

PROFILE

Experienced Data Engineer with expertise in building scalable data pipelines, real-time and batch data processing using Apache Spark, Python, and SQL. Skilled in data modeling, ETL workflows, and deploying solutions on cloud platforms like AWS. Adept at supporting analytics and machine learning use cases, including applications in finance and trading.

CONTACT

EMAIL (Preferred):

shashireddy40@gmail.com

PHONE:

+27 664004232

Linked-In Profile:

<https://www.linkedin.com/in/dataengrsa/>

PORTFOLIO

<https://shashireddy40.github.io/mypage/>

CERTIFICATIONS

ISTQB-CTFL

Big Data Engineering with Hadoop and Spark from E&ICT, IIT Roorkee

PERSONAL PROFILE

Full Name:

PATOORI SHASHIDHAR REDDY

Date of Birth:

10th May 1990

Languages Known:

English, Telugu and Hindi

SHASHIDHAR REDDY (SHA-SHI)

Senior Data Engineer | Python | Spark | AWS

South Africa |  Remote

EDUCATION



Jawaharlal Nehru Technological University, Hyderabad, India

Bachelor of Technology in Aeronautical Engineering.(2007-2011)

WORK EXPERIENCE



SVSS TRADING (PTY)LTD

Python Data Engineer | Remote | Contract | POC for 6 months

November 2024-May 2025



OLDMUTUAL SA

Data Engineer | Johannesburg, South Africa | Contract | Hybrid

March 2023–Oct 2024(1.6 years)



JUMO.WORLD

Senior Data Engineer | Cape town, South Africa | Remote

Oct 2020-Feb 2023 (2+ years)



SVSS TRADING (PTY)LTD

Quant Data Engineer | Johannesburg, South Africa

March 2019-May 2020. (1.3 years)



Sonata Software

Big Data Engineer | Hyderabad, India

January 2015–January 2019 (4 Years)

SKILLS

Languages:

- Python
- SQL
- Shell Scripting
- Scala

Big data Technologies:

- Apache Spark
- Apache NiFi
- Apache Airflow
- Delta Lake, MySQL
- Pandas, NumPy
- Polars
- PostgreSQL
- Quest DB (Time series DB)
- Dashboards (Grafana)
- Datadog
- Spark Structured Streaming
- Terraform

Utilities/Tools:

- Docker
- Postman
- VS code, Postgres
- Web scraping
- Web Socket handling

Cloud Platforms:

AWS Stack:

- AWS S3,
- Redshift,
- EMR,
- EC2,
- SNS, SQS
- Lambda
- Athena, Crawlers

Upskilling In:

- DBT (Data Build Tool)
- Databricks
- Snowflake
- Apache Kafka
- Machine Learning

Quantitative finance:

- Developed trading data pipelines for real-time and historical market data.
- Experienced in quantitative trading across equities, derivatives, and futures.
- Proficient in quant development using Web Sockets, time-series databases, and Python for streaming Data Pipelines.

Leadership Skills

- Mentored and onboarded junior engineers
- Conducted code and design reviews
- Designed architectural diagrams
- Contributed to cloud data platform strategies
- Presented data solutions to cross-functional teams
- Coordinated with stakeholders and product teams
- Led technical discussions and sprint planning

- **PROFESSIONAL EXPERIENCE**

PROJECT #1:

Title: Derivative Data Pipeline | Proof of Concept (PoC) – 6 Months Contract
Location: Remote, Freelancer
Duration: Nov 2024 -April 2025
Tools Used: Python, WebSocket, API
Team Size: 6
Role: Python Data engineer

Description:

Developed a Trading Pipeline as a Proof of Concept (PoC) for systematic trading, involving data acquisition, strategy development, back testing, and deployment. Key aspects of the project included:

- Data Engineering: Collected and cleaned market data from various APIs (Equities, Futures, Derivatives). Stored processed data in AWS (S3, Redshift) and structured databases for efficient retrieval.
- Trading Strategy Framework: Designed and implemented an algorithmic trading framework to support multiple strategies, including trend-following and mean-reversion models.
- Back testing & Optimization: Built a custom back testing engine to evaluate trading strategies using historical data, incorporating transaction costs and slippage.
- PNL Reports & Dashboards: Generated daily PNL reports and performance analytics dashboards to track trade efficiency, risk exposure, and strategy effectiveness.

PROJECT #2:

Title: Two-Pot retirement system
Location: Old Mutual SA, Johannesburg, South Africa
Duration: March 2023 to October 2024
Tools Used: Spark, Python, SQL, Postgres, AWS, GCP, Pandas
Team Size: 8
Role: Data Engineer

Description:

The two-pot retirement system from Old Mutual in South Africa allows members to access a portion of their retirement savings for emergencies while keeping the rest invested until retirement.

Responsibilities:

- Design, create, build & maintain data pipelines
- Gathering data from various sources and Sending to the Two Pot system data warehouse.
- Transforming data with spark and writing the data with various data formats
- Creating Dashboards with Dynatrace and creating notifications for the failures
- Automating the CI/CD Pipeline for the entire data pipeline
- Mentored and onboarded junior data engineers, conducted code reviews, and fostered a culture of knowledge sharing and best practices.
- Influenced architectural decisions and contributed to the design of data platform strategies across cloud-based environments.

- Facilitated sprint planning, task delegation, and progress tracking in Agile environments to ensure timely delivery of data projects.

PROJECT #3:

Title: Data Systems (EDW)
 Partners: Tigo, Airtel and MTN
 Location: Cape town, South Africa
 Duration: October 2020 – February 2023
 Tools Used: AWS, Spark, Python, SQL, Airflow
 Team Size: 5
 Role: Senior Data Engineer

Description:

JUMO is the market leading banking as a service platform, enabling real-time access to funds at the lowest possible operating cost.

Integrating into JUMO's platform enables our partners to offer loans, savings and a wide range of financial choices to a new group of customers.

The first ecosystem partnerships were established with Tigo, Airtel and MTN to bring short-term loan products to people and small businesses in Kenya, Zambia and Uganda.

Responsibilities:

- used various operators ETL pipelines that extract data from multiple sources, Transforming and stored in Delta Lake in optimized way
- Focus on simplified data models that represent business processes, operational excellence, systems & services resilience and efficiency, availability, maintainability, security.
- Building and handling successful designs, pipelines and development using Spark and AWS for large data processing systems
- build and manage workflows programmatically with Airflow
- Working in agile software-development methodology and utilizing technologies such as Jira, Git, Jenkins, Spinnaker
- Improvement of engineering processes (i.e., design reviews, build, operations, testing, release/deploy, monitoring etc.).

PROJECT #4:

Title: Trade Data Pipeline & Predictive Modelling for Algorithmic Trading
 Client: Individual Clients (NSE/BSE Indian Markets).
 Location: Sandton, Johannesburg.
 Duration: March 2019 to May 2020
 Tools Used: Python, SQL, Postgres, AWS S3
 Team Size: 4
 Role: Quant Data Engineer

Description:

Developed an automated data pipeline for gathering and processing trade and stock option data to support algorithmic trading strategies. Implemented robust back testing frameworks to validate strategies using historical data, ensuring optimal performance and risk management. Collaborated with machine learning engineers by providing accurate datasets to train predictive models for price action, enabling data-driven decision-making in real-time trading environments

Responsibilities:

- Designed and developed robust data pipelines to gather, clean, and process large volumes of trade data from various financial markets and exchanges.
- Implemented and optimized back testing frameworks for testing and validating algorithmic trading strategies using historical market data.
- Collected and processed stock option data, analysing trends and patterns to support trading strategy development and optimization.
- Collaborated with machine learning engineers to provide accurate and timely datasets, enabling predictive models for price action forecasting.
- Automated data gathering for real-time market data streams and option chains, integrating with internal analytics platforms to improve decision-making.
- Conducted exploratory data analysis (EDA) and statistical modelling to identify profitable opportunities and enhance trading algorithms.
- Utilized tools like Python, SQL, Pandas, and cloud platforms (AWS) for scalable data storage and analysis, ensuring efficient and low-latency data access for trading operations.

PROJECT #5:

Title	: TUI Travel Group
Brands	: Thomson Airways, Amadeus System, British Airways.
Client	: TUI UK, British Airways.
Location	: Hyderabad, India.
Duration	: March 2015 to December 2018
Tools Used	: Hadoop, Hive, Spark, Sqoop, Cloudera, Tableau
Team Size	: 10
Role	: Big Data Engineer.

Description:

TUI Group is a multinational travel and tourism company. It is the largest leisure, travel and tourism company in the world, and owns travel agencies, hotels, airlines, cruise ships and retail stores. TUI Travel Group have several UK-based tour operators Thomson Holidays and First Choice, by using this brands Passengers able to Book Flight tickets, Hotels and Holiday Packages with Cheapest Prices to Major European Countries

Responsibilities:

- Designing and implementing semi-structured data analytic platform leveraging Hadoop.
- Extensively used Hive/HQL or Hive queries to query or search for a particular string in Hive tables in HDFS.
- Worked on analysing Hadoop stack and different big data analytic tools including Pig, Hive, and Sqoop.
- Implemented Spark using Scala and Spark SQL for faster processing of data.
- Exported the analysed data to the relational databases using Tableau for visualization and to generate reports for the BI team.