

QIUSHI YAN

qiushi.yan@vanderbilt.edu ◇ 629 239-9151

[Website](#) ◇ [Github](#) ◇ [Linkedin](#)

EDUCATION

Vanderbilt University - Data Science Institute

Master of Science in Data Science

Nashville, TN

Expected May 2023

Communication University of China

Bachelor of Arts in Data Journalism

Beijing, China

September 2017 - June 2021

EXPERIENCE

51Talk

Analytics Engineer

Beijing, China

Intern, September 2020 - June 2021

- Identified key factors in teaching performance and ROI with regression modeling and hypothesis testing
- Developed monitoring dashboard for admins based on previously identified metrics using Django and Vue
- Standardized R Markdown report and presentation template with parameterization and custom themes

Analysys

Data Analyst

Beijing, China

Intern, May 2020 - August 2020

- Innovated Echarts-based data visualization components in web analytics service, increased loading speed by 10
- Designed tracking events for retention and churn analysis for ecommerce clients

DiDi

Data Mining Engineer

Beijing, China

Intern, March 2019 - August 2019

- Reorganized Shiny dashboards into reusable modules and Docker images, added CI/CD workflow
- Conducted overseas geo-marketing research based on cluster analysis and correspondence analysis

PROJECTS

Analysis of Airbnb Rentals in NYC - R

[app](#), [package](#)

- Developed an R package and shiny app to analyze Airbnb rentals in NYC, featuring spatial and text analysis of housing description, finished with a multinomial logistic model predicting pricing ranges.
- Promoted best practices, e.g., modular app, dev/prod stages, Rmd-first approach, pkgdown website.

Overwatch League Statistics API - Python, R

[app](#), [source](#)

- Created a self-documented esports match and player statistics API with FastAPI and AWS RDS
- Used R + Python to their advantage: R for data cleaning and Python for engineering and deployment

Trading Data Pipeline - Python, Docker

[source](#)

- Engineered an ETL pipeline in pandas for processing real-time trading data from the Deutsche Börse Group
- Designed OOP, well-tested, and configurable jobs on AWS S3 and Glue to produce parameterized reports

PySpark Change Data Capture - Python, AWS

[source](#)

- Implemented change data capture and feature engineering with pyspark jobs
- Designed full data pipeline with AWS RDS, S3, Lambda functions and Glue

ADDITIONAL

Certification

[AWS Certified Developer](#)

Languages

R, Python, TypeScript, SQL, Shell

Frameworks

tidyverse, Shiny, FastAPI, PySpark, React, D3

Tools

AWS, Github Actions, R Markdown