QuerySurge™

Automate your Data
Warehouse & Big Data Testing

The data testing software of choice
for Testers, Managers, Developers,
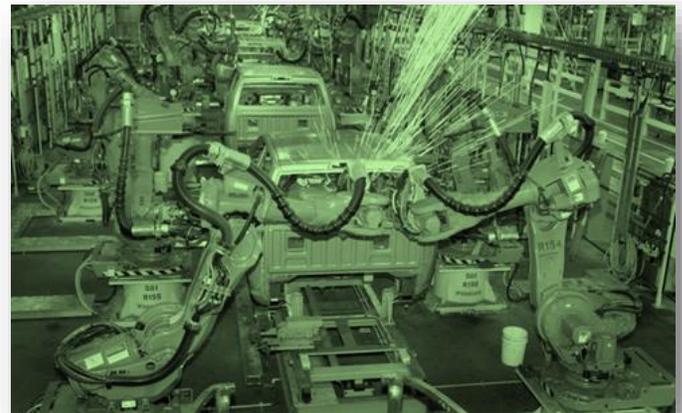Analysts, and Operations teams

Truck Manufacturer utilizes QuerySurge to achieve 100% data validation

A leading global industrial truck manufacturer was seeking an automated approach for its data warehouse testing initiative. They purchased QuerySurge and were able to increase testing coverage from 10% to 100% while greatly reducing testing resource hours.

Background:

A leading global truck manufacturer needed to test and validate its new data warehouse. Manual efforts by the internal testing team could only validate 10% of the data in the time allocated. The manufacturer searched for a more efficient automated testing process and discovered QuerySurge. After a successful 15-day free trial they purchased QuerySurge.

Challenges:

- The manufacturer needed to guarantee that the business-critical data was transformed from the source system to the target system according to mapping specifications.
- Bad data would cause incorrect Business Intelligence reporting information and could cause millions of dollars in lost revenue. The manufacturer needed to increase the amount of validated data from 10% to 100%.
- Manual testing, using minus queries and excel comparison, was taking too long and costing too much in test resources.

Solution:

Using QuerySurge, an automated approach to data validation and testing, the manufacturer was able to increase the amount of data tested tenfold, and decrease the cost of testing resources. QuerySurge was used to create automated tests (query pairs) over several weeks. A total of 20 query pairs were authored (20 queries against the legacy source system and 20 queries against the target data warehouse). These query pairs were logically grouped into test suites for scheduling test executions.

QuerySurge™

Automate your Data
Warehouse & Big Data Testing

The data testing software of choice
for Testers, Managers, Developers,
Analysts, and Operations teams

Execution time took 12 hours to execute the 20 query pairs and validate 400 million individual data cells. At the end of execution an email notification from QuerySurge was automatically sent to the team identifying the status (pass/fail) and completion time. Several test cycles were needed to clean the data, since defects (transformation & special character defects) were discovered.

This automation process resulted in verification of 100% of transformed data. The following project-level requirements were fulfilled:

- 100% of data was tested and validated
- Verification that the ETL code was developed correctly according to mapping specifications
- Verification that the ETL process did not improperly correct, reject or substitute data

Benefits:

The 20 automated QuerySurge tests discovered numerous critical defects that possibly could have cost the manufacturer millions of dollars. The defects discovered over the course of the ETL build process included special character defects. QuerySurge also provided metrics that were not previously available such as data reliability, using detailed reports generated by the tool. An additional value-add was the ability to prove to the manufacturer's customers that the data provided by the customers was incorrect - not the manufacturer's ETL process.

> **" We were able to prove to the customer that the data was successfully transformed per specification. The bad data we were seeing was due to bad data entry – not bad ETL transformations "**