QIUSHI YAN

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

Website ⬦ Github ⬦ Linkedin

EDUCATION

Vanderbilt University - Data Science Institute **Nashville, TN**
Master of Science in Data Science Expected May 2023

Communication University of China **Beijing, China**
Bachelor of Arts in Data Journalism September 2017 - June 2021

EXPERIENCE

51Talk **Beijing, China**
Analytics Engineer 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

Analysys **Beijing, China**
Data Analyst Intern, May 2020 - August 2020

- Innovated Echarts data visualization components in web analytics service, improved fetching speed by 10%
- Designed tracking events for retention and churn analysis in ecommerce

DiDi **Beijing, China**
Data Mining Engineer Intern, March 2019 - August 2019

- Refactored Shiny dashboards into modules and custom JavaScript templates, added testing and CI/CD workflow
- Conducted overseas geo-marketing research based on cluster and correspondence analysis

PROJECTS

Analysis of Airbnb Rentals in NYC - R shiny app, R package

- Developed a Shiny app and R package to analyze Airbnb rentals in NYC, featuring spatial and text analysis of housing description. Uses custom HTML templates, shiny-js handlers and Sass.

Overwatch League Statistics API - Python, R API, repo

- Created data ingestion and cleaning pipelines for analyzing player performance in the Overwatch League with R's data.table. Developed restful API serving cleaned data with Python's FastAPI and AWS RDS.

ETL Pipeline of Stock Trading - Python, R repo

- Engineered an ETL pipeline in pandas for processing real-time trading data from the Deutsche B¨orse Group. Designed automated, tested and configurable jobs on AWS S3 and Glue to produce parameterized reports.

JavaScript and TypeScript Notebook - React, TypeScript, Redux app, npm library

- Developed a jupyter-like coding environment for js & ts, with bundling and transpiling built into the browser. Provided command line npm library, run `npx javascript-notebook serve` to try it out!

PySpark Change Data Capture and Feature Engineering - Python repo

- Designed a data pipeline of change data capture and feature engineering with pyspark jobs, using AWS RDS, S3, Lambda functions and Glue.

ADDITIONAL

Languages	R. Python, SQL, TypeScript, Shell, Elixir
Frameworks	tidyverse, shiny, pandas, pyspark, scikit-learn, React, Svelte, GraphQL
Certification	AWS Certified Developer