

BASEBALL DATA ANALYST

EDUCATION

Wake Forest University School of Business, Winston-Salem, NCMaster of Science in Business Analytics, May 2026

Honors & Activities: Wake Forest Baseball Team Graduate Analyst and Statistician, Management Consulting Club, Program Ambassador

University of Alabama, Tuscaloosa, ALBachelor of Science in Management Information Systems, December 2024

Honors & Activities: Alabama Information Management Society (AIMS), UA Data Analytics Club, Phi Sigma Kappa

AREAS OF EXPERTISE

Quantitative Analytics | Data Science | Predictive Modeling | Data Analysis | Statistical Analysis | Shiny Apps | Software Development | Player Development

Data Visualization | Automation | Process Optimization | Algorithms | Leadership | Mathematic Analysis | Database Management | AI Development

Data Pipelines | Systems Engineering | Database Architecture | Supervised/Unsupervised Learning | Model Deployment | Trackman | KinaTrax

TECHNICAL SKILLS

Programming Languages: Python, SQL, R, C#, HTML, JavaScript, CSS, Dax, Node.js, .Net, React

Technical Tools: Power BI, Tableau, Plotly, Matlib, Ggplot2, Seaborn, Streamlit, Scikit Learn

Machine Learning: XGBoost, Bayesian, Random Forests, Ensembling, GAMs, Lasso-Ridge, Stacking, Time Series, Neural Networks, Monte Carlo

Operating Environments: R Studio, VS Code, Jupyter, MySQL, SQLite, DBeaver, PgAdmin4, GitHub

PROFESSIONAL EXPERIENCE

GRADUATE BASEBALL DATA ANALYST / STATISTICIAN, Wake Forest University BaseballJun 2025 – Present

- Developed 3 predictive gradient boosting machine learning pipelines from 900,000 data observations across multiple pitch types; whiff outcomes, groundball tendencies, and BABIP models and created a pitcher stuff+ model for fastball data using gradient boosting and stacking
- Generated actionable insights through pitcher performance leaderboards, delta analysis identifying over/under-performers, and comprehensive visualization analysis including correlation matrices, histograms, scatterplots, and feature importance plots
- Built a Shiny App for each predictive model to allow real-time interactive metric calculators for pitchers on their raw ball flight characteristics
- Operated portable and stadium Trackman units, KinaTrax units, performing routine data collection tasks

DATA ANALYST / INFORMATION SYSTEMS ENGINEER CO-OP, Mercedes-Benz USAug 2023 – Sep 2024

- Collaborated with manufacturing leadership, implementing data-driven optimization strategies and Python code that increased cart space efficiency by 15% which also reduced operating material costs, demonstrating commercial awareness
- Managed complex deployment of predictive scheduling optimization solutions, coordinating cross-functional teams applying time management methodologies to deliver 3 modified plant performance models that improved productivity through scheduling changes
- Developed client-facing dashboards using collaborative development processes for operational decision-making and cost optimization using the full software development lifecycle for production management actively, while actively seeking guidance and feedback

BASEBALL DATA ANALYST / STATISTICIAN, University of Alabama Baseball TeamSep 2022 – Apr 2023

- Designed a Python codebase with Plotly and Matplot libraries to manipulate large amounts of data and visualize effectively
- Constructed Shiny Apps for detailed advanced scouting reports on league play and reviewed SQL databases for data processing at game conclusion
- Data collection through Trackman for scrimmages, bullpens, and games.

ANALYTICAL PROJECT EXPERIENCE

Bayesian Hierarchical Pitcher Aging & Decline Detection Project:

- Architected a 3-component methodology combining Bayesian hierarchical modeling, PELT changepoint detection, and risk assessment frameworks to quantitatively separate natural aging from injury effects in MLB pitcher velocity across 10 years of Statcast data covering 5,500+ observations
- Validated approach through retrospective case study analysis of 8 MLB pitchers demonstrating distinct aging and injury pattern trajectories, and additional machine learning to produce uncertainty estimates of player future risk and velocity to build a reliable early warning system for player development and roster planning decisions
- Implemented changepoint detection algorithms to identify precise decline onset timing with uncertainty quantification, enabling front office personnel to distinguish between natural aging, mechanical issues, and injury-related performance degradation with statistical confidence intervals

Multi-Machine Learning Model Financial Services Fraud Detection Project:

- Performed data preparation and cleaning on a dataset of 900,000 observations through Python to prepare for modeling and learning phase
- Constructed multiple machine learning classification models (linear regression, logistic regression, weighted logistic, random forests, XGBoost) achieving 88% fraud detection rate and identifying key risk predictors through proper visualization and undersampling techniques
- Implemented a 3-tier system for dealing with fraud applications based on optimized threshold levels; high, medium, and low

Ensemble Model Future OBP Performance Prediction Project:

- Engineered sophisticated ensemble machine learning model combining Ridge Regression, Gradient Boosting, and Random Forests to predict 2021 OBP's for 550+ MLB players using 5 years of historical performance data
- Built advanced features including weighted historical OBP with recency bias, age-based performance curves, and regression-to-mean methodology, implementing uncertainty quantification with confidence intervals accounting for player experience levels to provide actionable risk assessments
- Delivered comprehensive documentation with feature importance rankings, model validation visualizations, and player-specific insights identifying high-confidence predictions versus regression candidates for roster construction applications

Monte Carlo Markov Chain MLB Playoff Probability & Simulation Project:

- Developed MCMC frameworks in R solving baseball probability problems including playoff matchup scenarios (100,000+ simulations) and walk-off win probability optimization, using binary search algorithms to identify thresholds needed to achieve specific game outcome targets
- Engineered multi-stage probabilistic models incorporating regular season outcomes, divisional standings, tiebreaker protocols, and best-of-three wild card series while implementing sequential dependency modeling to track state transitions across outs, base runners, batting order, and scoring scenarios for realistic game dynamics

LEADERSHIP EXPERIENCE

Eagle Scout: Led a cross-functional team of 30 volunteers as a project manager, designing and building the local K9 agility course reconstruction

Apex Outreach Service Project: Volunteering for the local church to help rebuild homes in surrounding communities annually