Haochen Miao

Seattle, WA | email | LinkedIn | 949-572-7244| GitHub

Motivated and detail-oriented Data Analysis student with experience in collecting, analyzing, and interpreting complex data sets using various tools such as SQL, Excel, and Tableau. Skilled in statistical modeling, data visualization, and programming languages such as Python and R.

Skills

Technical: Python, R, SQL, Java, JavaScript, HTML, MATLAB, GIS
Languages: English, Mandarin Chinese
Other: SkiLearn, Tensor Flow, Tableau, Apache Spark.

Education

University of Washington March 2024
Bachelor: Data Science & Geographic Information Systems GPA: 3.91 Seattle, WA
Relevant Coursework: Data Mining, Machine Learning, Database management, Statistical Analysis
Certificates: IBM Data Science Professional Certificate

Project Experience

Chatbot *(ML/ Jupyter, JSON, TensorFlow, matplotlib, numPy)* March 2023
The Chatbot project employs machine learning techniques, including natural language processing (NLP) and deep learning, to train models for generating responses to user questions.

- Implemented the Chatbot project, resulting in 25% improvement in customer satisfaction and an 82% accuracy rate in responding to customer inquiries.

Global Weather Manager *(Python, Apache Spark, SQL, Excel)* February 2023
The Global Weather Manager project leveraged global weather data to generate data-driven insights and predictions for future weather patterns. Processed 3M+ data points.

- Automated analysis of 150MB Excel weather data with Python, Apache Spark, and statistical techniques such as linear regression and Bayesian statistics.
- Managed with SQL DB, utilized multivariate and descriptive statistics for improved data modeling and query optimization, achieving 85% accuracy and 40% reduction in execution time.

COVID-19 Cases and Rates in the US *(Data Visualization/ JSON, HTML5, CSS, JS)* December 2022
The COVID-19 Cases and Rates project involved conducting data analysis and mining on a large JSON data file containing COVID-19 cases and rates for all US counties.

- Conducted data analysis and mining on 4M row COVID-19 JSON file, reducing processing time by 50%, improving accuracy.
- Utilized Tableau for data visualization and interpretation to create detailed reports and pivot charts, and developed interactive web map using Mapbox gl, HTML5, CSS, and JavaScript.

Lead Exposure & Cardiovascular Disease in Washington (Data Analysis/QGIS, JSON, SQL) July 2022
The Lead Exposure and Cardiovascular Disease identifies areas of high risk for lead exposure and cardiovascular disease mortality per 100k people in WA census tract.

- Created maps to visualize the results of the data analysis and utilized buffers and spatial queries to determine the impact of hospital proximity on cardiovascular disease mortality.
- Found a significant correlation between lead exposure risk and cardiovascular disease mortality , with a 25% increase in cardiovascular disease mortality rate in Southeast regions with high lead exposure risks.

Further Experience

MERCYWORLDWIDE™ October 2019 – Present

Assistant Coordinator Seattle, WA

collaborated with a team of 15 members to coordinate community services, including providing yard work assistance to the elderly and cleaning support to homeless shelters. Led the team in their efforts to make a positive impact on the local community.