Sherry Wang

sherrywangyx@gmail.com | (314) 762-8788 | linkedin.com/in/sherrywangyx

EDUCATION

Columbia Business School | New York, NY — Expected December 2025
MS, Marketing Science, Business Statistics (Advanced Analytics Track)
GPA: 9.5/10; Coursework: Multivariate Statistics (PhD), Machine Learning (PhD), Generative AI (PhD), Python for Data Science

Washington University in St. Louis | St. Louis, MO — May 2024
BSBA, Marketing & Data Analytics
GPA: 3.92/4.00; Dean's List (7 semesters)

WORK EXPERIENCE

Amazon — Seattle, WA
Business Intelligence Engineer Intern — June 2025 – August 2025

- Built an end-to-end **Python ETL** workflow to detect and fix 7 data quality issues impacting downstream stakeholder reporting; deployed it with **AWS S3**, **Lambda**, and **CDK** from development to production
- Implemented **LLM**-powered text classification via **Claude API** in AWS Bedrock for data quality checks, increasing accuracy by 50%, and created a QuickSight dashboard to track remediation progress
- Engineered an ETL pipeline in **Redshift SQL** to process 460K+ operational records and define new performance KPIs, visualized in QuickSight dashboards to inform operational strategy and guide monthly business reviews
- Developed a CatBoost **ML** model in **SageMaker** with Optuna hyperparameter tuning to predict missing values, improving KPI reliability by 8.5% and reducing reporting gaps by 1 week

Columbia Business School — New York, NY
Data Analyst — January 2025 – May 2025

- Scraped Google Scholar data to track 150+ faculty's year-over-year publication trends, driving insights into research impact
- Programmed advanced Qualtrics surveys with JavaScript for a research comparing sentiment analysis by LLMs vs. humans

NBCUniversal — Los Angeles, CA
Data Analyst Intern, Measurement Strategy — June 2024 – August 2024

- Built statistical and **machine learning** models (ARIMA and LSTM) to forecast Peacock's annual streaming minutes for the next 5 years, achieving 6.5% MAPE; forecasts informed content investment and distribution strategies
- Designed an **ETL pipeline** using Python and SQL on **Databricks** to transform 7.7B+ raw Nielsen streaming measurement data; automated workflows with **Snowflake** and **AWS S3**, driving a 70% boost in efficiency
- Led cross-functional development of the core streaming platform ratings **Tableau** dashboard, delivering audience and engagement metrics used by executives to assess show performance

NBCUniversal — New York, NY
Data Analyst Intern, Broadcast Programming Research — June 2023 – May 2024

- Produced daily and weekly insights reports on Nielsen TV ratings across 100+ programs, pinpointing drivers for growth
- Led comparative analysis of linear and digital audiences, sharing insights with senior executives to support brand partnerships
- Created Tableau dashboards from scratch and employed VBA Macros to streamline data visualization, saving 3 hours a week

PROJECTS & AWARDS

Movie Recommendation System with Neural Collaborative Filtering — December 2024

- Trained a Neural Collaborative Filtering (NCF) model using deep learning in Python with 28M+ movie ratings, optimizing user-item interactions via TensorFlow and achieving an RMSE of 0.818
- Integrated metadata into the model to deliver personalized top movie recommendations tailored to users' preferences

***3rd Place Finalist* | Manhattan College Business Analytics Competition** — April 2022

- Conducted in-depth analysis of 53 countries' COVID-19 resilience performance from the Bloomberg database using regression and clustering models in R; visualized results with Tableau and presented strategic recommendations based on findings

SKILLS

Programming Languages: SQL, Python (Pandas, NumPy, Scikit-learn, PyTorch, TensorFlow, PySpark), R
Tools: Tableau, AWS (Redshift, S3, SageMaker, Lambda, QuickSight), Excel, Databricks, Snowflake, Google Analytics, Qualtrics
Skills: Data Science & Analytics, Data Visualization, ML/Statistical Models, A/B Testing, Consumer Insights, Product Management