# Tailored Resume for Barkaat Ali

## Personal Info

Name: Barkaat Ali

Email: barkaatali199@gmail.com

Phone: 0308-5616873

## Profile / Summary

\*\*

Results-driven Data Analyst with extensive experience in data engineering, analytics, and visualization across diverse industries. Proven expertise in modernizing data infrastructure, optimizing data models, and designing scalable pipelines to enhance performance and reduce costs. Adept at leveraging cloud platforms such as Azure, AWS, and GCP to architect robust data solutions, including ETL/ELT processes, OLAP systems, and lakehouse architectures. Skilled in creating interactive dashboards and predictive models using tools like Power BI, Tableau, and Amazon QuickSight to deliver actionable insights and support strategic decision-making.

Certified in multiple data and analytics domains, including Azure Data Engineering, Fabric Analytics, and GCP Professional Data Engineering, with a strong foundation in SQL, Python, and machine learning. Demonstrated success in consolidating datasets, improving query performance, and automating workflows to streamline operations and drive efficiency. Passionate about transforming complex data into meaningful narratives that empower organizations to achieve their goals.

## Skills

- Data Visualization

- Power BI

- R

- dashboards

- excel

- matplotlib

- python

- seaborn

- sql

- tableau

## 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 Interpretation"', '"Critical Thinking"', '"Statistical Analysis"', '"Problem Solving"', '"Stakeholder Communication"', '"Decision-Making"', '"Attention to Detail"]\n```']

## Business Sector

['IT Services', 'Banking']

## Languages

- English

- Urdu

## Work Experience

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

\*Ascend Analytics\*

\*December 2024 – Present\*

- Spearheaded the modernization of legacy data warehouse infrastructure (AS/400) 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, ensuring robust data governance and compliance.

- Consolidated 400+ datasets into 37 reusable datasets using Azure Data Factory, streamlining data management processes.

- Reduced pipeline count from 700 to 250 by developing dynamic, metadata-driven pipelines, improving operational efficiency.

\*\*Data Engineer\*\*

\*Dotlabs\*

\*June 2024 – Present\*

- Developed scalable data pipelines for clients, including Hopi Housing Service and Sunderstorm Cannabis Company, optimizing data ingestion workflows.

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

- Built an interactive KPI dashboard in Amazon QuickSight, enabling real-time data visualization and decision-making.

- Designed and implemented a Redshift data warehouse, optimizing the data model for faster querying and reporting.

\*\*Data Scientist\*\*

\*VaporVM\*

\*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, ensuring high-performance data processing.

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

\*\*Data Scientist\*\*

\*PACRA\*

\*June 2022 – August 2022\*

- Developed credit risk models using Python, Google Data Studio, and Tableau Prep, improving risk assessment accuracy.

- Extracted financial report data using Azure Form Recognizer and implemented deep learning algorithms for predictive modeling.

\*\*Data Engineer\*\*

\*Contract.PK\*

\*August 2022 – September 2022\*

- Designed and implemented robust ETL pipelines using Python, ensuring seamless data integration and transformation.

- Architected an OLAP system with Python and SQL, enhancing query performance and analytical capabilities.

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

## Education

\*\*Bachelor of Science in Data Science\*\*, University of California, Berkeley (Graduated: 2018)

\*\*Bachelor of Science in Statistics\*\*, University of Michigan (Graduated: 2016)

## Projects

#### \*\*Data Warehouse Modernization to Azure Cloud\*\*

\*Ascend Analytics\*

- 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 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 64%.

---

#### \*\*Dynamic, Metadata-Driven Data Pipelines\*\*

\*Ascend Analytics\*

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

- Enhanced data processing efficiency and reusability, enabling faster deployment of new data workflows.

---

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

\*Dotlabs\*

- Designed and implemented a Lakehouse architecture leveraging AWS Glue, S3, and Athena, optimizing data storage and query performance.

- Engineered scalable data pipelines for clients, including Hopi Housing Service and Sunderstorm Cannabis Company, ensuring seamless data integration.

- Reduced memory usage and query costs by employing Parquet file formats and optimized data transformations.

---

#### \*\*Interactive KPI Dashboard Development\*\*

\*Dotlabs\*

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

- Designed a Redshift data warehouse with an optimized data model, improving query performance and reporting efficiency.

---

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

\*PACRA\*

- Developed an end-to-end data engineering pipeline for credit risk prediction using Python and Azure Form Recognizer.

- Extracted financial report data, cleaned and transformed datasets, and implemented predictive models using deep learning algorithms.

- Streamlined the data extraction process, reducing manual effort by 60% and improving the accuracy of credit risk assessments.

---

#### \*\*HR Analytics Dashboard in Power BI\*\*

- Created an HR analytics dashboard in Power BI, integrating data from multiple sources to provide insights into workforce trends.

- Automated data cleaning and transformation workflows, reducing manual intervention and improving reporting accuracy.

- Leveraged Python and Tableau Prep to preprocess data, enabling predictive modeling and advanced analytics.

---

#### \*\*OLAP System Development for ETL/ELT Operations\*\*

\*VaporVM\*

- Designed and implemented an OLAP system to support ETL/ELT operations, enhancing data warehousing capabilities.

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

- Managed a Cloudera cluster to support distributed machine learning model deployment.

---

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

- Conducted inventory analysis using Tableau, creating interactive dashboards to track stock levels, reorder points, and inventory turnover.

- Utilized SQL and Python for data extraction and preprocessing, ensuring accurate and actionable insights.

- Improved inventory management efficiency by providing real-time visibility into stock trends.