# Ravin G

**Senior Data Engineer | Data Architect | AI & Big Data Engineer**

ravin34.cloud@gmail.com | (201)685-4245 | Jersey City, NJ, USA

## Professional Summary

Accomplished Senior Data Engineer with 7+ years of experience in designing and building scalable, high-performance data solutions. Proven expertise in Data Architecture, Big Data engineering, and AI-driven data platforms. Adept at transforming complex data systems into actionable insights through cloud-native architectures, real-time processing, and machine learning. Strong background in delivering business value through innovative solutions, enhancing data security, and ensuring compliance (GDPR, HIPAA, SOC2). Skilled in leading cross-functional teams, optimizing costs, and enabling predictive and prescriptive analytics.

## Core Skills

- **Data Engineering & Architecture:** ETL/ELT with Airflow, dbt, Talend, CDC; large-scale data pipeline automation; dimensional modeling (Star, Snowflake, Data Vault).

- **Big Data & Cloud Platforms:** Azure (Synapse, Data Factory, Databricks, Data Lake), AWS (Redshift, Glue, Lambda, EMR, Kinesis), GCP (BigQuery, Dataflow, Vertex AI).

- **AI & Machine Learning:** ML model deployment, automated pipelines, AI-based anomaly detection, RAG (Retrieval-Augmented Generation), vector search (FAISS, Pinecone).

- **Analytics & BI:** Power BI, Tableau, Looker; self-service dashboards, KPI reporting, advanced analytics.

- **Programming & Automation:** Python, SQL, Scala, R, Java, Bash; Terraform; CI/CD with Jenkins, GitHub Actions, Azure DevOps.

- **Data Security & Compliance:** IAM, encryption, tokenization, regulatory compliance (GDPR, HIPAA, SOC2).

- **Cloud-Native Infrastructure:** Kubernetes, Docker, Helm, serverless architecture; observability with Prometheus, Grafana, Datadog, Monte Carlo.

- **Business-Focused Delivery:** Translation of technical solutions into business value, enhancing ROI and compliance for US clients.

## Professional Experience

**Senior Data Engineer | DataNova Solutions**                    July 2023 – Present

- Designed and deployed real-time streaming data pipelines with Kafka and Flink, enabling predictive analytics and enhancing decision-making for US-based healthcare and retail clients.

- Led migration efforts from on-premise to cloud-native platforms (Azure, AWS), resulting in a 40% reduction in operational costs and a 30% improvement in scalability.

- Developed end-to-end data governance frameworks ensuring compliance with HIPAA and SOC2, significantly improving security and regulatory adherence.

- Enhanced ETL/ELT pipeline performance using dbt and Apache Airflow, resulting in a 25% increase in processing speed and reliability.

- Collaborated with business and technical stakeholders to implement secure, cost-effective data solutions, driving a 15% reduction in cloud storage costs.

- Spearheaded the development of a data pipeline for processing over 10TB of data daily, supporting real-time analytics for a global e-commerce client, reducing data latency by 50%.

- Automated data quality checks and validation using Python and SQL, reducing manual data correction efforts by 75% and ensuring high accuracy in downstream systems.

- Developed a secure data-sharing framework using end-to-end encryption and tokenization, ensuring compliance with data privacy regulations (GDPR, HIPAA).

## Data Engineer | Illumina Technology Solutions          July 2020 – June 2023

- Led the modernization of data lakes using Delta Lake and Apache Iceberg, improving data storage efficiency by 50% and query performance by 35%.

- Designed and implemented OLAP solutions with Snowflake and ClickHouse to enable near real-time data analysis, improving reporting speed by over 60%.

- Developed and deployed self-service BI dashboards using Power BI and Looker, empowering business users to generate insights without relying on IT.

- Integrated ML-based anomaly detection into data pipelines, improving data quality and reducing data errors by 40%.

- Led the transition of legacy systems to cloud-native environments (Azure, AWS), enhancing agility and reducing operational costs.

- Integrated Apache Kafka and AWS Kinesis for real-time data processing, enabling faster analytics and reducing data retrieval time by 60

- Created and optimized automated data validation frameworks, significantly improving data consistency and quality, which led to a 30% reduction in data errors.

## Associate Data Engineer | Transcure          April 2018 – June 2020

- Developed automated ETL workflows using Airflow and dbt, improving data processing efficiency and reporting speed.

- Designed KPI dashboards for finance and healthcare clients, enabling actionable insights for better decision-making.

- Enhanced query performance by 30% through indexing, partitioning, and caching, optimizing data retrieval.

- Contributed to data governance framework design, ensuring regulatory compliance and audit readiness.

- Created alerting and monitoring tools with Prometheus and Grafana, reducing pipeline down-time by 20%.

- Built custom API services for seamless data integration between external systems and internal databases.

# Education

**B.Sc. in Computer Science**