Geetika Sood

gtksood@gmail.com | https://www.linkedin.com/in/geetika-sood-07 | +1 312-522-4908 | Chicago, IL

Professional Summary

Results-driven Data Engineer with 5+ years of experience in data integration, ETL, and database management. Adept at working with SQL, Python, and data visualization skills to extract, transform, and analyse complex datasets. Experienced in collaborating with cross-functional teams to implement data solutions that optimize business operations. Passionate about leveraging data to drive decision-making and operational efficiency in the healthcare and financial domains.

Professional Experience

Data Engineer   **Dublin, IE**
Insight SFI Research Centre, DCU   11/2022 - 10/2024

- Created an end-to-end digital data asset for Irish hospital systems by establishing a unified database for patient detail that reduced the data retrieval time by 30% for system users.

- Utilized PostgreSQL and Python to integrate large and diverse data sources, including patient records, financial transactions, and operational metrics for data extraction and processing, while ensuring accuracy and consistency.

- Utilized Neo4j to analyse entity relationships across entities in heterogeneously sourced data sources to better model the unified database and led a POC to host the initial model onto Snowflake platform to achieve data lineage and change data capture for better CI/CD.

- Collaborated with interdisciplinary researchers and healthcare professionals to identify gaps in the system and translate complex healthcare domain requirements into actionable data insights, contributing to advancements in public healthcare data integration.

- Co-authored academic preprints and a proposal that secured a position in the final phase of the National Challenge Fund (NCF), winning a grant of €2.5 million for the project.

Data Engineer (Technology Analyst)   **Bangalore, IN**
Infosys Ltd   02/2017 - 09/2021

- Re-engineered and migrated legacy applications from IBM mainframes to Micro Focus mainframes, ensuring seamless transition and improved system performance through detailed planning and execution.

- Developed robust data pipelines using Python and Pandas to migrate, transform, and validate data based on complex business logic, ensuring data integrity and alignment with business requirements.

- Automated trigger-based report classification and custom report generation using Python, reducing manual intervention by 50% and enabling business stakeholders to access accurate, timely information for daily operations and decision-making.

- Acted as L3 SPOC for resolving critical production issues, working closely with L1/L2 teams and business stakeholders. Participated in Agile/Scrum processes to ensure timely feedback and alignment during system migration and POC phases.

Education

Master of Science | Dublin City University | Dublin, IE

Bachelor of Engineering | Chitkara University | Chandigarh, IN

Key Skills & Technologies

- SQL (PostgreSQL, MySQL), Vector database

- Relational Database & Datawarehouse concepts

- Data Integration & ETL (Extract, Transform, Load)

- Python (Pandas, NumPy) & Data Processing

- Microsoft Excel (Advanced Functions, Pivot Tables)

- Data Visualization (Tableau, Looker)

- Agile & Project Management (JIRA, Git)

Key Projects & Achievements

- **Secured €2.5 million in funding** as a contributor and co-author of a research project proposal in the National Challenge Fund (NCF), demonstrating expertise in strategic data-driven solutions.

- **Recognized as a 'Most Promising Fresher' at Infosys** for outstanding performance in data engineering and analytics.

- **Selected for Infosys ACON Leadership Program**, receiving Harvard Business Soft Skills and Financial Domain Training.

Publications & Certifications

- **Exploring Chronic Disease Trends among Adults in the USA: A Statistical Analysis with Visual Insights**.
  The study analyses U.S. adult chronic disease trends extracting data from OLTP systems of CDC(U.S.) and forming OLAP system using PostgreSQL for data management and Tableau for visual insights. It reveals rising rates of diabetes and obesity, with significant disparities across demographics. Findings emphasize the need for targeted interventions, preventive care, and informed policies to address growing public health challenges.

- **Query Based Construction of Chronic Disease Datasets**
  It presents a method for building chronic disease datasets using SQL-based querying techniques. Emphasizing automation and accuracy, it showcases skills in SQL, data modelling, and ETL processes, enabling efficient data extraction, integration, and preparation for healthcare analytics and chronic disease trend analysis.

- **Certifications:** SQL for Data Science (Coursera), Data Visualization with Tableau (Udemy).

Additional Information

Authorized to work in the U.S. without visa sponsorship