Rachel (Qingrui) Li

778-321-3286 | qingruili3@gmail.com | https://github.com/qingruili | linkedin.com/in/qingruili

HIGHLIGHTS

Full-stack Data Scientist skilled in Python, SQL, ML, and data pipelines. Fast learner, self-motivated, and strong communicator, delivering data-driven insights and AI solutions while excelling in collaboration and problem-solving

SKILLS

- **Programming Languages & Frameworks:** Proficient in Python, SQL, C/C++, R, SAS, MATLAB, VBA, HTML/CSS, React.js, Node.js, PyTorch
- **Data Analysis & Visualization:** Pandas, NumPy, Matplotlib, Seaborn, Tableau, PowerBI, Machine Learning, A/B Testing, Hadoop, Spark, PySpark, Cassandra, MySQL
- **Cloud & Database Technology:** AWS (EC2, S3, Lambda, Redshift), Firebase, Alibaba Cloud, Google Cloud

EXPERIENCES

Data Analyst Intern, Kenda Rubber (China) Co. Ltd – Jiangsu, China May 2023 - Oct. 2023

- Designed and optimized a scalable data pipeline using **Python (Pandas, NumPy) and Alibaba Cloud**, integrating live-streaming sales data across multiple sources.
- Automated data cleaning for missing values and formatting errors, boosting dataset accuracy by **20%**
- Developed interactive **Tableau dashboards** with predictive analytics, enabling data-driven decisions that increased supply chain efficiency by **10%** and reduced operational costs by **5%**
- Partnered with 5 cross-functional teams to deliver software solutions and actionable insights, enhancing decision-making in logistics and operations

Business Analyst Intern, Kenda Rubber (China) Co. Ltd – Jiangsu, China May 2021 – Aug. 2021

- Gathered, analyzed and cleaned market data from **6** diverse sources using **Excel** and **Python**, improving data reliability by **30%** for strategic analysis
- Worked in a 3-member team to assess economic metrics and sales trends during COVID-19, forecasting market scenarios
- Crafted dynamic **Tableau** reports and presented data-driven recommendations to senior management, impacting key business strategies and improving operational efficiency by **15%**

PROJECTS

Song Trends Analysis Project Sept. 2024 - Dec. 2024

- Implemented a **PySpark-based data pipeline** on **AWS EMR**, efficiently processing **200K+** song data points alongside economic indicators and happiness metrics
- Automated **lyric retrieval via Genius API**, optimizing data storage in **AWS S3** and improving retrieval speed by **40%**
- Leveraged **Hugging Face NLP** for multilingual sentiment analysis, identifying emotional patterns in global music trends.
- Constructed an interactive **Power BI** dashboard to illustrate the relationship between economic indicators, song popularity, and sentiment dynamics

Hotel Price Optimization May 2023 – Aug. 2023

- Cleaned and preprocessed **100,000+ hotel pricing records** using **Python (NumPy, Pandas)**, improving data quality by 30%
- Developed **machine learning models** (Random Forest) to forecast hotel prices based on key factors

Pet Hospital Appointment System May 2022 – Aug. 2022

- Led a team of **3** to develop a comprehensive website by **HTML and CSS**, including a secure **SQL database** for storing medical records, a login system, and an appointment scheduler
- Established dynamic updates for doctor and hospital details, providing accessible medical information to help pet owners make informed decisions

EDUCATION

Simon Fraser University (SFU) **British Columbia, Canada**
Master of Science in Big Data - GPA: 4.22/4.33 Sept. 2024 – Present (Expected graduation: May. 2026)
Bachelor of Science in Data Science Sept. 2019 – Aug. 2023