

# APPLIED DATA SCIENCE & AI

One-Year Comprehensive Program | 8 Modules | Hands-On Portfolio Projects

52 Weeks

104 Hours

SUN 5-7 PM

8 Projects

LEAD TRAINER



FOUNDER - TECHMA ZONE

## MUHAMMAD DANIAL GAUHAR

Data Scientist | AI Engineer | Data Engineer | BI Expert

Transforming businesses through data-driven solutions. Founder of Techma Zone, dedicated to empowering the next generation of data professionals in Pakistan.

7+

YEARS EXP

1000+

STUDENTS

100+

PROJECTS

[linkedin.com/in/danialgauhar](https://www.linkedin.com/in/danialgauhar)

## ABOUT THE TRAINER

## PASSIONATE ABOUT DATA

Muhammad Danial Gauhar is a Data Scientist and AI Engineer based in Karachi, Pakistan. With over 7 years of experience, he has helped businesses transform operations through intelligent data solutions and AI-powered systems.

Currently working as a Data Engineer/Data Scientist at Traxccel, he specializes in building scalable data pipelines, implementing ML models, and developing AI solutions. He is a graduate of the University of Hertfordshire.

As the founder of Techma Zone, he is committed to bridging the gap between academic knowledge and industry demands.

University of Hertfordshire

Traxccel - Data Engineer/Data Scientist

Founder - Techma Zone

Karachi, Pakistan

## EXPERTISE

## SKILLS &amp; TECHNOLOGIES

## PYTHON

Programming, scripting &amp; automation

## MACHINE LEARNING

Scikit-learn, modeling, algorithms

## DEEP LEARNING

TensorFlow, Keras, neural networks

## NLP &amp; GEN AI

LLMs, LangChain, RAG systems

## DATA ANALYSIS

Pandas, NumPy, statistics

## VISUALIZATION

Power BI, Tableau, Plotly

## DATA ENGINEERING

ETL, Databricks, pipelines

## CLOUD &amp; MLOPS

AWS, Azure, deployment

## CAREER PATH

## PROFESSIONAL EXPERIENCE

## 2023 - Present DATA ENGINEER / DATA SCIENTIST

## Traxccel

Building scalable ML frameworks, industrial AI pipelines, and operational AI reasoning systems with Databricks.

## 2020 - Present FOUNDER &amp; LEAD INSTRUCTOR

## Techma Zone

Founded Pakistan's premier data science training institute. Trained 1000+ students with 95% placement rate.

## 2019 - 2023 DATA ANALYST / DATA SCIENTIST

## Previous Roles

ML projects, BI solutions, and data-driven strategies. Built expertise in Python, SQL, and visualization.

## CURRICULUM

# APPLIED DATA SCIENCE & AI

A comprehensive one-year journey transforming beginners into industry-ready data professionals. 8 carefully structured modules with hands-on portfolio projects at every stage.

#	MODULE	WEEKS	HOURS	KEY DELIVERABLE
01	Python Programming	1 – 8	16	Exploratory Data Analysis Project
02	SQL & Database Fundamentals	9 – 14	12	Complete Database Design & Query Project
03	ETL / ELT Using Python	15 – 20	12	End-to-End ETL Pipeline with Airflow
04	PySpark & Big Data Processing	21 – 27	14	Big Data Pipeline with PySpark & Delta Lake
05	Statistics & Probability	28 – 33	12	Statistical Analysis Report
06	Machine Learning	34 – 41	16	ML Model Training & Deployment
07	Deep Learning	42 – 47	12	DL Application (CV/NLP/Time Series)
08	Generative AI	48 – 52	10	Gen AI Application (RAG / AI Agent)

## MODULE 1

## M01

# PYTHON PROGRAMMING

Weeks 1 – 8 | 16 Hours | 2 Hours/Week

Master Python from scratch — the #1 language for data science, AI, and automation. Build a rock-solid programming foundation that powers everything you'll do in this program.

### Core Python

Variables, loops, functions, OOP

### Data Libraries

NumPy, Pandas, data wrangling

### Visualization

Matplotlib, Seaborn, Plotly

WEEK	TOPIC	DETAILS & LEARNING OUTCOMES
1	<b>Python Fundamentals</b>	Setup (Anaconda, VS Code, Jupyter), variables, data types, type casting, input/output, operators
2	<b>Control Flow &amp; Functions</b>	Conditional statements, loops, break/continue, functions, *args/**kwargs, lambda, scope
3	<b>Data Structures</b>	Lists, tuples, sets, dictionaries, comprehensions, nested structures, common methods
4	<b>Object-Oriented Programming</b>	Classes, objects, inheritance, polymorphism, encapsulation, abstract classes, magic methods
5	<b>File Handling &amp; Exceptions</b>	Reading/writing files, context managers, try/except/finally, custom exceptions, logging
6	<b>NumPy</b>	Arrays, operations, broadcasting, indexing/slicing, math functions, linear algebra, random
7	<b>Pandas</b>	Series/DataFrames, data I/O, selection (loc/iloc), filtering, groupby, merge/join, pivot tables
8	<b>Visualization &amp; Project</b>	Matplotlib, Seaborn, Plotly. Mini Project: End-to-end EDA on a real dataset



### MODULE PROJECT

Exploratory Data Analysis Project on a Real-World Dataset

## MODULE 2

## M02

# SQL & DATABASES

Weeks 9 – 14 | 12 Hours | 2 Hours/Week

Learn to speak the language of data. SQL is essential for every data role — from analyst to engineer. You'll design databases and write powerful queries that extract real business insights.

### Database Design

ER diagrams, normalization, schemas

### Query Mastery

Joins, subqueries, CTEs, window functions

### Performance

Indexes, execution plans, optimization

WEEK	TOPIC	DETAILS & LEARNING OUTCOMES
9	SQL Basics & DDL	Relational DB concepts, RDBMS overview, CREATE, ALTER, DROP, data types, constraints
10	Data Manipulation (DML)	INSERT, UPDATE, DELETE, SELECT, WHERE, ORDER BY, LIMIT, DISTINCT, CASE
11	Joins & Subqueries	All join types, self-joins, correlated/non-correlated subqueries, EXISTS/IN, CTEs
12	Aggregations & Windows	GROUP BY, HAVING, aggregates, window functions (ROW_NUMBER, RANK, LEAD, LAG)
13	Advanced SQL	Views, stored procedures, indexes, execution plans, normalization, denormalization
14	SQL Project	ER diagrams, schema design, complex analytical queries, performance tuning



### MODULE PROJECT

Complete Database Design & Query Project

## MODULE 3

## M03

# ETL / ELT USING PYTHON

Weeks 15 – 20 | 12 Hours | 2 Hours/Week

Build production-grade data pipelines. Learn to extract data from APIs, databases, and files — transform it into gold — and load it into warehouses. Automate everything with Apache Airflow.

### Data Extraction

APIs, web scraping, databases, files

### Transformation

Cleaning, wrangling, Pandas at scale

### Orchestration

Apache Airflow DAGs, scheduling, monitoring

WEEK	TOPIC	DETAILS & LEARNING OUTCOMES
15	<b>ETL Fundamentals</b>	ETL vs ELT, pipeline architecture, batch vs streaming, data lakes vs warehouses
16	<b>Data Extraction</b>	APIs (requests, REST), web scraping (BeautifulSoup), databases (SQLAlchemy), file formats
17	<b>Data Transformation</b>	Cleaning, type conversions, string manipulation, date/time, regex, Pandas at scale
18	<b>Data Loading &amp; Storage</b>	Bulk inserts, upserts, cloud storage (Azure Blob, S3), Parquet/Delta, partitioning
19	<b>Pipeline Orchestration</b>	Apache Airflow (DAGs, tasks, operators), scheduling, error handling, monitoring
20	<b>ETL Project</b>	Complete ETL pipeline: extract, transform, load, schedule with Airflow



### MODULE PROJECT

End-to-End ETL Pipeline with Apache Airflow

## MODULE 4

## M04

# PYSPARK & BIG DATA

Weeks 21 – 27 | 14 Hours | 2 Hours/Week

Process massive datasets that don't fit in memory. Master Apache Spark, Databricks, and Delta Lake — the industry-standard tools used by Netflix, Uber, and Airbnb for big data analytics.

<b>Spark Core</b> RDDs, DataFrames, Spark SQL	<b>Delta Lake</b> ACID transactions, time travel, MERGE	<b>Optimization</b> Partitioning, caching, Spark UI tuning
--	--	---

WEEK	TOPIC	DETAILS & LEARNING OUTCOMES
21	<b>Spark Fundamentals</b>	Big data concepts, Hadoop, Spark architecture, RDDs, cluster modes, PySpark setup
22	<b>DataFrames &amp; SQL</b>	Creating DataFrames, schema definition, data I/O, operations, Spark SQL
23	<b>Transformations &amp; Actions</b>	Narrow vs wide, lazy evaluation, UDFs, column expressions, handling nulls
24	<b>Joins &amp; Aggregations</b>	Join types, broadcast joins, aggregations, window functions, pivot/unpivot, data skew
25	<b>Optimization</b>	Catalyst optimizer, partitioning/bucketing, caching, broadcast vars, Spark UI
26	<b>Databricks &amp; Delta Lake</b>	Workspace, Delta Lake (ACID, time travel, MERGE), Unity Catalog, medallion architecture
27	<b>PySpark Project</b>	End-to-end big data pipeline: ingest, transform, SCD logic, Delta tables, optimize


**MODULE PROJECT**

Big Data Pipeline with PySpark & Delta Lake

## MODULE 5

## M05

# STATISTICS & PROBABILITY

Weeks 28 – 33 | 12 Hours | 2 Hours/Week

Data science without statistics is just guessing. Build the mathematical intuition to understand distributions, test hypotheses, and make data-driven decisions with confidence.

### Descriptive Stats

Central tendency, spread, distributions

### Hypothesis Testing

t-tests, chi-square, ANOVA, p-values

### Regression

Correlation, linear regression, R-squared

WEEK	TOPIC	DETAILS & LEARNING OUTCOMES
28	<b>Descriptive Statistics</b>	Central tendency, spread measures, skewness, kurtosis, distributions, five-number summary
29	<b>Probability Theory</b>	Sample spaces, probability rules, conditional probability, Bayes' theorem, combinations
30	<b>Distributions</b>	Binomial, Poisson, Normal, Uniform, Central Limit Theorem, Z-scores, Q-Q plots
31	<b>Hypothesis Testing</b>	Null/alternative hypotheses, Type I/II errors, p-values, t-tests, chi-square, ANOVA
32	<b>Correlation &amp; Regression</b>	Pearson/Spearman correlation, linear/multiple regression, R-squared, residuals
33	<b>Statistics Project</b>	Real-world dataset: EDA, hypothesis testing, correlation, regression, visualizations



### MODULE PROJECT

Statistical Analysis Report with Visualizations

## MODULE 6

## M06

# MACHINE LEARNING

Weeks 34 – 41 | 16 Hours | 2 Hours/Week

The heart of data science. Train models that predict, classify, and cluster. From linear regression to XGBoost — master the algorithms that power recommendation engines, fraud detection, and more.

### Supervised ML

Regression, classification, ensembles

### Unsupervised ML

Clustering, PCA, dimensionality reduction

### Deployment

Model evaluation, Flask API, production ML

WEEK	TOPIC	DETAILS & LEARNING OUTCOMES
34	<b>ML Fundamentals</b>	Supervised vs unsupervised, workflow, feature engineering, encoding, scaling, cross-validation
35	<b>Linear Models</b>	Linear/polynomial regression, regularization (Ridge, Lasso, ElasticNet), logistic regression
36	<b>Tree-Based Models</b>	Decision trees, entropy, Gini, pruning, Random Forest, feature importance, bagging
37	<b>Advanced Ensembles</b>	Gradient Boosting, XGBoost, LightGBM, CatBoost, hyperparameter tuning, bias-variance
38	<b>SVM &amp; KNN</b>	SVM (linear/kernel), margin maximization, KNN, distance metrics, curse of dimensionality
39	<b>Unsupervised Learning</b>	K-Means, hierarchical clustering, DBSCAN, PCA, t-SNE visualization
40	<b>Evaluation &amp; Deployment</b>	Confusion matrix, precision/recall/F1, ROC-AUC, model selection, Flask API
41	<b>ML Capstone Project</b>	End-to-end: problem definition, EDA, feature engineering, training, tuning, deployment



#### MODULE PROJECT

ML Model Training, Evaluation & Deployment

## MODULE 7

## M07

# DEEP LEARNING

Weeks 42 – 47 | 12 Hours | 2 Hours/Week

Go beyond traditional ML. Build neural networks that see images, understand text, and predict time series. Master TensorFlow, CNNs, RNNs, and the Transformer architecture that powers modern AI.

### Neural Networks

Perceptrons, backpropagation, DNNs

### Computer Vision

CNNs, transfer learning, image classification

### Sequences

RNNs, LSTM, GRU, time series, NLP

WEEK	TOPIC	DETAILS & LEARNING OUTCOMES
42	<b>Neural Network Basics</b>	Perceptron, activation functions, forward/backpropagation, gradient descent, loss functions
43	<b>TensorFlow/Keras DNNs</b>	Sequential/Functional API, dense layers, optimizers, learning rate, dropout, callbacks
44	<b>CNNs</b>	Convolution, filters/kernels, pooling, architectures (LeNet, VGG, ResNet), transfer learning
45	<b>RNNs &amp; Sequences</b>	RNN fundamentals, LSTM, GRU, bidirectional RNNs, time series, text classification
46	<b>Advanced DL</b>	Attention, Transformers, autoencoders, GANs overview, interpretability (Grad-CAM, SHAP)
47	<b>DL Project</b>	Capstone: image classification, NLP, or time series. Full pipeline & presentation



### MODULE PROJECT

Deep Learning Application (CV / NLP / Time Series)

## MODULE 8

## M08

# GENERATIVE AI

Weeks 48 – 52 | 10 Hours | 2 Hours/Week

The future of AI is generative. Build applications with GPT, Claude, and LangChain. Create RAG systems, AI agents, and production-ready Gen AI apps — the most in-demand skill of 2025 and beyond.

### LLM APIs

OpenAI, Claude, prompt engineering

### RAG Systems

Embeddings, vector databases, semantic search

### AI Agents

LangChain, function calling, tool integration

WEEK	TOPIC	DETAILS & LEARNING OUTCOMES
48	Intro to Gen AI & LLMs	Evolution of AI, Transformer deep-dive, attention, tokenization, GPT vs BERT
49	LLM APIs	OpenAI/Claude APIs, prompt engineering (zero-shot, few-shot, CoT), structured outputs
50	RAG & Vector DBs	RAG architecture, chunking, embeddings, ChromaDB/Pinecone/FAISS, semantic search
51	LangChain & Agents	LangChain framework, conversational AI, function calling, multi-step agents, tools
52	Gen AI Capstone	Build RAG Q&A system or AI agent. Deployment, evaluation (RAGAS) & presentation



#### MODULE PROJECT

Generative AI Application (RAG System / AI Agent)

## TECHNOLOGY STACK

## TOOLS &amp; TECHNOLOGIES

## LANGUAGES

Python 3.x, SQL

## DATA LIBRARIES

NumPy, Pandas, Matplotlib, Seaborn, Plotly

## DATABASES

PostgreSQL / MySQL, SQLite

## BIG DATA

Apache Spark (PySpark), Databricks, Delta Lake

## ETL / ORCHESTRATION

Apache Airflow, SQLAlchemy, pyodbc

## ML / DL FRAMEWORKS

Scikit-learn, XGBoost, LightGBM, TensorFlow, Keras, PyTorch

## GEN AI / LLMs

OpenAI API, Claude API, LangChain, ChromaDB, FAISS

## CLOUD / DEVOPS

Azure Blob Storage, Git &amp; GitHub, Docker

## IDEs &amp; NOTEBOOKS

VS Code, Jupyter, Google Colab, Databricks Notebooks

## EVALUATION

## ASSESSMENT &amp; CERTIFICATION

30%

## Weekly Assignments

Hands-on coding via GitHub

40%

## Module Projects

8 end-of-module capstones

10%

## Quizzes

Weekly quizzes &amp; participation

20%

## Final Capstone

Portfolio presentation

Complete all 8 module projects + 80% attendance = **Techma Zone Professional Certificate in Applied Data Science & AI**



# START YOUR DATA SCIENCE JOURNEY TODAY

Whether you're looking for data science training, AI solutions, or want to learn from industry experts, Techma Zone is your launchpad to a data-driven career.

PHONE / WHATSAPP

[+92 326 0376076](tel:+923260376076)

EMAIL

[info@techmazon.com](mailto:info@techmazon.com)

WEBSITE

[www.techmazon.com](http://www.techmazon.com)

[ENROLL NOW →](#)

LinkedIn: [/school/techmazon](https://www.linkedin.com/school/techmazon)

Facebook: [/techmazon](https://www.facebook.com/techmazon)

Instagram: [@techmazon](https://www.instagram.com/techmazon)

WhatsApp: [+92 326 0376076](https://www.whatsapp.com/+923260376076)

Office 16, RJ Mall 4th Floor, Johar Mor, Karachi, Pakistan