

# MUHAMMAD AHTISHAM

## DATA ENGINEER

✉ mahtisham.work@gmail.com | 📞 (+92) 3045164228 | 🌐 [LinkedIn Profile](#) | 📍 Islamabad, Pakistan

### PROFESSIONAL SUMMARY

Experienced Data Engineer with 5+ years of expertise in designing and implementing scalable data solutions across cloud platforms. Proven track record of delivering high-impact data infrastructure projects for international clients in utilities, real estate, and financial services. Skilled in modern data stack technologies including Databricks, AWS, and Azure, with demonstrated ability to optimize data processing efficiency by up to 40% and reduce operational costs by 50%. Pakistani citizen with ability to obtain EU work authorization within 5-8 weeks

### PROFESSIONAL EXPERIENCE

#### Data Engineer | Blutech Consulting

March 2022 – Present | Islamabad, Pakistan

##### Databricks Lakehouse Architecture – Water Utilities, UK

- Architected comprehensive Databricks Lakehouse solution integrating 5+ heterogeneous data sources, enabling unified analytics platform
- Orchestrated 20+ Azure Data Factory pipelines processing terabytes of daily data with 99.9% uptime, ensuring real-time analytics availability
- Developed specialized geospatial data processing capabilities using Apache Sedona, Fiona, and GeoPandas for Client's Merge Sewer Risk Model, creating master pipeline to process complex spatial datasets and deliver risk insights to frontend applications
- Implemented ACID-compliant Delta Lake tables, achieving 20% improvement in query performance and data consistency
- Designed cost-optimized external table architecture utilizing Azure Data Lake Gen2, reducing storage and query costs by 50%
- Established CI/CD practices with Azure DevOps, enhancing team collaboration and reducing production deployment errors by 30%
- Developed comprehensive monitoring framework using Databricks notebooks and Azure Monitor, decreasing pipeline failures by 25%

Technologies: Microsoft Azure, Databricks, Azure Data Factory, Azure Data Lake Gen2, Apache Sedona, Fiona, GeoPandas, Power BI, SQL Server, Azure DevOps, PySpark, SQL

##### Marketing Intelligence Platform – Real Estate, Canada

- Built end-to-end AWS Glue ETL pipelines integrating 4 diverse data sources into Amazon Redshift, maintaining 99.8% data integrity
- Collaborated with Marketing and Product teams to establish data governance framework and define actionable KPIs for campaign optimization
- Developed automated financial reporting dashboards for 20+ properties with integrated ML-powered pricing models, improving occupancy forecasting accuracy by 40%
- Conducted advanced tenant demographic analysis using Python & SQL, resulting in targeted marketing strategies increasing lease conversions by 20%
- Delivered executive-level comparative analytics and strategic reporting, directly supporting \$50M+ investment decisions
- Led data quality initiatives improving accuracy by 25% through automated validation and governance protocols

Technologies: Python, PySpark, SQL, AWS Lambda, Amazon Athena, AWS Glue, Amazon Redshift, S3, Power BI

##### Enterprise Data Warehouse Migration – Bank, South Africa

- Led critical migration of enterprise data warehouse from on-premises Sybase to Amazon Redshift, achieving 40% improvement in processing capabilities
- Transformed and optimized 800+ database objects (stored procedures, views, UDFs) ensuring zero data loss and 20% performance enhancement
- Implemented automated testing framework achieving 95% validation accuracy for data integrity across migration phases
- Managed project timeline and stakeholder communication, delivering migration 2 weeks ahead of schedule with minimal business disruption

Technologies: AWS, AWS Glue, Amazon Redshift, PySpark, SQL, PL/SQL, S3, CloudFormation, RDS, Sybase, Amazon Athena

##### Natural Language to SQL Application – Internal Project

- Developed comprehensive NL2SQL web application leveraging local LLM capabilities to transform natural language queries into executable SQL statements for SQL Server databases
- Architected fully open-source solution using Streamlit for user interface, Mistral LLM via Ollama for local inference, and LiteLLM for standardized model API integration
- Implemented dynamic database connectivity using pyodbc to execute AI-generated queries on restaurant ratings dataset with real-time result visualization
- Engineered sophisticated prompt templates and schema-aware query generation, ensuring accurate SQL translation based on database structure and user intent
- Delivered interactive analytics experience with live SQL preview functionality, demonstrating self-service business intelligence capabilities
- Achieved completely offline, cost-free deployment suitable for air-gapped environments and organizations with data privacy constraints

Technologies: Python, Streamlit, Ollama, Mistral LLM, LiteLLM, SQL Server, pyodbc, Pandas, VS Code

## Data Engineer | Bytewise Limited

September 2020 – March 2022 | Islamabad, Pakistan

### Cloud Data Infrastructure – American Environmental Solutions

- Designed scalable 3-tier AWS architecture implementing modern cloud-native data processing solutions
- Built robust ETL processes using AWS Glue for structured and semi-structured data, improving processing efficiency by 30%
- Implemented multiple loading strategies (Incremental, Full Load, SCD Type 1 & 2) ensuring optimal data freshness and historical tracking
- Developed automated data pipelines using AWS Glue Studio and Lambda, reducing processing time by 20%

Technologies: Python, PySpark, AWS, AWS Glue, Amazon Redshift, SQL, Amazon Athena, S3, Power BI

## TECHNICAL EXPERTISE

---

**Programming & Scripting:** Python, SQL, PySpark, PL/SQL

**Data Engineering:** ETL/ELT Pipelines, Apache Spark, Apache Airflow, AWS Glue, Azure Data Factory

**Cloud Platforms:** Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP)

**Data Warehousing:** Amazon Redshift, Google BigQuery, Delta Lake, Databricks Lakehouse

**Geospatial & Advanced Analytics:** Apache Sedona, Fiona, GeoPandas, Natural Language Processing

**AI/ML Integration:** Local LLM Deployment (Ollama), Mistral, LiteLLM, Prompt Engineering

**Analytics & Visualization:** Power BI, Tableau, Looker Studio, Streamlit

**Databases:** SQL Server, PostgreSQL, Amazon RDS

**DevOps & Tools:** Docker, Git, Jupyter Notebook, Azure DevOps, AWS CloudFormation

**Methodologies:** Agile, Data Modeling, Data Governance

## EDUCATION

---

### Bachelor of Science in Computer Science (BSCS)

National University of Modern Languages, Islamabad, Pakistan

CGPA: 3.50/4.00

## PROFESSIONAL CERTIFICATIONS

---

Databricks Certified Data Engineer Associate

AWS Partner: Technical Accreditation

AWS: Object Storage Certified

## LANGUAGES

---

English: Proficient (C2 Level) - Professional working proficiency

Urdu: Native speaker

German: Elementary (A1) - Currently studying

## ADDITIONAL INFORMATION

---

### Professional Development:

- Former member of Google Developers Student Club
- Former participant in Facebook Developer's Circle