Joey Kaminsky

jkaminsky@ucsd.edu • (224) 500-1842 • linkedin.com/in/joey-kaminsky • jkaminsky2.github.io/portfolio-website

EDUCATION

Bachelor of Science in Data Science, *University of California San Diego*   *Anticipated Graduation: June 2025*
    Minors in Cognitive Science and Finance
GPA: 3.99 / 4.00
- Awarded Provost Honors for all *10 completed quarters*.
- *Relevant Courses:* Data Structures & Algorithms, Blockchain, Scalable Analytics, Data Mining, Database Management, Advanced Probability, Web Design, Data Visualization, Data Science Foundations, Linear Algebra, Machine Learning, Business Analytics.

EXPERIENCE

National Aeronautics and Space Administration | *Data Analytics Intern*   *May 2024 – Present*
- Engineered an **R**-based fraud detection ETL pipeline in **Azure Databricks** analyzing *$7B+* in grants, integrating *25 dynamic data sources* and identifying *15 distinct fraud patterns*, resulting in *90% faster analysis time*.
  - Uncovered *20 potential fraud cases*, leading to investigations with value of *$5M+*.
- Created a comprehensive internal dashboard with **Power BI**, **Python**, and **SQL**, offering real-time, actionable insights and analysis.
- Developed and held an innovative **Power BI Python/R** scripting workshop to *100+ colleagues*, achieving *90% positive feedback*.

Bioregional Center at UC San Diego | *Data Analytics Specialist*   *October 2023 – September 2024*
- Spearheaded 4 *analytical projects* as technical lead for team of *5 urban studies professors and scientists*; supervised *3 team members* and served as the primary contact for stakeholders.
- Created a machine learning model using **ArcGIS**, **SQL**, and **Python**; identified *50 locations* in San Diego for community gardens.
- Analyzed *1,000+ acres* of agricultural land in the San Joaquin Valley with image recognition algorithms, locating *300+ hedge rows*.
- Executed multi-faceted field research projects regarding vacant lots, urban forestry, and farmers' markets, synthesizing findings using **Python** and **SQL** to create actionable insights.
- Delivered technical presentations to *300+ UCSD students* across *3 classes* and created educational content reaching *150+ students*.

San Diego Supercomputer Center | *NSF Research Experience for Undergraduates Intern*   *December 2022 – October 2023*
- Architected and deployed a **Python**-based analysis platform on **Kubernetes**, enabling *20+ analytical workflows* for processing complex datasets—namely image, network, and spatial analysis.
- Engineered a tailored image collection system using **Python**, implementing web scraping and facial recognition machine learning models, which processed *75,000+ images* with **Spark**; research presented at the *Our Water Data Future* conference.
- Streamlined COVID-19 screening operations for San Diego Unified School District through data-driven optimization, achieving *99.8% accuracy in medical recommendations* for *10,000+ users*.
- Wrote technical documentation for **SuAVE** visualization software and aforementioned projects using **JavaScript**, **HTML**, and **CSS**.

MedSpan Research | *Database Programmer*   *Aug 2021 – July 2022*
- Implemented row-level security protocols for sensitive client databases containing *30,000+ records*, achieving *0 security issues*.
- Leveraged **Salesforce** and **Power BI** to analyze customer engagement data, resulting in *25% increase in successful client outreach*.
- Designed and executed *90+ custom surveys* using **Qualtrics**, implementing complex **Excel** mapping algorithms that circumvented the need for data processing.

EXTRACURRICULAR INVOLVEMENT

Deloitte Mentorship | *Project Lead*   *March 2024 – June 2024*
- Engineered a geographic product recommender system using **Python** and machine learning, analyzing *$2M+ cash flow* data points to generate *$200K+ in projected revenue* opportunities.
- Received executive recognition for "flawless, realistic execution and valuable insights".

Data Science Student Society Consulting | *Data Consultant*   *March 2024 – June 2024*
- Migrated data from legacy system to relational database; created and utilized **Python** script to prevent possible conversion issues.
- Initialized a dynamic athlete performance rating system within **AWS** using a data-update scheduler for automatic updates.

San Diego Supercomputer Center | *Content Writer, Intern Team Lead*   *January 2023 - December 2023*
- Published *3 technical articles* on data science and technology, garnering *1,000+ views*.
- Mentored *7 high school interns* in data science methodologies, resulting in *2 publications* and *3 interns further pursuing the field*.

ARTICLES AND PUBLICATIONS

Articles
- *COVID-19 e-Decision Tree Continues Keeping K-12 Students Safe*   *June 2023*
  - https://www.sdsc.edu/News Items/PR20230712_e-Decision_Tree.html

Publications

- Fernald, A., Atkins, A., Kirkpatrick, C., Tang, C., **Kaminsky, J.**, & Zaslavsky, I. "The connectedness of transboundary groundwater researchers: from network analysis to network-building." The International Association of Hydrogeologists (2023): Page 135.
  - https://pureportal.spbu.ru/files/114659659/2_IAH_Programme_Oral_Abstracts.pdf

PERSONAL PROJECTS

BlockMarket: Ethereum Data & Data Analysis Marketplace   *2024*

- Decentralized marketplace for data and analytics scripts, solving longstanding seller safety and sensitive information leak issues.
- Integrated a data protection system using **AWS API Gateway** and **Lambda**, scrubbing *20 types of sensitive and malicious data*.
- Developed a comprehensive reliability framework incorporating blockchain-based reputation scoring, automated data quality checks, and smart contract enforcement, resulting in higher buyer trust ratings.

NBA 2K Rating Predictor - Does All-Star Voting Impact Player Rating?   *2023*

- Wrote a **Python** predictive analytics model using web-scraped NBA data, achieving *86% accuracy* in modeling NBA 2K player ratings through advanced linear regression; supported claim of a correlation between player popularity and game ratings.

Python Personalized Stock Tracker   *2022*

- Developed an automated stock analysis platform leveraging **machine learning** and **NLP**, processing financial articles daily from *3 sources* to identify market trends with *62% accuracy*.
- Configured a serverless **AWS** pipeline (**S3**, **Lambda**, **CloudWatch**) to deliver hourly portfolio analytics.

SKILLS & INTERESTS

Programming Languages: ACL, CSS, HTML, Java, JavaScript, Python, R, Solidity, SQL
Tools & Technologies: AWS (Cloutdwatch, EC2, Lambda, and S3), Azure Databricks, Blockchain, Data Visualization (ArcGIS, Power BI, SuAVE, and Tableau), Docker, Excel, Git, Github, Kubernetes, Limesurvey, Linux, MongoDB, Node.js, Polygon, Qualtrics, React.js, Salesforce
Python Packages: Apache Spark, Dask, GeoPandas, Keras, Matplotlib, NetworkX, NumPy, OpenCV, Pandas, Seaborn, Sklearn, Tensorflow
Interests: Weightlifting, Reading, Football, Basketball (Coaching and Playing), Softball (3x Intramural Champion), Spikeball, Investing