


**TEK06579 (SQL - 8 + Year)** Bachelor of Engineering

## Highlights

---

1. Having **8 years** of professional experience with Technical Expertise
2. As a Data Engineer, I have extensive knowledge working with Data warehousing, ETL & multiple SQL Databases. I specialize in designing, developing, and maintaining scalable data pipelines and infrastructure to ensure efficient data processing, storage, and analysis. My contribution mainly focused on fetching the data from multiple sources and converting into purposeful insights for Data Visualization purposes.
3. I excel in collaborating with cross-functional teams to understand data requirements.
4. Excellent team player with an ability to perform individually, quick-learner, meeting deadlines, good analytical and communication skills. Strong ability to manage & motivate the team, work well under pressure.

## Skills

### Primary Skills

- SQL

### Secondary Skills

- JIRA
- Python

### Other Skills

- Databases : SQL, Snowflake, Teradata, Oracle, Hive, PL/SQL
- Data warehousing : ETL (Extract Transform Load), Data Warehousing, Data Analysis, Data Modelling
- ETL Tools : IBM DataStage v11.3 / v11.7.1, Snap Logic, Qlik Enterprise Manager
- Reporting Tools : Power BI
- Deployment Tools : Git, GitLab
- Scheduling Tools : Tidal Automation, Control-M
- Big Data : Hadoop, Hive, Spark (PySpark)
- Programming Languages : Python, UNIX
- Other Tools : Jira, Azure Boards, PyCharm, Confluence
- Upgrading Technologies : Google Cloud Platform, Python

## Projects

### Project 1- American Express - IT Industry (23 months)

- Skill Set : Teradata SQL, PySpark, Hive, Hadoop, Python, Git, Confluence
- Team Size : 10

#### Description:

It involves finding solutions to Key Risk Identifier (KRI) for American Express and generating alerts to the Product Owner using Archer Dashboard using Big Data technologies like Hadoop, Hive and Programming languages like Python and PySpark. The Code has been thoroughly tested before deploying into production. The key objective is to arrive at a solution that is efficient and within the expected SLA.

#### Roles and Responsibilities:

1. Developing PySpark and Hive SQL code for Key Risk Identifier to create alerts and get pushed in Archer Dashboard.
2. Co-ordinated with product Owners to understand the requirements and implemented the code effectively
3. Worked on code conversion from Teradata to PySpark / Hive
4. Tested the implemented code with all possible test cases to avoid defects in production
5. Worked on Query Optimization techniques on long running batch codes to reduce the execution time
6. Reliability and Performance Test have been performed to test the stability of code under various dates and filter conditions
7. Ensuring the deliverables are done within the timely manner
8. Mentoring the new team members to explain the relevant technologies
9. Preparing the project status reports to be submitted to the senior management

### Project 2- Ally Financial - IT Industry (11 months)

- Domain : Auto Finance, Insurance
- Skill Set : Teradata SQL, PL/SQL, Snowflake, IBM DataStage, GitLab, Jira, Control-M, QEM Tool, Confluence, Python
- Team Size : 7 No's

#### Description:

Ally is one of the largest car finance companies in the U.S headquartered in Detroit, Michigan. The company provides financial services including car finance, online banking via a direct bank, corporate lending, vehicle insurance, mortgage loans, sale and lease agreements. We as Data processors team involved in development and enhancement of DataStage Jobs, Oracle, PL/SQL, Snowflake scripts, Performance Tuning in long running batch cycles

#### Roles and Responsibilities:

1. Participating in design discussions and analyze the problem in detail to understand the key points
2. Worked on code conversion from Oracle to Snowflake using IBM DataStage
3. Maintain historical data, which allows analysis of business events over time.
4. Debugging the issues in existing production script to avoid junk data
5. Scheduled jobs cycles using Control-M tool
6. Worked on enhancement of PL/SQL scripts to reduce batch cycle timings

### Project 3- Caterpillar - IT Industry (23 months)

- Skill Set : Snowflake & Teradata SQL / IBM DataStage / Tidal Automation
- Team Size : 15

### Description:

Caterpillar Inc. (often shortened to CAT) is an American Fortune 100 corporation that designs, develops, engineers, manufactures, markets, and sells machinery, engines, financial products, and insurance to customers via a worldwide dealer network. It is the world's largest construction-equipment manufacturer. The project involves maintaining and processing of dealer Data, Employee work hours, Freight recovery, Calculation of charges., etc.

### Roles and Responsibilities:

1. Worked on migrating **Teradata SQL** codes to **Snowflake Queries**
2. Involved in **Performance Tuning** and **Optimization of Snowflake queries/ DataStage Jobs** to reduce runtime in daily batch cycle
3. Worked on design development using **SCD type 1 and type 2**
4. **Taken responsibility** of complete DataStage Jobs & Snowflake SQL queries from development to successfully executing in Production
5. Attended Scrum meetings with on-shore business analysts to get specific user requirements and analyzing the mapping document
6. **Lead a small Team of 5 members** on Mini project for creating Power BI reports on Sales
7. Extended on weekend for supporting and monitoring production batch cycles for the New releases and code enhancements

## Project 4- Intesa Sanpaolo Bank - Finance Industry (20 months)

- Environment : Teradata SQL, Data Stage v11.3
- Team Size - 30

### Description:

Intesa Sanpaolo is the banking group, which was formed by the merger of Banca Intesa and Sanpaolo IMI, [Turin, Italy](#) The merger brought together 2 major Italian banks with shared values. This project involved new design from scratch to load the Business tables from Data Lake using DataStage and Teradata. Agile methodology is followed during the entire development process.

### Roles and Responsibilities:

1. Used **Teradata utilities** to extract the data from source and load to temporary tables
2. Worked on handling the CR's and implementing them on the already existing **Teradata** and DataStage objects
3. Worked on **Delta Logic** to fetch the Latest data from the Data Lake
4. Designing and developing various DataStage **Parallel and Sequence** jobs using the **SCD Type 1 and Type 2** logic and Worked on **Change Data Capture** to implement SCD - 2
5. Developing parallel jobs using DataStage stages like Sequential file, Aggregator, Copy, Funnel, Join, Lookup, Transformer, Merge, Remove duplicates, Filter, Change Data Capture, Peek, Surrogate Key, Row Generator, Column Generator, Teradata Connector, Sort
6. Grooming & Mentoring new team members in team to enhance the quality of code
7. Prepared RTC for DataStage jobs during package associating with Change Console

## Project 5- Wynn Resorts - IT Industry (13 months)

- Environment : Teradata SQL

- Team Size - 5

### **Description:**

Wynn is a luxury Resort and Casino located on the Las Vegas and It has two buildings adjacent to each other, Wynn Las Vegas & Encore Las Vegas. Both have more than 2000 rooms each and entire. Wynn owns another third property name as Encore Boston Harbor in Boston. Gaming and Lodging are the major areas have been contributed.

### **Roles and Responsibilities:**

1. Worked on Teradata **BTEQ, FLOAD, MLOAD, TPT** utilities scripts from source for ETL transformation
2. Defect analysis and resolution support for system, integration and User acceptance testing
3. Preparing Test cases/scenarios and capturing test results into documentation
4. Provided Production support for the Code releases and monitored the production runs and closed tickets raised.

## **Project 6- Operational Data Store - IT Industry (13 months)**

- Environment : Data Stage v11.3, Teradata SQL
- Team Size – 15

### **Description**

Operational Data Store(ODS) is a consolidation point for regions of an insurance provider having multiple Source transactions relating to the policy and claim. The ODS holds data from Subject Area like party, object, and reference data as well.

### **Responsibilities:**

1. Analysis, Design & Development of the DataStage parallel and Sequence jobs
2. Designed and developed the Sequencer jobs which will in turn trigger the Parallel jobs.
3. Involved in the Unit / System testing of the ETL and Database code.
4. Worked on identifying the defects and fixing them in a timely manner.
5. Quickly learnt the existing system from Insurance Team and transited to team members.
6. Planned the activities and met the Iterations cycles deadlines effectively.

### **Awards**

---

- Certified as IBM Infosphere DataStage Developer for Version 11.3