# PETE KUZMA

Tokyo, Japan ◇ Phone: +81-080-7791-0524
E-mail: pbkuzma@gmail.com ◇ : `github.com/PeteKuzma` ◇ LinkedIn: linkedin.com/in/pete-kuzma/

## PROFESSIONAL SUMMARY

Data scientist with 10+ years of experience in quantitative analysis, leveraging Python (NumPy, SciPy, Pandas, scikit-learn), Bayesian inference, and machine learning to extract insights from large datasets. Skilled in SQL querying, A/B testing, and building scalable data pipelines. Proven track record of leading international teams, solving complex problems, and publishing 10+ peer-reviewed journal articles.

## WORK EXPERIENCE

**JSPS International Postdoctoral Fellow** October 2023 - Present
*National Astronomical Observatory of Japan* *Osawa, Tokyo, Japan*

- Applied unsupervised machine learning and A/B testing in Python to analyze millions of data points, improving target identification efficiency.
- Developed multivariate regression models to extract meaningful patterns from astronomical datasets, ensuring data integrity and optimizing selection criteria.
- Automated SQL queries via Python API integrations, improving data retrieval efficiency by 80% and streamlining large-scale data processing.
- Led a 10+ member international research team, managing cross-functional collaborations across four countries.

**Postdoctoral Research Associate** April 2020 - October 2023
*University of Edinburgh* *Edinburgh, Scotland, UK*

- Designed and implemented four Bayesian inference models to analyze large datasets, integrating Python, SQL, and Unix-based tools for statistical evaluation.
- Conducted rigorous artificial A/B testing to assess model accuracy and optimize data-driven decision-making processes.
- Built and maintained a structured database of 100,000+ entries, improving data accessibility and enabling efficient query execution for international research projects.

**Commonwealth Rutherford Fellow** March 2018 - February 2020
*University of Edinburgh* *Edinburgh, Scotland, UK*

- Developed a scalable data analysis pipeline, integrating multiple Unix-based programs with Python wrappers for automation and efficiency.
- Implemented bootstrap analysis and logistic regression to assess data quality and validate model predictions.
- Managed database curation for two international collaborations, ensuring structured, high-quality datasets met rigorous research criteria within tight deadlines.

## EDUCATION

**PhD - Astronomy and Astrophysics** *April 2013 - December 2017*
*Australian National University, Canberra, Australia*

## TECHNICAL STRENGTHS

| | |
|---|---|
| **Programming** | Python (Numpy, Pandas, Sci-kit Learn, Matplotlib), SQL, R (Beginner: C/C++, FORTRAN) |
| **Data Science and ML** | Bayesian Inference, A/B testing, machine learning (Sci-kit Learn), data regression and analytics |
| **Databases** | SQL, TopCat, API data integration, ETL pipeline development |
| **Tools & Environments** | Unix/Linux, Jupyter, Git, Emacs, Vim, VS Code Studio |