Akash Yanpure

London, UK | +44 7586 487162 | Akash.yanpure@gmail.com

PROFESSIONAL SUMMARY

Data Engineer with 7+ years industry experience designing and operating scalable, secure data platforms across real-time, near-real-time, and batch systems. Skilled in Python, SQL, Apache Spark, Airflow, and dbt, and in automating cloud data platforms with Docker, Kubernetes, CI/CD, and Terraform. Delivers well-governed, high-quality data services across data platform architecture, governance, and security for analytics, machine learning, and GenAI.

TECHNICAL SKILLS

Programming: Python, SQL, Node.js, JavaScript, jQuery, AJAX

Data & Machine Learning: Spark, PySpark, Pandas, NumPy, TensorFlow, ETL / ELT, Data Lake, GenAI / LLMs

Cloud: AWS (Certified), Azure

Data Platforms & Warehouses: Snowflake (Certified), Redshift, PostgreSQL

Workflow / ETL: dbt, Airflow, Jenkins

DevOps: Docker, Kubernetes, Terraform

BI & Dashboards: Tableau, Power BI, Excel

Project Management & CI/CD: Git / GitLab / Azure DevOps, JIRA, Confluence

UI / UX: HTML / CSS / Bootstrap, Figma, Adobe Photoshop

PROFESSIONAL EXPERIENCE

Data Engineer

Norstella | London, UK | Nov 2022 – Present

- Designed and operated scalable, secure real-time, near-real-time, and batch data architectures using Apache Airflow, AWS Glue (Spark), and dbt, powering advanced analytics, machine learning, and GenAI use cases across teams.
- Built automated data integration and transformation frameworks to ingest data from multiple sources across raw, trusted, and curated layers on Snowflake, Redshift, and PostgreSQL, maintaining governance, lineage, and data quality.
- Implemented containerised, parallelised ETL/ELT pipelines on Kubernetes (Amazon EKS) with Docker, automating deployment and configuration via CI/CD and Terraform to improve performance, scalability, and reliability.
- Enabled self-service reporting by transforming raw data into curated, analytics-ready datasets with dbt, improving data discoverability and accessibility for end users.
- Led data governance and security work — data cataloguing, access controls, and lineage tracking — to ensure compliance and traceability across the organisation.
- Mentored junior engineers through peer reviews and high-quality Python code, and evaluated GenAI/LLM tools to speed up analytics tasks and boost team productivity.

Data Scientist

NayaOne | London, UK | Apr 2021 – Nov 2022

- Developed and deployed an NLP chatbot using Python (Flask) and a GPT-3 large language model on AWS.
- Built machine-learning classification and anomaly-detection models to flag outliers in lending applications and detect loan defaults.
- Engineered SQL pipelines and triggers to power analytics dashboards in Metabase and Power BI.
- Deployed machine-learning model APIs on AWS ECS using Docker.
- Partnered with fintech clients to gather and communicate business requirements.

Software Developer

FinIQ Consulting Pvt. Ltd. | Pune, India | Jan 2019 – Aug 2020

- Developed Node.js applications and custom REST APIs, optimising for high-traffic performance.
- Built responsive user interfaces with HTML, Bootstrap, React, and jQuery.
- Created and deployed C# backend services.
- Developed a financial order equity-basket prediction system in Python.

Game Tester

Ubisoft | Pune, India | Jun 2018 – Jan 2019

- Executed structured test plans across the full game development lifecycle — alpha, beta, and release — logging and tracking defects in JIRA within Agile sprints.
- Analysed gameplay and defect data to surface recurring issues and quality trends, giving developers clear, data-backed feedback to prioritise fixes.

EDUCATION

MSc in Data Science

Kingston University, London | Sep 2021

B.E. Computer Engineering

D Y Patil Institute of Engineering, Pune, India | May 2018

CERTIFICATIONS

AWS | Snowflake | Python | Power BI | Machine Learning | Web Scraping