

Cover Letter - Senior Data Engineer Position

Dear Hiring Manager,

I am writing to express my strong interest in the Senior Data Engineer position. With 6+ years of experience architecting enterprise-scale data platforms at Pharmaceutical Digital Manufacturing, I have a proven track record of building robust, scalable data infrastructure that enables business-critical decision making. I am excited to bring my expertise in real-time data pipelines, graph databases, and cloud-native architectures to your team.

Building Enterprise Data Platforms at Scale

At Pharmaceutical, I have designed and implemented data systems processing 500K+ daily transactions across pharmaceutical manufacturing operations:

Real-Time Data Platform: I architected a real-time data refresh system integrating 8 source systems (SAP, LIMS, Snowflake, Oracle, SQL Server) into a unified PostgreSQL data warehouse. This platform reduced data latency from 24 hours to <5 minutes, enabling near-real-time manufacturing visibility for 200+ stakeholders. The system processes 100K+ records daily with automated error handling, concurrency control, and comprehensive monitoring.

Enterprise Knowledge Graph: I built Pharmaceutical's first pharmaceutical manufacturing knowledge graph on Neo4j, managing 1M+ nodes and 5M+ relationships. This involved developing 50+ Airflow DAGs orchestrating complex ETL workflows, creating graph transformation pipelines using Cypher, and implementing data lineage tracking across 100+ tables. The platform enables semantic queries and graph-based analytics impossible with traditional relational databases.

Data Quality Excellence: I developed automated data quality frameworks with 25+ validation checks across critical datasets, reducing data issues by 40% and achieving 99.5% data quality scores. My alert system provides real-time HTML-formatted monitoring reports and immediate error notifications, improving issue detection time by 3 hours.

Technical Excellence & Best Practices

My approach to data engineering emphasizes reliability, performance, and maintainability:

Cloud-Native Architecture: I led the migration from on-premises infrastructure to AWS cloud, reducing operational costs by 30%. I containerized 20+ data applications using Docker, deployed to AWS ECS with automated CI/CD pipelines, and implemented infrastructure as code using Terraform.

Performance Optimization: I consistently deliver high-performance solutions, including 60% query performance improvement through indexing strategies and SQL optimization, and <5 minute data latency in real-time pipelines processing 100K+ records.

Orchestration & Monitoring: I have deep expertise in Apache Airflow, managing 50+ production DAGs with complex dependencies, dynamic task generation, and automated error recovery. My ETL logging framework tracks 200+ daily jobs with detailed performance metrics.

Data Governance: I established data lineage tracking, created data cataloging systems, implemented data reconciliation processes, and built comprehensive data profiling tools ensuring compliance and data quality.

Technical Stack & Expertise

Databases: PostgreSQL, Snowflake, Neo4j, Oracle, SQL Server, Redshift, pgvector

Programming: Python (Expert), SQL (Expert), Cypher, Bash, PySpark **ETL & Orchestration:** Apache Airflow, dbt, Custom Python frameworks, Pandas, SQLAlchemy **Cloud & DevOps:** AWS (S3, ECS, RDS, Lambda), Docker, Kubernetes, Terraform, Jenkins **Data Quality:** Great Expectations, Custom validation frameworks, Automated testing

Measurable Business Impact

My work has delivered significant value to Pharmaceutical:

- **Reduced data latency** from 24 hours to <5 minutes for critical metrics
- **Decreased cloud costs** by 30% through architectural improvements
- **Achieved 99.5% data quality** across production datasets
- **Improved query performance** by 60% through optimization
- **Automated 80+ manual processes**, saving 100+ hours monthly
- **Reduced debugging time** by 50% through graph-based lineage

Leadership & Mentorship

Beyond technical contributions, I have mentored 5+ junior data engineers, leading to 2 promotions. I created educational frameworks, conducted code reviews, and established best practices that elevated the entire team's capabilities.

Why?

What differentiates me is my ability to bridge complex technical implementation with business value. I don't just build pipelines—I architect data platforms that enable organizations to make faster, better decisions. My experience spans the full data engineering lifecycle: requirements gathering, architecture design, implementation, testing, deployment, monitoring, and continuous optimization.

Next Steps

I would welcome the opportunity to discuss how my experience in enterprise data platforms, real-time ETL, and graph databases can contribute to the company's data infrastructure. My GitHub repositories demonstrate my technical capabilities and commitment to engineering excellence.

Thank you for considering my application. I look forward to the possibility of joining your team and contributing to the company's data engineering success.

Sincerely,

Serkan Coskun