# Kieran Blackwood
*Data Engineer*

Phone: **(555) 847-2963**
Address: **Seattle, WA**
Website: **https://github.com/kblackwood**
Email: **kieran.blackwood@email.com**

- Experienced Data Platform Engineer with 5+ years building scalable data infrastructure processing 50TB+ daily across cloud and hybrid environments
- Specialized in real-time streaming architectures using Kafka and Spark, delivering sub-second analytics for business-critical applications
- Proven track record of migrating legacy systems to modern data platforms, reducing costs by 40% while improving performance 3x

## WORK EXPERIENCE

### CloudTech Solutions
March 2022 - Present

Senior Data Engineer

- Architected multi-tenant data platform on AWS processing 75TB daily data from 200+ sources, supporting 500+ concurrent users with 99.95% uptime
- Designed real-time fraud detection pipeline using Kafka Streams and Apache Flink, reducing detection time from 24 hours to 30 seconds
- Optimized Snowflake data warehouse performance through intelligent clustering and materialized views, cutting query costs by 45% while maintaining sub-second response times
- Led migration of 15 legacy ETL jobs to modern ELT architecture using dbt and Airflow, reducing maintenance overhead by 60%
- Implemented data quality framework with Great Expectations, achieving 99.8% data accuracy across critical business metrics

### FinanceFlow Inc
June 2020 - February 2022

Data Engineer

- Built end-to-end analytics pipeline processing 2M financial transactions daily using Apache Spark and Delta Lake, enabling real-time regulatory reporting
- Migrated on-premise Teradata warehouse to AWS Redshift, reducing infrastructure costs by $300K annually while improving query performance by 4x
- Developed automated data reconciliation system using Python and Pandas, reducing manual validation time from 8 hours to 15 minutes daily
- Created self-service analytics platform using Apache Superset, empowering 50+ business users to generate reports independently
- Established CI/CD pipelines for data workflows using GitLab and Terraform, reducing deployment time from 4 hours to 20 minutes

### RetailMetrics Corp
August 2019 - May 2020

Junior Data Engineer

- Developed ETL pipelines using Python and Apache Airflow to process e-commerce data from 25+ sources, supporting $50M annual revenue analytics
- Optimized PostgreSQL database queries reducing average report generation time from 45 minutes to 5 minutes
- Built automated data monitoring system using Prometheus and Grafana, reducing data incident response time by 70%
- Collaborated with data science team to productionize ML models, implementing feature stores serving 100+ model predictions per second

## TECHNICAL SKILLS

**Programming Languages:** Python, Scala, SQL, Java, Bash

**Big Data Technologies:** Apache Spark, Hadoop, Kafka, Flink, Storm

**Cloud Platforms:** AWS (S3, EMR, Glue, Redshift, Lambda), GCP (BigQuery, Dataflow, Pub/Sub), Azure (Synapse, Data Factory)

**Databases:** PostgreSQL, MySQL, MongoDB, Cassandra, Redis, Snowflake, BigQuery

**Data Orchestration:** Apache Airflow, Prefect, Luigi, dbt

**Data Modeling:** Dimensional Modeling, Data Vault 2.0, Star Schema, Kimball Methodology

**DevOps & Tools:** Docker, Kubernetes, Git, CI/CD, Terraform, Jenkins, Prometheus

## EDUCATION

### University of Washington
May 2019

Bachelor of Science in Computer Science | GPA: 3.8/4.0

**Relevant Coursework:** Database Systems, Distributed Computing, Big Data Analytics, Machine Learning, Data Structures & Algorithms, Statistics
**Capstone Project:** Built real-time recommendation engine processing 1M user interactions daily using Apache Kafka and collaborative filtering algorithms

## CERTIFICATIONS

### AWS Certified Data Analytics - Specialty
March 2023

Amazon Web Services

### Google Cloud Professional Data Engineer
January 2023

Google Cloud

### Databricks Certified Associate Developer for Apache Spark
November 2022

Databricks

### Snowflake SnowPro Core Certification
September 2022

Snowflake

## HONORS

### 1st Place, AWS re:Invent Hackathon
December 2023

Amazon Web Services

Designed serverless data lake solution processing IoT sensor data with 99.9% accuracy. Implemented cost-optimization algorithm reducing storage costs by 55% using intelligent tiering

### CloudTech Innovation Award
September 2023

CloudTech Solutions

Recognized for developing automated data lineage tracking system used across 15+ data teams

## PUBLICATIONS

**Optimizing Spark Performance for Large-Scale ETL Workloads**

Data Engineering Weekly

Technical article demonstrating 60% performance improvement through advanced partitioning strategies. Featured solution now implemented by 500+ data engineers based on community feedback

## OPEN SOURCE CONTRIBUTIONS

### Apache Airflow

2022-2023

Contributor

Contributed 3 bug fixes and performance improvements

### dbt-utils

Contributor

Developed custom macro for data quality validation, adopted by 200+ projects