# MIN (MIA) SHI

Dallas, TX ◇ 469·403·7557 ◇ Min.Shi@utdallas.edu ◇ minmiashi@gmail.com

LinkedIn Profile ◇ Github Personal Website

## Education

**The University of Texas at Dallas** **August 2019 – May 2024** *(Expected)**
Ph.D. Candidate in Political Science, Major International Relations, Minor Political Institutions and American Politics *GPA: 3.936/4.0*

**The University of Texas at Dallas** **August 2022 – May 2024** *(Expected)**
M.S. in Business Analytics *GPA: 4.0/4.0*

**The University of Texas at Dallas** **August 2021 – May 2024** *(Expected)**
M.S. in Social Data Analytics and Research *GPA: 3.936/4.0*

**The University of Texas at Dallas** **August 2019 – May 2022**
M.A. in Political Science *GPA: 3.917/4.0*

**Shandong University** **September 2016 – June 2019**
M.L. in International Politics *GPA: 88.78/100*

**Daito Bunka University** **September 2017 – August 2018**
Exchange Student in Political Science

**Shandong University** **September 2012 – June 2016**
B.A. in Japanese *GPA: 87.37/100*

## Research Experience

**School of Economic, Political and Policy Sciences, UTD** *Research Assistant* **May–August 2022**
↪ Prof. Jessica Hanson-Defusco

- Explored the effects of the 2014-16 Ebola Crisis on WHO-reporting Nations' Systemic Adaptations and 2020-21 COVID-19 Response; Collaborated with my coworkers in generating original data for 245 WHO-reporting nations, conducting statistical analytics, writing reports, and submitting to journals
- Accomplished data cleaning, transformation, and feature extraction for a collection of 1212 cross-country surveys using Python, utilized R in doing correlation and regression analysis

**School of Economic, Political and Policy Sciences, UTD** *Research Assistant* **May–August 2021**
↪ Prof. Thomas Gray, Prof. Banks Miller

- Performed data collection of 1291 supreme court cases using web-scripting
- Utilized time-series models in analyzing time gaps among case's schedules

**School of Economic, Political and Policy Sciences, UTD** *Research Assistant* **May–August 2020**
↪ Prof. Jonas Bunte

- Collaboratively researched on the benefits connection among U.S. government officers, senators, representatives, and U.S. firms
- Conducted detailed data analysis to detect potential financial and social connections

## Conferences

**2022 APSA Annual Meeting & Exhibition — Montreal, Quebec, Canada** **September, 2022**
Framing 2018 U.S.-China Trade War during the Trump and Biden Eras

**2022 ISDSA Meeting — Notre Dame, IN, USA.** **May 31-June 1, 2022**
Modeling U.S.-China Trade Relations: A Time Series Machine Learning Approach Using MNC Stock Data

## Publications

Yang Luhui, Shi Min. 2020. An Analysis of the Causes of Shinzo Abe's Policy Evolution and Adjustment towards China. *Journal of China's Neighboring Diplomacy.* Vol.7, No.2.

Yang Luhui, Shi Min. 2019. China Policy Adjustment or Changes by the Abe Administrations and Its Impacts. *Peace and Development.* No.3, pp.66-84.

## Data Analytic & ML Projects

**Geospatial Truck Fleet Big Data Analytics and Visualization** **August 2022 ˘ November 2022**

- Used big data Hadoop ecosystem to process geospatial data ingestion, transformation, and database creation
- Performed data exploration and visualization in Tableau by connecting to Hadoop ecosystem server
- Modeled how factors affect the truck driver risk factor, drew a final report and proposed suggestions on how to lower the probability of large trucks accidents

**Payroll Management System Database Design via MySQL** **June 2022 - August 2022**

- Led a group of five in conducting business requirements analysis and designing a payroll management database with MySQL consisting of 13 tables
- Created stored functions, procedures, and triggers to calculate employees' payroll per two weeks, fill in new employee's information, send PTO reminders automatically
- Performed extract-transform-load, data cleaning, and query optimization

**Modeling U.S.-China Trade War's effect on U.S. Firms using ML and Time Series** **January 2022 - May 2022**

- A project aimed at exploring how the U.S.-China trade war affects Multinational Corporations (MNCs) through a ML content analysis of policy changes and a time series GARCH modeling approach using stock data
- Utilized Pandas, NumPy, Matplotlib & Seaborn in data cleaning, visualization, and transformation
- Leveraged sentiment analysis to explore how the U.S. frame 2018 U.S.-China trade war
- Applied regression analysis in exploring the causal mechanism between trade war and S&P 500 revenues
- Built machine learning (ML) models in predicting the profound influence of the trade war on U.S. firms
- Used time-series GRACH models to evaluate MNCs' revenue & volatility quantified via stock data in Stata
- Presented at 2022 International Society for Data Science and Analytics Conference

**Content Analysis of News Coverage about U.S.-China Trade War** **August - May 2022**

- Led an analysis on how news organizations frame the 2018 U.S.-China trade war during the 2018-2022 period
- Leveraged machine learning skills such as top modeling and sentiment analysis to explore a collection of over 500 news articles
- Implemented time-series analysis and chi-squared test in modeling sentiments change tendencies among news coverage
- Selected as iPoster and expected to be presented at 2022 APSA Annual Meeting  Exhibition

**COVID-19 Worldwide Cases Synchronous Dashboard using Tableau** **December 2021 - January 2022**

- Designed a synchronous Tableau dashboard with advanced interactive functions to explore the COVID-19 severity
- Built a Tableau story to dig into the factors affecting the severity of COVID-19 by country and found out the deep connection between multiple aspects of factors with COVID-19 severity

**Data Visualization and Correlation Analysis with Multiple Tools** **September - December 2021**

- A project aimed at exploring the factors that affect World Happiness Index by country
- Utilized Python and R in data collection and data cleaning processes
- Deployed Python, R, R Shiny and Plotly Dash in exploring correlation among variables and visualizing the correlations

## Selected Course Work

| Data Science | Data Management | Data Modeling |
|---|---|---|
| Natural Language Processing | Big Data | Modeling for Business Analytics |
| Causal Analytics and A/B Testing | Database Fundations for Business Analytics | Regression and Multivariate Analysis |
| Programming for Data Science | Information Management | Applied Data Analytics with Python |
| ML for Socio-Eco and Geo Data | Cloud Computing Fundamentals | Applied Regression |
| Content Analysis using ML | Data Visualization | Introduction to Quantitative Methods |
| OOP in Python | Data Collection | Social Science Research Methodology |

## Technical Skills

| | |
|---|---|
| **Programming** | Python, R, SQL, Stata, SAS |
| **Tools** | Alteryx, Tableau, Jupyter Notebook, Excel Charts, R Shiny, LaTeX & TeX |
| **Database & Big Data** | MySQL, PostgreSQL, Mango DB, Amazon RDS, Hadoop, Sqoop, Hive, Impala, Pig, Spark |
| **Certificates** | Graduate Certificate in Applied Machine Learning at UTD, AWS Certified Cloud Practitioner<br>Alteryx Designer Core |
| **Languages** | English, Chinese, Japanese |

## Career Goals

**Being equipped with data analytic skills using Python, R, Stata, SAS & SQL, familiar with multiple industry analytical visualization tools, e.g., Tableau, Shiny, R Markdown Dashboard, and having abundant experience with statistical research methods, I focus on utilizing machine learning and quantitative statistical research skills to explore the mutual effect between the U.S. trade policies and the big firms' operations within the context of U.S.-China trade. My career goal is to become a researcher in this area or a professional data scientist in the industry.**