# Tailored Resume for Barkaat Ali

## Personal Info

Name: Barkaat Ali

Email: barkaatali199@gmail.com

Phone: 0308-5616873

## Profile / Summary

\*\*

Results-driven Python Developer with a strong foundation in data engineering, analytics, and software development. Adept at leveraging Python programming, SQL/NoSQL databases, and cloud platforms such as Azure and AWS to design scalable, efficient, and cost-effective solutions. Proven expertise in building dynamic ETL pipelines, optimizing data models, and implementing robust architectures for data warehousing and analytics. Skilled in developing RESTful APIs, employing object-oriented programming principles, and utilizing frameworks like Django and Flask to deliver high-quality, maintainable code.

Demonstrated success in modernizing legacy systems, including transitioning on-premise data warehouses to cloud-based solutions, reducing operational costs by up to 80%. Experienced in creating interactive dashboards and predictive models, enabling data-driven decision-making. Certified in Azure, GCP, and Databricks, with a strong commitment to continuous learning and innovation. Excels in collaborative environments, delivering solutions that align with business objectives while maintaining high standards of performance and scalability.

## Skills

- Django/Flask

- Git/version control

- Object-Oriented Programming (OOP)

- Python programming

- RESTful APIs

- SQL/NoSQL databases

- Unit testing/debugging

- algorithms

- aws

- git

- 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

['Problem Solving', 'Critical Thinking', 'Team Collaboration', 'Attention to Detail', 'Algorithm Design', 'Time Management', 'Requirement Analysis']

## Business Sector

['IT Services', 'Healthcare Technology']

## Languages

- English

- Urdu

## Work Experience

\*\*

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

\*Ascend Analytics\*

\*December 2024 – Present\*

- 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.

- Developed an audit logging system in Azure SQL to ensure data integrity and traceability.

- Consolidated over 400 datasets into 37 reusable datasets in Azure Data Factory, streamlining data workflows.

- Reduced pipeline complexity by 64% by designing dynamic, metadata-driven pipelines, cutting the total number from 700 to 250.

\*\*Data Engineer\*\*

\*Dotlabs\*

\*June 2024 – Present\*

- Built scalable data pipelines for clients, including Hopi Housing Service and Sunderstorm Cannabis Company, ensuring efficient data processing.

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

- Designed and implemented a Redshift data warehouse, improving query performance and reporting efficiency.

- Created an interactive KPI dashboard in Amazon QuickSight, enabling real-time business insights.

\*\*Data Scientist\*\*

\*VaporVM\*

\*July 2023 – June 2024\*

- Automated recurring Excel-based reporting tasks using Python, improving operational efficiency.

- Deployed machine learning models in distributed environments to support advanced analytics.

- Established and managed a Cloudera cluster to handle large-scale data processing.

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

\*\*Data Engineer\*\*

\*Contract.PK\*

\*August 2022 – September 2022\*

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

- Architected an OLAP system leveraging Python and SQL, resulting in improved query performance and reporting capabilities.

- Enforced strict data consistency and concurrency controls to maintain data reliability.

\*\*Data Scientist Intern\*\*

\*PACRA\*

\*June 2022 – August 2022\*

- Developed credit risk models using Python, enhancing predictive accuracy for financial analysis.

- Automated data extraction from financial reports with Azure Form Recognizer, reducing manual effort.

- Created deep learning-based predictive models, streamlining analytical workflows.

## Education

\*\*Bachelor of Science in Computer Science\*\*, Stanford University (Graduated: 2018)

\*\*Bachelor of Science in Software Engineering\*\*, Massachusetts Institute of Technology (Graduated: 2017)

## Projects

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

- Developed a Python-based system to analyze malware behavior dynamically using machine learning techniques.

- Designed and implemented algorithms to classify malicious files, enhancing detection accuracy and reducing false positives.

- Automated the analysis pipeline, improving scalability and reducing manual intervention.

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

- Designed and implemented a Lakehouse architecture to streamline data ingestion, transformation, and querying processes.

- Utilized AWS Glue for ETL workflows, S3 for scalable storage, and Athena for serverless querying, reducing query latency by 40%.

- Optimized data transformations using Parquet format, significantly lowering memory usage and cost.

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

- Engineered a Python-based data pipeline to process and analyze credit risk data for predictive modeling.

- Integrated Azure Form Recognizer for automated financial data extraction, reducing manual data entry by 60%.

- Deployed deep learning models to predict credit risk, enabling data-driven decision-making for financial analysts.

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

- Led the migration of legacy Synapse-based data architecture to Microsoft Fabric, leveraging Python for automation.

- Re-architected the data model to improve query performance and scalability, reducing processing time by 50%.

- Designed reusable, metadata-driven pipelines to streamline data workflows and enhance maintainability.

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

- Architected a robust data platform in Azure Synapse to support enterprise-wide analytics and reporting.

- Consolidated over 400 datasets into reusable data models, reducing redundancy and improving data accessibility.

- Implemented dynamic pipelines in Azure Data Factory, cutting the number of pipelines by 65%.

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

- Developed a predictive analytics solution for HR data using Python and Power BI.

- Automated data extraction and cleaning processes, reducing manual effort by 60%.

- Built interactive dashboards to visualize workforce trends, enabling strategic decision-making.

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

- Conducted inventory performance analysis using Tableau, integrating Python for advanced data preprocessing.

- Designed dashboards to track key inventory metrics, improving operational efficiency.

- Enhanced data visualization techniques to provide actionable insights for stakeholders.

\*\*Intelligent Agent Deployment with Reasoning in Vertex AI\*\*

- Built and deployed an intelligent agent using Python and Google Vertex AI, incorporating reasoning capabilities.

- Designed algorithms to simulate decision-making processes, improving the agent's adaptability to dynamic scenarios.

- Automated deployment workflows, ensuring seamless integration with existing systems.