# Tailored Resume for Barkaat Ali

## Personal Info

Name: Barkaat Ali

Email: barkaatali199@gmail.com

Phone: 0308-5616873

## Profile

\*\*

Results-driven Python Developer with a strong foundation in data engineering, analytics, and software development. Proficient in designing and implementing scalable data pipelines, modernizing data architectures, and developing robust APIs using Python frameworks such as Django and Flask. Experienced in cloud platforms (AWS/Azure), SQL database management, and RESTful API development, with a proven track record of optimizing workflows and delivering high-performance solutions. Adept at leveraging advanced analytics, machine learning, and ETL/ELT processes to solve complex business challenges. Certified in Azure, GCP, and Databricks, with hands-on expertise in building dynamic, metadata-driven systems and delivering actionable insights through data visualization tools like Power BI and Tableau. Recognized for a meticulous approach to debugging, unit testing, and ensuring data consistency across diverse projects.

## Skills

- Cloud platforms (AWS/Azure).

- Django/Flask frameworks

- Git/version control

- Python programming

- RESTful API development

- SQL and database management

- Unit testing and 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', 'Analytical Thinking', 'Team Collaboration', 'Attention to Detail', 'Time Management', 'Adaptability', 'Requirement Analysis']

## Business Sector

['IT Services', 'Healthcare Technology']

## Languages

- English

- Urdu

## Work Experience

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

\*Ascend Analytics\*

\*December 2024 – Present\*

- Spearheaded the modernization of data warehouse infrastructure by migrating to Azure Cloud, leveraging Python to re-architect the Enterprise Data Model for improved scalability and performance.

- Designed and implemented an audit logging system to ensure data integrity and traceability across pipelines.

- Consolidated diverse datasets using Azure Data Factory and developed dynamic, metadata-driven ETL pipelines to streamline data processing workflows.

\*\*Data Engineer\*\*

\*Dotlabs\*

\*June 2024 – Present\*

- Designed and deployed scalable data pipelines for real-time updates and transformations, ensuring seamless integration of structured and unstructured data.

- Optimized data transformations using Python and AWS Glue for the \*\*Hopi Housing Service\*\* project, enabling efficient KPI dashboard creation in Amazon QuickSight.

- Engineered a Redshift-based data warehouse with historical tracking (SCD Type 2) for the \*\*Sunderstorm Cannabis Company\*\*, improving query performance and reporting capabilities.

\*\*Data Scientist\*\*

\*VaporVM\*

\*July 2023 – June 2024\*

- Automated reporting processes by replacing manual Excel tasks with Python-based solutions, reducing operational overhead.

- Developed and deployed machine learning models for predictive analytics, enhancing decision-making capabilities.

- Conducted ETL/ELT operations to support data warehousing initiatives and managed a Cloudera cluster to ensure high availability and performance.

\*\*Data Engineer\*\*

\*Contract.PK\*

\*August 2022 – September 2022\*

- Engineered robust ETL pipelines to process and transform large datasets, ensuring data consistency and concurrency.

- Designed and implemented an OLAP system to support advanced analytics and reporting needs.

\*\*Data Scientist (Intern)\*\*

\*PACRA\*

\*June 2022 – August 2022\*

- Developed credit risk models using Python, leveraging deep learning algorithms for predictive analytics.

- Automated financial data extraction from reports and implemented data preprocessing pipelines to support modeling efforts.

## Education

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

\*\*Bachelor of Science in Software Engineering\*\*, Massachusetts Institute of Technology (Graduated: 2016)

## Projects

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

- Designed and implemented a Python-based system to analyze malware behavior dynamically using machine learning techniques.

- Automated feature extraction and classification processes, improving detection accuracy and reducing manual intervention.

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

- Led the migration of data architecture from Azure Synapse to Microsoft Fabric, optimizing performance and scalability.

- Developed Python-based ETL pipelines to streamline data ingestion and transformation, ensuring seamless integration with the new platform.

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

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

- Developed Python scripts to automate data transformation workflows, enabling real-time analytics and reducing latency.

\*\*Hopi Housing Service\*\*

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

- Utilized AWS Glue and DynamoDB triggers to optimize data transformations and created KPI dashboards in Amazon QuickSight for actionable insights.

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

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

- Developed Python-based pipelines using AWS Glue and implemented Slowly Changing Dimensions (SCD Type 2) for historical data tracking.

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

- Built an end-to-end Python-based pipeline for credit risk prediction, integrating data preprocessing, feature engineering, and machine learning models.

- Leveraged SQL and Python to automate data ingestion and transformation, improving the accuracy of predictive analytics.

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

- Architected a robust data model in Azure Synapse to support enterprise-level analytics and reporting.

- Developed Python scripts for metadata-driven pipeline automation, enhancing data consistency and reducing manual effort.

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

- Conducted advanced HR data analysis and visualization using Python for preprocessing and Power BI for dashboard creation.

- Delivered actionable insights into employee performance and retention trends.

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

- Automated data extraction and transformation using Python to support inventory analysis in Tableau.

- Created interactive dashboards to monitor stock levels, sales trends, and supply chain efficiency.