Samaneh Hajigholam . Jarvis Consulting

I'm a software developer with a technical foundation in databases, Python, and cloud-based data systems. My background combines practical experience analyzing healthcare data with hands-on work in data analytics and visualization. I also bring applied software development experience, providing a solid understanding of data engineering concepts such as ETL pipelines, RESTful APIs, and containerized environments. I'm passionate about transforming data into insight and building scalable, data-driven solutions that improve decision-making and deliver real-world impact.

Skills

Proficient: Python (Pandas, NumPy, Matplotlib), RDBMS/SQL, Linux/Bash, Git, Agile/Scrum

Competent: ETL Processes & Data Pipelines, Cloud Platforms (AWS, Azure), RESTful APIs, C/C++, Java, MERN Stack (React.js, Node.js, MongoDB)

Familiar: PySpark, C#, React Native, Docker, Machine Learning

Jarvis Projects

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

Cluster Monitor [GitHub]: Developed a Linux-based cluster monitoring system to collect and analyze CPU, memory, and disk usage across multiple servers running Rocky Linux 9. Implemented automated metric collection using Bash scripts and cron jobs, storing time-series data in a Dockerized PostgreSQL database for consistency and portability. The solution centralized system performance data and enabled efficient tracking of resource usage across nodes.

SQL Analytics [GitHub]: Implemented data modeling and SQL queries (joins, aggregation, filtering, string operations) using a Dockerized PostgreSQL instance. Reinforced analytical concepts through hands-on practice in relational database design and performance optimization.

Core Java Apps [GitHub]: Grep App: Built a Java-based CLI application similar to Linux grep that recursively searches directories using regex patterns and outputs matching lines. Applied object-oriented design, file I/O with buffered readers, and logging via SLF4J/Log4j. Packaged the app as a Fat JAR with Maven and containerized it using Docker for consistent deployment.

Python Data Analytics [GitHub]: Analyzed two years of real-world retail transaction data to uncover customer purchasing patterns and support data-driven marketing strategies. Built a reproducible analytics workflow using Dockerized PostgreSQL and Jupyter Notebook, and performed data cleaning, validation, and feature engineering with Python, Pandas, and NumPy. Implemented customer-level metrics and RFM segmentation while handling production data challenges such as cancellations, missing values, and duplicated frequency distributions. Delivered actionable insights to identify high-value and at-risk customers, enabling targeted promotions and retention strategies.

Equity Analytics Data Pipeline [GitHub]: Implemented an end-to-end batch data pipeline to ingest, validate, and analyze equity market data using Apache Spark and Apache Airflow. Built a medallion architecture with Spark transformations that enforce financial data quality rules. Orchestrated the workflow in Airflow with task dependencies, retries, failure callbacks, and data quality gates to prevent invalid data. Produced analytics-ready gold datasets for sector- and stock-level performance analysis, with automated validation and Pytest coverage to ensure the reliability and reproducibility of investment insights.

Highlighted Projects

PL/SQL Data Management System [GitHub]: Built and optimized stored procedures, triggers, and reporting views across multiple relational tables (Players, Teams, Rosters). Demonstrated proficiency in SQL logic, data relationships, and analytical reporting automation.

Analysis of Student Hope and Future Plans: Analyzed survey responses from 350 medical students to explore their hopes and expectations for their future careers, using Python (Pandas, NumPy) for data cleaning, transformation, and statistical analysis. Built an ETL workflow to extract raw survey files, handle missing values and duplicates, standardize categorical responses, and structure the data for analysis. Conducted correlation analysis to identify patterns in career aspirations, stress factors, and perceived challenges, revealing key trends such as the top three areas of concern affecting student optimism. Created interactive visualizations with Plotly and Seaborn to communicate findings to academic advisors.

Full-Stack E-Commerce Website [GitHub]: Built an e-commerce web app using ASP.NET Core MVC, SQL Server, and Azure deployment. It allows administrators to manage orders, users, and products efficiently.

Professional Experiences

Data Engineer, Jarvis (2025-present): Contributed to the design and development of data-driven software solutions within an Agile/Scrum environment, collaborating with engineers and technical leads to deliver backend and data-focused projects. Participated in sprint planning and daily stand-ups to support high-quality delivery. Worked hands-on with Python, SQL, and Java to build and maintain backend components, data processing workflows, and analytical data pipelines. Applied best practices in version control, testing, and code reviews using Git, Docker, Linux, and distributed data tools.

Full-Stack Developer, Inlearna(Startup) (Nov 2024-Jan 2025): Developed a learning platform using Node.js, Express.js, MongoDB, and React.js. Designed and implemented RESTful APIs for efficient data management, ensuring seamless communication between front-end and back-end systems. Built and optimized database schemas and queries in MongoDB, improving data retrieval speed and reliability for the platform.

Frontend Developer, McMaster University(iGem Group) (Jun 2024-Oct 2024): Built a responsive website using React.js, TypeScript, and Tailwind CSS for an international research competition. Collaborated remotely with researchers and developers to deliver a functional, visually clear, and scalable academic web solution.

Research Analyst, Rajaei Heart and Research Center (Sep 2019-Apr 2020): Collected, cleaned, and transformed clinical and research datasets to ensure data accuracy and consistency. Wrote SQL queries to extract and aggregate data from relational databases and used Python (Pandas, NumPy) for data preprocessing, validation, and exploratory analysis. Built repeatable data preparation workflows and produced visualizations and analytical reports to communicate insights on patient outcomes and research findings to clinical and research stakeholders. Maintained data quality standards and ensured proper handling of sensitive healthcare data.

Education

Seneca College (2022-2024), Diploma, Computer Programming & Analysis - Dean's Honour List - GPA: 4.0/4.0

Concordia University (2021-2022), Bachelor of Science, Computer Science (Transferred Program) - GPA: 3.86/4.0 - Transferred out

Tehran University (4 years), Bachelor of Science, Health

Miscellaneous

- IBM Data Analyst Professional Certificate(Jun 2025 - Sep 2025)
- Microsoft Azure AI Fundamentals(Jul 2025)