

Olivia Jackson Lambert

San Francisco, CA | oliviarosejackson@outlook.com | linkedin.com/in/olivia-jackson-lambert | github.com/olivia-jackson-lambert

Relevant Experience

Head of Data

Hockeystick Data (part of Liftoff Campaigns) | January 2026 – Present

- Managed 4 direct reports and 3 contractors, growing the team from 3 to 7 across data engineering, analytics engineering and data science functions. Mentored junior engineers in best practices, building a culture of ownership by empowering team members to question assumptions and investigate trade-offs and risks for projects.
- Collaborated with leaders across the organization to align on strategy and roadmaps.
- Tripled the team's monthly output by introducing AI workflows with Claude Code using Claude Tag, CLAUDE.md files and skills to accelerate work on repetitive tasks, bug fixes and simple feature requests. Spearheaded an initiative for an internal chatbot that interfaced with ChatGPT's API to answer FAQs and give tutorials on our internal tools, drawing from a markdown context repository.
- Reduced platform spend by \$60,000 a year and saved 50 hours per month spent on pipeline maintenance and workarounds by leading the migration from Civis data platform to Google Cloud and BigQuery, optimizing architecture for scale and 99.9% uptime.
- Reduced data tickets from ~300 to 5 per month by developing Next.js self-serve internal tooling and analytics products for other teams.
- Led the data science initiative for donor propensity modeling, overhauling donor segmentation and targeting which increased revenue from data acquisitions by 121%.
- Operationalized decision making and strategy for other teams by driving initiatives for analysis and dashboarding.
- Created an interactive app for the team to show the impact of PRs on the dbt model and in-app dependencies.

Senior Analytics Engineer

Democratic National Committee | August 2024 – December 2024 | Election Cycle Contract

- Deployed machine learning models serving 100K+ monthly users by building SQL models in DBT and orchestrating DAGs in Airflow
- Executed gradient-boosted XGBoost model on 160M+ voter records with 500+ engineered features to predict donor propensity scores, optimizing targeted direct mail fundraising campaigns
- Refactored data delivery pipeline by migrating to DBT and GCP, reducing API load by 99.999% and eliminating throttling issues
- Developed real-time election night dashboards in Looker with live PostgreSQL and BigQuery data sources, delivering precinct-level insights to senior stakeholders before major media outlets

Data Scientist – Consultant

Kenvue – Johnson & Johnson | September 2023 – March 2024

- Built global demand forecasting pipeline using Prophet and PySpark on Databricks, generating 10K+ monthly time series forecasts across 5 continents for SKU planning, inventory optimization, and financial budgeting
- Developed supervised learning classification module with sktime and scikit-learn to categorize time series, saving hundreds of hours of work a year
- Productionized ETL data pipeline with Redshift, Snowflake, and Streamlit, creating automated SQL-driven interactive dashboards for executives

Data Scientist – Consultant

Fidera Group | March 2023 – September 2023

- Architected automated ETL workflows with Python and Airflow to integrate datasets from Salesforce CRM and internal databases, processing 100K+ records daily and improving data accessibility for stakeholders
- Engineered data validation tests and introduced Git version control and Docker containerization for improved reliability

Technical Skills

Programming Languages: Python, R, SQL, C++, LaTeX

Data Science & ML: scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, Prophet, statsmodels, pandas, NumPy, SciPy

Data Engineering: Airflow, DBT, Spark, PySpark, ETL Pipelines, Docker, Git

Cloud & Databases: AWS (CCP Certified), GCP, Snowflake, Redshift, PostgreSQL, BigQuery

Visualization & BI: Next.js, Streamlit, Looker, Matplotlib, Tableau

Education

Master of Information & Data Science

University of California, Berkeley | Expected 2027 | Part-time Professional Degree

Relevant Coursework: Machine Learning, Distributed Computing (PySpark), Deep Learning, Data Engineering, Statistical Modeling

Master of Physics, Astrophysics | First Class with Honours

University of Edinburgh | 2017 – 2022

President, Physics & Astronomy Society, Carnegie Scholarship Recipient, Edinburgh Award for Leadership & Development

Technical Projects & Publications

“The Need for Large Sample Numbers to Demonstrate that Martian Environments Are Lifeless” – Nature Astronomy (12/2024)

Receipt Entity Extraction with Transformer NLP

- Built multi-task transformer pipeline using Hugging Face and PyTorch to classify and extract fields from retail receipts
- Technologies: Transformers, Named Entity Recognition (NER), Multi-Task Learning, CUDA, Tokenizers

High Energy Particle CNN Classifier

- Developed convolutional neural network to predict particle types from detector tracks, achieving 93% accuracy
- Technologies: Deep Learning, Hyperparameter Optimization, TensorFlow, Keras, scikit-learn

Research Experience

Research Assistant – Computational Astrophysics

Institute for Astronomy | May 2020 – September 2020

- Extracted and analyzed 500+ FITS images from Sloan Digital Sky Survey using complex SQL queries on terabyte-scale database and implemented C++ pixel-z algorithm for supernova analysis
- Built automated data processing pipelines in Python with Astropy, SciPy, pandas, and Matplotlib for feature extraction and visualization, reducing processing time from days to hours through optimization

Research Assistant – Computer Vision & Machine Learning

National Physical Laboratory | May 2019 – September 2019

- Developed autonomous sky imaging software for distributed network of Raspberry Pi devices in remote locations, implementing adaptive exposure control algorithms for months of unattended operation
- Applied machine learning with OpenCV and scikit-learn to classify cloud cover patterns using texture feature extraction (GLCM, HOG descriptors), improving satellite re-entry prediction accuracy by 23%