Rui Qiu

RQ47@GEORGETOWN.EDU | (571)363-5917 | LINKEDIN/RQIU | GITHUB/REXARSKI | REXARSKI.COM

Education

• **Georgetown University**	Washington DC, United States
Master of Science in Data Science and Analytics. cGPA: 4.0	*Aug. 2021 – May 2023 (Est.)*
• **The Australian National University**	Canberra, Australia
Master of Science in Statistics	*Feb. 2017 – Dec. 2018*
• **University of Toronto**	Toronto, Canada
Honors Bachelor of Science with Distinction (Maths, Statistics and Computer Science)	*Sept. 2011 – June 2016*

Skills

- **Programming**: SQL, Python, R, JavaScript, HTML, CSS, Java
- **Machine Learning**: tidyverse, tidymodels, caret, jupyter-notebook, scikit-learn, numpy, pandas
- **Visualization**: ggplot2, R Shiny, leaflet, d3.js, matplotlib, seaborn, bokeh, plotly
- **Data Science**: A/B testing, ETL, pipeline (cleansing, wrangling, visualization, modeling, interpretation), scraping, feature engineering, NLP, cluster analysis, time series analysis, PCA, stochastic process, experimental design, hypothesis testing
- **Big Data & Miscellaneous Technologies**: AWS, Spark, MongoDB, Docker, Linux, Shell, Hugo, Figma

Experience

- **International Finance Corporation** — Washington, D.C.
  Data Scientist Intern — *May 2022 – Present*
  - Craft the right test harness to ensure models are tested using appropriate model evaluation techniques (bias detection)
  - Refine MALENA (Machine Learning ESG Analyst) Natural Language Processing & Natural Language Understanding models
  - Work with CEG and IFC stakeholders to understand requirements and propose innovative dashboards for ML inference data visualization
  - Support CEG and IFC stakeholders in conducting experiments and implementing an Explainable AI Framework for ML models
- **Georgetown University** — Washington, D.C.
  Teaching Assistant / Bootcamp Manager — *Aug. 2021 – Present*
  - Work as a graduate teaching assistant for ANLY-503 (Advanced Data Visualization) and ANLY-560 (Time Series)
  - Deliver weekly office hours and assignment grading for a group of 30+ students
  - Assist in sharing course-related materials, managing course Slack channel and GitHub organization
- **K2L Canberra** — Canberra, Australia
  Product Data Analyst — *July 2019 – June 2021*
  - Utilized Python and SQL to track and analyze product data, resulting in a 15% increase in recommendation accuracy
  - Built data visualization with MySQL and R for product KPIs that reduced manual reporting time by 3 hours per week
  - Improved UX measures by 20% through increased user feedback collection and frequent iteration to customer needs
  - Collaborated with the engineering team to launch an internal kanban tool for a better work-from-home experience
  - Launched 6 products (websites, mobile apps, and web apps) for clients in 2 years for a total 65% revenue gain
- **K2L** — Canberra, Australia
  Co-founder/Data Analyst — *Feb. 2019 – June 2019*
  - Led a team of 12 to develop an online Q&A community combining the idea of StackOverflow and TikTok which grew to 1,000 users in 3 months
  - Performed market competition analysis based on scraped data to direct the initial product planning
  - Analyzed and strategized regarding business goals, schedules, and budgeting in the early stage of the development
  - Created functional prototypes that increased communication efficiency by 30% for clients

Projects

- **What Ingredients Are You Tasting In Authentic Japanese Flavors?**: A comprehensive interactive visualization project on Japanese cuisine with D3, plotly and leaflet; Website
- **Basketball Analytics and Beyond: Analysis of 188,946 Shots**: A project on various basketball related topics, both on-court and off-court; Website
- **No Pie Newsletter**: A newsletter curating the latest articles, packages, tools and podcasts about data visualization; Archive, Telegram
- **xG Visualization Pipeline (R, sports analytics, visualization)**: An automation script to parse and clean raw data from soccer games and visualize the details in 3 different aspects; Tweet, archive, repository
- **tarantino (R package development)**: An R package that generates color palettes inspired by Quentin Tarantino's movies for visualization; Featured in R-Weekly newsletter; Repository
- **@canberramapbot (R, API)**: A Twitter bot posts bird's-eye view over Canberra periodically; Repository, Twitter
- **sentRy (R Server, AWS, API)**: A DevOps monitoring tool that automatically sends Django error logs to subscribers' Telegram account; Reduced the total amount of notifications by 40% via aggregation and trimming redundant information; Blog, repository