

Irene Zheng . Jarvis Consulting

Data Analyst with over 2 years of experience and a Master's in Statistical Modelling. Skilled in transforming data into actionable insights through automation, visualization, and storytelling. Experienced in SQL, Python, and BI tools such as Looker Studio, Tableau, and Power BI. Adept at collaborating with cross-functional teams to streamline analytics workflows and ensure data integrity. Passionate about designing scalable pipelines, enabling self-service analytics, and applying statistical modelling and machine learning to optimize business performance and inform strategic decisions.

Skills

Proficient: SQL (BigQuery, PostgreSQL, MySQL), Python (Pandas, NumPy, Matplotlib), Data Visualization (Looker Studio, Tableau, Power BI), Data Cleaning & Transformation (ETL/ELT), Statistical Modelling and Analytics

Competent: R (Statistical Computing and ggplot2), Cloud Platforms (AWS, GCP), Data Warehousing (BigQuery, Snowflake), Git & Version Control (GitHub), Dashboard Design and KPI Reporting

Familiar: PySpark (Distributed Data Processing), Docker & CI/CD, Machine Learning (Regression, Clustering), SAS (Statistical Programming), Java, HTML/CSS/JavaScript (Basic Scripting)

Jarvis Projects

Project source code: https://github.com/jarviscanada/jarvis_data_eng_IreneZheng

Cluster Monitor [GitHub]: Designed and implemented a distributed Linux cluster monitoring agent to collect real-time hardware and usage metrics from multiple hosts. Built Bash scripts using vmstat, lscpu, and system utilities to gather CPU, memory, disk, and network statistics, and automated data ingestion with cron-based scheduling. Dockerized a PostgreSQL database and designed normalized schemas to store both static host metadata and streaming usage data. Developed analytical SQL queries for capacity planning, trend analysis, anomaly detection, and performance reporting. Ensured reliability through error handling, input validation, and modular script structure, enabling scalable monitoring across multiple nodes in the cluster.

SQL Data Analytics [GitHub]: Designed and developed relational database schemas and optimized SQL workflows to support analytics and business intelligence reporting. Created normalized tables and enforced data integrity through keys, constraints, and indexing strategies. Wrote complex SQL queries involving multi-table joins, window functions, aggregations, and subqueries to answer analytical questions and generate actionable insights. Built reusable reporting queries for performance tracking, cohort analysis, and trend identification. Improved query efficiency through execution plan analysis and optimization techniques, enabling faster data retrieval, reduced processing time, and more reliable insights for downstream decision-making.

Spring Boot App [GitHub]: Developed a RESTful microservice using Spring Boot and Docker to demonstrate modern Java backend development. Implemented CRUD endpoints, layered architecture, and API testing for robust functionality and maintainability.

Highlighted Projects

E-commerce Performance Dashboard: Designed and developed a full-stack e-commerce performance dashboard using Looker Studio and BigQuery. Automated SQL pipelines for KPIs such as conversion rate, revenue per hour, repeat purchase rate, and customer lifetime value. Reduced manual reporting time by 30% and improved decision-making efficiency through real-time analytics.

SDLC Process Standardization Framework: Developed and standardized SDLC documentation and delivery workflows during my Bytedance internship. Created reusable compliance templates, requirement checklists, and process guidelines adopted by 20+ R&D teams. Improved project delivery consistency, reduced onboarding time for new engineers, and strengthened product quality and regulatory readiness across multiple business units.

Marine Conservation Analytics (OCC): Consolidated and analyzed multi-source marine datasets to study debris distribution and ecosystem impact. Built Python-based data cleaning and visualization workflows, applying time-series and spatial analyses to identify key patterns. Delivered findings to support sustainable marine ecosystem management and policy development.

Professional Experiences

Data Analyst Consultant (Contractor), Branchy Solution (May 2025 - Jul 2025): Designed and developed an end-to-end e-commerce performance dashboard using Looker Studio and BigQuery, enabling business users to monitor key metrics in real-time and supporting data-driven decision-making. Built SQL pipelines and automated metric calculations for conversion rate, revenue/hour, repeat purchase rate, and customer lifetime value, improving marketing resource allocation and reducing reporting time by 30%. Created visual insights, including time series trends, geo heatmaps, cohort analysis, and product funnel drop-offs, to identify sales bottlenecks and high-performing categories. Partnered with marketing and product stakeholders to align dashboard metrics with business goals and implemented filters and drill-down features to improve usability and adoption.

Technical Program Manager Intern, Bytedance (TikTok) (Feb 2023 - Nov 2023): Led cross-functional planning and execution of four TikTok data compliance initiatives, collaborating with 14+ R&D teams to ensure regulatory adherence across data ingestion and access workflows. Integrated 3,000+ Policy and Security Modules (PSMs) into the Bytemesh system to strengthen internal traffic monitoring and control, improving audit readiness. Standardized project delivery using Meego, building reusable SDLC compliance templates adopted by over 20 teams across different business units. Delivered internal training, documentation, and built an on-call issue tracking platform to support smooth adoption and continuous improvement of compliance processes.

Risk Control Product Manager Intern, Sina Weibo (May 2021 - Aug 2021): Participated in the launch and testing of key risk control features across Weibo's public comment system; independently validated business logic and ensured consistency in moderation outcomes. Coordinated engineering and product teams to deliver anti-abuse modules on time, resolving 80% of launch blockers before deployment. Conducted UAT testing, created documentation for feature rollouts, and delivered product demos to internal stakeholders. Contributed to a cross-departmental effort to reduce harmful content, helping decrease inappropriate posts in high-traffic areas by 20% using improved detection rules.

Education

Northeastern University (Sept 2022 - May 2025), Master of Professional Studies, Statistical Modelling - Capstone: Sponsor-driven analytics dashboard/model - Coursework: Data Warehousing & SQL, Machine Learning, Data Engineering

University of Toronto (Sept 2018 - May 2022), Honours Bachelor of Science, Statistical Machine Learning and Data Mining - Received UofT Entrance Scholarship of CAD \$7,500 - Graduated with Honours

Miscellaneous

- Academy Accreditation - Databricks Fundamentals
- Skiing and outdoor recreation
- Badminton
- Community volunteering