

Steven Macauda

Data Scientist

📍 New York, NY, US

📞 +1 708-674-5912

✉ steven.macauda@gmail.com

🌐 [linkedin.com/in/steven-macauda](https://www.linkedin.com/in/steven-macauda)

🐙 [smacauda.github.io](https://github.com/smacauda)

Skills

- **Programming:** SQL | Python | C++ | Java | JavaScript | R | HTML | CSS
- **Communication/ Visualizations:** Tableau | Matplotlib | Seaborn | Jupyter
- **Data Science/Engineering:** Scikit-Learn | TensorFlow | PyTorch | Spark | Hadoop | Pandas
- **Cloud Technologies:** AWS | Azure | Google Cloud Platform

Experience

Data Science Fellow

May 2022 - present

Sharpest Minds

- Collaborating with Head of Data Science at COMMB.

Mathematics and Physics Tutor

Sep. 2017 - present

Multiple Companies

- Led laboratory sessions of 25 students for 2 years. Individual instruction provided to 11 students resulting in average improvement of one grade point. Created weekly lesson plans for students.

Research Associate

Dec. 2014 - Sep. 2017

University of Illinois at Chicago

- Developed and maintained Python and C++ scripts to monitor data quality.
- Analyzed LHC detector data at the Fermilab LPC to detect and document anomalies.
- Automated and performed calibration testing using deconvolution algorithms to find the optimal setting for pixel detector modules leading to a 10% improvement in the detector tracking parameters.
- Utilized ROOT, Matplotlib, and Seaborn libraries to document and present at weekly meetings.
- **Publications:** [JINST \(2018\)](#), [JINST \(2019\)](#)

Projects

Sales Prediction Using Time Series Analysis

<https://bit.ly/3P1hTLx>

- Time series analysis to forecast future sales of a US superstore using ARIMA model.
- Forecast future sales with mean absolute error (MAE) of 163.30.
- Dashboard created using Tableau to help store prepare for higher future demand.

Bike Demand Prediction

<https://bit.ly/3BDLmYN>

- Bike sharing data from Washington, DC. Trained random forest, lasso, decision tree, support vector, and k-neighbors regressors to find best performance with random forest with a root mean squared error of 0.324 bikes/hour.
- Hyper-parameter optimization using Optuna leading to a 1% performance increase.
- Tableau dashboard created to help forecast demand and compare station data.

Air Quality Forecasting

<https://air-quality-app-xl4lmy5tzq-uc.a.run.app>

- End-to-end project using Google Cloud Platform to produce a time series forecast of air quality by month using Prophet with root-mean-squared-error of 0.01 or less.
- Web app created to visualize results.

Education

PhD Student - Physics

Sep. 2017 - Jan. 2019

University of California

University of California Sep. 2017 - Jan. 2019

PhD Student in Physics Davis, CA

- **Key Courses:** Quantum Field Theory, Statistical Mechanics, Data Science for Particle Physics, Mathematical Methods.

BS - Physics

Aug. 2013 - May 2016

University of Illinois Chicago

University of Illinois at Chicago May, 2016

BS, Physics

- **Seymour Margulies Scholarship:** awarded to the student who receives the highest grade in upper level electromagnetism course.
- Completed **CMS Data Analysis School (2016)** at the Fermilab LPC.