Jacky Lam | Data Engineer

https://www.jackylam.io

Email : ny.jackylam@gmail.com
Mobile : 646-287-5229

PROFESSIONAL SUMMARY

Data Engineer with 8 years' experience building scalable data platforms, transforming analytics pipelines, and deploying modern cloud-first data solutions. Proficient in Python, SQL, and Spark across AWS and GCP. Proven ability to collaborate across Finance, Product, Marketing, and Ops to deliver actionable insights and data infrastructure. Experienced in leading CI/CD implementations, data modeling with dbt, and cost-optimized ETL development.

TECHNICAL SKILLS

- **Languages**: Python, SQL, R
- **Big Data Tools**: Databricks, Apache Spark, PySpark, Spark SQL
- **ETL/Orchestration**: dbt, Apache Airflow, Dagster, Airbyte, DMS, Fivetran, Panoply
- **Data Warehousing**: Redshift, BigQuery, Snowflake
- **BI/Visualization**: Looker, Tableau, Domo
- **Cloud Platforms**: AWS, GCP
- **DevOps/CI-CD**: Git, Github Actions, Poetry, Jira
- **Databases**: PostgreSQL, MySQL

EXPERIENCE

Nuts.com   Jersey City, NJ
Data Engineer (Promoted from Senior Data Analyst)   *January 2023 - Present*

- **ETL Pipeline Refactor & Cost Optimization**: Refactored and optimized the core ETL data pipeline ingesting all website sales data—rewrote transformations and re-architected batch jobs to reduce infrastructure costs by 63% and cut job run times by 50%.
- **Product Cost Metrics**: Drove a company-wide initiative to unlock visibility into product cost metrics by leading cross-functional collaboration between Finance, BI, and DevOps. Spearheaded project scoping, requirements gathering, and the end-to-end development of scalable data pipelines in Databricks and reporting infrastructure in Looker—enabling more accurate, actionable financial insights.
- **Databricks CI/CD Implementation & Job Reorganization**: Architected and implemented CI/CD pipelines for Databricks by migrating notebooks and scripts into Git-based version control and reorganizing jobs by functional domain—standardizing deployments.
- **Data Modeling & dbt ELT Pipeline Architecture**: Architected and modeled raw sales data into efficient fact and dimension tables using dbt—optimizing schema design to reduce query times, unlock deeper visibility for business users, and accelerate data-driven insights.
- **Looker Style Guide Development & Standardization**: Partnered with BI and Analytics stakeholders to develop and roll out a comprehensive Looker style guide—standardizing terminology, LookML structure, and design patterns to boost collaboration, reduce onboarding time, and ensure high-quality, consistent dashboards across the organization.

Senior Data Analyst   *June 2021 - January 2023*

- **Stabilized BI Infrastructure with Blue/Green Deployment**: Engineered and deployed a Blue/Green data deployment strategy to stabilize BI infrastructure, reducing query latency errors by over 80% and significantly improving dashboard reliability and user trust.
- **Launched Foundational Analytics for New Digital Gifting Product**: Partnered with Product, Gifting, and Engineering teams to architect and implement the foundational reporting infrastructure for "Gift Drop", a new product category—delivering end-to-end visibility into adoption, performance, and growth.
- **Established Scalable Metrics Framework for New Product Category**: Defined a unified, company-wide framework for the Gifting product category and collaborated with stakeholders to develop a comprehensive suite of performance metrics—establishing a consistent foundation for strategic analysis and long-term growth tracking.
- **Uncovered Key Customer Segments Through Journey Analytics**: Designed and executed customer journey analytics to identify key gifting segments, leveraging behavioral and transactional data. Enabled the commercial team to uncover loyalty gaps and develop data-driven strategies for targeted customer engagement and retention.

EXPERIENCE (CONTINUED)

Motivate LLC. — New York, NY
Senior Data Analyst — *July 2019 - June 2021*

- **Modernize Data Warehouse Architecture**: Architected and implemented a scalable data warehouse solution following Lyft's acquisition, aligning legacy data systems with Lyft's modern data infrastructure and ensuring seamless migration and integration.
- **Led Company-Wide BI Migration to Tableau**: Led enterprise-wide BI tool migration from Domo to Tableau, designing and deploying optimized, consolidated dashboards that reduced redundancy and enhanced reporting accuracy. Established and managed a dedicated internal Slack channel to promote collaboration and rapid feedback with field leadership.
- **Enabled Near Real-Time Field Operations with GIS Dashboards**: Developed geospatial heat map visualizations to support E-Bike Battery Swapping operations, providing operational leaders with real-time insights into field activity and improving logistical decision-making.
- **Defined KPIs and Dashboards Across 8 Cities**: Collaborated with field teams across 8 North American cities to define and track KPIs, ensure compliance with city and service agreements, and build operational dashboards that empowered senior leadership with data-driven insights.
- **Managed a Cross-Functional Data Team**: Managed and mentored a cross-functional data team of one Data Engineer and one Data Analyst. Facilitated daily stand-ups and sprint planning to ensure timely project delivery and alignment with organizational goals.

NYC Taxi & Limousine Commission — New York, NY
Senior Policy Analyst (Promoted from Policy Analyst) — *March 2017 - July 2019*
Policy Analyst — *March 2016 - March 2017*

- **Forecasted Ride Fare Dynamics Across Rideshare Providers and Taxis**: Developed a multivariate linear regression model to forecast ride fares across Taxi, Uber, Lyft, Via, and Juno. Conducted comparative rate analysis to support a landmark for-hire driver income policy, influencing regulatory decisions.
- **Launched High-Impact GIS Tool Viewed Over 2.3M Times**: Redesigned and digitized the Department of Transportation's For-Hire Vehicle Relief Stand database using GIS tools, producing an interactive, driver-friendly map viewed over 2.3 million times since launch—dramatically improving public accessibility.
- **Defined Regulatory Data Standards and Vendor Compliance**: Authored regulatory data specifications and led vendor compliance efforts. Partnered with IT to implement data ingestion pipelines, ensuring accurate and timely loading of for-hire vehicle datasets into production systems.

NYC Department of Parks & Recreation — Queens, NY
Street Tree Census Assistant Coordinator — *March 2015 - March 2016*

J. Trucking Inc. — Queens, NY
Operations Manager — *August 2014 - March 2015*

EDUCATION

Georgia Institute of Technology
Master of Science in Analytics

Queens College, City University of New York
Bachelor of Business Administration

MILITARY SERVICE

New York Army National Guard — New York, NY
Air Battle Management System Operator — *Mar 2013 - Mar 2016*