Zhihui (Elisa) Zhang

elisa99@bu.edu | 510-717-1453 | https://github.com/Elisa1999

Education

Boston University Boston, MA
MS in Statistical Practice Sep. 2021- Dec. 2022

Shanghai University of Finance and Economics (SUFE) Shanghai, China
B.Mgt. in Information Management and Information System Aug. 2017- Jul. 2021
Honorable Mention of Interdisciplinary Contest In Modeling in 2019

University of California, Berkeley Berkeley, CA
Exchange program Aug. 2019- Aug. 2020

Skills: Python, SQL, R, Java; Visualization tools: Tableau, Power BI

Professional Experiences

Boston University
Statistical Consultant Sep.2021 – present

- Assisted researchers from across the Boston University community to define and resolve clients' problem from a statistical perspective
- Performed **data intensive transformation and aggregation** in R
- Conduct **statistical analysis in R** on topics involving **data visualization**, **hypothesis testing** and **model fitting**
- Generate **insightful reports** and conduct **presentation** to clients

Research Assistant at Department of Dermatology
Data Assistant Feb.2022 – present

- Created efficient **ETL pipeline in Python**, leveraging library such as **hic-straw** for HiC data loading and **data normalization**
- Visualized promoter-centered chromatin interaction(**interaction between genos**) using **Python Plotly**
- Built **dynamic plots** where user can pass in different parameters at **runtime** to generate different **metrics** using **Plotly in Dash**

Fidelity Investments - Data Science Center of Excellence
Student partner Sep.2021 - present

- Created an **automated Data Science Pipeline** that leverage **AWS Sagemaker** for Model training and **AWS S3** for data storage
- Analyze operation data from **Snowflake warehouses**, help to determine the efficient and inefficient aspects of the warehouses
- Conducted **Data Warehouse cost forecast** using historical usage data to derive strategic query usage plan and **limit unnecessary costs in database maintenance**

Yo-ren Limited
Data Analyst Intern Jan.2021 – Apr.2021

- Supported daily strategy operation work for Lawson convenience stores in China by heavily utilizing **Postgresql** for **data extraction, inquiry, transformation, and aggregation**.
- Diagnosed problems of Lawson Station Application using **data in various formats such as data collection, user surveys and feedbacks** to explore possible solutions to different customers facing problems
- Identified high-potential customer base, key advertising channels, and key areas by **analyzing user purchase history**; facilitated precision marketing and **improved Gross Merchandise Volume by 5%**

Minsheng Securities Co.,Ltd. - Quantitative Trading Department
Quantitative Intern Nov.2020 - Jan.2021

- Used **Python** to collect stocks data from different sectors, selected top 10 stocks with the highest ROE(return on equity) in each sectors and **visualized their ROE** changes in three years;
- Created different factors to build **multi-factor stock selection models** in TMT(Technology, Media & Telecommunications) sector; Used **PCA** to reduce factors

Course Projects

Hurricane exposure visualization

R, Shiny App Nov. 2021

- Built a **shiny dashboard** to visualize hurricane Ike-2008 exposure, **combined rainfall data with geography information** to kriging rainfall in county-level after the landfall of the hurricane
- Check out the App at : **https://elisa1999.shinyapps.io/hurricane/**

SUFE Connect+ Project

Python, Flask, NER Sep. 2020 - Jan. 2021

- Built a **web application** to provide a platform for students at SUFE to find internship and share interview experiences
- Implemented the **matching process** of job applicants' resumes with the job information by extracting information on the resume using **NER (Named Entity Recognition) model**

Iowa Liquor Sales Data Warehouse Project

SQL server, PowerBI Dec. 2020

- Used Iowa Liquor sales data from 2012 to 2020 to implement a **data warehouse** in **four dimensions** from a perspective of a vendor; **imported the cleaned data into SQL server**
- Visualized the data warehouse in **Power BI** following the business needs of the vendor

Financial Fraud Data Mining Project

Python, LightBGM Aug. 2020

- Established a user portrait model by **mining the data** from online microlending platforms; **Predicted the probability of fraud** based on user's purchasing habit and credit score
- Filled the missing values using **Random Forest** and reduced the number of features with **PCA**; Used **AUC** as an indicator to select suitable models and reached an AUC score of 0.77