Pramukh Shankar

858-261-4166 | pramukhshankar28@gmail.com | GitHub | Portfolio | Medium | linkedin.com/in/pramukhshankar/

Education & Skills

New York University, Tandon School of Engineering — Starting Aug. 2026
M.S. Urban Data Science — *New York, NY*

University of California, Santa Barbara — Sept. 2021 – June 2025
B.S. Statistics & Data Science; B.A. Economics; Minor in Philosophy — *Goleta, CA*

Languages & Tools: Python, R, SQL, SAS, GIS, HTML, STATA, Tableau, SPSS, Excel
Libraries & Frameworks: NumPy, Pandas, Scikit-learn, PostgreSQL, Langchain, Next.js, FastAPI, TensorFlow, PyTorch, Keras

Experience

Team Leader — May 2025 – Nov. 2025
AmeriCorps NCCC Forest Corps — *Sacramento, CA*

- Led and coordinated a 7-person field team through demanding conservation projects, managing daily operations, task delegation, scheduling, housing logistics, team budgeting, and communication to ensure projects were completed efficiently and safely
- Navigated challenging team dynamics and high-pressure field conditions by addressing conflicts professionally, facilitating team meetings, maintaining morale, and enforcing safety protocols, PPE compliance, and operational guidelines across all project activities
- Collaborated with nonprofits and federal environmental agencies to coordinate project logistics, communicate progress, and support event and restoration planning while leading physically intensive work including invasive species removal and natural forest regeneration efforts

Research Assistant — Sept. 2024 – June 2025
Health Innovation Lab, UCSB — *Goleta, CA*

- Analyzed 100K+ cancer-trial and carcinogen-exposure records using Python (Pandas, NumPy), and statistical modeling to quantify facility-level environmental exposure and its association with county-level cancer incidence
- Conducted comparative analytics on 500+ U.S. vs. China drug approvals using SQL, timeline analysis, and visualization tools to model differences in FDA review durations.
- Identified statistically significant correlations for 10+ chemicals linked to county-level cancer rates using regression models; narrowed dataset from 200+ candidate chemicals to a shortlist of 14 high-risk compounds

Research Assistant — June 2023 – Sept. 2023
American Institute for Behavioral Research and Technology — *San Diego, CA*

- Applied effect size analysis (Cohen's d), z-tests, and Likert-scale modeling to experimental data from 400+ participants to quantify demographic susceptibility to online influence; engineered an R- and Excel-based processing pipeline for CSV and survey data, cutting analysis time by 40%
- Aided in development of APA-submitted manuscript; contributed to literature review, journal selection, and revision workflows
- Built a vote-shift metric and reproducible calculator using Python and SQL, showing how search results can sway up to 20% of votes

Data Intern — June 2022 – Sept. 2022
Mucci Assessment — *San Diego, CA*

- Developed Python pipeline (NumPy, Pandas) to scrape and clean 500+ pediatrician contacts, reducing manual data entry by 10+ hours per week
- Developed and deployed a backend scheduling application using Python and SQL, reducing manual processing time by 60% and saving 15+ staff hours per week.

Projects

Run For Office with Nationbuilder

- Built a domain-whitelisted Playwright crawler for county sites nationwide, capturing 1K+ HTML/PDF links per run and replacing Google Search API workflows—cutting data collection time by 80%.
- Integrated PyMuPDF and PDFPlumber to standardize PDF processing, facilitating LLM extraction across diverse document formats and reducing data errors by 15% company-wide
- Built an automated web scraper to aggregate county-level URLs, enabling efficient data collection and analysis workflows in Python
- Designed grounded LLM extraction at the county level for open offices (OpenAI o1), cutting nationwide collection costs by 70% ($35,000+)

Pokémon Type Prediction & Analysis

- Trained Random Forest, Decision Tree, LDA, and KNN models on 800+ entries with 20+ attributes, achieving ROC_AUC of 0.69
- Achieved 28.2% prediction accuracy, outperforming the 5.6% random baseline by 5×
- Implemented an interactive Tableau dashboard using Kaggle datasets to dynamically visualize and rank Pokémon type matchups

Classifying Tennis Player Styles

- Created a Render-hosted interactive web app that scrapes and analyzes pro tennis match data, classifies player styles from rally length, shot frequency, and winner types, and benchmarks performance against tour averages for real-time insights
- Applied k-means, DBSCAN, and fuzzy c-means clustering to segment player styles from match features

Leadership

President — Oct. 2021 – June 2025
UCSB Club Tennis — *Goleta, CA*

- Managed club operations for 50+ students, coordinating fundraisers, social events, competitive practice sessions, travel logistics, meetings, and tryouts