

# Merykokeb Belay

[mb2478@georgetown.edu](mailto:mb2478@georgetown.edu) • <https://mkkbelay.github.io/> • [LinkedIn:merykokeb-belay](#)

## EDUCATION

**Georgetown University** | *M.S. Data Science for Public Policy* | Washington, DC 2020-2022

- **Relevant Coursework:** Data Mining; A/B Testing; Applied Statistical Learning and Modeling; Time Series Forecasting; Relational Database Systems and SQL; Data Structures and Algorithms; Data Visualization

**Manchester University** | *B.S. Biology and Int'l Studies* | North Manchester, IN 2013-2017

## PROJECTS

- **Predicting Access to Healthful Food Retailers:** trained supervised machine learning models (logistic regression; LASSO; random forest) to classify U.S. census tracts into one of three food accessibility categories.
- **Estimating the Causal Effect of Political Party on Ideology:** used a regression discontinuity design to explain congressional ideology as a function of political party.
- **Analyzing flight statistics from U.S. Transportation:** used Amazon RDS and EC2 instances to create a MySQL database with over 400,000 records. Explored/analyzed patterns of flight delays using MySQL Workbench.

## EXPERIENCE

**Amida Technology Solutions** | *Data Science Intern* | Washington, DC June 2021-Aug 2021

- Researched and built an ETL pipeline using Python and MongoDB to support a team of engineers building a healthcare quality forecasting application for hospitals
- Trained and validated time series forecasting models in Python; collaborated with a team of 4 engineers to prototype a dashboard to display model predictions
- Created presentation slides, API documentation, and boilerplate code to eliminate project backlogs

**Georgetown University** | *Research Assistant* | Washington, DC Jan 2021-June 2021

- Led data collection and analysis efforts for two research projects on school closures and welfare benefits
- Ideated a data analytics plan; identified key datasets and presented research strategy suggestions
- Analyzed school closure data; used Python and R to link publicly available datasets to census data on socioeconomic indicators

**American Institutes for Research** | *Research Assistant* | Washington, DC July 2018-June 2020

- Assisted researchers with developing a causal model to explain the institutional and behavioral patterns associated with HBCUs production of baccalaureate graduates in STEM who subsequently earn STEM doctorates.
- Retrieved, cleaned, and analyzed historical postsecondary data for 100 HBCUs using STATA and R; worked with interdisciplinary researchers to identify a feature set
- Assisted with implementing and interpreting causal and predictive models (i.e., structural equation modeling and CART decision tree)
- Partnered with a cross-functional team; drafted reports and visualizations to translate technical output into non-technical recommendations/insights

## SKILLS

- **Programming/Analytics:** Python, R, Stata, MySQL, MongoDB, Git
- **Statistics and Machine Learning:** Causal inference, hypothesis testing, experimental and quasi-experimental research designs, supervised and unsupervised machine learning
- **Data Visualization:** Tableau, ggplot, streamlit, RShiny