# MATTHEW KAYE

mrkaye97.github.io

@ mrkaye97@gmail.com

**(**646) 853-5997

in kayem20

nrkaye97

### **EXPERIENCE**

### **Data Scientist**

### **CollegeVine**

Sept 2020 - Present

- Cambridge, MA
- First member of the CollegeVine data science team. Responsible for establishing data science workflow best practices and providing easy-to-use data science tools and understandable explanations of data scientific concepts to other members of the team
- Clustered users with a modified K-Prototypes algorithm to allow for missing values in the cluster calculations
- Advised on CollegeVine survey conducting and analyzed results for use in published reports on college admissions
- Serving as a subject matter expert for everyday statistics- and data science-related questions

### **Baseball Operations Fellow**

#### **Baltimore Orioles**

- Mar 2020 Sept 2020
- Baltimore, MD
- Created a fully Bayesian, simulation-based projection system for MLB player performance over a six year time horizon
- Created an Gradient Boosting- and simulation-based projection system for a minor league player's MLB performance
- Created a gamma hurdle regression framework to predict free agent salaries
- Implemented a robust-to-multimodality version of the Metropolis-Hastings algorithm to determine the optimal way to position our defense against a specific hitter
- Worked on a variety of day-to-day data science tasks related to game strategy and player evaluation

### Research Intern

## Federal Reserve Bank of St. Louis

- **J**une 2019 Aug 2019
- St. Louis, MO
- Prepared new data series from the BLS, IMF, and more for upload to FRED, using pandas for data wrangling
- Fit neural network and multinomial naïve Bayes models to apply correct descriptive tags to FRED series to improve search performance

### **Econometrics Teaching Assistant**

### **Carleton College Department of Economics**

- **M**ar 2019 June 2019
- Northfield, MN
- Ran semiweekly sessions to help econometrics students master course material, improve their R and LATEX skills, and hone their statistical intuitions
- Helped students complete homework assignments and projects and prepare for exams
- Graded homework assignments

# **SKILLS**

### Languages:

Java Python R SQL

### Frameworks, Software, and Tools:



#### **Data Science:**

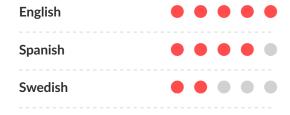
Bayesian Modeling and Inference

Data Wrangling Data Visualization

Machine Learning Statistical Inference

Time Series Modeling

# **LANGUAGES**



# **EDUCATION**

### B.A. Economics and Mathematics

### **Carleton College**

**Sept 2016 - Nov 2019** 

Economics Thesis: The Effect of Transit on Life Satisfaction: A Multilevel Modeling Exploration of Urban Happiness

Mathematics Capstone: Dynamic Linear Models and the Kalman Filter

### **Choate Rosemary Hall**

**Sept 2013 - June 2016** 

# **INTERESTS**

