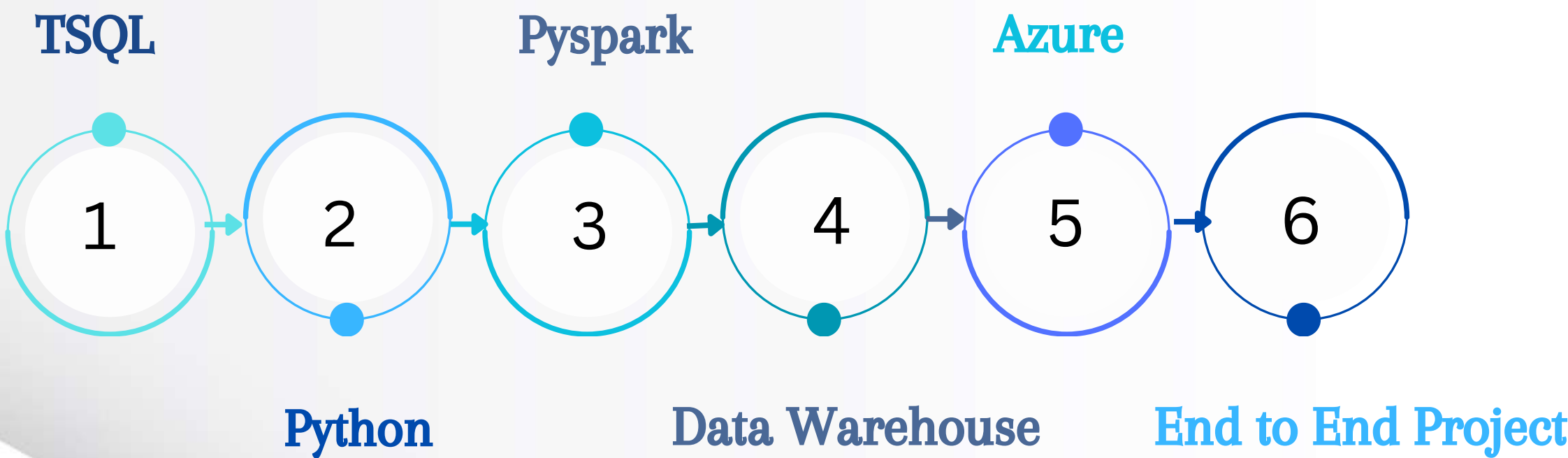
Azure Data Engineer

Course Duration
5.5 Months

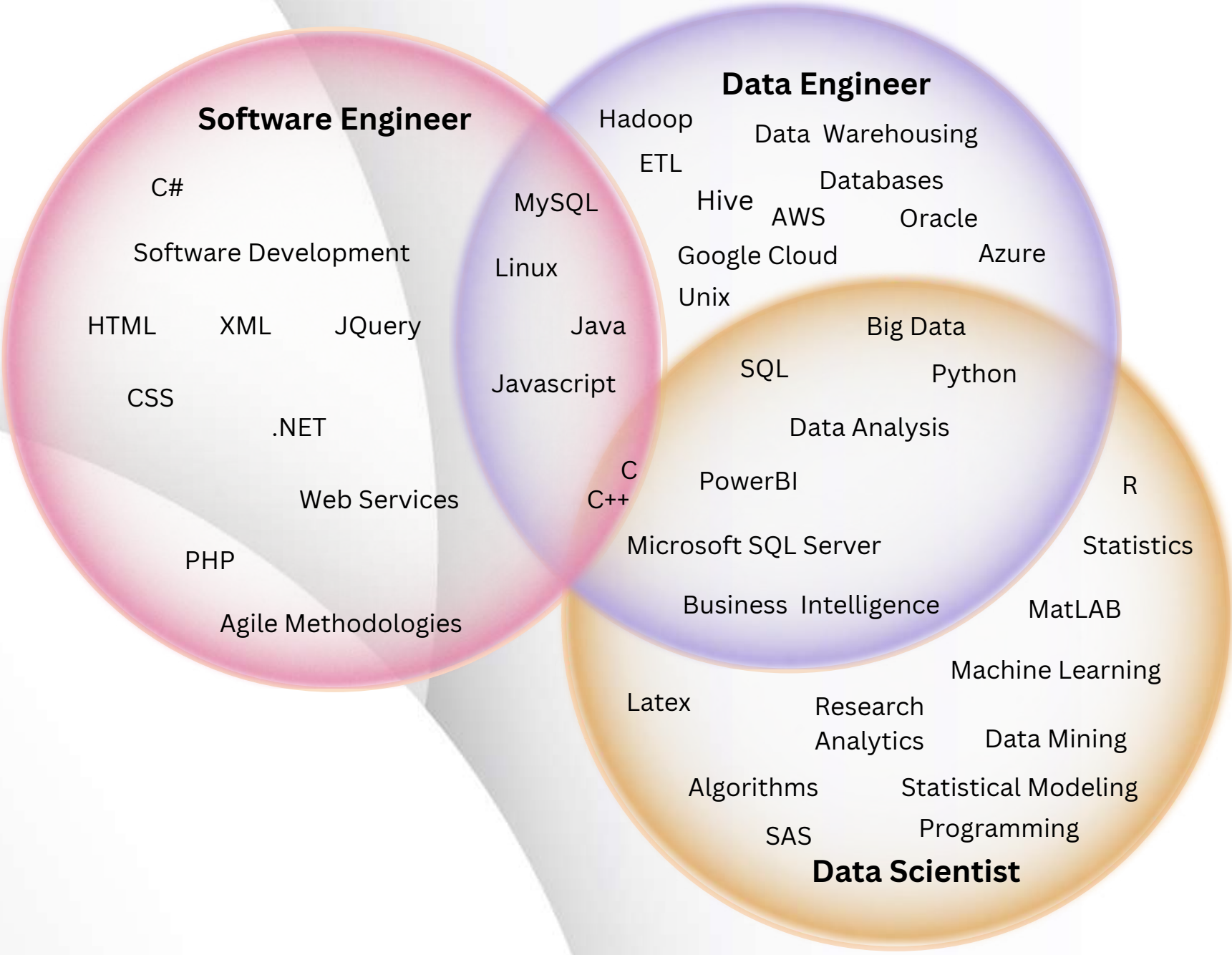
Batches
Weekend(Sat/Sun)

Mode
Online(Zoom)

Technologies



Skills And Qualifications



Introduction

SQL Type

- DDL
- DML
- TCL
- DCL
- DQL

Constraints

- Primary Key
- Unique
- Foreign Key
- Check
- Default
- NOT Null

Clause

- DISTINCT
- WHERE
- ORDER BY
- GROUP BY
- HAVING

Operators

- IS NULL
- LIKE
- BETWEEN
- IN

Functions

- String
- Date
- Numeric
- Conversion
- Aggregate

Joins

- Inner Join
- Left Outer Join
- Right Outer Join
- Full Outer Join
- Cross Join
- Self Join

Subqueries

- Single-row subqueries
- Multiple-row subqueries
- Correlated subqueries

Set Operator

- Union All
- Union
- Intersect
- Minus

Views

- Simple View
- Complex View
- Materialized views

Indexes

- B-Tree Indexes
- Bitmap Indexes
- Function Based Indexes

Analytical Functions

- RANK
- DENSE_RANK
- ROW_NUMBER
- FIRST_VALUE,
- LAST_VALUE
- LAG
- LEAD

Advanced SQL

- PIVOT
- UNPIVOT
- PARTITIONS

PLSQL

- Stored Procedure
- Stored Function

Python

- Setup Python
- "Hello world"
- Datatypes-Numbers, Strings, Boolean
- String Functions: Indexing, Slicing
- Print Formating with Strings
- Variables
- Operators
- List
- Dictionary
- Tuples
- Sets in Python
- If, Elif and Else
- For Loops
- While Loops (including break and continue, Pass)
- Functions and Recursions
- Local and Global variable
- Error and Exception Handling

Pandas And Numpy

- Create Pandas DataFrame
- Numpy Array Operations
- Pandas DataFrame Operations

Spark

- Why Spark?
- Spark Unified Stack
- Spark Architecture
- Spark APIs
- SparkSession
- DataFrames
- Transformations, Actions & Lazy Evaluation
- Spark UI

Spark SQL

- Spark Structured APIs
- DataFrames & Datasets
- Spark SQL Types
- Spark SQL Code Execution
- DataFrame Operations
- DataFrame Aggregation Functions
- Grouping, Windows & Joins
- Spark SQL Data Sources

Spark Streaming

- Introduction to Spark Streaming
- PySpark Streaming with Apache Kafka
- Real-world Practical use cases
- Operations On Streaming Dataframes and Datasets
- Window Operations

Datawarehouse

- Why Data warehousing?
- What is Data warehouse?
- DW Architecture
- OLTP vs OLAP
- ETL Basics
- Dimensional Modeling (Fact, Dimensions)
- Star Schema
- Snowflake Schema
- Slowly Changing Dimension

Cloud

- What is cloud?
- Why the cloud?
- Service Models
- Deployment Model
- Major Cloud Providers

Azure

- Setup Free Trial
- Azure Resource Management
- Subscription, Resource Group
- Regions and Availability Zones
- Azure Key Vault

Azure Storage

- Region Pairing
- Azure Blob Storage
- Azure Data Lake Gen 2 Storage
- Replication Options
- Access Tiers
- Azure Storage Explorer

Azure SQL

- What is Azure SQL?
- Azure SQL PaaS Options
- Single vs Elastic vs Managed Instance
- Delta Lake Tables, Instances
- DTU vs Vcore Purchase Model
- Setup Azure SQL +
Create Sample Tables

Azure Data Factory

- Azure Data Studio
- Azure Data Factory Overview
- Demo - Provision ADF
- Linked Services
- Integration Runtimes
- Datasets -Dynamic
- Pipelines and Activities
- Copy Data
- For Each
- Until
- Get Metadata
- Lookup
- If Activity
- Wait
- Fail
- Execute Pipeline
- Global Parameters
- Data Flows
- Schedule Trigger
- Event Trigger
- Monitor Pipelines

Azure Synapse

- Azure Synapse Analytics
- Overview Serverless, Dedicated SQL
And Spark Pool
- Access DataLake Storage Account
- Azure Synapse Notebooks, SQL
Scripts, Pipelines

Spark using Synapse/Databricks

- Spark Overview
- Spark Dataframes
- Reading CSV files from Data Lake
- Transform Data using (Spark SQL)
- Aggregate Dataframes
- Dataframe Joins and other transformations
- Spark UDF Functions

Databricks

- Spark Dataframe Write in Parquet Format
- Calling Spark Notebooks from ADF
- Databricks Workspace, Clusters, Notebooks
- Databricks Jobs Run
- Databricks Compute, SQL Scripts
- Access Data using DataLakeStorage, Storage Account Key
- Create DeltaLake Tables, Instances



Our Alumni Work in Top Companies

