

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

Data Analyst Intern

Beijing, China

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 Intern

Beijing, China

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 Intern

Beijing, China

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

Overwatch League Statistics API - Python, R

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

- Implemented a self-documented esports match and player statistics API with FastAPI and AWS RDS
- Used R plus 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 object-oriented and configurable jobs that interacts with multiple s3 buckets and save clean data to generate parameterized reports, included 100% test coverage, profiling and containerization.

JavaScript Notebook - TypeScript, React & Redux, Node.js

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

- Developed a jupyter-like notebook app for js/ts, providing command line interface available on NPM
- Enabled user to import any third party library on the fly, powered by custom esbuild plugins

Titanic Survival - R

[report](#), [source](#)

- Developed a binary logistic model to predict survival probability, including redundancy analysis, multiple imputation nad bootstrap validation

ADDITIONAL

Languages

R, Python, TypeScript, SQL, Shell

Frameworks

tidyverse, data.table, Shiny, FastAPI, pandas, React, Vue, pandas

Tools

AWS, Github Actions, R Markdown

Certification

[AWS Certified Developer](#)