

HongKai(Kai) Zhang

hongkaiz@seas.upenn.edu | +1 (626)-464-6404 | <https://www.linkedin.com/in/hongkaizhang-bd/>

EDUCATION

University of Pennsylvania, Philadelphia, PA

Expected May 2025

M.S.E in Data Science (STEM)

- **Coursework:** Machine Learning, Big Data Analytics, Programming Languages and Techniques, Database & Info Systems, Artificial Intelligence, Deep Learning for Data Science, Data and Analysis for Marketing Decisions

Johns Hopkins University, Baltimore, MD

Jan 2022 - May 2023

M.S. in Applied Economics with a concentration in Financial Economics, GPA: 3.97/4.0

University of California Santa Barbara, Santa Barbara, CA

Sep 2020 - Dec 2021

B.S. in Financial Mathematics and Statistics, GPA: 3.71/4.0

- Transferred from the University of California Santa Cruz in 2020.

EXPERIENCE

MiraclePlus (Former Y Combinator China)

July 2022 - October 2022

Investment Research Department Intern

Remote

- **Technology Research and Analysis:** Conducted in-depth research and analysis on emerging technologies in data-intensive industries, including Web3.0, Computer Vision, and AI4Science. Evaluated their potential impact on the investment landscape and market trends.
- **Leadership and Collaboration:** Led a cross-functional **team of 5** members to actively engage with over 70 front-end technology experts from global tech companies and universities on a weekly basis. Achieved a success rate **3%** above the average in advancing investment projects to the interview stage.

Beijing Dand Technology Co., Ltd

August 2019 - September 2019

Data Analysis Intern, Enterprise Resource Planning Department

Beijing, CN

- **Exploratory Data Analysis (EDA):** Utilized the Tidyverse package in R to analyze overall sales trends, regional sales patterns, and seasonal variations based on more than **200k** sales data. Identified and highlighted best-selling items for each season and products experiencing declining sales trends
- **Regression Analysis:** Utilized R to construct Lasso regression to examine the factors influencing sales and performed seasonal forecasts for specific products, in order to assist the company in making informed decisions regarding procurement and warehouse management
- **Data Visualization:** Created heatmaps to illustrate annual total sales, bar plots to showcase seasonal product sales, and model-specific plots in R. These visualizations were instrumental in generating and presenting the National Sales Status Report for the three seasons of 2019 to the Sales and Sourcing department manager in a clear and easily understandable manner

PROJECTS

Stock/Portfolio Quantitative Analyst Project

- **Optimal Portfolio Construction:** Applied Markowitz's portfolio theory and employed Portfolio Frontier Functions to determine the most efficient portfolio allocation base on six leading companies from diverse industries
- **Time Series Analysis and Forecasting:** Utilized time series analysis in R, including ARIMA and GARCH models, on preprocessed historical data for each stock and the investment portfolio spanning the past decade in order to accurate predictions of future performance, enabling informed investment decisions

U.S. Census and Educational Data Analysis Project

- **Data Preprocessing and EDA:** Leveraged R to clean and merge more than **60k** records from population census and education data sources. Checked collinearity and subsequently quantified population poverty indicators
- **Machine Learning Analysis:** Developed five distinct machine learning models to predict and analyze population poverty indicators. Evaluated model using Receiver Operating Characteristic (ROC) analysis
- **In-Depth Analysis:** Enhanced prediction accuracy by **2%** with only a **7%** training error using Random Forest models. Additionally, identified the top three crucial variables influencing the model's performance.

SKILLS AND CERTIFICATION

- **Programming & Software:** Python, R, Java, Spark SQL, PySpark, MATLAB, SAS, Stata, AWS, Tableau
- **Machine Learning:** Scikit-learn, PyTorch, TensorFlow, SciPy, NumPy, Pandas
- **Analytics Skill:** A/B testing, time series analysis, regression, predictive model, data visualization, probability
- **Certificates:** Google Data Analytics Professional Certificate