Leon Shpaner

Data Scientist, Educator

lshpaner@ucla.edu | +1 (310) 721-7126 | www.leonshpaner.com
github.com/lshpaner | linkedin.com/in/lshpaner | Google Scholar

Summary

Data scientist and adjunct faculty with 15+ years of experience in healthcare, education, and finance. Focus on predictive modeling to support clinical decision making, improve patient outcomes, and reduce mortality. Maintain open source libraries and contribute to peer-reviewed work in critical care.

Education

M.S., Applied Data Science, **University of San Diego** (GPA: 4.0/4.0) — Aug 2022
Executive Certificate Programs, **Cornell University** — 2018 – 2025
B.A., Economics, **University of California, Los Angeles (UCLA)** — Sep 2007

Professional Experience

UCLA Health, ***Data Scientist*** — Apr 2023 – Present

- First author of two *American Society of Nephrology* Kidney Week abstracts on kidney failure risk prediction using internal EHR data, both achieving AUCs in the 80th percentile: one reproducing gold-standard risk equations, and a second extending prediction across CKD stages 1–4.
- Built the `kfre` Python library and `kfre` R package, implementing Tangri et al.'s kidney failure risk equations for broad use in predictive healthcare analytics.
- Co-developed Python libraries `model_tuner` (machine learning workflow optimization, hyperparameter tuning, and pipeline automation) and `equiboots` (bias and fairness auditing), leveraging bootstrapping techniques for equitable model evaluation across diverse groups, and served as a speaker for the EquiBoots tutorial at JupyterCon 2025 `[YouTube]`.
- Produced an XGBoost Regression model used in production to predict high-cost Medicare Advantage patients year-over-year; implemented system-wide, with a forthcoming paper detailing the project's impact and methodology.
- Developed an XGBoost model used in production to predict one-year mortality risk, facilitating timely transition to supportive (palliative) care.

Children's Hospital Los Angeles, ***Data Scientist*** — Mar 2022 – Oct 2022

- Developed an Extra Trees machine learning model in the pediatric ICU to predict post-terminal extubation mortality (AUC 0.88), resulting in a co-authored paper in *Pediatric Critical Care Medicine*.
- Maintained data quality and PHI compliance across a multi institution pediatric research collaborative.

Real Social Dynamics/Self Mastery Co., ***Sales Data Analyst*** — May 2014 – Dec 2023

- Automated sales reports in Excel using VBA and built SQL queries for revenue aggregation and KPI analysis.

The Los Angeles Film School, ***Financial Analyst*** — Sep 2013 – Oct 2020

- Integrated financial reporting into CPM software to streamline data flow.

- Automated reports with VBA macros, cutting 20 staff hours per month.

Teaching Experience

UCLA Extension, ***Instructor*** Aug 2023 – Present
- Instruct fundamentals of data science, statistics, and machine learning in Excel and Python.

University of Maryland Global, ***Adjunct Professor*** Apr 2023 – Present
- Teach undergraduate courses in data science and statistics with Python.

Purdue University Global, ***Adjunct Faculty*** Oct 2022 – Feb 2024
- Instructed database fundamentals, entity relationship diagrams, tables, and SQL in Access.

University of San Diego, ***Adjunct Professor*** Dec 2020 – Present
- Co-designed and instruct graduate statistics with Python course for the Applied AI program.
- Designed and instructed an Excel-based statistics course for the Data Analytics Certificate.

Cornell University, ***Course Facilitator*** Nov 2018 – Nov 2025
- Facilitated data analytics and statistics certificate courses.

Open Source Software

[1] **kfre** (Kidney Failure Risk Equation) [PyPI] [CRAN]
Python and R libraries implementing Tangri kidney failure risk equations.

[2] **model_tuner** (Workflow Optimization) [PyPI]
Helpers for ML pipeline optimization and tuning.

[3] **equiboots** (Fairness Auditing) [PyPI] [JupyterCon 2025]
Bias and fairness auditing with bootstrapping.

[4] **eda_toolkit** (Exploratory Data Analysis) [PyPI] [JupyterCon 2025]
Tools for streamlined exploratory data analysis.

[5] **model_metrics** (Performance Assessment) [PyPI]
Evaluation tools for classification and regression models.

Skills

Programming	Python, R, SQL, JavaScript, LaTeX, HTML, CSS
Libraries	Python (NumPy, Pandas, Scikit Learn, Statsmodels, SciPy, MLflow, AutoKeras, Aequitas, Flask, Dash Plotly), R (Tidyverse, Ggplot, Caret, e1071)
Tools and Platforms	Git, Azure, Databricks, BigQuery, Tableau, Looker Studio, Excel, Access
Languages	English (fluent), Russian (fluent), Spanish (basic)

Honors and Awards

JupyterCon 2025 | Official Speaker Credential, The Linux Foundation November 2025
- Earned the official speaker credential for delivering accepted presentations at JupyterCon 2025.

UCLA Health | Staff Appreciation & Recognition (STAR) Award June 2025

- Recognized for supporting UCLA's clinical analytics & operations, demonstrating strong dedication to collaborative care and organizational excellence.

Society of Critical Care Medicine | **Early-Stage Researcher Award (Co-author)** January 2024

- Co-author on the abstract *"Death One Hour After Terminal Extubation (DONATE): A Multi-site Cohort Study,"* selected for SCCM's 2024 Critical Care Congress Early-Stage Researcher Award and Star Research presentation.

Talks & Workshops

JupyterCon 2025 (The Linux Foundation) | **Conference Speaker** November 2025

- EquiBoots: Fairness Focused Model Evaluation [Recording] [Schedule]
- EDA Toolkit: Streamlining Exploratory Data Analysis in Jupyter [Recording] [Schedule]

Centers for Disease Control and Prevention (CDC)/NIOSH | **Data Science for Everyone Workshop** Jan 2022

- Designed and delivered a workshop on Python and R, focusing on data preparation, exploratory analysis, and supervised learning, including a practical case study on traffic accident data. [GitHub Repository] [Workshop Details]

University of California, Los Angeles (UCLA) | **Guest Lecturer – Economics Department** 2018 – 2023

- Guest lectured in ECON 104: Data Science for Economists (2023). [Recording]
- Instructed a summer workshop on Financial Modeling in Excel (2018). [Original] [Revamped]

Publications

[1] Shpaner, L., Petousis, P., Nicholas, S. B., Bui, A., Duru, O., Kornowske, L. M., . . . (2025). Improving Kidney Failure Risk Predictions for Clinical Trials Across CKD Stages 1–4: TH-PO1195. *Journal of the American Society of Nephrology*, 36 (10S), 10.1681.

[2] Zemke, A. M., Mayhand, K., Alicic, R. Z., Kornowske, L. M., Jones, C. R., Daratha, K. B., . . . Shpaner, L. (2025). Predicting Kidney Function Decline in American Indian and Alaska Native Populations with Diabetes: TH-PO1004. *Journal of the American Society of Nephrology*, 36 (10S), 10.1681.

[3] Tuttle, K. R., Kornowske, L. M., Zuckerman, J. E., Shpaner, L., Petousis, P., Nast, C. C., . . . (2025). Risk of Kidney Failure by Diagnosis and Histology in Diabetes: FR-PO0338. *Journal of the American Society of Nephrology*, 36 (10S), 10.1681.

[4] Winter, M. C., Zhou, A. X., Laksana, E., Aczon, M. D., Ledbetter, D. R., Avesar, M., . . . Shpaner, L. (2025). Death One Hour After Terminal Extubation in Children: Validation of a Machine Learning Model to Predict Cardiac Death After Withdrawal of Life-Sustaining Treatment. *Pediatric Critical Care Medicine*, e997–e1008.

[5] Winter, M., Zhou, A., Laksana, E., Aczon, M., Ledbetter, D., Avesar, M., . . . Shpaner, L. (2024). 50: Death One Hour After Terminal Extubation (DONATE): A Multisite Cohort Study. *Critical Care Medicine*, 52(1), S25.

[6] Shpaner, L., Petousis, P., Duru, O., Daratha, K. B., Norris, K. C., Tuttle, K. R., . . . (2023). Kidney Failure Risk Equation Prediction in a Real-World Population with CKD: FR-PO936. *Journal of the American Society of Nephrology*, 34 (11S), 665.