Richie Rivera

Distinguished Data Science Engineer

551-404-8805
New York, NY
Richie@RiveraDC.org
riverar9.github.io

Product-focused data engineer with a background in data science and 6+ years of experience building and scaling the data infrastructure that powers ML models and enterprise-wide business intelligence. Proven ability to "ship" and maintain reliable data products, from architecting scalable data marts (50+ GB/day) on GCP to owning data observability and transformation frameworks. Passionate about bridging the gap between data science and engineering to build the platforms that accelerate product innovation.

PROFESSIONAL EXPERIENCE

Verizon | New York, NY

Distinguished Data Science Engineer (Aug 2025 - Current)

- Defining the strategic technical roadmap for the next generation of marketing data products, focusing on scalability and multi-source data integration.

- Leading the end-to-end migration of all core marketing data (media, KPI, sales) from on-prem systems to a unified GCP architecture.

- Expanding the use cases and technical capabilities of the MDS data mart, enabling new cross-functional analytics and ML-driven insights.

Principal Data Science Engineer (Oct 2021 – Aug 2025)

- Architected and "shipped" the marketing-wide MDS data mart, a scalable data product processing 50+ GB daily that standardizes data access for 100+ users across marketing, finance, and data science

- Owned data observability for all marketing data pipelines; implemented robust data quality, freshness, null, and augment checks, enabling business KPI reporting in minutes instead of days.

- Reduced analytical dataset build time from 8 weeks to 3 weeks by engineering a standardized, automated data validation and transformation framework in Python and SQL.

- Engineered a CI/CD-based orchestration system using GitLab and custom cron jobs to manage all ELT/ETL tasks, reducing media data delivery timelines from 6 weeks to 2 days.

- Accelerated ML workflows by designing and maintaining standardized training datasets, directly supporting data science modeling and A/B test analysis for the CMO department.

Marketing Effectiveness Consultant (Jan 2020 – Oct 2021)

- Translated stakeholder requirements from business units, external agencies, and internal teams into technical specifications for data engineering and validation.

- Conducted and designed A/B and Causal Impact experiments to determine incremental sales and traffic changes, providing benchmarks used for enterprise-level decisions.

- Led data validation meetings to clarify and contextualize media advertising data anomalies, improving clarity and trust in insights delivered to leadership.

American Airlines | Dallas, TX

Predictive Maintenance Engineer (May 2019 – Jan 2020)

- Engineered feature-extraction pipelines from raw aircraft sensor data, creating the standardized datasets for predictive maintenance models, reducing unexpected delays.

- Developed a scalable process to identify and action predictive maintenance findings, "shipping" a data-driven workflow that ensured clear communication and swift action.

- Liaised between maintenance operations and engineering teams to productionize predictive models, resulting in the development of the core Predictive Maintenance system.

Material Planning Analyst (Sep 2017 – May 2019)

- Engineered a comprehensive material forecast data model to identify supply chain risk, enabling proactive analytics and improving service levels.

- Shipped a cross-compatibility data system to document and certify parts for use across multiple aircraft, improving maintenance efficiency and fleet serviceability.

- Managed data integrity for 600,000+ SKUs supporting the entire American Airlines fleet.

EDUCATION

City University of New York | New York, NY

Master of Science, Data Science (Expected May 2026)

Embry-Riddle Aeronautical University | Daytona Beach, FL

Bachelor of Science, Aerospace Engineering (May 2017)

SKILLS

- **Data Engineering:** Data Modeling, Data Warehousing, ETL/ELT Pipelines, Data Observability, CI/CD, Docker

- **Languages & Databases:** SQL, Python (pandas, PySpark) , GCP (BigQuery) , Snowflake, Teradata

- **Pipeline & Orchestration:** GitLab, git, Cron, VM-based orchestration, Python/SQL transformation scripts

- **Data Science & ML:** Scikit-learn, fastai , A/B Testing , Causal Impact , Predictive & Prescriptive Maintenance, Feature Engineering, Time Series Analysis

- **Visualization & BI:** Tableau, Power BI, Looker, Matplotlib

CERTIFICATIONS

- **Alteryx Designer Core Certification (2022)**
- **Alteryx Foundational (2022)**
- **IBM Data Science Professional (2021)**
- **Snowflake Data Warehouse (2020)**