Xiyu Li

2807 Cresmont Ave, Baltimore, Maryland, 21211
• 865-403-9637 • xiyuli027@gmail.com • www.linkedin.com/in/xiyu-li-4628561a1 • https://rita-li527.github.io/

PROFESSIONAL SUMMARY

Data Analyst who drives results by transforming complex data into clear insights. Skilled in R, Python, SQL, Tableau, and Power BI, with expertise in statistical analysis, predictive modeling, and data storytelling.

SKILLS

- **Programming**: Python(Pandas, NumPy, Scikit-learn, SciPy, Matplotlib), R, Java, SQL, GitHub
- **Data Analysis & Visualization**: Tableau, PowerBI, MS Excel (PivotTables, Macros, VBA), Dashboard Development
- **Data Processing & Modeling**: Data Cleaning, Data Transformation, Predictive Modeling (Regression Analysis, Classification, Clustering, Dimension Reduction, Machine Learning)

WORK EXPERIENCE

Research Data Analyst   **May 2023 - Aug 2023**
Zhengzhou University, Department of Computer Science   *Zhengzhou, China*
- Supported development of real-time ECG recognition application, collaborating with cross-functional teams; accurately preprocessed data for **80+** patients using Python (NeuroKit), ensuring data quality and enhancing application accuracy
- Designed PowerPoint presentations and marketing posters, effectively communicating technical product details to diverse audiences and attracting potential customers, increasing engagement and interest in product

Data Analyst Intern   **Jun 2022 - Aug 2022**
Midwest Reliability Organization   *Saint Paul, MN*
- Curated and managed **5-year** operational records in Microsoft Excel, performed data refinement and analysis to enhance accuracy, and redesigned charts and graphs for 3 energy data systems, improving data integrity
- Built innovative **PowerBI dashboard** to visualize key performance metrics across three energy systems, streamlining communication with stakeholders and supporting data-driven decision-making within organization

PROJECT EXPERIENCE

Baltimore Crime Prediction and Visualization Project   **Jan 2025 - April 2025**
- Conducted end-to-end analysis of **14 years 750,000+** crime report data across 55 neighborhoods in Baltimore city
- Collected, cleaned and integrated millions spatial-temporal data from police reports, census sources, and NOAA weather files; applied spline-transformed time features to capture non-linear crime trends different type of incidents
- Assessed **KNN**, **Random Forest** and **XGBOOST** to predict monthly crime counts, explaining 85% of variation while addressing model complexity and avoiding overfitting through model tuning and validation
- Built interactive interactive **Tableau dashboard** to visualize crime types, trends, and area-level risk comparisons, capturing key spatial-temporal patterns to support targeted public safety planning

Flight Delay Analysis and Prediction Project   **Sep 2023 - May 2024**
Project Leader   *Baltimore*
- Led **team of 5** in comprehensive analysis of **100 million+** flight delay and cancellation data in US, coordinating weekly meetings, task distribution, and progress tracking to ensure efficient execution
- Examined **40+** predictors related to flight schedules, airport conditions, and weather through data visualization in Python and R, pinpointing key factors contributing to flight delays
- Utilized **logistic regression**, decision tree, and random forest models to forecast delays, achieving **72% accuracy** with **46%** true positive rate, helping airlines improve efficiency and reduce delay-rated costs
- Optimized results using Bayesian logistic regression with MCMC simulation and hierarchical models, increasing accuracy to **78%** on a refined 500-sample dataset

EDUCATION

JOHNS HOPKINS UNIVERSITY   **Aug 2023 - May 2025**
Master of Science in Engineering, Data Science   *Baltimore, MD*
- **Coursework:** Computing in Applied Math, Data Mining, Time Series Analysis, Bayesian Statistics, Parallel Computing

MACALESTER COLLEGE   **Sep 2019 - May 2023**
B.A., Statistics, Computer Science   *Saint Paul, MN*
- **GPA:** 3.73/4.0
- **Achievements:** ASA DataFest 2023 Awards - Best Interactive Visualization, Department Teaching Assistant
- **Coursework:** Statistical Modeling, Machine Learning, Database Management, Data Structure, Multivariable Calculus