

# Yuba Amoura, PhD

5+ years of experience in **Python, data science, and statistical analysis**. Skilled at building scalable solutions and turning complex data into clear insights. A strong team player with a talent for making technical concepts easy to understand and apply. Uniquely able to bring **adaptability, resilience, and a global perspective** to any team. Experienced in presenting complex technical concepts to technical and non-technical audiences including senior executives.

## EXPERIENCE

### Data Science Analyst May 2024 – Current

*Alberta Investment Management Corporation (AIMCo), Toronto, Canada*

- Developed a **recession indicator** using macroeconomic and financial data with machine learning techniques (**Python**).
- Designed and implemented a robust **database solution** for storing and accessing large-scale economic and financial data (**SQL, Python, Pandas, Databricks**).
- Built tools to automate the creation and updating of **100+ macroeconomic charts** in a dynamic chartbook app (**Python, Streamlit, Plotly**).

### Doctoral Researcher May 2019 – Aug 2023

*University of Waterloo, Waterloo, ON, Canada*

- Led a novel research project designing an original method for analyzing **galaxy cluster data**, advancing our understanding of the Universe without additional costs.
- Processed and transformed **TBs of raw simulation data** from multiple sources into a standardized format, enabling meaningful analysis (**Python, Numpy, SQL, Matplotlib, Pandas**).
- Predicted **cosmological properties** using cleaned data and advanced regression techniques.
- Developed and simulated **25 unique cosmological models**, generating 100TB of data that will be the foundation for future research over the next decade (**Cloud computing, Linux, Bash, C++**).

## MACHINE LEARNING PROJECTS

- Power Outage Prediction:** Built a predictive model using **severe weather data** to forecast power outages ([Project link](#))
- Windmill Detection** (Classification Competition): Achieved 91% accuracy in detecting windmills from satellite images using a superlearner model (**R, Python, OpenCV**)
- Image Size Reduction via K-means Clustering:** Reduced image color palettes with minimal quality loss using K-means clustering (**Python, Pycharm, OpenCV**)
- Cancer Gene Prediction with PCA:** Applied **Principal Component Analysis (PCA)** to identify gene combinations most correlated with various cancer types (**R**).
- Galaxy Cluster Analysis with PCA:** Used **PCA** on galaxy cluster data to explore structural properties and age correlations ([GitHub](#))

## CONTACT

- Toronto, Canada
- +12269783575
- [amourayuba@gmail.com](mailto:amourayuba@gmail.com)
- [Linkedin](#)
- [GitHub](#)
- [Personal website](#)

## TECHNICAL SKILLS

**Advanced** Python: Numpy, Scipy, Matplotlib, Jupyter, LaTeX, Linux

**Intermediate** OpenCV, Scikit-learn, SQL, Pandas, Git/GitHub, Plotly, Streamlit

**Familiar** R, C/C++, Matlab, TensorFlow, PyTorch

## EDUCATION

- PhD in Astrophysics**

May 2019 – June 2023  
*University of Waterloo, Canada*

- Data Science Bootcamp**  
Jan 2024 – April 2024

*The Erdos Institute*

- Masters in Statistics Modelling-ML**  
Sept 2018 – March 2019  
*Université Paris Descartes, France*

Relevant coursework:  
Optimization, Stochastic Algorithms, Classification, High Dimension Learning, Poissonian Processes  
Ranked first in the masters

- Masters in High Energy Physics**  
Sept 2014 – Mar 2016  
*Sorbonne Université, France*

## COMMUNICATION

Fluent in English, French, Berber and Arabic

## OTHER

- Tutored 100+ students (high school and university) in Math, Physics, Statistics
  - Teaching chess to visually deficient students using original and innovative learning techniques adapted to the students