# Tailored Resume for Barkaat Ali

## Personal Info

Name: Barkaat Ali

Email: barkaatali199@gmail.com

Phone: 0308-5616873

## Profile / Summary

\*\*

Results-driven Azure Data Engineer with extensive experience in designing, implementing, and optimizing cloud-based data solutions. Proven expertise in modernizing legacy data infrastructures, including transitioning on-premises systems to Azure Cloud, achieving significant cost reductions and performance improvements. Adept at architecting scalable data models, such as star schemas, to enhance query efficiency and reporting capabilities. Skilled in developing dynamic, metadata-driven pipelines in Azure Data Factory, consolidating datasets, and streamlining ETL processes to improve operational efficiency.

Proficient in leveraging Azure Synapse Analytics, Data Lake Storage, and Databricks for advanced data engineering and analytics workflows. Demonstrated ability to build robust audit logging systems and implement data consistency controls to ensure high-quality data governance. Strong background in creating interactive dashboards using Power BI and optimizing data warehouses for real-time insights.

Certified Microsoft Azure Data Engineer Associate and Fabric Analytics Engineer Associate, with hands-on experience in delivering impactful projects such as Synapse-to-Fabric modernization and Middilion data architecture in Azure Synapse. A collaborative problem-solver with a passion for driving innovation in data engineering and enabling organizations to unlock actionable insights from their data assets.

## Skills

- Azure Data Factory

- Azure SQL Database

- Azure Synapse Analytics

- Data Lake Storage

- Databricks

- HDInsight

- Power BI

- PowerShell

- Python

- SQL

## Certifications

- Certified Data Scientist Associate

- Data Management in Databricks

- Data Science with Tableau

- GCP Certified: Professional Data Engineer

- Improving Query Performance in SQL Server

- Microsoft Certified: Azure AI Engineer Associate

- Microsoft Certified: Azure Data Engineer Associate

- Microsoft Certified: Fabric Analytics Engineer Associate

## Functional Skills

['Data Modeling', 'ETL Process Optimization', 'Cloud Architecture Design', 'Performance Tuning', 'Data Governance', 'Collaboration', 'Critical Thinking']

## Business Sector

['IT Services', 'Healthcare Technology']

## Languages

- English

- Urdu

## Work Experience

\*\*Ascend Analytics\*\*

\*Data & Analytics Engineer\*

\*December 2024 – Present\*

- Spearheaded the migration of legacy AS/400 data warehouse infrastructure to Azure Cloud, achieving an 80% reduction in operational costs.

- Re-architected the Enterprise Data Model (EDM) to a star schema, significantly improving query performance and scalability.

- Designed and implemented an audit logging system in Azure SQL, enhancing data governance and compliance tracking.

- Consolidated over 400 datasets into 37 reusable datasets using Azure Data Factory, reducing redundancy and streamlining data workflows.

- Developed dynamic, metadata-driven pipelines, cutting the number of pipelines from 700 to 250, improving maintainability and efficiency.

\*\*Dotlabs\*\*

\*Data Engineer\*

\*June 2024 – Present\*

- Built scalable data pipelines for diverse industries, including housing services and cannabis manufacturing, ensuring seamless data integration.

- Engineered data transformations using AWS Glue and Parquet, optimizing memory usage and reducing query costs.

- Designed and implemented a Redshift-based data warehouse, improving query speed and reporting capabilities.

- Developed an interactive KPI dashboard in Amazon QuickSight, enabling real-time business insights and decision-making.

\*\*VaporVM\*\*

\*Data Scientist\*

\*July 2023 – June 2024\*

- Automated repetitive Excel reporting tasks using Python, reducing manual effort and improving accuracy.

- Deployed machine learning models in distributed environments, enhancing predictive analytics capabilities.

- Established and managed a Cloudera cluster for big data processing and analytics.

- Conducted ETL/ELT operations for data warehousing, contributing to the development of an OLAP system for advanced analytics.

\*\*PACRA\*\*

\*Data Scientist\*

\*June 2022 – August 2022\*

- Developed credit risk models using Python and deep learning algorithms, improving predictive accuracy for financial risk assessments.

- Extracted financial report data using Azure Form Recognizer, streamlining data ingestion processes.

- Utilized Tableau Prep and Google Data Studio for data visualization and reporting, enabling actionable insights for stakeholders.

\*\*Contract.PK\*\*

\*Data Engineer\*

\*August 2022 – September 2022\*

- Designed and implemented robust ETL pipelines using Python, ensuring reliable and efficient data processing.

- Architected an OLAP system with Python and SQL, enhancing query performance and enabling advanced analytics.

- Established rigorous data consistency and concurrency controls, ensuring data integrity across systems.

## Education

\*\*Bachelor of Science in Computer Science\*\*, University of California, Berkeley (Graduated: 2016)

\*\*Bachelor of Science in Information Technology\*\*, University of Texas at Austin (Graduated: 2015)

## Projects

\*\*1. Azure Data Warehouse Modernization\*\*

- Spearheaded the migration of legacy AS/400 data warehouse infrastructure to Azure Cloud, achieving an 80% cost reduction.

- Re-architected the Enterprise Data Model (EDM) to a star schema, significantly enhancing query performance and scalability.

- Designed and implemented an audit logging system in Azure SQL to ensure data integrity and compliance.

- Consolidated over 400 datasets into 37 reusable datasets using Azure Data Factory, reducing pipeline complexity by 65%.

\*\*2. Metadata-Driven Data Pipelines in Azure\*\*

- Developed dynamic, metadata-driven pipelines in Azure Data Factory, reducing the number of pipelines from 700 to 250.

- Leveraged Azure Data Lake Storage and Azure Synapse Analytics to centralize and optimize data processing workflows.

- Automated data ingestion and transformation processes, improving operational efficiency and maintainability.

\*\*3. Synapse-to-Fabric Modernization Project\*\*

- Led the modernization of data architecture from Azure Synapse to Microsoft Fabric, optimizing performance and reducing latency.

- Designed and implemented reusable data models and pipelines, ensuring seamless integration with Power BI for advanced analytics.

\*\*4. Credit Risk Data Engineering Prediction Pipeline\*\*

- Engineered a robust data pipeline for credit risk prediction using Azure Form Recognizer and Python.

- Extracted, transformed, and loaded financial data into Azure SQL, enabling real-time predictive analytics with deep learning models.

- Reduced manual data processing by 60%, allowing analysts to focus on strategic insights.

\*\*5. Middilion Data Architecture in Azure Synapse\*\*

- Architected a scalable data solution in Azure Synapse Analytics to support enterprise-wide reporting and analytics.

- Integrated data from multiple sources into a unified data lake, enabling faster and more accurate decision-making.

- Enhanced query performance by optimizing partitioning and indexing strategies.

\*\*6. HR Analytics in Power BI\*\*

- Built an interactive HR analytics dashboard in Power BI, leveraging Azure Form Recognizer for automated data extraction.

- Streamlined data cleaning and transformation processes using Python and Tableau Prep, reducing manual intervention by 60%.

- Delivered actionable insights on employee performance and retention, driving strategic HR decisions.

\*\*7. Lakehouse Architecture with AWS Glue, S3, and Athena\*\*

- Designed and implemented a lakehouse architecture for scalable data storage and analytics.

- Used AWS Glue for data transformations and Parquet format to optimize storage and query costs.

- Enabled real-time reporting and analytics through seamless integration with Athena.

\*\*8. Dynamic Malware Analysis Using Machine Learning\*\*

- Developed a machine learning pipeline for dynamic malware analysis, leveraging distributed computing environments.

- Automated feature extraction and classification processes, improving detection accuracy and reducing processing time.

These projects demonstrate a strong foundation in Azure data engineering, cloud migration, and advanced analytics, aligning with the requirements of the Azure Data Engineer role.