Jason Chongkyung Kim

267-453-6620 ckim5@swarthmore.edu LinkedIn Github Personal Website

Education

Swarthmore College **Aug. 2018 – May 2022**
Bachelor of Science in Computer Science and Statistics **Cumulative GPA: 3.67; Major GPA: 3.92**

Technical Skills

Languages: Python, SQL, R, C, C++, HTML/CSS/JavaScript
Technologies: Docker, Kubernetes, Redis, AWS, Grafana, Apache Airflow, Jupyter Notebook, PostgreSQL, Domo, Plotly

Experience

Telnyx **June 2021 – August 2021**
Software Engineer Intern **Chicago, IL**

- Built a new **ETL pipeline** for primary customer data warehouse and developed **DAG**s using **Python** and **SQL**.
- Made the code for creating and configuring new DAGs **scalabe** by using **Python decorators**.
- Wrote unit tests for DAGs using **Python** to prevent data corruption and SQL injection attacks.
- Deployed new commits to development and production and oversaw the deploy process by assessing **Apache Airflow** UI, **Grafana**, and **Kubernetes**.

Hucu.ai **April 2021 – May 2021**
Data Science Intern **Chicago, IL**

- Created an analysis report template using **Python** and Jira's **REST API** to assess workflow of engineers at Hucu.
- Built a statistical metric to study engineers' workflow speed and efficiency and created interactive graphs using **Plotly**.
- Brainstormed with CEO and program managers to find best engineering and workflow practices.

Swarthmore College Statistics Department **May 2020 – May 2021**
Data Science Intern **Swarthmore, PA**

- Designed a **logistic regression model** and built **parallelized algorithm** in **R** to estimate the survival probability of whales in the dataset and to discern **statistically significant variables** that affect a whale's survival rate.
- Reported data findings to the Swarthmore math department of 10+ professors and 30+ students. Paper to be finalized and published by October 2021.

LotusFlare **August 2020 – November 2020**
Data Science Intern **Sunnyvale, CA**

- Used **Python** and Jira's **REST API** to find trends in the Jira dataset and automate data analysis reports for the program managers.
- Developed an automated email notification system that alerts engineers about due Jira tickets that improved ticket completion rate (before due date) from 63% to 91%.
- Coordinated with product managers and data scientists to understand existing management behavior, suggest best practices, and highlight potential risks.
- Reported data analysis findings in bi-weekly meetings to 30+ engineers and product managers.

Grader for Mathematical Statistics Course **August 2021 – Present**

Projects

Stein's Estimate: Predicting the Premier League (WIP) (Github) | *Python* **July 2021**

- Develop a web scraper using **Python** to retrieve soccer data from TransferMarkt.
- Use Stein's Estimate to predict English Premier League positions for the 2021-2022 season.

Average face of Euro 2020 (Github) | *Python, facer* **June 2021**

- Built a web scraper using **Python** and **Beautiful Soup** to extract 618 player profile pictures from the official Euro 2020 website Official Euro 2020 Website.
- Used **facer** package to calculate the average face of all 618 players' faces and the average face of all 24 teams.

Personalized Spotify Report (Github) | *Python* **June 2021**

- Created a personalized Spotify listening data report using **Python** and Spotify's Web **API**.
- Visualized the popularity of the personal Top 50 artists and songs alongside most represented artists in the Top 50 songs.

Coursework

Computer Science: Data Structures and Algorithms, Artificial Intelligence, Discrete Mathematics, Theory of Computation, Intro to Computer Systems
Math: Linear Algebra - Honors, Honors Several Variable Calculus, Probability, Differential Equations, Statistical Methods II, Mathematical Statistics I, Mathematical Statistics II