

**TEK011671** (Data Engineering - 5 + Year) BTech

Highlights

- Experienced Data Engineer and Business Analyst with 5+ years in management consulting and Life Sciences domain.
- Expertise in ETL, MDM, data analysis, SQL, Python, AWS, and Azure.
- I have a proven track record of driving business growth through technical solutions and data-driven insights.

Skills

Primary Skills

- AWS
- Azure
- Databricks

Secondary Skills

- Snowflake

Other Skills

- SQL
- Python and PySpark
- Apache Spark
- Data Lake and Data Mart
- Data Warehousing
- Data Design and Semantics
- Data Profiling and Data Quality
- Risk Identification and Mitigation
- Data Management and Analysis
- ETL Technologies
- DevOps and CI/CD process
- Descriptive and Predictive
- Analytics
- Data Visualization and Reporting
- Azure
- AWS
- Spark
- Databricks
- ADLS
- Looker
- Jira
- Jupyter
- Pycharm
- Matillion
- Snowflake
- Anaconda
- Git
- Fivetran
- Hive
- Control M
- MS Office tools

Projects

Data Engineer (57 months)

- Engineered data solutions using the Microsoft Azure stack (Azure Data Factory, Databricks, Data Lake Storage, Functions, etc) for a European client, leveraging Git for version control and seamless deployment to production environments. o Impact - Successfully onboarded multiple brand's data with robust ETL pipelines and reduced client cost by up to 60% by utilizing Azure cloud and its services.
- Delivered three consecutive projects to production in 2022 as a SQL/Python developer, utilizing SQL Workbench, Jupyter, and ETL tools like Matillion and Snowflake to streamline data processing and integration. o Impact - Projects Renewal and 100% client satisfaction with implementation and output.
- Optimized data pipelines with Apache Spark, Delta Lake, and Data Factory pipelines, achieving a 30% processing speed increase and a 15% boost in user retention.
- Developed and optimized Python-based BI reports, aligning with KPI requirements and business rules, to generate ready-to-use data for MSTR dashboards, enhancing decision-making capabilities.
- Migrated over 50+ SAS scripts to Python and SQL, significantly improving performance and reducing data processing time from hours to minutes, enabling faster analysis of large datasets.
- Bridged business need gaps with technical data solutions, enhancing workflows through effective design and management.
- Created comprehensive documentation for data mapping and conducted profiling and lineage analysis to ensure data accuracy.
- Proactively addressed data risks by implementing controls and strategies to maintain data integrity and reliability.

Associate Data Engineer/Analyst (28 months)

- Project Management and Stakeholder Coordination: Led end-to-end project delivery from requirement gathering to final delivery, managing interactions with multiple stakeholders and securing necessary approvals to provide data to clients. This involved technical support for new software systems, creating database structures, handling terabytes of sensitive client data using data warehousing, and monitoring ETL processes for efficiency.
- Data Analysis and Integration: Analyzed new business rules and developed enhancements for data requirements, including unit testing and performance optimization of code. Managed and improved existing data warehouse systems, worked with diverse pharmaceutical datasets (e.g., Exponent, DDD, APLD) and directly interfaced with third-party source vendors (e.g., IMS, IQVIA, Veeva) to ensure high-quality data integration and delivery.

Awards

- Certified in Python - PCEP
- Bravo Award for outstanding individual contribution to a project.
- Best project Award under the 5 million USD category.
- National Talent Search Examination (NTSE) Scholar