

Tiffany Wang

Objective: Senior Data Engineer - Thrive Market

Focus Areas: Python⁺ | AWS Big Data⁺ | Distributed Systems | ETL Optimization

Phone: (626)-223-6123 | Location: Redondo Beach, CA 90277

LinkedIn: [linkedin.com/in/tiffanywangengineer](https://www.linkedin.com/in/tiffanywangengineer) | Email: tiffany.wang.engineer@gmail.com

SUMMARY

Big Data Architect with 8+ years of expertise in **petabyte-scale data processing**, **AWS/GCP cloud-native solutions**, and mission-critical ETL pipelines. Reduced data warehouse latency by 83% through Spark⁺ optimizations and achieved 99.99% SLA compliance for 1.6M+ user e-commerce platforms. Certified in Python/SQL with proven success in building enterprise-grade data lakes.

SKILLS

Core Technologies: Python⁺, Spark⁺ (PySpark/SQL), AWS⁺ (Redshift/Glue/Kinesis), Kafka⁺, Airflow⁺

Data Platforms: Snowflake⁺, BigQuery⁺, Hadoop⁺ (HDFS/YARN), Databricks⁺

DevOps: Terraform⁺, Jenkins⁺, Docker⁺, Kubernetes⁺, CI/CD Pipelines

Modeling: Medallion Architecture, Slowly Changing Dimensions (SCD), GDPR/HIPAA Compliance

EXPERIENCE

Senior Data Engineer

Capital Group | Irvine, CA | 2019–2023

Key Achievements:

- **AWS Redshift⁺ Data Warehouse Modernization**

- Migrated 50TB+ legacy EDW to **Spark⁺-on-EMR** with Delta Lake, reducing query latency by 83% and cutting \$1.2M annual costs.
- **Technical Decision:** Implemented **Kafka⁺** over Kinesis for real-time order stream processing (5ms P99 latency vs 15ms).
- Built automated data quality gates using Great Expectations, resolving 95% of integrity issues pre-production.

- **Member Behavior Analytics Platform**

- Designed **Snowflake⁺** medallion architecture for 1.6M+ user profiles, enabling real-time personalization with <200ms API response.
- **Hook:** Developed Python⁺ UDFs for SCD Type 2 handling, reducing ETL complexity by 40%.

Cloud Data Architect

Aerospace Innovation Lab | Remote | 2023–Present

Key Projects:

- **Real-time Inventory Optimization System**

- Architected **Kafka⁺-Spark⁺** streaming pipeline processing 1M+ events/min, reducing stockout incidents by 65%.
- Implemented **Terraform⁺**-managed AWS infrastructure with 99.99% uptime SLA.

- **GDPR-compliant Data Lake**

- Built **Hadoop⁺**-based data lake with column-level encryption, achieving 100% audit compliance for EU user data.
- **Technical Hook:** Optimized Parquet file partitioning strategy, cutting storage costs by 30%.

EDUCATION

MS Software Engineering | Embry-Riddle Aeronautical University | 2015–2017

BS Computer Science | Nanjing University of Aeronautics | 2012–2016

Technical Impact: 83% latency reduction | \$1.2M cost savings | 99.99% SLA