GitHub: github.com/dheerajchand

 $[\mathsf{RESEARCH}, \mathsf{ANALYSIS}, \mathsf{ENGINEERING}] \to \mathsf{UNDERSTANDING}$ 

Austin, TX (30.2672°N, 97.7431°W)

# PROFESSIONAL SUMMARY

Data scientist and software engineer with 15+ years building systems that matter. Discovered 2.7M misclassified voters, saved organizations \$4.7M through better algorithms, and built platforms used by thousands of analysts nationwide. Expert in translating complex analytical requirements into scalable technical solutions.

### **KEY ACHIEVEMENTS AND IMPACT**

Discovered 2.7M misclassified Democratic voters through data analysis • Handled billions of records with millions of columns for tens of thousands of users

### **CORE COMPETENCIES**

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

#### PROFESSIONAL EXPERIENCE

# Siege Analytics | Partner - Austin, TX 2020 - Present

# **Data Science & Political Analytics**

- Lead comprehensive research initiatives for presidential, gubernatorial, and congressional campaigns affecting millions in strategic spending decisions
- Uncovered decades of demographic miscoding in voter files, discovering 2.7M previously mischaracterized Democratic voters
- Developed Python boundary estimation algorithm enabling mapping and analysis at every level of election in the United States
- Algorithm reduced mapping costs by 73.5%, saving campaigns and organizations \$4.7M and enabling smaller nonprofits to conduct redistricting analysis
- Architect enterprise-scale cloud data warehouse solutions on AWS processing billions of records for electoral analytics
- Design and implement scalable ETL pipelines using PySpark, dbt, and PostgreSQL/PostGIS for large-scale geospatial datasets
- Develop advanced analytical tools and machine learning algorithms using Python, Pandas, NumPy, and Scikit-learn for fraud detection

# Lake Research Partners | Senior Data Scientist - Washington, DC 2018 - 2020 Political Research & Analytics

- Built the first collaborative and multi-actor contributed poll of polls used by the Democratic Party
- Developed system that later became the Polling Consortium Database at The Analyst Institute
- Worked on all aspects of questionnaire design, sampling, reporting and analysis for Congressional, Senate and Presidential elections affecting millions in campaign spending decisions
- Conducted statistical modeling and analysis using SPSS, ArcGIS, Quantum GIS, GRASS, Stata, PostgreSQL, PostGIS, and Oracle
- Pioneered integration of advanced mapping techniques into standard reports including choropleths and hexagonal grid maps
- Trained staff on building Python tooling for report generation and analysis

GitHub: github.com/dheerajchand

[RESEARCH, ANALYSIS, ENGINEERING]  $\rightarrow$  UNDERSTANDING

Austin, TX (30.2672°N, 97.7431°W)

# The Praxis Project | Technical Director - Oakland, CA 2015 - 2018

# **Technology & Data Infrastructure**

- Led technology operations for multi-million dollar organization while assisting in search for full-time CTO
- Directed all technology decisions and practices for massive multinational non-governmental organization
- · Led training initiatives for beneficiaries on spatial and Census data analysis for public health research
- Conducted training programs for NGO staff in web development using Drupal, PHP, and MySQL
- Managed technology infrastructure supporting community health initiatives across multiple countries
- Architected and developed 25 Drupal sites to integrate with membership databases, activism CRMs and government agencies, under guidelines from Kellogg Foundation and Robert Wood Johnson Foundation

# Salsa Labs | Senior Software Engineer - Washington, DC 2012 - 2015

# **CRM & Data Platform Development**

- Maintained and extended comprehensive geospatial analysis and reporting tools for Java-based CRM system used by tens of thousands of users simultaneously
- Developed custom tile server for Web Map Service (WMS) integration using GeoTools and OpenLayers,
  enabling interactive visualization of CRM and Census data that improved 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 interfacing with Government and Activism APIs
- · Handled billions of records with millions of columns in high-performance CRM system

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

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

- Conceived, architected, and engineered FLEEM web application using Twilio API handling tens of thousands of 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 tabular and graphical reporting system with Python, GeoDjango, PostGIS, and Apache webserver
- Designed survey deployment system facilitating thousands of simultaneous phone surveys, saving \$840K in operational costs plus millions in avoided software licensing through completely self-built system
- Managed comprehensive research operations for progressive political initiatives and candidates

# The Feldman Group | Data Analyst - Washington, DC 2008 - 2010

### **Political Data Analysis**

- Administered all quantitative and qualitative research operations for presidential, gubernatorial, congressional, and senatorial campaigns affecting millions of dollars in spending decisions
- Managed team of 6 research analysts and field staff for comprehensive survey fielding at multi-million dollar research firm
- 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
- Enhanced value of research deliverables through advanced analytical techniques using SPSS, OSCAR, PHP, and MySQL

# **KEY PROJECTS**

GitHub: github.com/dheerajchand

 $[\mathsf{RESEARCH}, \mathsf{ANALYSIS}, \mathsf{ENGINEERING}] \to \mathsf{UNDERSTANDING}$ 

Austin, TX (30.2672°N, 97.7431°W)

# **Ballista Redistricting Platform (2020 - Present)**

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 Survey System (2010 - 2012)

Web application using Twilio API for regulated political surveys, handling tens of thousands of simultaneous calls

Technologies: Twilio API, PHP, MySQL, JavaScript, Drupal

Impact: Saved PAC \$840,000 annually in polling costs

### Multi-tenant Data Warehouse (2015 - 2018)

Longitudinal analysis platform across attitudinal, behavioral, demographic, and geographical dimensions

Technologies: PostgreSQL, Python, Django, D3.js, Tableau

Impact: Processed billions of records for electoral analytics and demographic analysis

# 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, creating unified analytical framework

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

### **TECHNICAL SKILLS**

**PROGRAMMING AND DEVELOPMENT** *Python* (Django/GeoDjango, Flask, Pandas, PySpark, SciKit-Learn, TensorFlow); *JVM Languages* (Scala (Spark), Java, Groovy); *Web Technologies* (JavaScript, React, d3.js, PHP, HTML/CSS); *Database Languages* (SQL, T-SQL, PostgreSQL/PostGIS); *Statistical Computing* (R, SPSS, SAS, Stata)

**DATA INFRASTRUCTURE** *Cloud Platforms* (AWS (EC2, RDS, S3), Google Cloud Platform, Microsoft Azure); *Big Data* (Apache Spark, PySpark, Hadoop, Snowflake, dbt); *Databases* (PostgreSQL/PostGIS, MySQL, Oracle, MongoDB, Neo4j); *Geospatial* (ESRI ArcGIS, Quantum GIS, GeoServer, OSGeo, GRASS); *DevOps* (Docker, Git, CI/CD pipelines, automated testing, version control)

**RESEARCH AND ANALYTICS** Survey Methodology (Design, sampling, weighting, longitudinal analysis); Statistical Analysis (Regression, clustering, classification, time series); Machine Learning (Supervised/unsupervised learning, neural networks, NLP); Geospatial Analysis (Spatial statistics, cartography, remote sensing); Research Management (Team leadership, methodology design, stakeholder communication)

**PLATFORMS** *Tableau* (Advanced (Data Visualization, Dashboards, Server)); *Power BI* (Intermediate (Business Intelligence, Reports)); *Jupyter* (Advanced (Data Science, Notebooks, Visualization))

TRACK Git (Expert (Version Control, Collaboration, CI/CD)); Jira (Advanced (Project Management, Issue Tracking))