

# MIN (MIA) SHI

469-403-7557 ♦ [minmiashi@gmail.com](mailto:minmiashi@gmail.com) ♦ [Personal Profolio](#) ♦ <https://www.linkedin.com/in/min-mia-shi/>

## EDUCATION

### The University of Texas at Dallas

Doctor of Philosophy in Political Science

Master of Science in Social Data Analytics and Research (STEM)

Master of Science in Business Analytics (STEM) – Data Science Track

Expected May 2024

GPA: 3.942/4.0

GPA: 3.942/4.0

GPA: 4.0/4.0

## SKILLS

**Programming & Tools:** Python, R, SQL, SAS, Stata, Tableau, Power BI, Alteryx

**Database & Big Data:** MySQL, PostgreSQL, Hadoop, Sqoop, Hive, Impala, Pig, Spark

**Certificates:** Graduate Certificate in Applied Machine Learning, AWS Certified Cloud Practitioner, Google Analytics, Alteryx Designer Core Certificate, Appian Certified Associate Developer, ACCELQ Automation Engineer

**Languages:** English, Chinese, Japanese

## WORK EXPERIENCES

### Data Researcher Assistant

*The University of Texas at Dallas*

May 2020 - Present

*Richardson, TX*

Took responsibility for data manipulation and model building for 10+ global health and policy analytics projects.

- Directed the data gathering processes, utilizing diverse methods like sampling, surveys, and web scraping.
- Developed robust statistical models, including multi-variable regression, fixed-effect regression, difference-in-difference, and time-series models, to facilitate correlation and causal inference studies.
- Oversaw a team of over five junior research assistants, ensuring smooth collaboration and timely completion.

### Data Science Student Consultant

*Working for Onyx CenterSource through The University of Texas at Dallas*

Aug. 2023 - Dec. 2023

*Dallas, TX*

Lead a team in creating an AI-driven chatbot to enhance customer engagement for online interactions.

- Employed NLP and MySQL for analyzing and querying an extensive database containing over 10 million entries.
- Achieved 25% improvement in response efficiency and provided 99% accurate predictions using XGBoost model.
- Contributed to a 15% rise in user engagement, increasing customer satisfaction and bolstering company's image.

### Marketing Data Analyst

*Lucion Technology Corp., Ltd.*

Jul. 2017 - Aug. 2017

*China*

Served as a Data Analyst Intern responsible for data management, data visualization, and business analysis.

- Improved the efficiency of data extraction by 40% through data optimization in MySQL.
- Employed Microsoft Visio to visualize intricate network structures and aided in product comprehension.
- Produced Business Intelligence (BI) reports, offering insights based on user structures and competitor analysis.

## PROJECTS

### Forecasting Stock Prices Through NLP Examination of Newspaper Articles

May 2023 - Dec. 2023

- Developed automated web scraping for 7,000+ WSJ articles, increasing data acquisition efficiency by 30%.
- Employed various vectorizers for WSJ article analysis, such as Tfidf Vectorizer, n-grams Count Vectorizer, etc.
- Utilized Naïve Bayes and Random Forest models, enhancing S&P 500 prediction accuracy by 12%.

### Analytical Insights and Marketing Strategy Guidance for a Food Company

Feb. 2023 - May 2023

- Handled data cleaning over 1.3 million raw data records using Python, ensuring data quality and accuracy.
- Developed interactive dashboards in Tableau, enhancing data accessibility and supporting business analytics.
- Employed SAS to construct regression and time series models, leading to a 15% increase in forecasting accuracy.

### Big Data Risk Analysis and Data Visualization for a Trucking Company

Aug. 2022 - Dec. 2022

- Processed and analyzed geospatial data with Hadoop, Hive, Impala and Spark, reducing processing time by 40%.
- Developed Tableau visualizations linked to Hadoop and built interactive dashboards for business matrix analysis.
- Conducted linear regression and multivariate analysis, contributing to predictive accuracy by 15%.

## ORGANIZATIONS

- Member - Data Science Club, UT Dallas
- Member - Texas Dragon Toastmaster's Club

## ELIGIBILITY

F1 VISA: Eligible to work in the U.S. on OPT for up to 36 months.