

Yanyan (Dora) Deng

New York, NY | 516-652-8341 | ydeng79@fordham.edu | <https://doradeng629.github.io>

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, Web applications development, Information system strategy, Deep Machine Learning

HOFSTRA UNIVERSITY, FRANK. ZARB SCHOOL OF BUSINESS, NY

Aug.2017 - May.2019

MS, Finance GPA: 3.3/4.0

Relevant Coursework: R Studio, Financial Analysis and Planning, Advanced Statistical Modeling in Finance, Advanced Applications in Risk

Xi'an SHIYOU UNIVERSITY, China

Aug.2012 - May.2016

BS, Geological Engineering GPA: 3.7/4.0

Relevant Coursework: Advanced Mathematics, Liner Algebra, Probability and Mathematics Statistics

HIGHLIGHTS

Language Skills: SQL, Python, Tableau, R, JavaScript, HTML, AWS, GCP, SPSS, MS Office

Certificates: Gabelli Leadership Certification, Google Analytics Certification

EXPERIENCE

NYU Langone Fenyo Lab, Data Science Research Assistant Intern, New York, NY

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

Gabelli School of Business, Data Analysis Teaching Assistant, New York, NY

Dec.2019 - Dec.2020

- Collected, aggregated, and analyzed **10GBs** 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, Data Analyst Intern, Qinghai, China

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
- Improved click-through rate **130%** by revising product keyword index, while ensuring the continuation and enhancements of online sales event

PROJECTS

Deloitte March Data Crunch Madness

Jan.2020 - May.2020

- Predicted winning team of each game in 2019 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 Random Forest
- Earned Honorable Mention (top10%) from judging panel

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
- Contributed to reducing homelessness and crime rate, improving social status and city appearance

Predict Traffic Accident Severity in U.K.

Feb.2020 - May.2020

- Utilized Pyspark to gather 1.6 Million UK Traffic Accidents, summarized 23 variables to related the severity of traffic accident, 3 statistical models like Decision Tree, Random Forest to analysis accuracy and created symbol map Tableau chart
- Provided 10+ traffic safety recommendations to reduce traffic accidents and predict accidents

Brazil Medical Appointment Absentee Rate Analysis

Oct.2019 - Mar.2020

- Emphasized data cleaning and data analysis based on **100k** medical appointment dataset using Python package such as **Pandas** and **Numpy**
- Analyzed and visualized critical factors to prediction of patients' no show, providing 3 plans to increase clinic appointment attendance rate and guiding people to promote health examinations