

Viren Sasalu

+1(520)524-8272 | virensasalu@gmail.com | [LinkedIn](#) | [GitHub](#)

SUMMARY

Results-driven data professional with 5+ years of progressive experience as a Data Analyst and Data Engineer, delivering impact across analytics, pipeline optimization, and real-time data systems. Proven track record of driving business value through end-to-end data solutions, forecasting models, and cross-functional initiatives in fast-paced, data-centric environments. Skilled in enhancing site work and functionality to ensure optimal usability and web experience, while adhering to data governance standards, including the handling of personally identifiable information.

EXPERIENCE

Data Engineer III, Kanira Consulting Inc, United States **Jun 2025 — Present**

- Designed, built, and optimized **large-scale distributed ETL pipelines** using **Python, SQL**, and advanced data validation techniques, reducing overall data processing time by 40% across critical reporting workflows and analytics modules, in alignment with agile methodologies.
- Ingested structured and unstructured data into **Amazon S3** using scalable ingestion patterns via **AWS Glue** and **Lambda**, enhancing data accessibility and accelerating analytics for 3+ business units, ensuring high **availability** and fault tolerance.
- Developed automated and reusable **data validation** scripts using advanced **Python** and SQL triggers, improving data accuracy by 25% and significantly reducing manual correction efforts across teams, supporting cross-functional **initiatives**.
- Transformed complex operational data into analysis-ready formats for **Power BI** dashboards, improving refresh reliability, stakeholder trust, and reducing downtime by 60%, aligning with **Domain Driven Design** principles.
- Orchestrated daily incremental loads using **AWS Glue Workflows** and **Step Functions** with built-in monitoring via **CloudWatch**, ensuring 98% SLA adherence for timely availability of production data, supporting **solution-oriented** data governance.

Data Engineer, Arukas Tech **Jan 2021 — Jul 2023**

- Engineered scalable ETL pipelines using **Python** and **Apache Airflow**, automating ingestion of complex HR datasets and cutting manual data handling by 60%, boosting payroll system reliability through agile **projects**.
- Managed implementation of enterprise grade **access control and API security protocols**, ensuring role based permissions, preventing unauthorized access, and achieving **zero security breaches** across mission critical systems
- Designed modular, scalable, and reusable data models with **DBT in Snowflake**, enabling faster and more consistent workforce performance reporting and reducing turnaround time for internal analytics by 35%, supporting **application data** enhancements.
- Built real-time streaming pipelines with **Apache Kafka** and **PySpark**, delivering up-to-the-minute metrics on employee engagement and attendance for time-sensitive HR operations, aligning with **API first** principles.
- Deployed robust predictive analytics infrastructure by integrating **Looker ML models and pipelines**, significantly improving accuracy of employee attrition and wellness forecasts by 25%, supporting **microservices** architecture.
- Orchestrated secure, high-throughput, and fault-tolerant data flows using **AWS** and **Databricks frameworks**, accelerating financial and clinical data deployments for enterprise-level analytics teams, ensuring **business domain** alignment.
- Boosted data quality and compliance by automating validation checks with **AWS Glue** and **Databricks**, ensuring 35% improvement in integrity for HIPAA and SOX-regulated data.

Senior Data Analyst, WNS Global Services **Feb 2019 — Jan 2021**

- Led and Managed** a team of analysts in creating impactful, data-driven visualizations and business reports, identifying KPIs and diagnosing performance gaps, resulting in a **25% improvement** in reporting efficiency and a **15% faster** decision-making process
- Directed enterprise-wide **data governance** initiatives, establishing policies, standards, and **stewardship programs** that ensured **100% regulatory compliance**, reduced **data inconsistencies by 30%**, and strengthened executive confidence in strategic analytics
- Streamlined reporting processes by developing automated dashboards in **Power BI**, reducing manual reporting time by 40% and enabling real-time decision-making for sales and operations teams, supporting **cross-functional initiatives**.
- Automated repetitive data tasks using **Python scripts**, improving workflow efficiency and saving over 10 hours per week in manual data manipulation and validation efforts.

Data Analyst, Cutso Foods LLP **Dec 2017 — Feb 2019**

- Developed and optimized ETL pipelines using **Python, SQL**, and **AWS Glue**, improving transaction data processing speed by 25% and ensuring data integrity in high-volume banking operations, aligning with agile **methodology**.
- Designed interactive **Tableau dashboards** powered by advanced time-series forecasting models to enhance credit risk analysis workflows, resulting in 20% more accurate and reliable trend predictions.
- Collaborated with cross-functional stakeholders to define KPIs and presented findings through **PowerPoint** storytelling and interactive data visualizations, supporting strategic planning sessions, aligning with **organization** goals.
- Automated dynamic **SQL-driven Power BI** reports for Medicaid and insurance claim tracking, minimizing manual effort on 30+ recurring reports and substantially improving end-to-end reporting efficiency, supporting **business domain** enhancements.

CERTIFICATIONS

AWS Certified Data Engineer - Associate (DEA-C01)

Google Data Analytics Professional Certificate

Tableau Business Intelligence Analyst Professional Certificate

Google AI Essentials

PROJECTS

Text Sentiment - Emotion detection using Tensorflow | Won First Place, [Link](#)

- Developed a sentiment-emotion detection model employing TensorFlow, RoBERTa, CNN, and RNN, enhancing multi-label text classification accuracy by 43% through custom focal loss and hyperparameter optimization, yielding a 0.85 F1 score.

FashionFlux (2025) – Stable Diffusion-based Virtual Try-On App | Won First Place, [Link](#)

- Developed **FashionFlux**, a real-time virtual try-on pipeline that fuses **Segment Anything, OpenPose**, and **ControlNet-guided Stable Diffusion** to render photorealistic garments on user photos.

SKILLS

Analytics & Visualization	: Power BI, Tableau, Looker, Excel, SAS
Data Management	: SQL, MongoDB, Snowflake, Redshift, Data Warehousing, Data Mining
ETL & Big Data Technologies	: ETL Pipelines, Apache Airflow, Informatica, Hadoop, Apache Spark, MapReduce, Google Dataflow.
Cloud & Infrastructure	: AWS (S3, Lambda, Glue, Redshift, Athena, CloudFormation), Azure Bicep, Kubernetes, Terraform
Machine Learning & NLP	: Scikit-Learn, TensorFlow, PyTorch, Keras, RNN, CNN, Transformers, SpaCy, NLTK, Hugging Face, GloVe, CUDA
Programming & DevOps Tools	: Python, Scala, Git, GitHub, GitHub Actions
Relevant Skills	: Data Visualization, Data Warehousing, Statistical Analysis, Data Management, Data Governance

EDUCATION

University of Arizona, Arizona, United States - Master of Science, Information Systems, Machine Learning

Aug 2023 — May 2025

Visvesvaraya Technological University, India - Bachelor of Engineering, Electronics and Communications