

# Kelly Lam

Portfolio: [codingwithkelly.github.io](https://codingwithkelly.github.io) • (408) 355-8601 • [kelly.kn.lam@gmail.com](mailto:kelly.kn.lam@gmail.com)

## EDUCATION

### GEORGIA INSTITUTE OF TECHNOLOGY

*M.S. in Analytics, Computational Data Analytics Track*

Expected Fall 2021 - Fall 2023

### SAN JOSE STATE UNIVERSITY

*B.S. in Applied Mathematics, Concentration in Computational Mathematics*  
*Minor in Economics*

Graduated December 2020

- **Honors:** Cum Laude, Dean's Scholar (Fall 2017/Fall 2019)
- **Relevant Coursework:** Introduction to Database Management Systems, Data Structure and Algorithms, Mathematical Modeling, Numerical Analysis, Probability Theory

## EXPERIENCE

### STEM Mentor Intern

*Girlstart*

January 2019 – May 2019

- Organized weekly STEM activities with 2 other mentors for 45 young girls (grades 3-5) to bring awareness to career options in STEM
- Conducted a survey at week 18 and achieved a 93% improvement in their knowledge of STEM and a 533% increase in their desire to pursue STEM careers.

## TECHNICAL SKILLS

**Languages:** SQL, Python, R, Java

**Libraries:** Pandas, NumPy, Seaborn, Matplotlib, Sklearn, Tidyverse, BeautifulSoup, Selenium

**Software:** MS Excel, MySQL, Jupyter Notebook, Tableau

**MOOCs:** (Udemy) R Programming A-Z™: R For Data Science With Real Exercises!

(Udemy) Machine Learning A-Z™: Hands-On Python & R In Data Science

## PROJECTS

### Glassdoor Data Analyst Essential Knowledge Guide

- Web scraped 20 pages of Glassdoor using **BeautifulSoup** and **Selenium** to obtain 200 interview questions from to create a study guide and scraped 1000 data analyst job postings to explore with **Pandas** and **Seaborn**.
- Extracted essential job skills with **NLP** techniques: Word2Vec, Topic Modeling/LDA, and Rule-Based Matching.
- Developed a **Support Vector Regression** model to estimate data analyst salaries with a MAE of \$12,700 and deployed the model using **Streamlit**.

### Spotify Optimal Musical Flow Playlist

- Analyzed and wrangled 7,000+ rows of personal Spotify data using **Pandas**, **NumPy**, and **Seaborn** to visualize streaming trends and correlation between each extracted "audio feature" from Spotify's API.
- Coded a **Python** script to automate curated songs from a .csv file into a Spotify playlist.
- Deployed web application to **Streamlit/Heroku** and presented analytical findings on a **Tableau** dashboard.

### Airline Reservation System

- Spearheaded project with 2 software engineering students utilizing **Java** and **MySQL** to allow users interact with the airline database that supports archiving with triggers and stored procedures.
- Created a database such that its relations were in 3NF or BCNF to eliminate data redundancies.

### Spatial Analysis on Homelessness in the Bay Area

- Used **ArcGIS** and **SQL** to create 18 choropleth graphs based on calculated Gini coefficients to infer homelessness in the Bay Area and map distributions of different races experiencing homelessness.
- Wrangled 200,000+ rows of data in **ArcGIS** and **Excel** by cleaning, structuring, and enriching the data.
- Discovered a 46% increase in homelessness and 49.64% reduction of shelter resources.