Eyal Kazin Ph.D.

Profile

Research and product driven Data Scientist and ex-cosmologist with 15+ years of experience applying machine learning, probabilistic modelling, and simulations to high-impact problems across academia and industry. Proven track record of translating noisy, high-dimensional data into production-ready models. Expertise in Bayesian inference, causal reasoning, optimisation, and scientific communication, with a passion for mentoring and cross-functional collaboration.

I have a permit to work in the UK and Europe.

Core Skills & Tools

Machine Learning	Deep learning (PyTorch), Gaussian processes, tabular data ML (scikit-learn), LLMs:; ML Interpretability: SHAP, RAG Triad
Optimisation	Pareto optimisation, evolutionary algorithms
Statistics	Bayesian & causal inference, hypothesis testing, design of experiment
Programming	Proficient: Python, Git, AzureML, SQL; Intermediate: R, DBT
Communication	Stakeholder-facing, mentoring, technical writing, public speaking (PyData, RSS)
Writing	Medium/Towards Data Science: @eyal-kazin, peer-review astronomical journals.

Recent Experience

Nov 2020- **Lead Data Scientist**, *Zimmer Biomet*, London, England, (AI Med Tech)

- Lead R&D on predictive models using wearables, biomarkers, and surgical robot data to optimise patient recovery workflows
- Developed a mechanism to identify surprising consistency in post-operative pain trajectories; awarded Duke Faculty poster and selected for upcoming surgical conference
- Co-developed a model to quantify knee joint geometry, enabling surgical process optimisation and novel visual aids to support clinical decision-making
- Developing tools and mentoring causality to improve data and model interpretability

Tools/methods: Python, git, AzureML, machine/deep learning, Bayesian/causal inference, LLMs/RAG Gaussian processes, optimisation.

Apr 2020-Aug 2022 **Staff Data Scientist**, *Babylon*, London, England, (AI Health Tech)

- Developed and optimised components of probabilistic diagnosis engine
- Collaborated with clinicians and engineers to deliver improvements of 30–50% in KPIs
- Delivered insights and tools for product and clinical stakeholders.

Tools/methods: Python, git, GCP, Bayesian/causal inference, DBT, Looker, Tableau

Apr 2018-Jan 2020 **Senior Data Scientist**, *LabGenius*, London, England, (BioTech - Drug Discovery)

- Built DNA sequence optimisation algorithms for therapeutic antibody design, leading to $\times$100 increase in protein stability and $\times$400 in potency
- Led design-of-experiment initiative for assay optimisation resulting in doubling DNA yield
- Delivered end-to-end modelling solutions for integration by engineers

Tools: Python, git, GCP, machine/deep learning, Bayesian inference, Pareto optimisation, Gaussian processes, design of experiment.

elzurdo • *eyal-kazin* • *7KJ4H4IAAAAJ* • *eyalkazin*
eyal-kazin.medium.com

Education

2005–2011 **Ph.D. Physics**, *New York University*, New York, NY, USA
Dissertation: *Large-Scale Clustering of Galaxies* bit.ly/2uffbuB
- Lead contributor to five peer-reviewed cosmology publications, with 960+ citations.
- Used data, simulations and statistical modelling to explore cosmic structure formation
- Developed new correlation estimators for galaxy clustering data for improved interpretation.

Tools/methods: Python, C, Bayesian inference, clustering

2001–2004 **B.Sc. Physics (with honors)**, *Ben Gurion University*, Be'er Sheva, Israel

Selected Publications & Research

Kazin (2025) in prep **Improved Sequential Hypothesis Testing**, Introducing a novel Bayesian heuristic for unbiassed data collection. Presented at Royal Statistical Society International Conferences, UK: contributed talk (25'), poster (24') (bit.ly/precision-goal-poster)

Kazin et. al (2025) in prep **Variabilities In Pain Trajectories Following Primary Total Knee Arthroplasty Segmented By Immediate Post-Operative Values.** - Demonstrating that self-reported pain is likely to be more consistent than commonly assumed. Selected for contributed talk a surgeon conference and awarded Duke Faculty Poster 2025.

Kazin et. al (2014) **Improved distance measurements to z=1 with reconstruction of the baryonic acoustic feature** - Demonstrated on real and simulated data that by using velocity fields to shift galaxies to their near-original positions one obtains more accurate estimates of dark matter and dark energy. Cited 435 times (256 since 2019).

Kazin et. al (2010) **The baryonic acoustic feature and large-scale clustering in the Sloan Digital Sky Survey luminous red galaxy sample** - By using simulations we demonstrate that an apparent abnormality in the real galaxy 2-point clustering is likely to be due to cosmic variance rather than "new physics". Cited 297 times (39 since 2019).

Public Writing & Speaking

May, 2025; Dec, 2024 **Causality - Mental Hygiene for Data Science**, PyData Global Conference (recording link), Data Science Festival (recording), TDS article: bit.ly/causal-hygiene

Feb, 2025 **Information Theory for People in a Hurry**, Medium series bit.ly/info-theory-hurry

May/Sep 21 Nov 2020 **Improved Decisions with Pareto Fronts**, PyData Global 2020, PyCon US, Australia recording: bit.ly/moo-youtube-intro, tutorial: bit.ly/improved-decisions-pareto

Mentoring & Teaching

2015 onwards Mentoring data scientists and researchers in causal inference and ML interpretation

2014 onwards **Adopt a Physicist** - Student correspondence regarding physics as a career.

Languages & Extras

Languages Proficient: English, Hebrew, Conversational: Spanish

Fun fact For 10 years I have maintained a daily video diary

Hobbies Tennis, pingpong, guitar, languages, short film production, basketball, cycling, traveling, surfing.

elzurdo • eyal-kazin • 7KJ4H4IAAAAJ • eyalkazin
eyal-kazin.medium.com