

Jason Chongkyung Kim

✉ ckim5@swarthmore.edu | 🌐 [Personal Website](#) | 🐙 [Github](#) | 💼 [Linkedin](#) | ☎ 267-453-6620

Education

Swarthmore College

Computer Science and Statistics

Aug. 2018 – May 2022

- **GPA:** 3.61/4.0
- **Major GPA:** 3.91/4.0
- **Relevant Coursework:** Data Structures and Algorithms, Software Engineering, Artificial Intelligence, Discrete Mathematics, Probability, Linear Algebra, Multivariable Calculus
- **Skills:** Python, pandas, C/C++, SQL, JavaScript, HTML/CSS/Bootstrap, Javascript, Plotly.js, R, Stata, Jupyter Notebook, Git, VS Code, Atom, APIs, Google Analytics

Coding Projects

Manchester United Diagnosis: What is the problem? (ongoing) | Python, Jupyter Notebook, Kaggle, Plotly

- Analysis of Manchester United data in order to diagnose the club's downfall since 2013.
- Retrieve data from Wikipedia using **pandas** to study Manchester United's average league position from 2009.
- Extract and visualize FIFA ratings across Premier League's top teams over the past five years using **plotly**

Twitter Sentiment Analysis Engine | Python

- A **natural language processing** project that uses **TextBlob** for detecting **polarity**, on top of pulling **Twitter API** using **tweepy**.
- Set up data pipeline to retrieve data from Twitter
- Developed a Twitter Sentiment Analysis Engine that studies tweets' polarity based on keywords
- Out of 10 keyword inputs, detects keyword with highest polarity score.

Experience

Software Engineering Intern

FinePrint Data

Feb 2021 – Present

Philadelphia, Pennsylvania

Machine Learning Research Assistant

Swarthmore College Statistics Department (Steve Wang Lab)

May 2020 – Present

Swarthmore, Pennsylvania

- Studied and parsed the North Atlantic Right Whale Consortium (NARWC) dataset pertaining to whale sightings to study individual whales' survival rates.
- Built parallelized algorithm in **R** using our logistic model to estimate the survival probability of every whale in the dataset.
- Ran a **logistic regression** model to discern significant variables that meaningfully affect a whale's survival rate.
- Cleaned dataset of 70,000+ entries, manually going through observations to remedy errors and report the issues to NARWC.
- Reported data findings to the Swarthmore math department of 10+ professors and 30+ students. Paper to be finalized and published by October 2020.

Data Science Intern

LotusFlare

August – Nov 2020

Sunnyvale, California

- Developed **Python** script using Jira's **REST API** to automate Jira issues' due date notification emails to LotusFlare employees.
- Build algorithms with **Python** and **R**, using **pandas** on **Jupyter Notebook** and working with Jira's **REST API**, to analyze Jira time-series data to improve product delivery efficiencies.
- Coordinate with product managers and the data science team to understand existing management behavior, to suggest best practices, and to highlight potential risks.
- Report data analysis results in bi-weekly meetings to 30+ engineers and product managers, visualizing the results with **Plotly**.

Economics Research Intern

Swarthmore College

May – August 2019

Swarthmore, Pennsylvania

- Debugged Stata codebase that analyzes US banks dataset of millions of observations and hundreds of variables