Yanyan (Dora) Deng

New York, NY | 516-652-8341 | mogudyy@gmail.com | https://www.linkedin.com/in/yanyan-dora-d-8177761a2/

SUMMARY

A data-driven team player skilled in gathering business questions, translating them into data analysis, and providing strategic recommendations. I have contributed to increasing forecast accuracy in financial planning by 65%. I excel in solving problems for stakeholders through efficient data reports that deliver reliable insights.

EDUCATION

FORDHAM UNIVERSITY, *GABELLI SCHOOL OF BUSINESS,* NY Aug.2019 - Dec.2020
MS, Information Technology GPA: 3.9/4.0
Relevant Coursework: Database Management, Data Mining, Business Analytics, Information system strategy

HOFSTRA UNIVERSITY, *FRANK. ZARB SCHOOL OF BUSINESS,* NY Aug.2017 - May.2019
MS, Finance GPA: 3.6/4.0
Relevant Coursework: Financial Analysis and Planning, Advanced Statistical Modeling in Finance, Advanced Applications in Risk

Xi'an SHIYOU UNIVERSITY, China Aug.2012 - May.2016
BS, Engineering GPA: 3.7/4.0
Relevant Coursework: Advanced Mathematics, Liner Algebra, Probability and Mathematics Statistics

HIGHLIGHTS

Language Skills: SQL, Tableau, PowerBI, Python, Relational Database, Machine Learning
Certificates: Gabelli Leadership Certification, Google Analytics Certification

EXPERIENCE

Siemens Healthineers (Top500 - Tier1 Healthcare) Tarrytown,NY
Data Analyst Aug.2021 – Aug.2023

- Assessed data models by mining datasets from **SAP** and querying an 800M dataset in **SQL** to enhance sales data accuracy
- Completed automated monthly reporting in **PowerBI**, streamlining ongoing reporting and analysis processes, leading to a **65%** increase in budget forecast accuracy for financial planning
- Led **AWS cloud data** migration by constructing **data pipelines** (extracting, transforming, and loading) in data management
- Collaborated closely with IT and finance teams to fulfill ad-hoc requests that feed into **business intelligence dashboard (Tableau),** enabling global stakeholders to make informed strategic decisions and reducing **1M** product risks

NYU Langone Fenyo Lab New York, NY
Data Science Research Assistant Intern Oct.2020 - Feb.2021

- Implemented **Correlation, Significant level in R studio** and **Pandas** in **Python** to formatted distributor data and recognize **450** mutant genes in **30k** raw dataset
- Created scientific basis for biologists, leading scientist to explore unique diseases by **visualization** tool in R studio

Deloitte - NCAA March Data Crunch Madness Competition New York, NY
Data Scientist Jan.2020 - May.2020

- Predicted winning team of each game in 2020 NCAA with accuracy over 75% and log loss of 0.439
- Trained 17-year historical NCAA dataset with more than **150 variables** before feature selecting, tested models using machine learning algorithms like ANN and earned Honorable Mention (top10%) from judging panel

Gabelli School of Business New York, NY
Data Analysis Teaching Assistant Dec.2019 - Dec.2020

- Collected, aggregated, and analyzed **11GBs** of Instagram data on Google Cloud to measure popularity of Product Tags using Pandas
- Built **automated visualization dashboards** by extracting data from **MySQL** server to track 50 fashion trends

Beidou Cloud Service Co. Ltd Qinghai, China
Data Analyst Intern Jan.2017 - Jan.2019

- Provided competitive prices of 16 different styles of couch by writing **100+** SQL queries on MySQL to extract hundreds of rows of historical price data for California clients
- Helped design **10+** couch material keywords index for online store, added additional feature to database with **6,000+** observations and performed A/B test detect online store revision
- **Doubled** client amounts in Thailand and **tripled** couch order sizes by presenting **Tableau** visualization

PROJECTS

Homeless and Hate Crime Incident on Victimization Jan.2020 - May.2020

- Conducted initial data exploration, filtered outliers on two datasets with **140K+** disorder data from 2007-2018
- Created 20 **Tableau** charts according to hypothesis for visual analysis, displayed correlation and Random forest model with SPSS to test hypothesis and contributed to reducing homelessness and crime rate, improving social status and city appearance