Joseph George

Github: @josephageorge | 732-689-0320 | joseph.a.george@gmail.com | Data Science Blog: @thezonemaster

EDUCATION

Rutgers University – New Brunswick   December 2025

Bachelor of Arts in Computer Science, Minor in Statistics   **GPA: 3.78/4.00**

Bachelor of Science in Business Analytics and Information Technology

- **Honors:** Dean's List(All Semesters), School of Arts and Sciences Honors Program, Scarlet Scholarship, Phi Beta Kappa (Top 10%)
- **Relevant Coursework:** Econometrics, Statistics (1-2), Regression Methods, Machine Learning Principles, Linear Algebra, Time Series Modeling, Calculus (I-III), Discrete Structures I (Proofs), Discrete Structures II (Probability), Introduction to Data Science, Data Mining, Design and Analysis of Computer Algorithms, Business Data Management (Databases), Bayesian Data Analysis, Theory of Probability, Systems Programming, Computer Architecture, Linear Optimization, Data Structures

SKILLS

Programming Languages & Data Technologies: Python, R, SQL, Stata, Java, JavaScript, MATLAB, Docker, Tableau
Libraries & Frameworks: Pandas, NumPy, Scikit-learn, ggplot2, R Shiny, Seaborn, Plotly, TensorFlow, PyTorch

WORK EXPERIENCE

Silver Bulletin (Nate Silver's Newsletter)   **Remote, United States**

Assistant Sports Analyst   *June 2025–Present*

- Publishing original data-driven sports analyses for a newsletter with over 300,000 subscribers as part of a small team of three with Nate Silver, blending longform writing, clean data visualizations using Datawrapper, and advanced statistical modeling.
- Building, rebooting, and maintaining popular and widely used models such as ELWAY (NFL Elo Model) and QBERT (Quarterback Rating Model) which have drawn thousands of subscribers to the site.
- Scraping, cleaning and normalizing large-scale databases for model production using reproducible, industry-standard data practices.

Mexico City Capitanes (NBA G-League Franchise)   **Mexico City, Mexico**

Basketball Analytics Associate   *April 2025–Present*

- Translate proprietary metrics (True Rebounding Rate and HUSTLE) to G-League/NCAA contexts, supporting over thirty scouting and recruiting decisions.
- Developing data application that allows scouting team to upload any dataset, filter by desired statistics, and create custom data visualizations without any previous programming background.
- Participate in team development ideation weekly with the general manager and scouting team by blending analytics, data visualization, and film study to guide player valuation under tight budget constraints.

Fannie Mae   **Washington D.C., United States**

Data Science Intern   *June 2025–August 2025*

- Delivered three ad-hoc analyses comparing secondary credit reporting with internal data sources, informing internal credit risk initiatives through complex SQL queries.
- Provided data documentation on numerous fields, aiding in the cleaning and scoping process for new dataset.
- Analyzed credit report dataset, generating visualizations and correlation matrices using ggplot2 across 50+ fields to guide income model feature selection.
- Prototyped income prediction model in R (caret), identifying key credit variables with early signs of lift.

Kiva.org   **San Francisco, California**

Data Science Co-Op, Product Analytics   *January 2025–May 2025*

- Led user analysis for MyKiva, a badge-based personalization initiative; quantified a 14% decrease in churn among high-badge users using SQL, R, and Amplitude, which led to further product design and testing.
- Diagnosed and resolved a Looker bug that misattributed corporate match eligibility, boosting match accuracy by 25% and restoring trust across 3+ partner programs.
- Collaborated with product and engineering to define KPIs, run A/B tests, and optimize dashboards, reducing time-to-insight by 30% via improved SQL and Amplitude reporting flows.

Wakefern Food Corporation   **Edison, New Jersey**
Data Engineer Intern   *June 2024–August 2024*
- Programmed a custom SQL-based store procedures application to pipeline data from store terminals into a unified DB2 relational database.
- Automated business processes across Human Resources, Payroll, Finance, and Supply Chain divisions by over 25% using JavaScript, Node.js, and SAPUI5 framework to develop applications within the SAP framework.

Scarlet Computing Solutions   **Princeton, New Jersey**
Data Science Intern   *March 2024– May 2024*
- Visualized OCR model performance with Matplotlib and Plotly, highlighting accuracy improvements and trends through dynamic dashboards.
- Optimized model performance by calculating Levenshtein distance and character error rates, using libraries such as NLTK and Scikit-learn.

Wakefern Food Corporation   **Edison, New Jersey**
Data Engineer Intern   *May 2023– August 2023*
- Built a real-time stats dashboard in React.js/Node.js, improving system response time by 40% at Wakefern's pilot automated store.
- Developed custom churn and retention metrics, evaluating store pilot performance over its first financial year.
- Built an Electronic Shelf Label monitoring tool with the MERN stack, reducing downtime alerts by 25%.

New Jersey Economic Development Authority   **Trenton, New Jersey**
Data Analyst Intern   *May 2022– August 2022*
- Built interactive Tableau maps for 10,000+ small business tax records; revealed spatial trends for leadership reports.
- Automated Python pipelines, cutting data wrangling time by 70% and completing full runs in under 5 minutes.

RESEARCH PROJECTS

Burning Glass Institute   *October 2025–Present*
- Migrated Labor Matters newsletter, which provides daily information on American labor markets, education, and automation, to a larger audience on Substack.
- Conducted sentiment analysis on LinkedIn summary data, to determine societal good patterns within specific industries or education levels.

The Zone Master: Basketball Analytics Blog   *July 2023–Present*
- Created proprietary advanced defensive versatility and hustle metrics using Python, Pandas, and NumPy, along with predictive models for contract extensions and career shooting, leveraging Scikit-learn, XGBoost, and Seaborn.
- Communicated comprehensive analyses through over fifty written blog posts, data visualizations, and podcasts.
- Led insights into sports analytics, exploring applications of TensorFlow, PyTorch, and deep learning frameworks in player tracking.
- Ranked in the top fifty of rising sports blogs on Substack, growing a subscriber base of 300+ including NBA analysts, machine learning engineers, and team executives.
- Presented my research at Yale University to an audience of professors, front office executives, researchers, and graduate students.

Behavioral Economics Lab - Rutgers University – New Brunswick
Research Assistant   *September 2024–December 2024*
- Built an introductory Prospect Theory–inspired model integrating temporal weighting (TW) and delay discounting (DD) to simulate human sales decisions under uncertainty, using nonlinear value functions, probability weighting, and loss aversion.
- Estimated individual-level parameters for over 150 subjects via differential evolution and custom incremental search, minimizing squared error between observed and predicted decisions.
- Developed diagnostic visualizations for parameter convergence and error progression using Matplotlib, which reduced misfitting cases and improved model interpretability in weekly research presentations.

INTERESTS

Interests: Writing, Basketball Analytics, Environmental Sustainability, Quiz Bowl, United States History