

## PROFESSIONAL SUMMARY

Senior data scientist and software engineer specializing in geospatial machine learning and large-scale demographic analysis. Developed algorithms that improved demographic classification accuracy from 23% to 64%, processed data across 178,000+ precincts, and built platforms serving thousands of analysts nationwide.

## KEY ACHIEVEMENTS AND IMPACT

Discovered systematic race coding errors affecting all Black and Asian-American voters

## CORE COMPETENCIES

Software Engineering • Data Engineering • Data Analysis • Geospatial / Demographic Expertise • Research & Analytics • Programming & Development • Data Infrastructure

## PROFESSIONAL EXPERIENCE

### **Siege Analytics | Founder & Principal Data Scientist - Austin, TX 2012 - Present**

#### **Data Science & Political Analytics**

- Identified decades of systematic demographic miscoding in national voter databases
- Developed geospatial machine learning algorithms that improved automated demographic classification accuracy from 23% to 64% (178% improvement)
- Applied meta-analytical approaches to detect and correct population-scale demographic errors
- Corrected systematic bias affecting millions of voters across all US electoral districts
- Built validation frameworks ensuring demographic accuracy across 178,000+ precincts

### **NGP VAN | Senior Software Engineer - Washington, DC 2012 - 2015**

#### **Political Technology & CRM Systems**

- Maintained geospatial analysis tools for Java-based CRM system used by tens of thousands simultaneously
- Developed custom tile server enabling interactive visualization improving contact rates by 53% and segmentation accuracy by 88%
- Built advanced geospatial analysis capabilities using Java, JavaScript, MySQL, and TileMill
- Integrated mapping and visualization tools for political campaign data analysis

### **PCCC | Research Director - Washington, DC 2010 - 2012**

#### **Political Research & Data Analysis (FLEEM System)**

- Conceived, architected, and engineered FLEEM web application using Twilio API
- Handled tens of thousands of simultaneous phone calls using emulated predictive dialer
- Built IVR polling system supporting Senators Martin Heinrich and Elizabeth Warren
- Developed survey deployment system facilitating thousands of simultaneous surveys
- Saved PAC \$840,000 annually in polling costs through automated infrastructure

## KEY PROJECTS

**National Redistricting Platform (2020 - 2021)**

Cloud-based GeoDjango platform for redistricting analysis with real-time collaborative editing and Census integration, used by thousands of analysts nationwide

Technologies: GeoDjango, PostGIS, AWS, Docker, React, Python

Impact: Reduced mapping costs by 73.5%, saving organizations \$4.7M in operational expenses

**FLEEM Political Polling System (2010 - 2012)**

Completely self-built IVR system using Twilio API that contacted tens of thousands of voters daily, replicated call center functionality to performance parity

Technologies: Twilio API, Python, Django, PostgreSQL, JavaScript

Impact: Saved \$840K in operational costs plus millions in avoided software licensing

**Geospatial Demographic Classification System (2013 - 2016)**

Machine learning platform that discovered systematic coding errors and improved demographic classification accuracy from 23% to 64%

Technologies: Python, Scikit-learn, PostGIS, GeoPandas, TensorFlow

Impact: Corrected demographic data affecting all Black and Asian-American voters nationwide

**Polling Consortium Dataset Meta-Analysis (2013 - 2016)**

Comprehensive meta-analysis of polling data from tens of polling and mail firms with different methodologies and encoding systems

Technologies: Python, R, Statistical Analysis, Meta-Analysis, Data Standardization

Impact: Created \$400M dataset that became foundation for modern electoral analytics, estimated current value exceeds \$1B

**EDUCATION**

Bachelor of Arts in Plan II Honors - University of Texas at Austin (Austin, TX) | 2008

Honors: Interdisciplinary liberal arts program

**TECHNICAL SKILLS**

**DATA ENGINEERING** *Apache Spark, PySpark, Dask* (Large-scale data processing and distributed computing); *dbt, Airflow* (Data transformation pipelines and workflow orchestration); *PostgreSQL/PostGIS, Snowflake* (Database design and geospatial data management); *AWS (EC2, RDS, S3), Docker* (Cloud infrastructure and containerization)

**PROGRAMMING** *Python* (NumPy, Pandas, Scikit-learn, Django, Flask (15+ years)); *SQL* (Complex queries, optimization, spatial analysis (15+ years)); *R* (Statistical modeling, ggplot2, spatial packages (12+ years)); *JavaScript* (React, D3.js, Node.js, real-time applications (10+ years))