

## PROFESSIONAL SUMMARY

Political research and redistricting expert with 15+ years building systems that matter. Discovered systematic demographic coding errors affecting all Black and Asian-American voters, developed geospatial ML algorithms improving classification accuracy from 23% to 64%. Expert in electoral analysis and redistricting optimization.

## KEY ACHIEVEMENTS AND IMPACT

Predictive excellence: Utilized advanced sampling methods to decrease survey margin of error from  $\pm 4.2\%$  to  $\pm 2.1\%$  • Increased voter turnout prediction accuracy from **71%** to **87%** • Executive authority: Briefed Presidents, Congressmen, Senators, Governors on election integrity, voter sentiment and postmortem analysis • Methodological advancement: Improved segmentation accuracy **34%** and survey incidence **28%**

## CORE COMPETENCIES

Programming and Development • Machine Learning & AI • Data Infrastructure

## PROFESSIONAL EXPERIENCE

### Siege Analytics | Partner - Austin, TX 2005 - Present

#### Data Science & Political Analytics

- Discovered systematic race coding errors affecting all Black and Asian-American voters, developed geospatial machine learning algorithms improving demographic classification accuracy from **23%** to **64%**
- Built redistricting platform used by thousands of analysts nationwide with real-time collaborative editing and Census integration
- Utilized advanced sampling methods to decrease survey margin of error from  $\pm 4.2\%$  to  $\pm 2.1\%$ , increasing voter turnout prediction accuracy from **71%** to **87%**, and ensuring survey results more closely reflected true population attitudes
- Trigonometric algorithm for boundary estimation reduced mapping costs by **73.5%**, saving campaigns and organizations **\$4.7M** and enabling smaller nonprofits to conduct analysis
- Built real-time FEC analysis systems using Python, Pandas and PySpark to detect likely fraud, money laundering and financial crimes across billions of records daily, performing time series analysis on trillions of records in the political spending sub-economy valued over **\$2 trillion**

### PCCC | Research Director - Washington, DC August 2011 - August 2012

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

- Conceived, architected, and engineered FLEEM web application using Twilio API handling tens of thousands of simultaneous phone calls using emulated predictive dialer for regulated political surveys
- Developed IVR polling system for early quantitative research supporting Senators Martin Heinrich and Elizabeth Warren
- Built comprehensive tabular and graphical reporting system with Python, GeoDjango, PostGIS, and Apache webserver

### Myers Research | Senior Analyst - Austin, TX 2012 - 2014

#### Political Research & Analysis

- Designed comprehensive survey instruments for specialized voting segments and niche markets
- Developed sophisticated analytical products and reports that delivered actionable insights to clients
- Co-developed a web application to manage all aspects of survey operations, from instrument design to data collection and analysis

## Lake Research Partners | Programmer - Austin, TX 2008

### Political Polling & Research

- Designed questionnaires and analyzed data for complex market research studies across diverse industries
- Conducted statistical modeling and analysis to address multifaceted consumer behavior questions
- Pioneered the integration of advanced mapping techniques into standard reports, including choropleths and hexagonal grid maps

## The Feldman Group | Field Director - Austin, TX 2011 - 2012

### Political Campaign Management

- Managed all aspects of survey fielding for a multi-million dollar research firm, including scheduling, oversight, sampling, and quality control
- Developed and implemented data warehousing solutions for efficient storage and retrieval of research findings
- Created custom reports and data visualizations based on specific client requirements

## KEY PROJECTS

### FLEEM Political Polling System

**About:** Web application using Twilio API for regulated political surveys, handling tens of thousands of simultaneous calls with predictive dialer functionality

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

**Impact:** Saved PAC \$840,000 annually in polling costs while significantly improving data collection efficiency

### National Redistricting Platform

**About:** Cloud-based GeoDjango platform for redistricting analysis with real-time collaborative editing and Census integration, used by thousands of analysts nationwide during 2021 redistricting cycle

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

**Impact:** Reduced mapping costs by 73.5%, saving organizations \$4.7M in operational expenses. Served 12,847 analysts across 89 organizations.

### Geospatial Demographic Classification System

**About:** Machine learning platform for demographic analysis that discovered systematic coding errors and improved classification accuracy from 23% to 64%

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

**Impact:** Corrected demographic data affecting all Black and Asian-American voters, improved electoral prediction accuracy by 22%

## TECHNICAL SKILLS

**PROGRAMMING AND DEVELOPMENT** *Python* (15+ years: NumPy, Pandas, Scikit-learn, TensorFlow, Django, Flask, GeoPandas, Asyncio); *R* (12+ years: Statistical modeling, ggplot2, dplyr, spatial packages (sf, sp), Shiny); *SQL/PostGIS* (15+ years: PostgreSQL/PostGIS, MySQL, complex spatial queries, optimization, database design); *JavaScript* (10+ years: React, D3.js, OpenLayers, Node.js, real-time applications, WebSockets); *Java* (8+ years: Enterprise applications, Spring framework, geospatial libraries (GeoTools)); *Other Technologies* (Shell scripting, Git, Docker, Kubernetes, infrastructure as code)

**MACHINE LEARNING & AI** *ML Frameworks* (Scikit-learn, TensorFlow, PyTorch, XGBoost, LightGBM); *Geospatial ML* (Spatial feature engineering, geographically weighted regression, spatial clustering); *Techniques* (Classification, regression, ensemble methods, time series, NLP, computer vision); *Validation* (Cross-validation, A/B testing, statistical significance, model interpretability)

**DATA INFRASTRUCTURE** *Processing* (Apache Spark, PySpark, Dask, parallel computing, distributed systems); *Pipelines* (Airflow, dbt, ETL design, data quality monitoring, automated testing); *Storage* (Data warehousing, data lakes, columnar storage (Parquet), data modeling); *Streaming* (Kafka, Redis, real-time processing, event-driven architecture)