Yi Su
+86 1320138138 | ysu020719@gmail.com

EDUCATION

University of Washington   Seattle, Washington
Bachelor's Degree in Data Science (GPA: 3.85/4.0)   *Sep 2020 - Jun 2024*

- **Core Courses:**Applied Statistics and Experimental Design | Statistical Computing | Databases And Data Modeling | Advanced Multivariable Calculus| Advanced Linear Algebra| Probability | Computer Language and Program Design

PROFESSIONAL EXPERIENCE

Suzhou Huichuan Technology Co.,Ltd.   Suzhou, China
Operations Management Engineer   *Jul 2023 - Sep 2023*

- Handled the statistical and analytical work of the Human Resources Department's operational data independently
- Generated and presented operational reports on a monthly basis to track and analyze departmental performance
- Leveraged the operational data effectively and effciently by using Business Intelligence tools(Power BI)
- Monitored the execution of the Human Resources Department's first-level organizational KPIs, including Recruitment success rate, turnover rate, overtime hours per capita, ensuring timely closed-loop corrections when necessary
- Played a pivotal role in maintaining and improving the department's operational efficiency and effectiveness

EMRTS   (Online) Seattle, USA
Data Analyst   *Apr 2023 - Jun 2023*

- Analyzed Medicaid Enrollment and Expenditure data in Washington State over the past ten years
- Employed data preprocessing algorithms with Python on billion size dataset, Enrollment Demographics, healthcare services Categories, Geographic, Income Levels, as well as Social Determinants of Health, improving the efficiency twice
- Created a visual narrative through box plots, time series to present trends and patterns, extracting valuable insights
- Utilized R and Python to perform statistical analysis and ensured data accuracy and consistency by performing data validation and collaborating with cross-functional teams

PROJECT EXPERIENCE

Visualize and analyze the Human Development Index (HDI) and social indicators dataset   *Mar 2023 - Jun 2023*

- Investigated longevity and education gaps reflected by   Switzerland and South Sudan to drive the health and education policy making.
- Utilized Python libraries such as pandas, seaborn, matplotlib, pyplot, scipy.stats, and numpy to clean a large and disorderly dataset.
- Employed data visualization techniques to extract valuable insights from the dataset,
- Applied statistical tests on mean years of schooling (MYS), life expectancy (LE), expected years of schooling (EYS), gender inequality index, and HDI to demonstrate the significance of association between longevity and education gaps .
- Presented our findings through visualizations and reports to summarize the results effectively. Prepared a concise 5- minute presentation to showcase our project and its outcomes

Designed a database for an online retailer   *Sep 2022 - Dec 2022*
*Achieved high grades, validating the effectiveness and quality of the designed SQL database

- Utilized SQL to build a robust database structure tailored to the retailer's needs
- Employed SQL windowing function, stored procedures, Subqueries, aggregate functions to optimize database performance and enhance functionalities
- Ensured efficient data retrieval and manipulation by implementing transactions and error handling technique

Conducted PCA analysis to abstract key contribution factors to the environmental performance index   *Jan 2022 - Mar 2022*
*Analyzed two extensive datasets, each containing over 30 variables, representing data from 180 countries worldwide.

- Utilized PCA (Principal Component Analysis) in R to identify and extract the most significant contributing factors to the environmental performance index
- Developed a user-friendly dashboard using Tableau to visualize and effectively communicate the findings and insights from the PCA analysis

Developed an R package comprising functions for K-Nearest Neighbors, Linear Model & One Sample T-Test *Jan 2022 - Mar 2022*

- Created a comprehensive R package with well-organized functions for K-Nearest Neighbors, Linear Model, and One Sample T-Test to facilitate statistical analysis and modeling
- Ensured a user-friendly experience by documenting the package extensively, providing clear explanations and examples for each function's usage
- Conducted rigorous testing to validate the package's functionality and accuracy, ensuring reliable and robust results for users

Developed a friend recommendation engine for a social network using Python   *Sep 2021 - Dec 2021*

- Implemented two distinct approaches within the recommendation engine:Utilized the number of mutual friends to suggest potential connections between users
- Leveraged a person's influence power within the network to make personalized friend recommendations
- Conducted extensive testing of the recommendation engine using Facebook graphs to validate its effectiveness
- Enhanced users' experience and fostering meaningful connections within the social network, by producing satisfactory results

LANGUAGES & SKILLS

- **Languages:** Chinese (Native); English (Proficient)
- **Programming Skills:** SQL, R, Python, Tableau, PowerBI