

HAKIM POCKETWALLA

SENIOR DATA ENGINEER

hakimpocketwalla@gmail.com | +91 8007000453 | Pune | <https://www.linkedin.com/in/hakim-najmi-pocketwalla>

I am a data engineering enthusiast who likes solving client problems using technology. I am effective at communication usually being the point of contact for most clients. I enjoy building automation and implementing new ideas in big data.

WORK EXPERIENCE

SENIOR DATA ENGINEER, TIGER ANALYTICS (Mar 2024 – present)

- Leading the end-to-end planning, design and implementation of big data pipelines for a major healthcare client optimized existing jobs to create a completely new metadata driven framework with multiple layers of processing. Reduced existing codebase by 35-40% and improved performance by 25%.
- Implemented various Spark optimization techniques to reduce job runtime and created a one-click script to trigger dynamic workflows based on which data files arrived, each processed differently based on metadata stored. Implementation Schema evolution dynamically using metadata and delta tables.
- Module lead for food industry client to enhance their current framework implementation to first implement observability as their requirement and then further created Databricks workflows to streamline individual tasks previously run manually by their team. Implemented backfill for failed pipeline or delayed file arrival.
- Delivered key components up to the production layer, such as JSON flattening with indexing, dynamically setting Spark properties through YAML configurations, and configuring lookback and offset values to control start and end times of incremental data at the day and hour levels.

SENIOR SOFTWARE ENGINEER DATA ENGINEERING, EPAM (Dec 2022 – Dec 2023)

- **Lead** implementation of an enterprise big data ingestion and processing solution for backend analytics of a customer portal for a major Product based client using **PySpark, Python and SQL** running on **AWS EMR** with **Airflow** for orchestration and scheduling. Processed large volumes of semi-structured data running into TBs.
- I was the primary person responsible in requirement gathering across cross functional teams **leading daily scrum calls** for updates of the data engineering team along with task allocation to all team members bridging the communication from onshore Product director based on changing priorities, capacity and bandwidth decisions.
- Responsible for core development, data quality and PySpark optimization effort along with maintaining a high level code quality of an important production pipeline for the client.

DATA ENGINEER, PAYTM (Feb 2022 – Dec 2022)

- Part of **Paytm's lending department** working with the core data engineering team as a **module lead**. Responsible for all requirements gathering, **design, implementation and complete ownership** of all data engineering solutions for Paytm's product on the app - Personal Loans.
- Interfaced directly with Paytm's stakeholders such as the product team, collections team, risk data team and Creditmate team. Lead calls discussing initiatives to understand their pain points, **scoping cloud architecture, creating technical designs** and complete **implementation and maintenance of Spark data pipelines** for product requirements.

BIG DATA AND CLOUD SENIOR DEVELOPER, ABZOOBA (Dec 2019 – Feb 2022)

- Built **scalable Big Data ETL** pipelines on both the **AWS and Azure** platforms for various clients in the Financial and Retail domains. Leading the collaboration with clients single-handedly as a senior member in the **Agile** methodology.
- Created a **Pyspark and Airflow based data ingestion accelerator** as an internal asset which was then deployed to two major clients.
- Worked with client data science teams to implement a feature store for financial crimes and AML domain.
- Regularly helped the recruitment team with **taking technical interviews** on Big Data, Cloud computing, Programming and SQL.

BUSINESS TECHNOLOGY ANALYST BIG DATA, ZS ASSOCIATES (Jun 2019 – Dec 2019)

- Worked on cloud migration creating ETL pipelines using **Azure Databricks** and **Azure Data Factory**.
- Created a custom buying recommendation engine for a Pharmaceutical Client using **Neo4j** and **Python Flask** APIs.
- Implemented Analytics reports on procurement based over time, spend, budget.

SOFTWARE ENGINEER BIG DATA, ACCENTURE (Jul 2017 – May 2019)

- Worked on projects using **Apache Spark**, Hadoop, **Hive**, Kafka, AWS. Helped build Analytical reports for clients by processing raw data using Spark and Hadoop processing.
- Worked on Alexa skills kit and **Google Dialogflow** to deliver **chatbots** and custom voice-based assistant using AWS Lambda functions and Echo Dot.

TECHNICAL SKILLS

- **Big Data Technologies:** PySpark, Spark Streaming, Kafka, Databricks, Snowflake, Airflow.
- **Microsoft Azure cloud:** Azure Data Factory, Databricks, ADLS, Azure Synapse, Azure SQL DB.
- **Amazon Web Services cloud:** EMR, Lambda functions, S3, RDS, EC2, Glue, IAM.
- **Programming:** Core Python, Python Flask, Pandas, Cypher QL.
- **Databases:** PostgreSQL, MySQL, RDS Aurora, Azure SQL DB, Snowflake, Azure Synapse, AWS Redshift.
- **Other tools and methodologies:** Git, JIRA, Bitbucket, Confluence, Jupyter, Agile, Sprint planning.

AWARDS AND RECOGNITIONS

- | | |
|--|--------------------------------------|
| • Microsoft Certified Azure Fundamentals | • AWS Certified Cloud Practitioner |
| • Accenture Outperform Award Winner | • Neo4J Certified Professional |
| • Databricks Academy Spark Training | • Cloudera Spark and Hadoop Training |

EDUCATION

- 2013 – 2017 | Pune, India
BACHELOR OF ENGINEERING (COMPUTER SCIENCE), VISHWAKARMA INSTITUTE OF INFORMATION TECHNOLOGY
- 2011 - 2013 | Pune, India
HIGH SCHOOL, S.M. CHOKSEY JUNIOR COLLEGE
- 2001 - 2011 | Muscat, Sultanate of Oman
SCHOOL, INDIAN SCHOOL AL WADI AL KABIR