mauricefarber@gmail.com
07405 639354

Maurice Farber

LinkedIn

Summary

Spatial data scientist with 7+ years of experience working on production-grade geospatial pipelines, ML-driven validation workflows, and data and analytics products for public sector, infrastructure, and sales use cases. Deep expertise in spatial data science, including PostGIS, GeoPandas, h3, and network analysis, underpinned by a Masters in Geographic Data Science from LSE. Proven track record of engineering robust, cloud-based data pipelines with BigQuery, Airflow, and Databricks that handle multi-source datasets and satisfy user requirements. Fluent in crafting enterprise strategy and guiding stakeholders through ideation to implementation. Grounded in the maths behind maps, including spatial indexing, graph-based reasoning, network analysis, and rigorous ML evaluation.

Skills

Spatial Data Science: PostGIS, GeoPandas, h3, network analysis, spatial indexing, topology validation, spatial ETL, ArcGIS, QGIS
Machine Learning: scikit-learn, XGBoost, random forests, feature engineering, hyperparameter tuning, embeddings, cross-validation
Pipelines: BigQuery, Apache Airflow, Databricks SQL, Azure DevOps, ETL design, QA/QC automation, reproducibility standards
Cloud & Tooling: Python, R, BigQuery, PostgreSQL, Azure (DevOps, Databricks), Git, Tableau, PowerBI

Education

London School of Economics
M.Sc. in Geographic Data Science (Merit) | Sep 2023 – Dec 2024

The University of Chicago
B.A. in Geography (Honors) & Statistics (Minor) | Sep 2013 – Jun 2017

Work Experience

LSE Cities — London, UK (Hybrid)
Data Scientist, Geospatial Analytics — *Jun 2025 – Present*

- Engineered large-scale spatial ETL pipelines (Python + Airflow) ingesting OSM, satellite imagery, and multi-source geodatabases, implementing automated QA/QC checks, anomaly detection and reproducibility standards to maintain data accuracy across pipeline runs.
- Designed ML-driven data validation and feature-generation workflows (scikit-learn) incorporating embedding-based similarity matching, spatial integrity checks, and iterative refinement, via Git-based version control to produce stable, generalisation-ready inputs.
- Built entity resolution and taxonomy harmonisation pipelines using text parsing, normalisation, and similarity scoring to clean and align semi-structured metadata across external datasets, improving model interpretability and downstream analytical validity.

ATK — New York, NY, USA (Remote)
Senior Consultant, Analytics — *Sep 2022 – Aug 2023*

- Owned data modernization initiative, defining product strategy, data models, and pipeline architecture for an operations monitoring & analytics platform processing large-scale agriculture datasets and leading implementation & onboarding in the executive suite.
- Configured multi-source spatial ETL pipelines (Databricks SQL + Azure DevOps) integrating geospatial and operational datasets with process automation to power reliable dashboards, simplifying operational overview & increasing executive usage of analytics by 250%.
- Conducted scenario testing on Python revenue forecasting models and location planning analysis with QGIS, supporting due diligence for a $10M capital investment decision with interactive dashboards and webmaps, and analytics walk-throughs for senior stakeholders.

Gartner, Inc. — New York, NY, USA (Remote)
Senior Consultant – Territory Planning & Analytics — *Nov 2021 – Aug 2022*

- Applied causal impact analysis (pre/post comparisons) to evaluate $4.5M territory redesign with Excel, using Agile Ceremonies and ArcGIS Online webmaps to coordinate sales leaders, iterate designs, and align with business development targets.
- Built Tableau revenue forecasting dashboards and ArcGIS Online webmaps to evaluate the impact of $15M in sales territory changes, optimising revenue-size cutoffs, and coordinating with product team to tune offerings, improving Local Gov. sales performance by 4%.
- Advised sales leadership on analytics-driven territory strategy, translating model assumptions and commercial impacts to non-technical stakeholders, guiding Go-to-Market strategy and launch of $15-20M Canadian Local Government sales team.

Wellesley Cove (Subcontracting Guidehouse) — New York, NY, USA (Remote)
Senior Consultant – Contractor — *Jul 2021 – Nov 2021*

- Managed $30M in FEMA COVID emergency aid for state governments, performing structured reviews of financial and emergency management workflows, identifying compliance gaps with evolving federal guidelines, and recommending operational improvements.
- Product owner of PowerBI compliance dashboard, querying financial data in Snowflake with SQL, conducting data analysis with Excel, and iterating product improvements with clients that streamline federal audit data preparation process.

Booz Allen Hamilton — Washington, DC, USA
Senior Management Consultant — *2016 & Aug 2017 – Jun 2021*

- Produced time-series analysis of military healthcare data queried with Databricks SQL and accurate spatial layers in ArcGIS Pro for Storymaps to guide stakeholders to support operational decision-making, discovering $1.25M in resource misalignment.
- Owned full lifecycle of data dashboard products in Tableau and PowerBI for multiple federal agencies, gathering requirements from stakeholders, aligning product with dashboard best practices, and ingesting Snowflake and DataBricks data lakes with SQL.
- Visualized & analyzed federal monetary collections infrastructure with ArcGIS Enterprise, regularly presenting Storymaps to stakeholders.

Selected Graduate Degree Projects

- **Masters Dissertation: "Pixels to Pathways"**: Designed a spatial ML pipeline to generate novel public transit routes from US Census data. Applied spatial interpolation, network analysis (street connectivity, walking isochrones), and space constraints to identify feasible alignments. Used random forests with spatial weights for variable selection and evaluated final routes using betweenness centrality and welfare analysis, combining graph-based spatial reasoning (analogous to GNN-style network evaluation) with interpretable ML outputs.
- **Machine Learning for Social Science Final Project**: Predicted U.S. voter turnout using a XGBoost model, selected via Grid Search and cross-validation. Tuned features and hyperparameters to optimise performance, achieving an F1 score of 0.825 and a recall of 0.983, demonstrating rigorous model evaluation, uncertainty quantification, and iterative refinement.