John Carter

Data Engineer

📍 Colombo, Sri Lanka

📧 john.carter.de@gmail.com

📞 +94 77 123 4567

🔗 linkedin.com/in/johncarter-de

🔗 github.com/johncarter-data

Professional Summary

Data Engineer with 4+ years of experience designing, building, and maintaining scalable data pipelines and analytics platforms. Strong background in ETL/ELT, cloud data warehouses, and big data processing. Experienced in transforming raw data into reliable, analytics-ready datasets to support business intelligence and machine learning use cases.

Technical Skills

Programming:

- Python, SQL, Bash

Data Engineering:

- ETL/ELT Pipelines

- Data Modeling (Star & Snowflake Schemas)

- Batch & Streaming Processing

Big Data & Streaming:

- Apache Spark, PySpark

- Apache Kafka

Databases & Warehouses:

- PostgreSQL, MySQL

- Snowflake, BigQuery

Cloud & DevOps:

- AWS (S3, Glue, Lambda, Redshift)

- Docker, Git, CI/CD

Orchestration & Tools:

- Apache Airflow

- dbt

- Great Expectations

Work Experience

Data Engineer

ABC Analytics Pvt Ltd — Colombo, Sri Lanka
March 2022 – Present

- Designed and maintained scalable ETL pipelines using **Python, Airflow, and AWS Glue**

- Built data ingestion pipelines processing **5M+ records daily** from APIs and databases

- Optimized SQL queries and data models, reducing dashboard load times by **40%**

- Implemented data quality checks using **Great Expectations**

- Collaborated with data scientists to provide clean datasets for ML models

- Migrated on-prem data warehouse to **Snowflake**, improving scalability and cost efficiency

---

Junior Data Engineer

DataWorks Solutions — Colombo, Sri Lanka
July 2020 – February 2022

- Developed batch data pipelines using **PySpark and PostgreSQL**

- Assisted in maintaining Kafka-based streaming pipelines

- Wrote complex SQL transformations for analytics reporting

- Automated daily data validation and monitoring scripts

- Supported BI teams with reliable and well-documented datasets

---

Projects

Real-Time E-Commerce Data Pipeline

- Built a Kafka → Spark → Snowflake pipeline

- Processed live order and customer events

- Enabled near real-time sales analytics dashboards

Data Warehouse for Job Recommendation System

- Designed star-schema warehouse for job and user interaction data

- Built ETL pipelines using Airflow and dbt

- Improved recommendation query performance by **35%**

Education

BSc (Hons) in Software Engineering
University of Colombo
2016 – 2020

Certifications

- AWS Certified Data Analytics – Specialty

- Apache Spark Developer Certification

Soft Skills

- Strong problem-solving mindset

- Excellent communication with cross-functional teams

- Documentation & data governance focused