JUSTIN CLANCY

Data Analyst | Data Scientist | Astrophysics Ph.D.

justinpclancy@gmail.com (+44)7386804709 justin-clancy github.com/justinc97

PROFESSIONAL SUMMARY

Data scientist with a PhD in astrophysics specialising in time-series analysis, anomaly detection, and large-scale Python data pipelines. Experienced analysing petabyte-sized datasets, developing machine learning models, and deploying automated workflows within international research collaborations. Strong communicator with a track record of translating complex quantitative analysis into actionable insights.

TECHNICAL SKILLS

Programming: Python, SQL, R, Bash
Data Analysis: Pandas, NumPy, SciPy, scikit-learn
Machine Learning: Regression, Classification, Clustering, XGBoost
Data Engineering: ETL Pipelines, Data Modelling, PostgreSQL, MySQL
Visualization & Reporting: Tableau, Plotly, Matplotlib, Seaborn, Dashboards, LaTeX
Tools: Git, Docker, Linux, Jupyter, VSCode
Methods: Time-series analysis, Hypothesis testing, Bayesian statistics, Forecasting

PROFESSIONAL EXPERIENCE

LLM Training Specialist *Oct 2025 – Present*
Mercor Intelligence Freelance Consultant

- Designed and evaluated complex quantitative reasoning tasks used to train large language models on scientific and data analysis workflows.
- Created labelled datasets and detailed solution pipelines to improve model performance in statistical reasoning and multi-step analytical tasks.
- Applied structured quality assessment processes to ensure consistency and accuracy across large-scale training datasets.

Research Data Scientist *Jan 2022 – Present*
Simons Observatory Collaboration Melbourne, VIC, Australia

- Built scalable Python data pipelines to ingest and analyse multi-million-record time-series datasets from 60,000+ sensors.
- Developed machine learning and statistical detection algorithms to detect anomalies and patterns within noisy observational data streams.
- Designed the collaboration's first automated anomaly detection system for high-frequency time-series, now integrated into production operations.
- Produced clear visual reports and presentations for technical and non-technical stakeholders across international teams.
- Coordinated shared data analysis tasks across a 50+ institution international collaboration, managing timelines and technical deliverables for shared pipeline infrastructure.

SELECTED PROJECTS

Active Anomaly Detection Pipeline

- Designed and implemented an end-to-end Python pipeline for real-time anomaly detection in high-frequency sensor time-series data, integrated into a production data processing environment.
- Applied matched filtering and density-based clustering to detect sub-second events across terabytes of data streams from a high-density sensor network, achieving $\geq 90\%$ detection rate.
- Characterised pipeline performance across a full parameter space via Monte Carlo simulation across 100+ injection trials per configuration, achieving $\sim 5\%$ false positive rate and source localisation to within 10% of the instruments resolution.

- Submitted to the Open Journal of Astrophysics (2025): The Simons Observatory: Development of a Pipeline to Detect Rapid Transients in Time-Ordered Data
- Co-authored a companion paper applying random forest classification to anomalous event detection: The Atacama Cosmology Telescope: Machine-learning-driven Tools for Detecting Millimeter Sources in Timestream Preprocessing

Large-Scale Simulation and Forecasting Analysis

- Generated and analysed full-sky simulated datasets across 6 frequency bands to quantify systematic bias introduced by signal model assumptions in a multi-frequency detection pipeline.
- Build statistical frameworks to compare signal suppression against noise scatter across 9,700+ detections, isolating a consistent $\sim 4\%$ systematic bias in recovered signal strength for high confidence detections
- Demonstrated that model assumptions introduce measurement biases exceeding random noise for the strongest signals, with findings informing pipeline design decision for upcoming observational surveys and instrument development.

Statistical Analysis of Satellite Data

- Aggregated and analysed a public catalogue of 6,282 sources across full-sky survey data, implementing a stacking pipeline to extract population-level signals too faint to detect in individual observations.
- Developed bootstrap resampling framework (2,500+ iterations) for uncertainty quantification, achieving an $11\,\sigma$ signal detection from sources individually below the detection threshold.
- Built and publicly released an open-source Python masking tool adopted by the research community for ongoing survey analyses, with configurable parameters for different use cases (Github).
- Published in Monthly Notices of the Royal Astronomical Society (2023): Polarization fraction of Planck Galactic cold clumps and forecasts for the Simons Observatory

TEACHING EXPERIENCE

Undergraduate Teaching Assistant — *Mar 2020 – October 2025*
The University of Melbourne — Melbourne, VIC, Australia

- Delivered lab instruction in statistical data analysis, error propagation, and experimental design to 400+ undergraduate physics students over 5 years.
- Developed assessment rubrics and marking schemes for data analysis assignments, ensuring consistency across large cohorts.
- Mentored students through project-based research tasks, supporting data collection, analysis workflows, and written interpretation of results.

EDUCATION

Doctor of Philosophy (PhD) - Science — *2022 – 2025*
The University of Melbourne — Melbourne, VIC, Australia

Master of Science (Physics) – With Distinction — *2020 – 2021*
The University of Melbourne — Melbourne, VIC, Australia

Bachelor of Science (Major in Physics) — *2016 – 2018*
The University of Melbourne — Melbourne, VIC, Australia

AWARDS & SCHOLARSHIPS

- **Dr Albert Shimmins Postgraduate Writing-Up Award**, The University of Melbourne (2025)
- **Laby PhD Travelling Scholarship**, The University of Melbourne (2024)
- **Research Training Program Scholarship**, The University of Melbourne & The Australian Government (2022–2025)
- **Melbourne Research Scholarship**, The University of Melbourne (2022)
- **Klein Prize in Experimental Physics**, School of Physics, The University of Melbourne (2022)