# Matthew **McAnear**

Senior Data Scientist

*Ann Arbor, MI*

+1 423-277-0411 | mcanearm@gmail.com | mcanearm

## Professional Experience

### Senior Data Scientist — *San Diego, CA*

Point Predictive — *2021 - Present*

- Fit statistical models for predicting early loan default and chargeoff for auto lenders
- Load customer data into multiple database environments
- Implemented DynamoDB-backed solution for simple database lookups to improve AWS Lambda cold start times
- Write performant SQL for real-time and aggregate historical borrower statistics
- Implemented a PostgresSQL-based probabilistic record linkage algorithm for retrieving prior credit applications of incoming borrowers

### Senior Machine Learning Engineer — *Reno, NV*

Clear Capital — *2020-2021*

- Served as principal maintainer for Clear Capital's AVM, a system that produces 150 million new predictions and 150GB of new data each week.
- Built a Bayesian hierarchical model and accompanying webservice to predict home complexity
- Implemented a model-based error prediction procedure to create value ranges around AVM estimates.
- Implemented reporting dashboards Amazon Quicksight dashboards for model analysis and validation.
- Designed and administer Redshift-based data warehouse to power machine learning models, ETL workflows, and ad-hoc analytical queries.

### Data Scientist II — *Reno, NV*

Clear Capital — *2017-2019*

- Led team of five to streamline original AVM model, reducing build time by over 90% and costs by $4,000 per month.
- Improved AVM performance from last place to industry leader in 6 months (based on absolute mean prediction error).
- Designed and deployed an S3 and DynamoDB backed application for photo cataloguing that manages 35+ terabytes of photos.

### Data Scientist I — *Reno, NV*

Clear Capital — *2015-2017*

- Built an automated valuation model (AVM) on commodity hardware to predict home prices using distributed, highperformance R and PostgreSQL
- Implemented multi-model aggregation system for final prediction of house values on a monthly refresh cycle.
- Built basic webservices in Python, Flask, and AWS Lambda for serving AVM model predictions and internal company data through GET requests.

### Data Scientist — *Philadelphia, PA*

Seer Interactive — *2014-2015*

- Designed and carried out web-based experiment on domain recognition using multivariate hierarchical regression.
- Wrote web-crawlers and multithreaded programs in Python and R for ETL jobs.
- Supported analytics account managers and external clients through project planning, KPI identification, Google Analytics implementation, and automated reporting

## Tools

Languages

- R, Python, PostgreSQL, bash, RMarkdown

Technologies

- AWS, DynamoDB, Redshift, Linux

Packages

- scikit-learn, PyMC3, numpyro, pandas, numpy, data.table, ggplot2, pyspark

## Education

### University of Michigan — *Ann Arbor, MI*

M.S., Applied Statistics — *2022-Expected 2025*

### University of Pennsylvania — *Philadelphia, PA*

M.S, Nonprofit/NGO Leadership — *2013-2014*

### Bucknell University — *Lewisburg, PA*

B.A., Mathematics & Economics, magna cum laude — *2009-2013*