

## **Michael A. Babb, PhD**

[babb\\_mike@outlook.com](mailto:babb_mike@outlook.com) | (206) 295 - 3612

### **Data Scientist**

**University of Washington**

**January 2016 – December 2021**

Design and implementation, in python and R, of data engineering pipelines using linear optimization and regression modelling resulting in a 30-percent increase in the study population (136M to 178M) and an 8-fold increase in complete cases.

Deployment of 155K generalized linear models in a parallelized R computing environment.

Design and implementation in R of parallelized data visualization pipelines.

Data engineering, explanatory regression modelling, and visualization of historical and contemporary data using python, R, and SQLite.

Consultation and instruction in feature engineering, data analysis, geospatial analysis, statistical analysis, machine learning, data visualization, and cartographic visualization using python, R, and ArcGIS.

### **GIS Analyst**

**Zillow.com**

**August 2013 – January 2016**

Authorship of geospatial feature engineering and machine learning algorithms in python and R.

Feature engineering and statistical analyses in support of live-site operations and maintenance using python and R.

Data-driven analysis of the technical feasibility and business-need suitability of dataset acquisition and incorporation into production, at-scale, machine learning pipelines.

Authorship of new geospatial feature engineering algorithms in python and R that led to a 10-fold increase over then-existing algorithm efficiency and coverage.

Data-driven business-needs assessment of the strengths and efficiencies of existing geospatial feature engineering algorithms.

Use of R to perform model diagnostics leading to the identification of biases in production machine learning models.

### **Geographer**

**U.S. Census Bureau**

**April 2012 – October 2014**

Coordinated federal-academic project proposal development, project management, and researcher access to confidential federal statistical data.

Site operation and administration of a Federal Statistical Research Data Center.

Completed build out and received certification to operate a federally compliant Federal Statistical Research Data Center (1 of 16 at the time of opening).

### **Education**

**Ph.D., Geography, University of Washington, 2021**

#### **Technical expertise**

Data Science  
Geospatial Analysis  
Data analysis  
Cartographic Visualization  
Research design and implementation

#### **Software proficiency**

R - 7 years  
Python - 7 years  
SQL - 10 years  
qGIS and ArcGIS - 15 years  
Git - 6 years