# Yueyue Min

Fort Lee, NJ | (949) 791-7258 | ym2980@columbia.edu

## EDUCATION

**Columbia University, Graduate School of Arts and Sciences**          **New York, NY**
*Master of Arts in Quantitative Methods in the Social Sciences (Data Science Track)*          Sep 2023-May 2025
- GPA**:** 3.8 / 4.0; GRE:336
- Relevant Coursework**:** Advanced Machine Learning, Timeseries Forecasting, Visualization, NLP, Bayesian Statistics

**University of California, Irvine**          **Irvine, CA**
*Bachelor of Science in Cognitive Sciences (Minor in Informatics & Computer Science)*          Sep 2019-Mar 2023
- GPA**:** 3.95 / 4.0, Magna Cum Laude
- Relevant Courses**:** Machine Learning, Computational Methods, Cognitive Robotics, Statistics, Neuroscience

## EXPERIENCE

**GradeMate**          **New York, NY**
*Data Scientist*          Sep 2024 – May 2025
- Designed and deployed **scalable ML models** using **LLaMA 3** and Python to automate grading and feedback workflows, reducing manual grading variance and increasing efficiency for educators.
- Engineered a time-aware scoring algorithm incorporating A/B testing and time series dynamics to adapt to teacher grading styles and student learning behavior.
- Led collaboration with ML pipelines with Databricks and Python, automating model deployment and reducing latency in result feedback loops by 40%.

**ADEL Group**          **Hong Kong, China**
*Business Intelligence Analyst*          Apr 2023 - Aug 2024
- Built forecasting models using XGBoost to analyze financial trends and regional demand signals, used SQL to extract and structure time-series data, supporting improved inventory planning and pricing strategies across Southeast Asia.
- Automated analytics workflows with pandas to replace manual reporting, reducing turnaround time for market performance insights and improving executive alignment.
- Developed KPI dashboards in Power BI translating operational and revenue data into actionable metrics, driving better allocation decisions during monthly strategy reviews.

**SF Technology**          **Shenzhen, China**
*Machine Learning Engineer Intern, AI and Visual Recognition Group*          May 2021 – Aug 2021
- Enhanced warehouse automation by refining object detection models with TensorFlow and augmenting edge-case images using OpenCV, reducing misclassifications by 10%.
- Designed modular data pipelines with validation and feature tracking layers, enabling model reusability across lighting and warehouse layouts.
- Conducted anomaly detection on route tracking data using statistical control charts, identifying inefficiencies and supporting rerouting process redesign.

## PROJECTS

**Sentiment Analysis of New York Times Articles** | Python, BERT, GCP          Sep 2024 – May 2025
- Designed and implemented NLP pipeline using BERT to perform sentiment analysis and topic modeling to analyze large-scale text data from the New York Times API, surfacing public attitude dynamics across sectors.
- Built and deployed a Streamlit dashboard on GCP, enabling real-time exploration of sentiment trends by keywords, publication dates, and article sections.

**Urban Wildlife Habitat Interaction Visualization in Central Park** | Python, R          Jun 2024 – Aug 2024
- Conducted geospatial analysis of squirrel behavior in Central Park, mapping interactions to tree density, time, and human
- Developed a **Shiny** app for dynamic exploration of wildlife patterns, enhancing user engagement and understanding.

**Credit Card Recommendation System for College Students**| Python, R, Scikit-learn, NLTK          Sep 2023 – Dec 2023
- Conducted market and user research for Consumer Financial Protection Bureau to identify key factors influencing credit card preferences among college students, leveraging surveys and CFPD annual reports and publicly available datasets.
- Built a recommendation system using scikit-learn, integrating NLP techniques to analyze user reviews.

## SKILLS

- **Programming:** Python, SQL, R, MATLAB
- **Tools**: Tableau, Power BI, Excel, MySQL, Databricks, AWS
- **Packages**: Pandas, NumPy, Scikit-learn, TensorFlow, XGBoost, Statsmodels, Matplotlib, Seaborn, OpenCV
- **Techniques:** Forecasting, AB Testing, Time Series, Clustering, Classification, Text Mining, Geospatial Analysis