Zachary Hafen-Saavedra

PhD, Theoretical and Computational Astrophysics
z.hafen.saavedra@gmail.com || zhafen.github.io || Chicago, IL ||

Summary

[click here for a work sample]

Data scientist with ten years of experience earned as a Northwestern University and UC Irvine computational astrophysicist. Extensive history leading interdisciplinary collaborations and communicating complex concepts. Seeking positions in data science and data analysis.

Skills

Techniques: data analysis (cleaning, sparse, big data, exploratory, time-series), dashboarding, machine learning (inc. NLP), frequentist/Bayesian statistics, GIS (georeferencing, mosaicking, etc.), code testing
Interpersonal skills: technical leadership and management, storytelling, mentoring
Tools: Python (numpy, pandas, matplotlib, scikit-learn), SQL, NoSQL, BI (Cognos BI, Streamlit), TensorFlow, nltk, Unix, C/C++, QGIS, parallel computing, git, Windows/Mac/Unix

Experience

Data Scientist September 2023 - Present
Adler Planetarium Chicago, IL

- In progress: Processed aerial images of city lights at night, enabling at least six new collaborative projects, by performing GIS georeferencing, mosaicking, and de-noising.
- Elevated data science capabilities across the planetarium (with 500,000+ visitors/year), by serving as the primary resource for data science education and consultation.

Business Data Analyst June 2023 - September 2023
Northwestern University, Center for Interdisc. Explor. and Research in Astrophysics Evanston, IL

- Empowered business staff to independently explore, update, and adjust visualizations, reducing the need for technical support, by building online BI dashboards.
- Re-enabled and extended financial reporting by replacing fragile Excel macros with robust and locally-maintainable Python and Shell solutions while minimizing dependencies.
- Extracted actionable insights for improving organization cultural climate, as measured by professionally-administered DEI surveys, by algorithmically extracting quantitative data from 100+ page reports.

McCue Prize Postdoctoral Fellow in Cosmology July 2020 - June 2023
University of California–Irvine, Department of Physics and Astronomy Irvine, CA

- Employed natural language processing to convert >200,000 scientific abstracts to quantitative data
- Trained an ensemble voting model to use abstract word content and paper metadata to predict citation count to within 3 citations per year for 75% of the validation set
- Performed complex filtering of $>$2 TB of remote data via the NASA astrophysics data system API
- Spearheaded an international team spanning eight institutions to validate Bayesian statistical models

National Science Foundation Graduate Fellow in K-12 Education June 2014 - July 2020
Northwestern University, Department of Physics and Astronomy Evanston, IL

- Processed tens of TB of $>$20-dimensional data using high-performance-computing resources, reducing to $<$100 GB of highly-interpretable data
- Utilized and upgraded a C code to generate >100,000-CPU-hour simulations of entire galaxies
- Performed time-series decision-tree classification to predict the extragalactic origin of Earth
- Introduced $>$100 students from underrepresented backgrounds to data science by leading one of Chicago's first high-school data-science education initiatives
- Collaborated with a multidisciplinary range of scientists to publish (to date) 36 papers, 7 as a lead author