

Eric

Data Analyst

Los Angeles, CA, United States | [LinkedIn](#) | [GitHub](#)

Data analyst with 4+ years of experience building predictive models, dashboards, and analytics pipelines. Delivered \$11M+ in annual savings at Hyundai through time-series analysis and operational KPI optimization.

EXPERIENCE

Handshake AI Fellow Sep 2025 — Dec 2025

Handshake

- Evaluated and validated LLM-generated outputs to ensure factual accuracy, logical consistency, and reduced bias, significantly improving the reliability of datasets for downstream analysis.
- Analyzed AI response patterns across diverse prompt scenarios to identify core error typologies, providing data-driven, actionable feedback to enhance overall model performance.
- Processed and structured complex unstructured data through rigorous annotation, transforming raw inputs into standardized formats ready for advanced analytics.

Research Data Scientist Aug 2025 — Dec 2025

University of Michigan - School of Information

Ann Arbor, Michigan, United States

- Conducted extensive data collection and preprocessing to enhance public datasets and partner organization data for research purposes.
- Designed and implemented predictive models and algorithms using advanced machine learning and AI techniques.
- Validated model performance through rigorous experimentation, ensuring alignment with research objectives.

Strategy & Operations Manager | Business Analytics Jan 2020 — Mar 2024

Hyundai Motor Group

San Jose, California, United States

Led data analytics initiatives across predictive maintenance, EV charging expansion, and operational optimization. Built time-series anomaly detection models that reduced unplanned downtime by 40% and saved \$10M+ annually. Analyzed 10K+ EV charging transactions to calculate station-level ROI metrics informing expansion of 5+ charging hubs. Designed Power BI dashboards tracking approval workflows and process KPIs, cutting lead times by 50% and annual costs by \$1.5M. Standardized performance data for 200+ startups and automated data consolidation pipelines, reducing manual work by 70%.

EDUCATION

University of Michigan Aug 2024 — Aug 2026

Master's Degree in Data Science

GPA 3.86

Hanyang University

Bachelor of Science in Mechanical Engineering

PROJECTS

E-commerce user session data analysis, Checkout A/B Test Analysis | E-Commerce User Session Data

Analyzed 50K+ e-commerce user sessions using Python and SciPy to evaluate checkout flow changes. Validated statistical significance and translated results into checkout optimization recommendations.

SaaS user activity tracking and visualization, Cohort Retention Dashboard | SaaS User Activity Data

Built SQL-based cohort analysis tracking 30/60/90-day retention across monthly user cohorts. Designed Tableau dashboard with retention heatmaps and churn-risk segmentation to identify drop-off points.

E-commerce data transformation and modeling, dbt Analytics Pipeline | E-Commerce Orders and Customer Data

Built end-to-end analytics pipeline in dbt and PostgreSQL, transforming raw order and customer data into 12 tested staging and mart models. Developed KPI-ready aggregate tables that accelerated Tableau reporting with consistent data marts.

NYC collision data analysis with geolocation, Rainy vs. Clear Day Accident Analysis – NYC Collision Data

Cleaned and integrated 196K+ NYC collision records with weather and geolocation data using Python Pandas. Visualized crash hotspots using Folium and identified weather-correlated risk patterns.

Integrating Multi-Source Movie Data for Rating Prediction and Recommendation

[Data Integration]

- Built a preprocessing pipeline integrating 'IMDb', 'OMDb', and 'MovieLens' movie datasets with Python, including missing value handling, feature standardization, and schema consolidation to create a unified dataset for NLP-based movie rating prediction and recommendation modeling.

[Modeling and Analysis]

- Developed supervised learning models (Logistic Regression, Random Forest, XGBoost, KNN) with cross-validation and

hyper-parameter tuning, achieving improved predictive performance through feature engineering and ensemble methods for movie rating prediction.

: movie rating prediction with improved performance (best model: XGBoost, ~69% accuracy / 0.67 F1)

[Result]

- Delivered a unified NLP pipeline that improved movie rating prediction accuracy and generated insights for content recommendation, highlighting the value of integrating multi-source data with machine learning.

Data-Driven Analysis of Nursing Home Funding Inequities

- Produced reproducible workflows enabling MEJI to refresh analyses with future cost report updates.

- Designed a data dictionary and merged financial records with census/NH demographic datasets at the ZIP-code level.

- Applied statistical testing, z-score-based outlier detection, and resampling to evaluate regional disparities.

Integrated Investment Data & Management Platform

CRM Data Integration Pipeline

- Built web-based dash board and Airtable data pipelines to consolidate CRM data and designed a cross-department approval workflow, reducing investment decision-making time by 50%.

Post-Investment Management System

- Architected and deployed a web-based monitoring platform with Power BI dashboards and Kanban boards, enhancing transparency for 20+ stakeholders and boosting efficiency by 30%.

Data-Driven Decision Enablement

- Standardized reporting and approval workflows, providing executives with real-time visibility into portfolio performance and risks to ensure faster, data-driven decisions.

CERTIFICATIONS

Claude Code in Action, Anthropic

Mar 2026

Requirement Analysis of Technology Foresight Process Design for Communication Systems in the Automotive Industry,

Nov 2020

Intermediate PostgreSQL, University of Michigan

Supervised Machine Learning: Regression and Classification, DeepLearning.AI

SKILLS

Analytical Skills · Python (Programming Language) · PostgreSQL · MySQL · Tableau · Jira · Microsoft Power BI · SQL · Git · R (Programming Language) · Seaborn · Pandas (Software) · Data Quality · Product Analysis · Statistical Data Analysis · Hypothesis Testing · Snowflake · Business Impact Analysis · Data Wrangling · Extract, Transform, Load (ETL) · Matplotlib · Business Strategy · Data Analysis · Machine Learning · NumPy · Feature Engineering · Data Integration · Time Series Analysis · Data Visualization · Data Pipelines · Information Visualization · BigQuery · dbt · GitHub · Docker · AWS S3 · AWS EC2 · Airtable · Excel · Datadog · SciPy · Cohort analysis · Retention analysis · Funnel analysis · Anomaly detection · EDA · Folium