Tony Shen

(202) 251-4340

Atlanta, GA 30332

tonyshen311@gmail.com

Select Experience

U.S., Department of Housing and Urban Development   Washington, DC

Program Data Analyst   Sep. 2023 – *present*

- Engineer machine learning framework using lasso **regression** on HMDA loan data to support fair housing investigations, enabling data-driven assessments of discriminatory underwriting practices
- Design and deploy a **Python**-based BISG model and internal tool, enabling field workers to statistically impute race-and-ethnicity from names and addresses, improving fair housing investigations by reducing individual manual analysis time by **80%**
- Develop and implement **R shiny** dashboards with geospatial visualizations to communicate data-driven insights, supporting leadership decision-making on investigations and policy enforcement
- Architect and maintain an **MS SQL Server** database to centralize disparate datasets, leveraging **Python** automation to generate compliance reports, reducing manual reporting time by **70%** and improving data consistency
- **Research** and identify key variables to strengthen the Education Quality Index used in housing policy, advising on new data sources to improve the assessment of school equity and accessibility
- Develop **PowerBI** dashboards to visualize key fair housing metrics, improving HUD leadership's ability to track disparities and enforcement actions
- Create interactive geospatial maps in **ArcGIS Pro** to analyze fair housing complaints, supporting regional investigations

The Civics Center   Los Angeles, CA

Research and Technology Analyst   Jun. 2023 – Sep. 2024

- Developed **Python** web-scraping scripts to automate voter file data collection, expanding research coverage across jurisdictions by **96%**, eliminating manual data extraction
- Engineered and managed a **Google BigQuery** database, migrating all organizational data to a centralized platform, enhancing accessibility, and enabling advanced analytics for policy research
- Designed and led **A/B testing** for voter registration outreach, optimizing postcard campaign strategies to maximize engagement

Summit Consulting LLC   Washington, DC

Data Science Consultant (Contestant)   Mar. 2023 – Apr.2023

- Optimized legal resource allocation by developing an **exponential smoothing time-series** model to predict MoM query volume, enabling a 30% improvement in volunteer lawyer scheduling efficiency
- Enhanced platform engagement by applying **natural language processing** to analyze 400,000+ attorney-client conversations, uncovering key sentiment trends and insights to improve service quality
- Designed and presented a **Tableau** dashboard to translate complex predictive modeling and sentiment analysis results into clear, actionable insights, enhancing stakeholder understanding of key legal service trends.

Education

Georgia Institute of Technology   Atlanta, GA

Master of Science in Analytics, Computational Data Analytics concentration, GPA: 4.00   part-time

University of Maryland   College Park, MD

Bachelor of Science in Government and Politics: Honors Program, GPA: 3.87   2023

Inaugural Data Curation Fellow | Undergraduate Teaching Assistant: *Scopes and Methods, Social Science Research*

Skills

Technical: R, Python (Pandas scikit-learn), A/B Test, SQL, Excel, ArcGIS Pro, Tableau, PowerBI, OOP, Git, BigQuery

Languages: English, Chinese Mandarin (Fluent)