Shriyansh Singh

+1-930-333-5141 | shriyansh.singh24@gmail.com | LinkedIn

SUMMARY

Healthcare Data Engineer with 3+ years of experience building **scalable ELT pipelines** and **data models** for analytics. Expertise in **SQL**, **Python**, and **cloud platforms** with strong focus on data quality and **healthcare analytics**.

PROFESSIONAL EXPERIENCE

Healthcare Data Engineer
April 2024 - Dec 2024

Hyphenova AI Health Division
Los Angeles, CA

- **Designed** and **implemented dimensional data models** using **star schema** that integrated clinical and claims data from disparate sources, enabling 360° patient analytics
- **Constructed** automated **ELT pipelines** using **Python**, **SQL**, and **Databricks** that processed terabytes of healthcare data with 99.9% reliability and compliance with HIPAA regulations
- **Developed** comprehensive **data quality frameworks** with proactive monitoring that reduced data issues by 78% and ensured accuracy of downstream analytics products
- **Collaborated** with clinical leaders to translate complex healthcare requirements into technical specifications, delivering intuitive data models trusted by stakeholders

Data Integration Specialist
May 2022 - Oct 2022

Enterprise Business Technologies Healthcare Solutions
Mumbai, India

- **Architected** scalable **ETL workflows** using **SQL** and **Python** that standardized healthcare metrics across multiple provider systems, improving reporting consistency by 63%
- **Implemented version control** with **Git** for all data transformation code, enabling collaborative development while maintaining code quality through systematic reviews
- **Created** optimized **data models** that simplified complex healthcare relationships, accelerating query performance by 47% while enhancing analyst productivity
- **Established data validation protocols** that automatically identified anomalies in clinical and claims data, preventing erroneous analysis and ensuring regulatory compliance

HEALTHCARE DATA PROJECTS

Patient Outcomes Analytics Platform | *Databricks, Python, Spark, SQL, Azure, Tableau*   *Jan 2024 – Apr 2024*
- **Engineered** a comprehensive **healthcare data platform** using **Databricks** that integrated clinical, claims, and social determinants data into unified patient views
- **Developed automated quality checks** that evaluated data completeness, consistency, and accuracy against healthcare industry standards, ensuring reliable analytics
- **Designed** intuitive **Tableau dashboards** that visualized key performance metrics and quality measures, enabling clinical leaders to make data-driven decisions

Healthcare Data Integration Framework | *Python, SQL, AWS, Git, Snowflake*   *Sep 2023 – Dec 2023*
- **Architected** a scalable **ELT framework** on **AWS** that standardized the integration of diverse healthcare data sources including EMRs, claims systems, and labs
- **Implemented dimensional modeling** techniques that transformed complex healthcare data into intuitive **star schemas** optimized for analytical queries
- **Created** comprehensive documentation and knowledge transfer materials that enabled analysts to effectively utilize healthcare data models without deep technical expertise

TECHNICAL EXPERTISE

Data Engineering: ELT/ETL Pipeline Design, **Data Modeling**, **Data Quality**, **Dimensional Modeling**, **Star Schema**, Snowflake Schema
Programming: **Python**, **SQL**, **Spark**, PySpark, Shell Scripting, **Excel**
Database & Cloud: **Databricks**, **Azure**, AWS, Snowflake, PostgreSQL, SQL Server, Oracle
Business Intelligence: **Tableau**, Power BI, Looker, Data Visualization
Software Development: **Git**, **SDLC**, CI/CD, Documentation, Code Reviews, Testing
Healthcare Domain: Clinical Data, Claims Processing, HIPAA Compliance, Quality Measures, Population Health

EDUCATION

Indiana University Bloomington
Aug 2023 – May 2025

Master of Science in Data Science
Indiana, United States

- Relevant Coursework: Database Design, Healthcare Informatics, Data Warehousing, Statistical Analysis, ETL Processing