Abtahee Islam

Email: abtaheeislam@gmail.com
Mobile: 0466065027

SUMMARY

Analytical, solutions-focused student who bridges technical capability and business outcomes and combining Python/R/SQL, information-systems thinking, and a genuine interest in how technology drives customer value. Experienced working in fast-paced operational environments, collaborating cross-functionally to solve real problems, and communicating insights clearly to different audiences. Eager to support customers through digital transformation and contribute to high-impact initiatives at scale.

EDUCATION

University of Sydney (Bachelor of Advanced Computing)
- Majors: Business Information Systems and Data Science

SKILLS

- **Process & Systems Thinking**: Map end-to-end workflows, elicit requirements, write user stories and use cases, and align information systems design to business objectives.
- **Statistical Modelling & Machine Learning:** Linear and logistic regression, tree-based models, model evaluation (RMSE, R^2, accuracy), cross-validation, and model selection techniques (AIC).
- **Data Cleaning & Feature Engineering:** Handling missing data, outlier detection, transformation, unit standardisation, and building reproducible data preparation pipelines.
- **SQL & Data Analysis:** Writing efficient analytical queries with joins, aggregations, window functions, indexing, and query optimisation for large datasets.
- **Data Visualisation & Communication:** Performing EDA and presenting insights clearly through dashboards and structured reports using Python, R, Tableau, and Power BI.

TECHNICAL SKILLS

- **Languages:** Java, SQL, R, Python, Java, JavaScript, Bash, Git, HTML, CSS

- **Framework/Tools:** Quarto, PostgreSQL, Jupyter Notebook, Tableau, Power BI, Jira, Confluence, Excel, Pytorch, Google Workspace

PROJECTS

Predicting House Prices in New York (Data Science Project)
- 1,700+ New York residential property sales to identify key drivers of housing prices using structured statistical modelling.
- Performed data cleaning, transformation, and exploratory data analysis, including correlation analysis and visual diagnostics.
- Built and compared multiple linear regression and random forest models, evaluating interpretability vs predictive performance.
- Applied stepwise model selection (AIC) to identify a stable seven-predictor model explaining ~64% of price variation.
- Interpreted model coefficients and non-linear effects to produce actionable insights for buyers, developers, and planners.

Greater Sydney Liveability Index (Spatial Data Science Project)
- Designed and implemented a normalised PostgreSQL database integrating multiple large public datasets at SA2 level (population, income, businesses, transport, schools).
- Extracted and ingested live geospatial data from a NSW government API, automating collection across regions.
- Used PostGIS spatial joins to accurately map transport stops, points of interest, and services to ABS SA2 boundaries. Engineered a composite "liveability score" using SQL-based feature aggregation and per-capita normalisation.
- Optimised analytical queries using indexes (including spatial indexes) to improve performance.
- Presented findings through tables, visualisations, and a structured analytical report, linking data outputs to policy-style conclusions..

Student Survey Data Analysis (Individual Project)
- Cleaned and prepared a messy, non-random survey dataset, addressing missing values, inconsistent formats, and outliers.
- Standardised real-world inputs including height (mixed units), free-text gender responses, social media platforms, commute modes, and time-of-day data.
- Applied statistical tests (chi-square, Wilcoxon rank-sum, permutation test) with appropriate assumption checks and visual diagnostics.
- Interpreted results clearly, accounting for bias, limitations, and uncertainty, and delivered a reproducible analytical report.

HOBBIES/INTERESTS

- Business Information System Association(BISA) – Events Coordinator, organising events.
- NRL/Football