# BEN MEIER

⚲ Chicago, IL
☎ 608-620-4620
✉ *meier_benjamin@yahoo.com*
🔗 *https://bennnyys.github.io/Ben-Meier/*

## SUMMARY

A data scientist passionate about leveraging data and technology to improve socioeconomic, environmental and political issues. I have 6 years of experience working with as programmer to leverage data and technology to inform economic research, policy and financial regulation.

## EDUCATION

**Master of Information and Data Science, University of California, Berkeley**    📅 2022 – Present
- GPA 4.0 of 4.0, Fall 2023 expected graduation

**BS, Applied Math, Industrial Engineering, Iowa State University**    📅 2011 – 2016
- GPA 3.9 of 4.0

## TECHNICAL SKILLS

**Programming**:  Python (pandas, NumPy, psycopg2, SciPy, Scikit-learn, Keras TensorFlow, PyTorch, Beautiful Soup), Stata, R (RStudio, tidyverse, ggplot2, dplyr, shiny), SAS, AWS sagemaker, AWS S3, Google colab
**Data skills**: Databases (PostgreSQL, SQL Server, SSIS), Anaconda, Git, VS Code, Visual Studio, Jira, Agile
**Data engineering:**  AWS, Linux, Docker, PostgreSQL, Neo4j, Mongo, Redis, Kafka, Zookeeper, NGINX
**Visualization**:  Tableau, Power BI, Python (Matplotlib, Seaborn, Plotly), HTML, CSS
**Methods**:  Machine learning (clustering, K means, K nearest neighbors, decision trees, random forest, naive Bayes, neural networks, CNN), Statistics (logistic/linear regression, gradient descent)

## PROJECTS

**Predicting Building Site Energy Usage**
- Predicting building site energy usage using building performance data. Linear regression and random forest models were trained, but a random forest showed the lowest mean absolute error (TensorFlow, sklearn)

**Exploratory Data Analysis of Los Angeles Traffic Collisions**
- Course final project demonstrating experience with EDA and visualization in python (pandas, Seaborn)

**Clustering Analysis of Holding Companies**
- Self-directed project to identify clusters of financial institution in U.S. using public FR Y-9C data. Uses EDA, K-means and KNN clustering, and visualization in python (pandas, NumPy, scikit-learn, Matplotlib, Seaborn)

## PROFESSIONAL EXPERIENCE

**Federal Reserve Bank of Dallas**    ⚲ Dallas, TX
*Programmer Analyst II, Statistics*    📅 2019 – Present
- Developed four extract-transform-notify python automation jobs to eliminate over 700 hours of staff work per year
- Developed 15 SSIS packages to transition survey collection platform to new PaaS environment
- Developed K-means clustering and KNN clustering notebooks to test proof of concept to pilot AWS Sagemaker environment
- Performed data migration and database design to on-board two new survey collections

## Federal Reserve Bank of Dallas, continued

*Programmer Analyst II, Statistics*                          2019 – Present

- Developed dashboard for preliminary survey data to enable on-demand customer data access eliminating a 1-day lead time
- Led requirements, development, testing, and deployment of five minor version software releases
- Developed python scripts to migrate survey data in real-time to new database environment

## Federal Reserve Bank of Dallas                              Dallas, TX

*Economic Research Assistant, Research*                      2018 – 2019

- Developed R script to automate the retrieval, cleaning, and validation of data for economic modeling
- Developed VBA excel workbook to automatically update FOMC regional data, eliminating 4 days of rework for every 6-week cycle
- Provided research support and analysis for real-time economic updates in the 11th District

## Federal Reserve Bank of Dallas                              Dallas, TX

*Business Operations Analyst, Houston Branch*                2016 – 2018

- Improved data collection and analysis of district cafeteria IRS de minimis calculation
- Analyzed and summarized regional economic conditions for the Board of Directors
- Supported executive leadership research requests during and after Hurricane Harvey
- Represented Houston Branch on the President's Sustainability Initiative Council

## VOLUNTEER

- Mentored 2nd grade students through coding challenge (Bold Idea/ideaSpark)
- Sustainability Initiative (project manager for waste management, data analyst for calculating travel emissions)
- Eagle Scout