# Leonid Shpaner

Los Angeles, CA | (310) 721-7126 | lshpaner@ucla.edu | linkedin.com/in/lshpaner | github.com/lshpaner | https://www.leonshpaner.com

#### **EDUCATION**

University of San Diego – Master of Science (M.S.) - Applied Data Science [GPA: 4.00/4.00] University of California, Los Angeles (UCLA) - Bachelor of Arts (B.A.) - Economics

**EXPERIENCE** 

Data Scientist | UCLA Health | Los Angeles, CA | Full-time

- First author of an abstract published by the American Society of Nephrology, successfully reproducing gold-standard kidney function risk equations on internal EHR data; achieved AUCs in the 80th percentile and optimized Brier scores.
- Developed open-source 'kfre' Python library to reproduce kidney failure risk equations by Tangri et al., designed for broad use in enhancing predictive analytics in healthcare.
- Co-developed the open-source 'model tuner' Python library to facilitate comprehensive machine learning workflows, including data splitting, pipeline construction, and model calibration, providing a robust toolset for optimizing and enhancing model performance.
- Developed an XGBoost Regression model in production to predict high-cost Medicare Advantage patients (bloomers) yearover-year; successfully implemented system-wide, with a forthcoming paper detailing the project's impact and methodology.
- Collaborating with a multidisciplinary team of faculty members and data scientists to extract meaningful and actionable insights from a variety of healthcare data, including but not limited to biomedical and patient-related information.

Data Scientist | Litecoin Foundation | Remote | Contractor

Leading a GitHub-based project to build a Python library for cryptocurrency price forecasting, with a supporting white paper under development: Litecoin Forecasting Project.

#### Data Scientist | Children's Hospital Los Angeles | Los Angeles, CA | Full-time

March 2022 – October 2022 Conducted research in pediatric ICU developing an Extra Trees machine learning model to predict mortality post-terminal extubation with 88% AUC; co-authored a paper for Critical Care Medicine and enforced rigorous data quality and PHI compliance standards, supporting the DONATE multisite cohort study and aiding the team's recognition with an Early-Stage Researcher Award.

## Sales Data Analyst | Self Mastery Co. | Remote | Contractor

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

Financial Analyst | The Los Angeles Film School | Los Angeles, CA | Full-time

- Integrated the financial reporting infrastructure into newly adopted CPM software (Prophix), streamlining data processes and contributing to a 20% reduction in expenses through enhanced reporting efficiencies.
- Automated financial reports using VBA macros, significantly reducing department workload by 20 hours monthly and facilitating more data-driven decision-making.

## **TEACHING EXPERIENCE**

**Instructor** | UCLA Extension | Remote | Contractor

Instructing data science fundamentals course that covers statistics and basic machine learning in Excel and Python.

Adjunct Faculty | Purdue University Global | Remote | Contractor

Instructed database fundamentals, entity relationship diagrams (Visio), table creation, and SQL queries in Access.

Adjunct Professor | University of San Diego | Remote | Contractor

- Co-designed and currently instruct graduate statistics using Python for the M.S. in Applied Artificial Intelligence program.
- Designed and instructed inaugural course on Excel-based statistics for the Data Analytics Certificate Program.

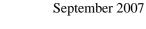
## Course Facilitator in Data Analytics & Statistics | Cornell University | Remote / Contractor

November 2018 – Present

- Facilitating courses with focused feedback and guidance on student assignments and projects
- Completed additional certificates in Data Science and Data Analytics through gift of learning program.

## **TECHNICAL SKILLS**

- Programming Language and Tools: Python (NumPy, Pandas, Scikit-Learn, MLflow, AutoKeras, Aequitas, SciPy), R (Ggplot, Tidyverse, Caret, e1071), SQL, JavaScript, Microsoft Azure, BigQuery, LaTeX, HTML, CSS, Access, Excel, Tableau, Google Data Studio (Looker), Git.
- Data Science: data analytics, visualization, hypothesis testing, data mining, time series, machine learning and predictive modeling (supervised/unsupervised: regression, classification, clustering).
- Languages: English (fluent), Russian (fluent), and Spanish (basic).



April 2023 – Present

August 2022

May 2014 – December 2023

September 2021 – Present

September 2013 – October 2020

August 2023 - Present

October 2022 – February 2024

December 2020 - Present