

Nathan Blumenfeld

 (201) 492-3944 |  nathanblumenfeld@gmail.com |  [nathanblumenfeld](https://www.linkedin.com/in/nathanblumenfeld) |  [nathanblumenfeld](https://github.com/nathanblumenfeld)

EDUCATION

Cornell University, College of Engineering Expected May 2023

B.S. in Information Science, Systems, and Technology

Primary Concentration in Data Science, Secondary Concentration in Networks, Crowds, and Markets

Selected Coursework: Machine Learning, Artificial Intelligence, Linear Algebra, Differential Equations, Object-Oriented Programming & Data Structures, Game Theory, Probability & Statistics, Network Theory, Market Design, Econometrics

PROFESSIONAL EXPERIENCE

Amazon, *Data Engineer Intern* Summer 2022

- Contributed to data lineage project for PerfectMile, the company's 3rd largest internal data provider
- Created SQL parser to determine table lineage of complex data processing jobs with AWS Lambda

BaseballCloud, *Research & Development Intern* Spring 2022

- Developed collegiate xStat nearest-neighbor models with scikit-learn (writeup available upon request)
- Created Flask microservice to productionize the above, wrote code to assist with model maintenance
- Authored park factor, wRC+, and other common advanced metric calculations to augment core products

Cornell Baseball Analytics Organization Fall 2021 - Spring 2022

- Attempted to start a student-run analytics organization for Cornell Baseball, modernize team operations
- Created interactive advance scouting reports with Streamlit using data and video from Synergy
- Assembled a centralized cloud database, successfully advocated for player access to Rapsodo data

Cameo, *Data & Analytics Intern* Summer 2021

- Designed and built ETL pipeline for Google Trends data with Airflow after researching data's potential utility
- Prototyped library to rapidly evaluate the predictiveness of third-party data against internal metrics
- Wrote intensive SQL queries for product A/B testing, talent classification, and other ad-hoc reporting

FEATURED PROJECTS

collegebaseball: an open-source college baseball analysis package with functionality for data acquisition and metric calculations, including a scraper for stats.ncaa.org. Accepted for poster presentation at the 2022 Carnegie Mellon Sports Analytics Conference, and currently undergoing inclusion into the [SportsDataverse](#)

Abstract: a public web application for D1 baseball stats with interactive leaderboards, percentile rankings, and more

SPECIALIZED SKILLS

Programming Languages: Python (proficient), SQL (proficient), R (familiar), Julia (Beginner), Java (familiar)

Libraries/Frameworks: scikit-learn, PyTorch, NumPy, Pandas, Streamlit, Matplotlib, Flask, Airflow, tidyverse

Computing/Software: Git, AWS, Figma, Microsoft Office Suite, Docker, LaTeX, Gurobi

Natural Languages: English (native), Spanish (intermediate), Latin (proficient)

LEADERSHIP EXPERIENCE

Sigma Alpha Mu Fraternity, Beta Chapter, *Exchequer, House Manager*

Men's Rugby Football Club, *Treasurer, Vice President for Public Relations*