Prayag Purohit

Toronto, Canada | +1 (647) 614 2425 | prayagpurohit1@gmail.com | https://prayag-purohit.github.io/portfolio/ | LinkedIn

Summary

Results-driven Data Professional with 2 years of hands-on experience translating complex business needs into actionable data solutions. Blending a strong foundation in Finance and Economics with advanced Data and Business Intelligence skills, I excel at leveraging modern data technologies to drive insights, support informed decision-making, and optimize operations. Proven ability in communication and business acumen to deliver impactful, data-driven strategies.

Skills

Programming Languages: Python | R | SQL | DAX

Tools: Tableau | Power BI | Advanced MS Excel | Office suite

Cloud & DevOps: Git | AWS (CPP Certified) | Google Cloud Platform | MongoDB

Techniques: Machine Learning fundamentals | ETL pipelines | KPI definition | Statistical analysis

Soft Skills: Analytical thinking | Requirement gathering | Problem Solving | Collaboration | Communication | Storytelling

Work Experience

Research Assistant – Humber College Research and Innovation – Toronto, Canada   Mar 2024 – Present

- Designed and implemented scalable data pipelines to automate the transformation and enrichment of unstructured datasets used in experimental research.
- Led end-to-end architecture planning, including version control, modular design, and CI-ready workflows to ensure reproducibility across multi-phase experiments.
- Built a production-grade codebase (2,000+ lines) in Python to automate data ingestion, preprocessing, and treatment generation — reducing manual effort by >90% and saving $5,000+ in labor costs.
- Integrated MongoDB and cloud storage solutions to manage high-volume JSON data with schema validation and error handling for robust pipeline performance.
- Enabled collaboration across technical and non-technical stakeholders by developing reusable, well-documented scripts and version-controlled environments using Git.

Freelance Analytics Engineer – Bondi Water Controls – Toronto, Canada   Dec 2024 – Present

- Developed and designed zero cost, scalable data engineering pipelines to collect and analyze data from 1,000+ IoT devices deployed across the Greater Toronto Area.
- Collaborated with stakeholders to define business requirements and build analytics solutions that drove operational efficiency and cost savings.
- Developed a KPI framework aligned with client goals to enable data-informed decision-making and real-time performance tracking.
- Refactored legacy codebases for clarity, security, and performance; implemented caching layers that reduced dashboard load times by 80%.
- Restructured the alert data model (from flat files to nested schemas) and built an interactive dashboard, significantly improving data accessibility and increasing client satisfaction.
- Automated monthly reporting workflows using AWS Lambda and Python, reducing manual reporting effort by over 90% and improving data delivery timelines.

Business Intelligence Associate – OPLinnovate Pvt Ltd – Ahmedabad, India   Aug 2022 – Aug 2023

- Collaborated with cross-functional teams in a fast-paced environment to analyze large datasets with Tableau, SQL and MS Excel.
- Initiated efficiency improvements by learning VBA and developing scripts, reducing manual efforts by 60% within 4 days.
- Produced monthly operational KPI reports, interpreting consumer behavior and data trends to guide promotional strategies, driving a 5% month-over-month increase in consumer acquisition.

Education

Humber Institute of Technology & Advanced Learning   Toronto, Canada
Post Graduate in Business Insights and Analytics   Graduated: Jul 2025
Awards: Dean's Honor Roll all semesters, Datathon Runner-up
Relevant Coursework: Machine Learning, Big Data Management, Research Methods, Canadian Business

Ahmedabad University   Ahmedabad, India
Bachelor's in Business (Dual major: Finance and Economics)   Graduated: Aug 2021

Project Experience

[Ultramar Inventory Tracking System](#)
Skills: Python | Streamlit | Git | Systems Planning | LLMs | Pandas

- Built an automated inventory reconciliation system for a convenience store in Aurora that ingests, cleans, and standardizes sales and invoice data from Excel and PDF files, enabling dynamic stock level updates through a Streamlit-based interface.
- Developed ETL pipelines using Python and Pandas to preprocess structured (Excel) and unstructured (PDF) data, including sales logs and beer store invoices, integrating a Gemini LLM API for PDF-to-CSV conversion.
- Engineered robust data cleaning and transformation logic, incorporating column standardization, type coercion, merging datasets, and generating analytical summaries (sales, restocks) to ensure inventory accuracy and traceability.

[Spending Tracker](#)
Skills: Python for finetuning LLM model & ETL | MS-SQL for data warehousing (SSIS & SSMS) | Power BI for dashboarding | Git | SQL

- Designed an end-to-end pipeline to categorize transaction data into 21 categories using a fine-tuned language model.
- Built automated data extraction scripts, integrated them with a data warehouse, and implemented a Power BI dashboard for real-time insights.
- Improved classification accuracy by 15% compared to a leading competitor, validated through chi-square testing on unseen data.

Languages

- English – Fluent – CLB 12 (Top 0.01% score in CELPIP)
- Hindi – Native speaker
- Gujarati – Native speaker

Volunteer Experience

Student Advisory Committee Member – Humber College   Sep 2024 – Dec 2024
- Collaborated with faculty and students to address concerns, propose new initiatives, and drive campus improvements.
- Demonstrated leadership and teamwork by breaking down complex problems and implementing actionable solutions.

Student Peer Mentor – Humber College   Jan 2024 – Apr 2024
- Guided first-year students academically and socially, easing their transition into college life.
- Organized workshops and discussion sessions, fostering a supportive and collaborative learning environment.