

# XIAOYUN WANG

[xiaoyun.wang00@gmail.com](mailto:xiaoyun.wang00@gmail.com) • 919-638-1232 • Durham, NC • Portfolio: [xwang.me](http://xwang.me)

## EDUCATION

DUKE UNIVERSITY, The Fuqua School of Business

Durham, NC

*Master of Science in Quantitative Management: Business Analytics*

May 2023

Coursework: Data Science for Business, Decision Analytics & Modeling, Data Visualization, Applied Probability and Statistics

FUDAN UNIVERSITY

Shanghai, China

*Bachelor of Laws: International Politics*

Jun 2022

Awards: First Prize of the Scholarship for Outstanding Student (top 5%), Graduate Scholarship (2022)

## SKILLS

**Databases & BI:** ETL, SQL, MariaDB, MySQL, AWS Cloud Service (IAM, EC2, S3, RDS, Lambda), Tableau, Power BI

**Technical:** R, Python, Scikit-Learn, TensorFlow, PyTorch, NLP (NLTK), Git, Wind, Bloomberg, Advanced Excel

**Statistical Methods:** A/B Testing, Regression, Classification, Clustering, Decision Tree, Monte Carlo, Optimization

## EXPERIENCE

Guotai JUNAN SECURITIES CO., LTD., Fixed-income Department

Shanghai, China

*Investment Intern*

Jan – May 2022

- Designed Python code and Excel database for investment **income analysis** and **risk estimation**.
- Delivered 4 monthly **reports** about **investment returns** to managers, presentations and deliverables for over 30 Asset Backed Securities using visualizations and clear communication.
- Serving as bridges between projects and investors, conducted due diligence and asset valuation based on company **Balance Sheet** and **liquidity stress testing** with other financial modules.

TENCENT, Public Strategy Research Center

Shanghai, China

*Data Analyst Intern*

Jul – Aug 2021

- Crafted a cohesive story**, interpreted **85K+** pieces of **operational data** in time series through **SQL**.
- Coordinated with cross-functional teams to identify key metrics and KPIs for the 2021 Tencent Global Digital Ecosystem Summit, launched **8 dashboards** into 22 industries and 355 cities to track them, increasing visibility into business performance and more informed decision-making.
- Automated **web crawler** with **Python** to search and organize public information of cloud services policies.
- Authored** analysis to 2021 “Cloud + AI” White Paper as a team worker and provided real-time insights.

## SELECTED PROJECTS

**NLP Analysis on ESG Performance with Wells Fargo (Capstone Project, Python, NLP, AWS RDS, S3)**

- Conducted **Natural Language Processing (NLP)** analysis to ESG reports with **Python NLTK** package.
- Implemented an **automated** approach to analyze large volumes of data using Python and **AWS RDS Database**, which improved the efficiency and accuracy of the analysis and had a positive impact on the organization's sustainability efforts.
- Built **word frequency** to identify key topics, **sentiment analysis** to assess stakeholder attitudes, **semantic similarity** analysis to compare ESG descriptions across reports and over time.
- Presented findings and recommendations to senior manager, developed sentiment and inconsistency scores to evaluate ESG performance, utilizing **Tableau Dashboard** and communication skills.

**Analyzing and Visualizing Duke Women's Lacrosse Data (ETL, SQL, Tableau)**

- Identified the need for better data analysis and visualization tools in the sports industry, developed an **ETL solution** based on **SQL** and **python** to address this need for Duke Women's Lacrosse staff to maximize team competitiveness.
- Extracted data** from NCAA website to compile most recent game data program by **R** code.
- Transferred and loaded data** by setting up **MariaDB** through **Python** and **SQL**, enabling more informed decision-making and improved business outcomes by implementing traceability in analysis.
- Generated **data report** through **Tableau Dashboards** to illustrate trends with coaching staff to plan practices.

**Human Resource Predictive Model for Reducing Employee Turnover (Machine Learning, TensorFlow)**

- Led a team of five to develop and implement a predictive model for **reducing employee turnover**.
- Improved computing efficiency by applying dimensionality reduction algorithms (**PCA**) to remove duplicated features.
- Predicted** an attrition rate of **19%**, employed **K-Fold** with a **logistic regression** model to identify features.