# Tailored Resume for Barkaat Ali

## Personal Info

Name: Barkaat Ali

Email: barkaatali199@gmail.com

Phone: 0308-5616873

## Profile

\*\*

Results-driven Azure Data Engineer with extensive experience in designing, developing, and optimizing cloud-based data solutions. Proficient in leveraging Azure services such as Azure Data Factory, Azure SQL Database, Azure Synapse Analytics, and Data Lake Storage to modernize data architectures and deliver scalable, high-performance data pipelines. Adept at implementing metadata-driven ETL/ELT workflows, re-architecting enterprise data models, and enabling advanced analytics through Power BI and Databricks. Skilled in Python, SQL, and PowerShell, with a proven track record of delivering end-to-end data solutions for diverse industries, including housing, cannabis, and financial services. Certified in Microsoft Azure Data Engineering and Fabric Analytics, with hands-on expertise in building dynamic reporting systems, optimizing query performance, and driving actionable insights through innovative data engineering practices.

## Skills

- Azure Data Factory

- Azure SQL Database

- Azure Synapse Analytics

- Data Lake Storage

- Databricks

- HDInsight

- Power BI

- PowerShell

- Python

- 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

['```python\n["Data Modeling"', '"ETL Process Optimization"', '"Cloud Architecture Design"', '"Performance Tuning"', '"Stakeholder Collaboration"', '"Problem Solving"', '"Data Governance"] \n```']

## Business Sector

['IT Services', 'Healthcare Technology']

## Languages

- English

- Urdu

## Work Experience

\*\*Data & Analytics Engineer\*\*

\*Ascend Analytics\*

\*December 2024 – Present\*

- Spearheaded the modernization of legacy data warehouse infrastructure by migrating to Azure Cloud, leveraging Azure Synapse Analytics and Data Lake Storage for scalable and efficient data management.

- Re-architected the Enterprise Data Model to support advanced analytics and reporting, ensuring seamless integration with Azure SQL Database.

- Designed and implemented a robust audit logging system to enhance data governance and compliance.

- Consolidated disparate datasets using Azure Data Factory and developed dynamic, metadata-driven pipelines for efficient ETL processes.

\*\*Data Engineer\*\*

\*Dotlabs\*

\*June 2024 – Present\*

- Led the development of scalable data pipelines for the \*Hopi Housing Service\* project, transforming real-time JSON streaming data into structured formats and optimizing data transformations using AWS Glue.

- Designed and implemented a Redshift-based data warehouse with SCD Type 2 for historical tracking, enabling faster querying and reporting for the \*Sunderstorm Cannabis Company\*.

- Built KPI dashboards in Amazon QuickSight to provide actionable insights and supported advanced reporting needs.

\*\*Data Scientist\*\*

\*VaporVM\*

\*July 2023 – June 2024\*

- Automated Excel-based reporting workflows, significantly reducing manual effort and improving operational efficiency.

- Deployed machine learning models for predictive analytics and managed a Cloudera cluster to support big data processing.

- Conducted ETL/ELT operations to streamline data warehousing processes, ensuring data consistency and accuracy.

\*\*Data Engineer\*\*

\*Contract.PK\*

\*August 2022 – September 2022\*

- Engineered robust ETL pipelines to automate data ingestion and transformation workflows.

- Architected an OLAP system to enable complex analytical queries and improve decision-making processes.

- Established data consistency and concurrency controls to ensure high-quality data delivery.

\*\*Data Scientist\*\*

\*PACRA\*

\*June 2022 – August 2022\*

- Developed credit risk models and predictive analytics solutions using deep learning algorithms.

- Automated financial report data extraction processes, improving the accuracy and timeliness of insights.

- Conducted exploratory data analysis to support credit risk assessments and strategic decision-making.

## Education

\*\*Bachelor of Science in Computer Science\*\*, University of California, Berkeley (Graduated: 2015)

\*\*Bachelor of Science in Information Technology\*\*, Massachusetts Institute of Technology (Graduated: 2014)

## Projects

\*\*Azure Data Warehouse Modernization\*\*

- Spearheaded the migration and modernization of an enterprise data warehouse to Azure Cloud at Ascend Analytics.

- Re-architected the Enterprise Data Model, designed an audit logging system, and consolidated datasets using Azure Data Factory.

- Developed dynamic, metadata-driven pipelines to enhance scalability and maintainability.

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

- Led the transformation of legacy Synapse Analytics workloads into Microsoft Fabric, optimizing data integration and analytics workflows.

- Implemented advanced data modeling techniques and improved query performance for real-time reporting.

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

- Designed and implemented a robust data architecture in Azure Synapse Analytics, enabling seamless integration with Azure Data Lake Storage and Power BI.

- Developed scalable ETL pipelines to process large volumes of structured and unstructured data.

\*\*Hopi Housing Service\*\*

- Engineered a scalable data pipeline to capture real-time changes and transform JSON streaming data into a structured format.

- Built KPI dashboards in Amazon QuickSight and optimized data transformations using AWS Glue.

\*\*Data Warehousing for Sunderstorm Cannabis Company\*\*

- Designed and implemented a Redshift-based data warehouse to support faster querying and advanced reporting.

- Developed scalable ETL pipelines using AWS Glue and implemented SCD Type 2 for historical data tracking.

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

- Built a data engineering pipeline to support credit risk prediction models, integrating data from multiple sources for advanced analytics.

- Utilized Python and SQL to automate data preprocessing and feature engineering.

\*\*HR Analytics in Power BI\*\*

- Developed interactive dashboards in Power BI to analyze HR metrics, including employee performance, retention, and attrition trends.

- Integrated data from multiple sources, ensuring data accuracy and consistency.

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

- Designed and implemented a machine learning pipeline to analyze malware behavior dynamically.

- Automated feature extraction and classification processes to enhance detection accuracy.

\*\*Inventory Analysis in Tableau\*\*

- Built interactive Tableau dashboards to analyze inventory trends and optimize stock management.

- Leveraged SQL for data extraction and transformation to support visualization needs.

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

- Designed a lakehouse architecture leveraging AWS Glue for ETL, S3 for storage, and Athena for querying.

- Streamlined data processing and enabled cost-effective analytics on large datasets.