

# XUDONG ZHANG

Cell: +65 XXXX XXXX  
Email: [xudong.zhg@gmail.com](mailto:xudong.zhg@gmail.com)  
MyPage: <https://xudong-z.github.io/me/>



## EDUCATION & SKILLS

**University of Southern California, Marshall Business School** **May 2018**

- Master of Management Studies, Graduate Certificate in Business Analytics GPA: 3.67/4.0
- Relevant Coursework: Database Management, Machine Learning, Fraud Detection, Hypothesis Test

**Shanghai Jiao Tong University, School of Agriculture and Biology** **Jun 2016**

- Bachelor of Science in Bio-technology GPA: 3.43/4.0
- Majored in Statistics and Genetic Engineering

### Technical Skills & Certificates

- Data Science: Python / R / SQL / Linux / Machine Learning / Git / MongoDB / PowerBI
- Others: AWS Certified Solution Architect Associate / Finance, CFA Level 1 Passed / Blockchain (web3.js)

## WORK EXPERIENCE

**Louis Dreyfus Company, Singapore** **May 2021 - Now**

*Data Scientist / Freight Researcher*

- Design and implement databases for hourly satellite messages, conduct feature engineering/enrichment for the convenience of analyzing real-time marine traffic and port operations (**MongoDB, SQL, geopandas, statsmodel**)
- Implemented ML models (KDTree, DBSCAN, Xgboost) to cluster and predict voyage destination and regional commodities trading volume, achieving 92.3% accuracy on classifier and 2.38% MPE on regressor (**sklearn**)
- Contribute to proprietary python SDK as data access layer, design tools such as SQL query builder, time series analyzer, string fuzzy matcher, and quick geo-spatial visualization (**Git, SQL, dash/plotly, folium**)
- Re-design legacy R models into Python, migrate job schedulers from local to cloud (**Azure DevOps, Jenkins**)

**Bloomberg LP, Singapore** **Jul 2018 - May 2021**

*Data Analyst*

- Build ETL pipelines to ingest streaming/batching data from various APAC data sources (SGX, SHFE, Mysteel etc) by using RESTful API, Python, cloud storage and in-house tech stacks, automate data quality over 60 million data points per day (**beautifulsoup4, AWS S3, PostgreSQL, Apache Airflow**)
- Automate Research/News team's regular fundamental research to improve operational efficiency, eg. building monitors for certain vessels entering user-defined geo-polygons, as well as coal/ore inventory WoW change at China ports, saving them ~600 working hours per year (**SQL, geopandas, bqplot**)
- Provide client-facing support(L1) and maintenance over terminal functions and database in-charge, also work as Data TechOps with Engineering team for continuous development (L2) and incident report&response

## OTHER TECH PROJECTS

**Credit Card Fraud Detection** ([shiny app on R Studio](#)) **R: caret/ggplot2/shiny/dplyr**

- Identified fraud transactions with 98.2% accuracy and 81.2% ARPRC on hugely imbalanced (0.17%) credit card transaction data using ML methods (Random Forest winning over Logit, Neural Networks etc)

**Summer Olympic Medalist Profile Analysis** ([dash app on Heroku](#)) **Python: dash/plotly/pandas/requests**

- Scraped all the wikipedia pages to get the basic info of 21000+ summer Olympic medalists, and conduct data profiling in a animated visualization tool over their age/nationality/sports/gender

**Blockchain Technology Project** **Python, Web3.js, ERC20**

- Built JS scripts to send and monitor ERC20 transactions over Ethereum mainnet & testnet ([github repo](#))
- Compiled a python tool to generate private/pubkey for Bitcoin/Ethereum/Tronweb wallet based on the corresponding crypto algorithm ([github repo](#))