David Almona

859-324-3853 almonadavid@gmail.com linkedin.com/in/davidalmona almonadavid.github.io

Education

Centre College **2022 – May 2026**

Bachelor of Science in Economics & Finance (Major), Data Science (Minor) *Danville, KY*

- Cumulative GPA: 3.62 | Major GPA: 3.81
- Relevant Coursework: Modern Calculus 1, Econometrics, Intro to Management, Data Visualization, Programming and Problem Solving, Economics of College Sports, Statistical Modeling, Managerial Finance, Data Analysis and Communication.

African Leadership Academy **2020 – 2022**

High School Diploma *Johannesburg, South Africa*

- Cambridge A Levels: Physics and Mathematics | Other Courses: Entrepreneurial Leadership and African Studies

Experience

Technical Assistance Pilot Program Fellow **Aug 2025 – Dec 2025**

Sustainable Business Network (SBN) | Incoming, August 25th 2025 *Cambridge, MA*

Undergraduate Research Fellow **Jun 2025 – Jul 2025**

Carnegie Mellon University - Statistics & Data Science | Capstone Project *Pittsburgh, PA*

- Developed machine learning models (XGBoost and logistic regression) to predict soccer pressing effectiveness with 77.1% AUC, analyzing 252,646 pressing sequences from 502 Major League Soccer matches using 10Hz player tracking data provided by SkillCorner.
- To be presented at the Carnegie Mellon Sports Analytics Conference 2025.

Data Analytics: Qualitative & Quantitative Insights Externship **Mar 2025 – Jun 2025**

Beats by Dr. Dre | Capstone Project *Remote*

- Analyzed 4,628 survey responses using Python (Pandas, NumPy) to recommend $150-250 wireless speaker launch, identifying 71% market penetration and 21% purchase intent.
- Segmented consumer data to reveal 68% of users are aged 18-24, using statistical analysis to validate Gen Z as primary target demographic with 64% daily usage rates.
- Applied NLP techniques to categorize open-ended responses on sound quality and features, extracting key pain points (battery life, sound quality) to guide product development.
- Built correlation matrices identifying price-brand reputation as key purchase drivers, visualized insights through interactive dashboards for executive decision-making.

Undergraduate Research **Sep 2024 – Nov 2024**

Centre College | Capstone Project *Danville, KY*

- Built RShiny dashboard analyzing shot data for Centre College Division III women's soccer team.
- Presented during poster session at the Carnegie Mellon Sports Analytics Conference 2024.

Projects

National Olympic Performance: Analyzing Economic, Demographic, and Historical Determinants of Summer Olympics Medal Counts | *Stata* | *View* **Sep 2024 - Dec 2024**

- Conducted a panel data econometric analysis of Olympic medal determinants using Stata, analyzing 716 observations across 185 countries from 2008 to 2024, achieving an adjusted R-squared of 0.98.
- Developed and compared multiple regression models to identify key predictors of Olympic success, incorporating fixed effects to control for country-specific characteristics.
- Performed statistical hypothesis testing and addressed violations of classical assumptions, revealing significant impacts of host country status and health expenditure.

Cincinnati Reds Baseball Analytics Student Hackathon | *R* | *View* **Jan 2025 – Feb 2025**

- Built elastic net regression models in R to predict 2024 MLB player usage (PA/BF), engineering 30+ features from 3 years of Savant tracking data including advanced metrics (SIERA, FIP, launch angles).
- Processed panel data for 1,000+ players across 2021-2023 seasons, implementing 10-fold cross-validation and optimizing variable transformations to improve prediction accuracy.
- Achieved model validation through temporal split testing (2021-2022 training, 2023 testing), handling edge cases, and delivering predictions for roster planning decisions.

Additional

Languages & Tools: SQL, R, Stata, Python, Tableau, Power BI, Git/Github

Activities: Student-Athlete - Track and Field, Academic Honor Roll, 2023 & 2024 • Vice President: Centre Brother-to-Brother club