

MAXIMIZING DATA'S POTENTIAL

Data Analytics in Data Storage Device Development and Test

Debin Wang
Sr. Engineer | Drive Design & Test Analytics Platform Lead
Drive Reliability Group
Seagate Technology (US)

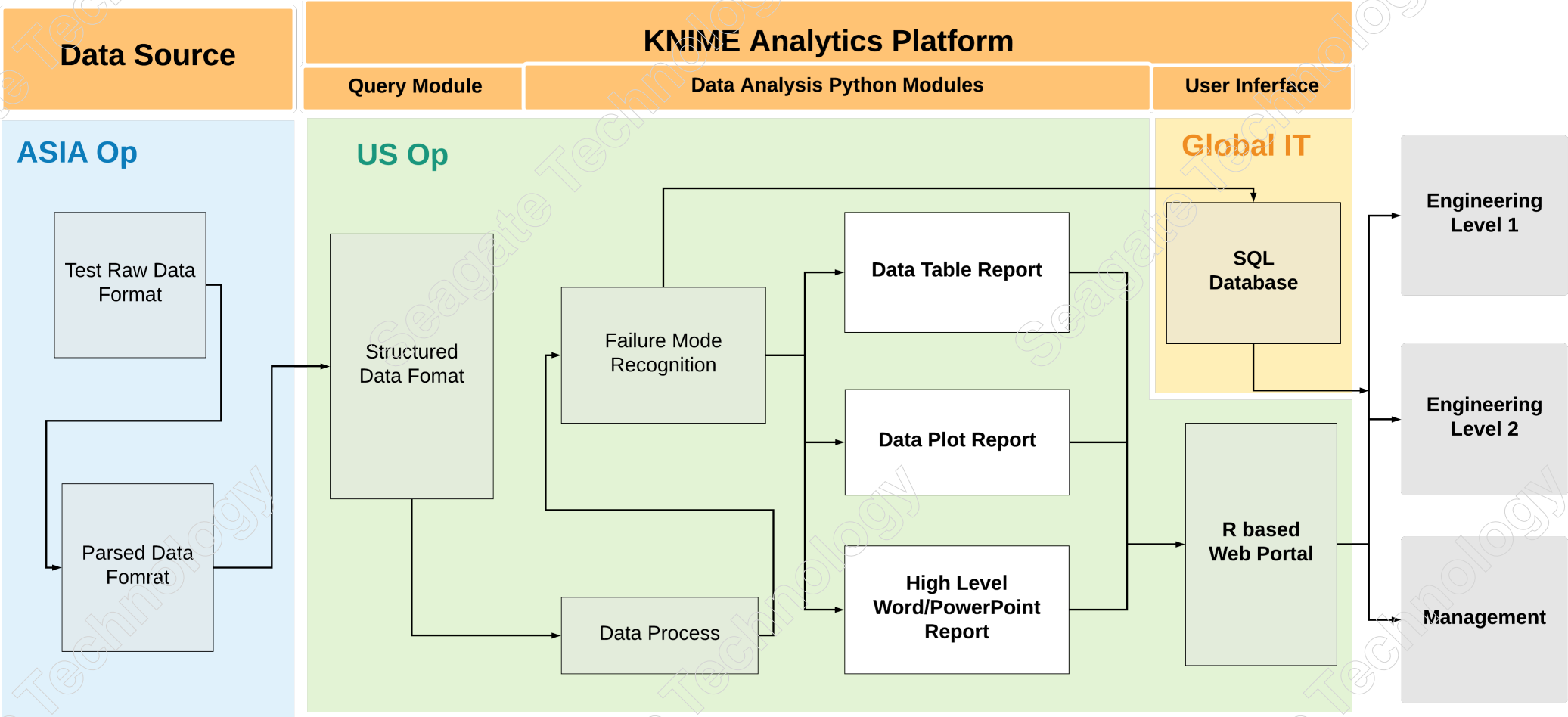


Agenda

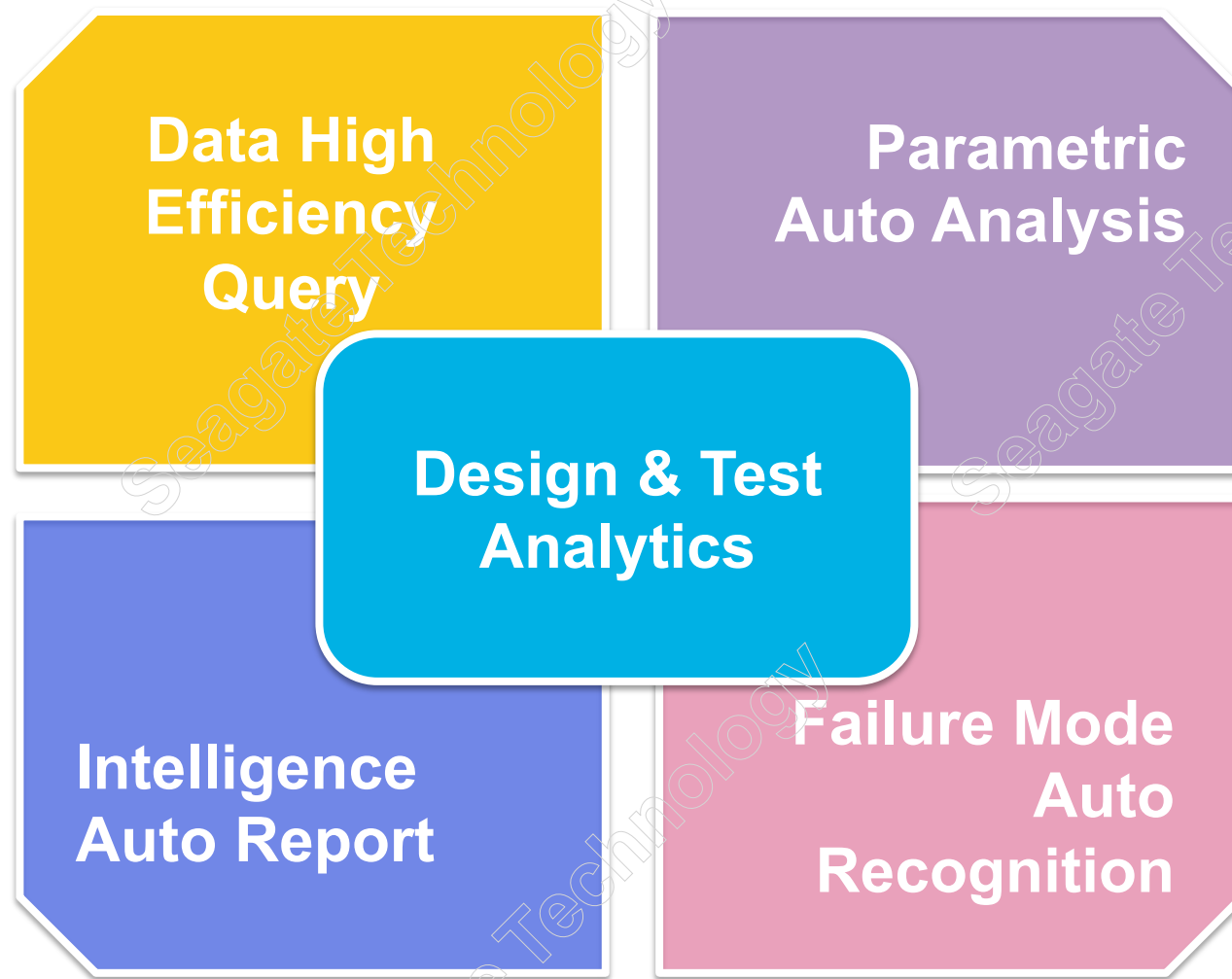
- ***About the Speaker***
- ***Drive Design & Development Analytics Platform***
- ***KNIME Applications & Working Cases***
 - *Integrated analytics platform*
 - *Parallel data query*
 - *Device failure mode recognition & prediction*



KNIME Deployment with Device Test Analytics Platform



Intelligence Platform for Device Design and Test



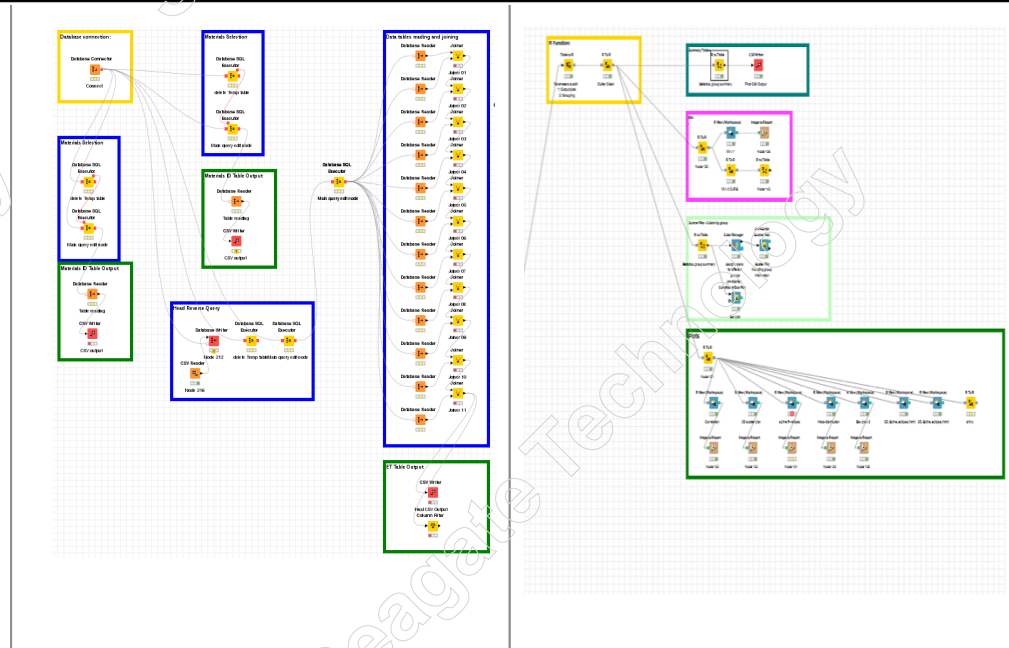
- **KNIME** as a key enabler to drive the entire analytical streamline process.



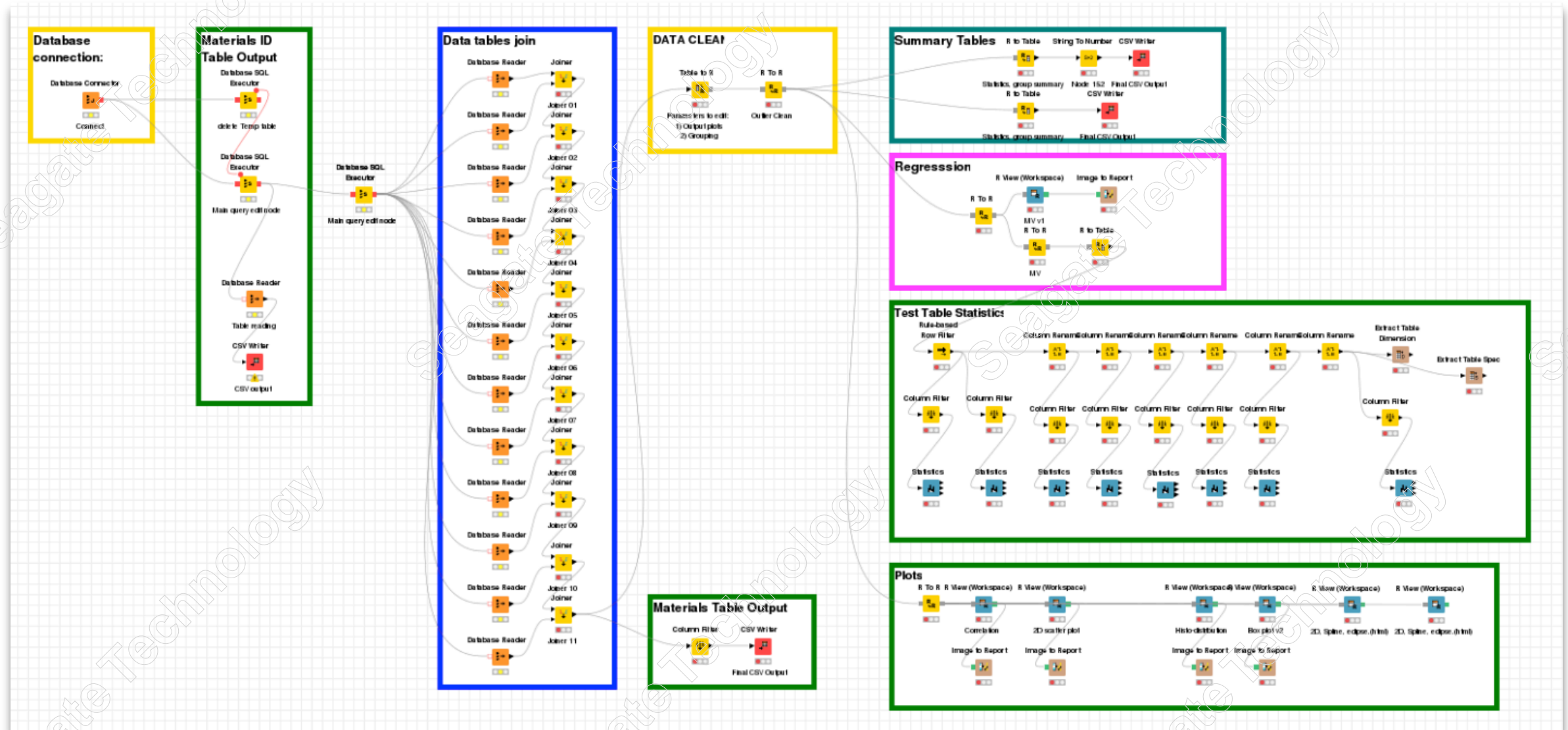
Integrated Drive Design and Test Analytics Platform

		<i>Data Query</i>	<i>Data Reports</i>
Data Source		Device test log files and tables, images, etc	Summary tables, plots, pareto analysis, and reports.
Platform	Before	Web Browser, SQL Developer.	Excel, SAS JMP, Rstudio, Python IDE, Word, PowerPoint
	Now	<i>KNIME</i>	
		SQL Nodes	Python, R, Java Nodes

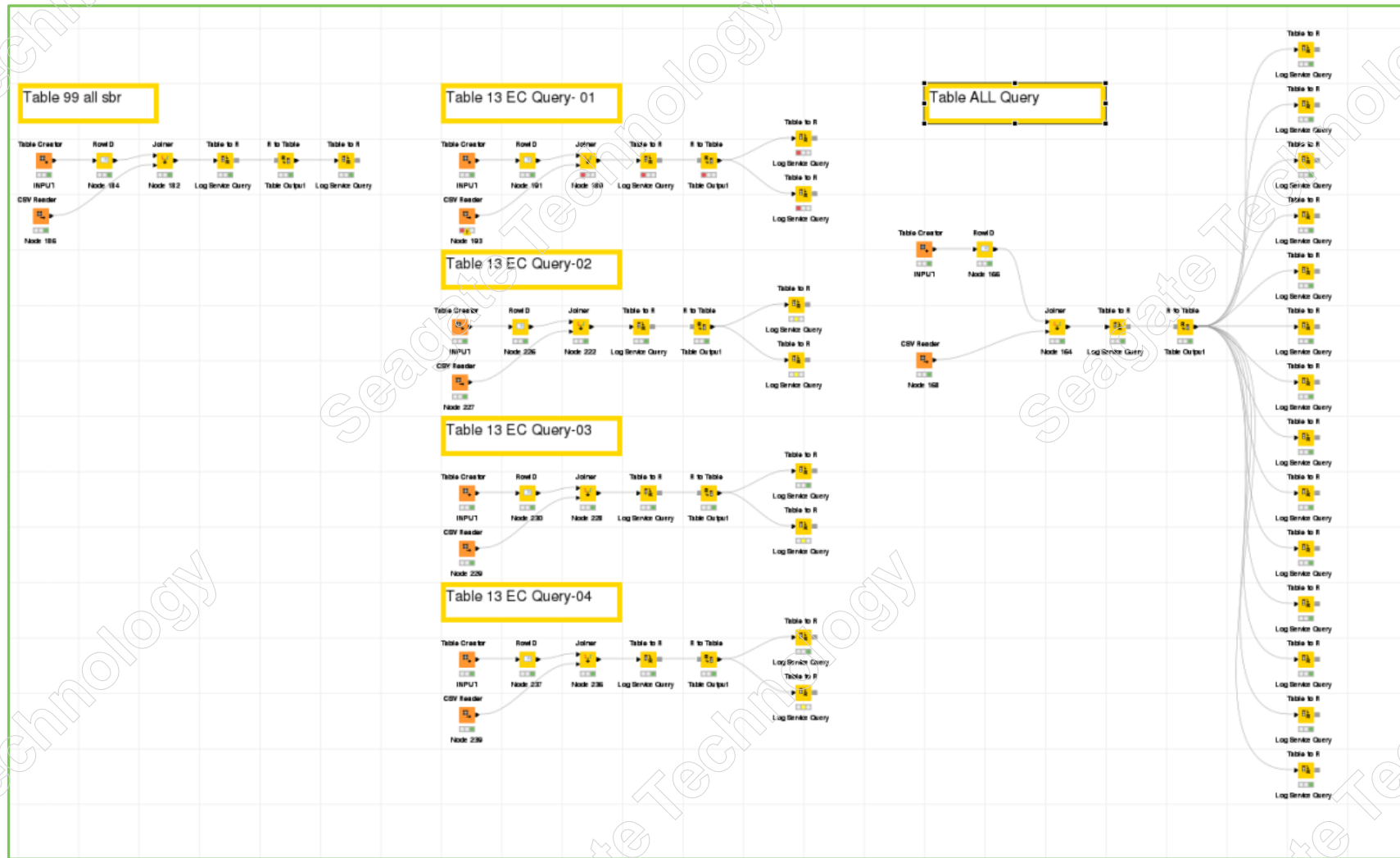
- **KNIME:** an integrated analytics platform that streamlines query, analytics, and reporting.
- **Cross platforms:** Windows, Mac, Linux (cluster server capable).
- **Open source** (cost reduction).
- **GUI:** visually create data flows, pipelines.
- **Integration:** Java, SQL, Python, Perl coding fragments, as well as R and ImageJ open source projects.



Work Case Scenario: Device Test Data Analysis



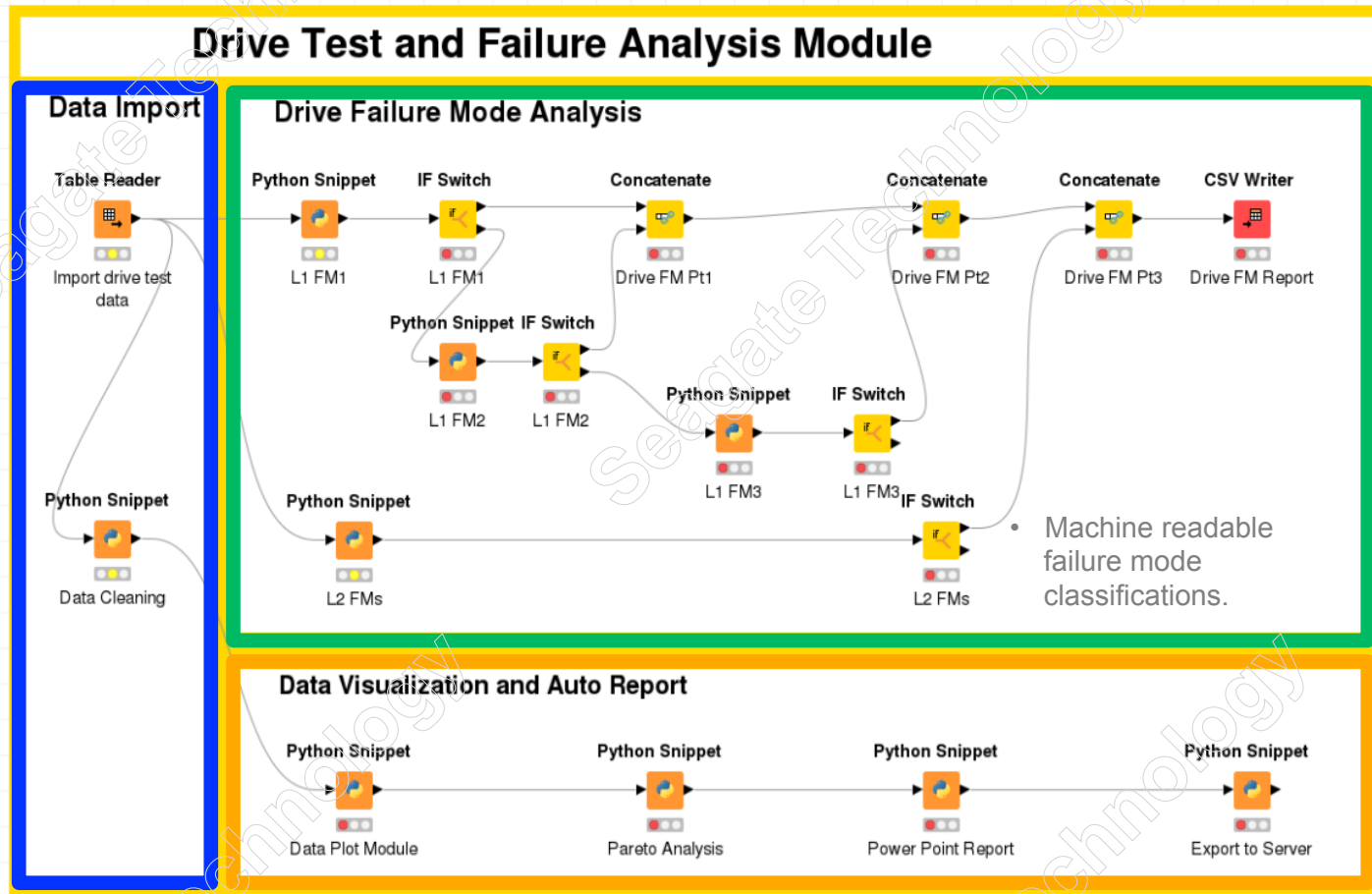
Work Case Scenario: Parallel Data Query



- KNIME enables powerful parallel query mechanisms to maximize data query efficiency.

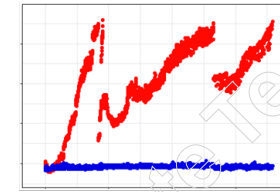


Work Case Scenario: Device Failure Prediction

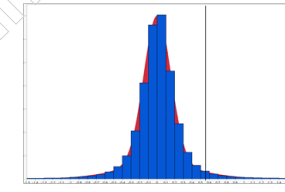


- Human readable plots and summary tables to digest and understand failure symptoms.

FM Regression & Prediction



Input (x1, x2)



Parameter (a1, a2, a3)

$$f = f(x_1, x_2, a_1, a_2, a_3)$$

Benchmark		Machine Recognition
1	✓	1
1	✓	1
1	✗	0
⋮		⋮
0	✓	0

Matching Score xx.x%

