

Sarah He

(626) 233-1717 | sarahhe725@gmail.com | sarahhe05.github.io/sarah-he/

EDUCATION

University of California, San Diego
Bachelor of Science in Data Science

Expected Graduation: June 2026
Provost Honors, GPA: 3.89/4.00

EXPERIENCE

Data Science Intern

August 2024 - Present

A Round Entertainment

Jersey City, NJ

- Scraped and parsed HTML data using the BeautifulSoup library to build a robust dataset from online sources, enabling in-depth analysis of concert industry trends
- Developed an interactive Tableau dashboard that allowed non-technical stakeholders to explore data by artist and time period, effectively visualizing key metrics and trends
- Trained and optimized predictive models (Linear Regression, Random Forest, XGBoost) to estimate concert revenue, providing valuable insights into factors affecting revenue generation

Data Visualization Staff, Photographer

October 2023 - Present

The UCSD Guardian

La Jolla, CA

- Collected and interpreted datasets to produce informative and interactive graphics
- Covered diverse content to produce photographs that complemented and enriched articles

Human Resources Intern

August 2024 - October 2024

A Round Entertainment

Jersey City, NJ

- Automated attendance tracking by processing employee attendance logs in Excel, streamlining analysis of performance trends to support data-driven decision-making
- Facilitated effective communication between employees and management

Webmaster

January 2024 - November 2024

The UCSD Guardian

La Jolla, CA

- Regularly updated The Guardian website with fresh content to promote online reach and audience engagement
- Contributed to redesign efforts for The Guardian website, enhancing user experience and modernizing site functionality

PROJECTS

Major Power Outage Analysis | *Python, HTML, CSS*

December 2024

- Conducted in-depth analysis of power outage data to identify regional and infrastructural factors influencing outage severity
- Applied feature engineering and hyperparameter tuning to develop a Logistic Regression model that forecasts severe power outages with 85% accuracy
- Conducted permutation testing to assess missingness dependency, model fairness, and analyze key trends in the dataset

EcoMenu | *Python, Node.js, Tableau, HTML, CSS, JavaScript*

August 2024

- Sourced, compiled, and cleaned datasets on the carbon footprint of over 350 food ingredients and 2100 recipes
- Developed web application enabling users to select ingredients and receive low-carbon recipe recommendations
- Presented visualizations to showcase insights from the recipe and ingredient data
- Boost Hacks II: Most Creative Hack, CirFin CREATE: Civic Achievement Award

San Diego Crime Analysis | *Python, Tableau*

July 2024

- Analyzed 4 years of crime data in San Diego, identifying trends and correlations for effective data presentation
- Built interactive dashboard with Tableau to visualize statistics and trends of various crime types

Dino Diet | *Python, HTML, CSS, Flask, JavaScript*

April 2024

- Implemented XGBoost classifier on dinosaur diets with 93% accuracy
- Built Flask application that allows users to select dinosaur features and view corresponding diet predictions
- Developed frontend web page design with data visualizations to showcase statistics

LEADERSHIP

Starry Sticker Shop

June 2020 – September 2023

Co-Founder

Temple City, CA

- Managed product design, customer service, and financial planning, growing the Etsy business to 270+ global orders
- Analyzed and visualized sales patterns and customer sentiment for actionable insights using Python and Tableau

Temple City Infinity Math Circle

March 2020 – August 2023

Co-Founder, Head of Planning Committee, Instructor

Temple City, CA

- Oversee summer program offering 200+ elementary and middle school students weekly math lessons
- Lead weekly committee meetings to develop class materials for 24 areas of study

SKILL SET

Technical: Python, Java, SQL, Tableau, R, HTML/CSS, JavaScript, Excel, LaTeX, Git, Jupyter Notebook, Figma, Flask

Production: Adobe Lightroom, Adobe Photoshop, CapCut, PowerPoint