KNIME in Seagate Springtown

DATA IS POTENTIAL

Scott Morrison – Data Analytics Engineer
Brendan Doherty – Data Analytics Engineer

SEAGATE

Outline

- Introduction to Seagate
  - Who we are within Seagate
  - Background information on the product
  - Data Challenges

- Embedding KNIME in a manufacturing environment
  - Equipment Tool Root Cause Diagnostic
  - Pathway Taken – Springtown N.Ireland

- Q&A

Who are we?

Seagate Analytics Group

People come from many different organizations within Seagate as well as the outside.

We have individuals located in most of the time zones where Seagate operates.

- Longmont (USA)
- Minneapolis (USA)
- Derry/Londondery – N.Ireland (Springtown)
- Wuxi (China)
- Singapore
- Korat and Teparuk (Thailand)

Languages

- English (American and Irish)
- Chinese (Mandarin and Cantonese)
- Thai
- Spanish
- Amharic (Ethiopia)
- Hindi, Telugu (India)
- Filipino

Actively participating in new college grad & intern programs to ensure we stay up to date on the latest advances in analytics.

Partnering with IT teams for data and technology delivery.

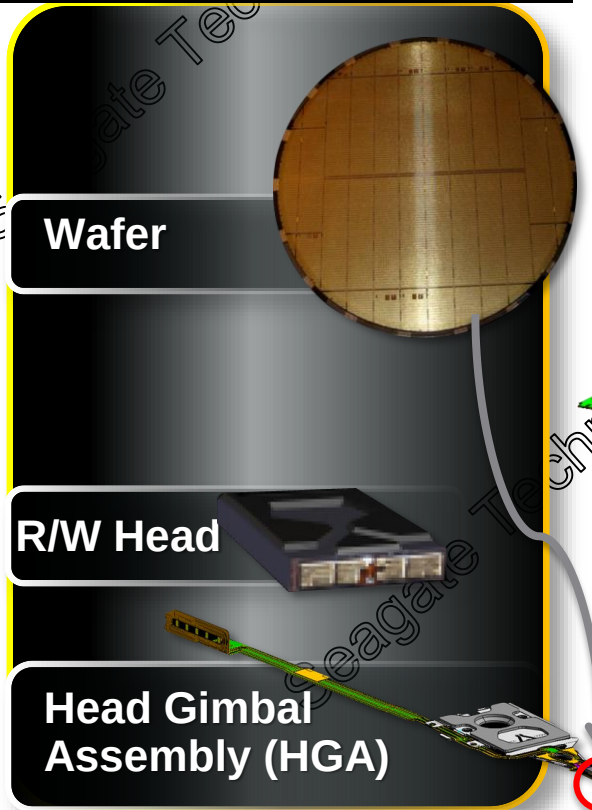
Hard Drive Scale Analogy

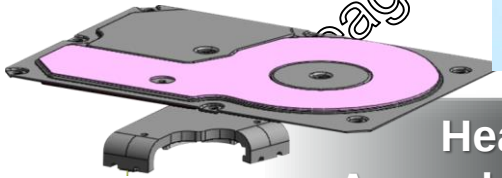
- **Boeing 747** flying **1mm** above the ground, Circling the earth once every **25 seconds**
  - Counting every **blade of grass** along the way
  - Shrink that down so it fits in the **palm of your hand**
- You have something **equivalent** to a **modern day hard drive**

Sourced from Ted.com 2015

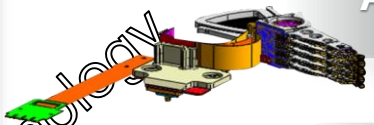
Disk Drive Components

Recording Heads

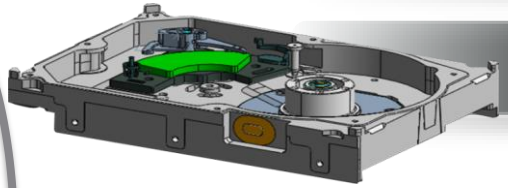




Head Stack Assembly (HSA)



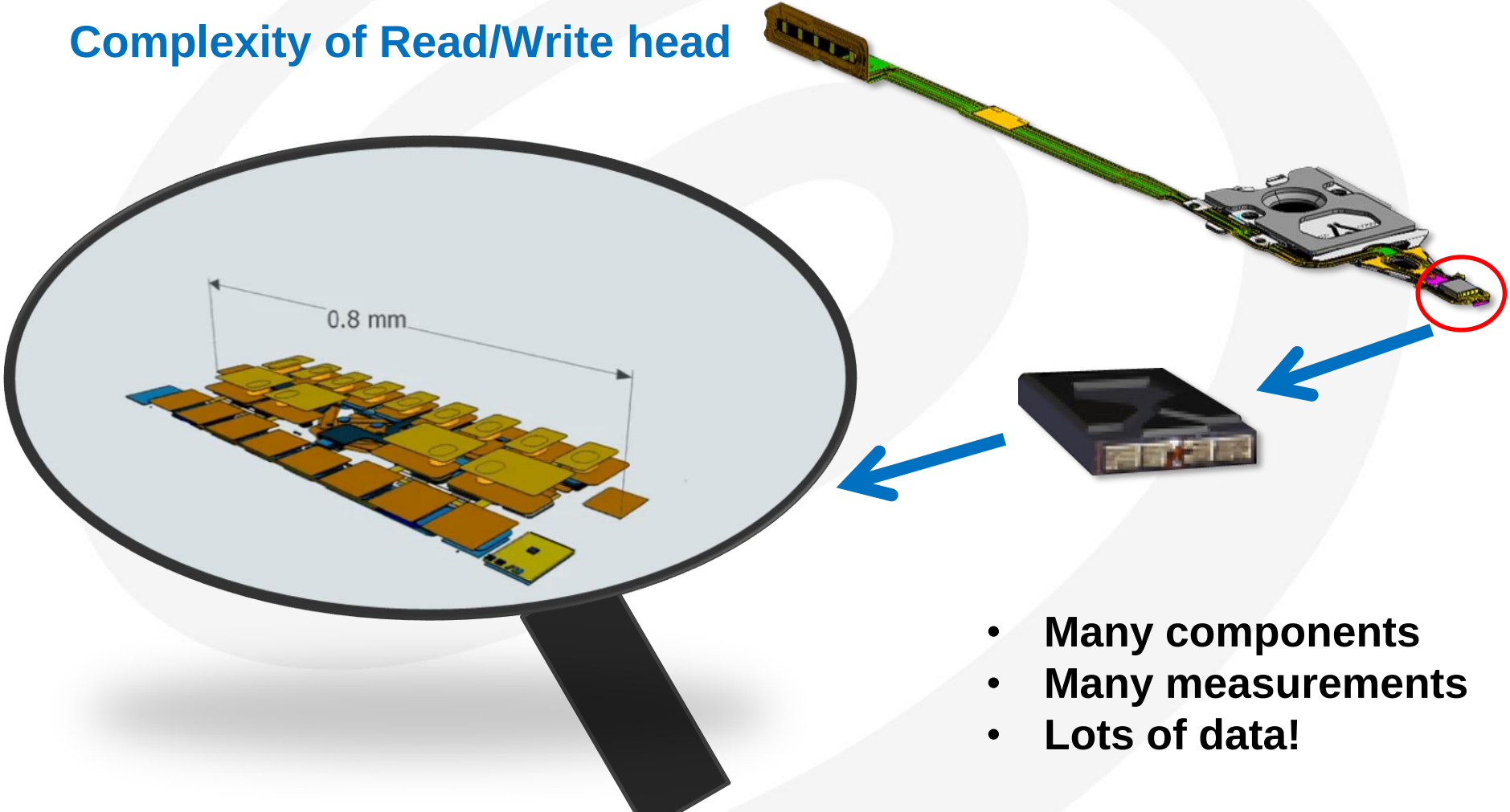
Media



Disk Drive

Vertically integrated company

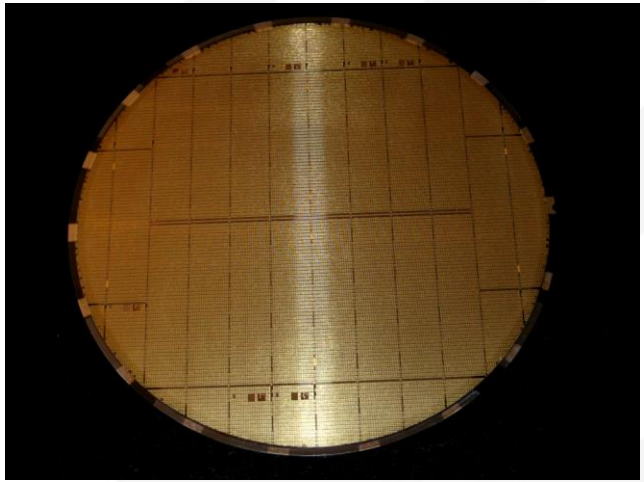
Complexity of Read/Write head



- Many components
- Many measurements
- Lots of data!

The Challenge

1 Wafer



X 1xxxx's HDD Drives

Wafer Process facts

- 100,000 complex micro-electronic devices per wafer
- Up to 1600 steps to build one wafer
- 100+ different types of chemical, compounds and elements

Critical Process Metrics

- Zero excursions
- Downstream customer performance/yield
- Process cycle time
- Process yields

Process Data

- 150 million Oracle records per day
- 1.3M individual data points per week – structured & unstructured data sets, image files, equipment data
- Over 400 million device electrical measurements per week

- ***Bottom line – High margin for error, early containment is vital***

DATA IS POTENTIAL

Embedding KNIME in a Manufacturing Environment

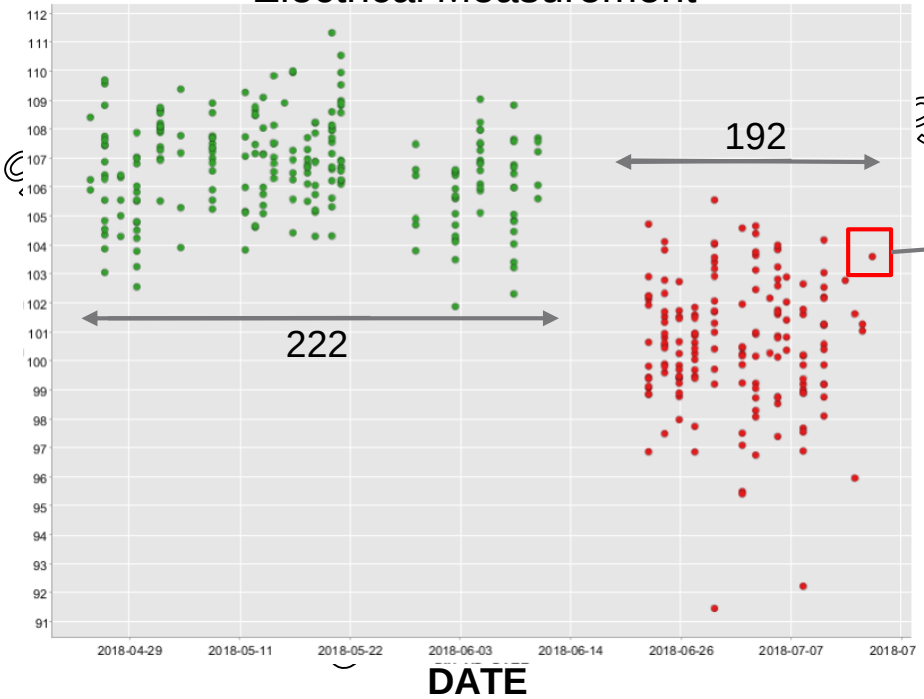
Equipment Tool Root Cause Diagnostic

SEAGATE

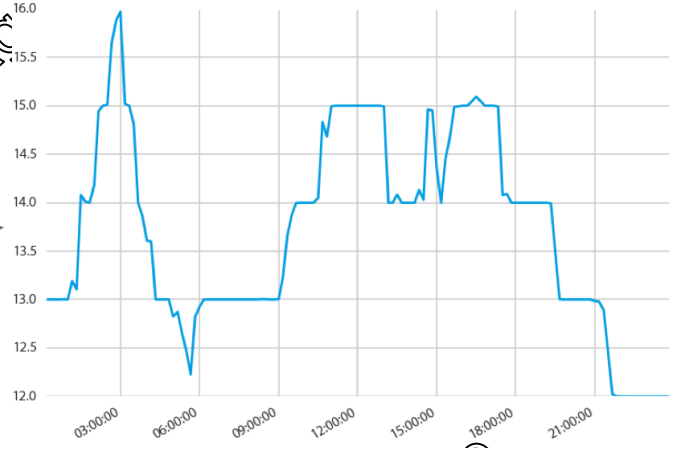
Tool Sensor Data Challenge

Problem

Electrical Measurement

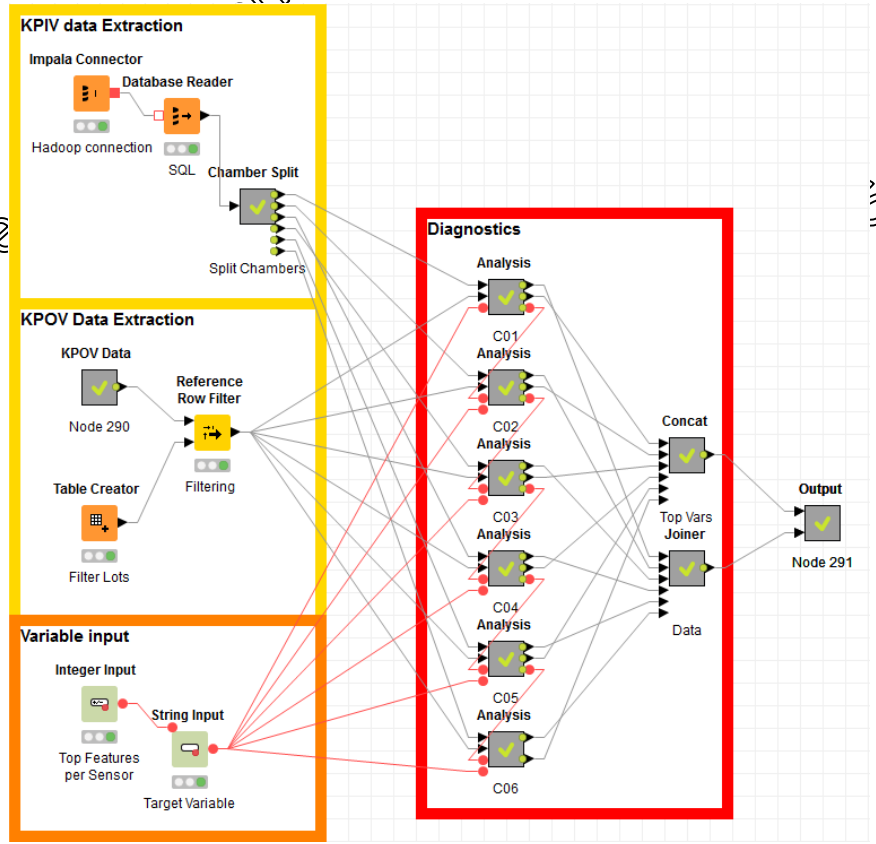


Tool Sensor Trace



X 398 Sensors

KNIME Workflow



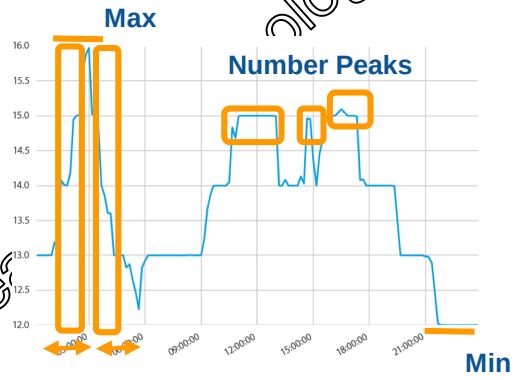
Data Extraction

- KPIV – Time series tool sensor data (Model inputs)
- KPOV – Electrical test data (Model target)

Diagnostics

- Data cleanse
- Feature extraction
- Classification
- Variable influence

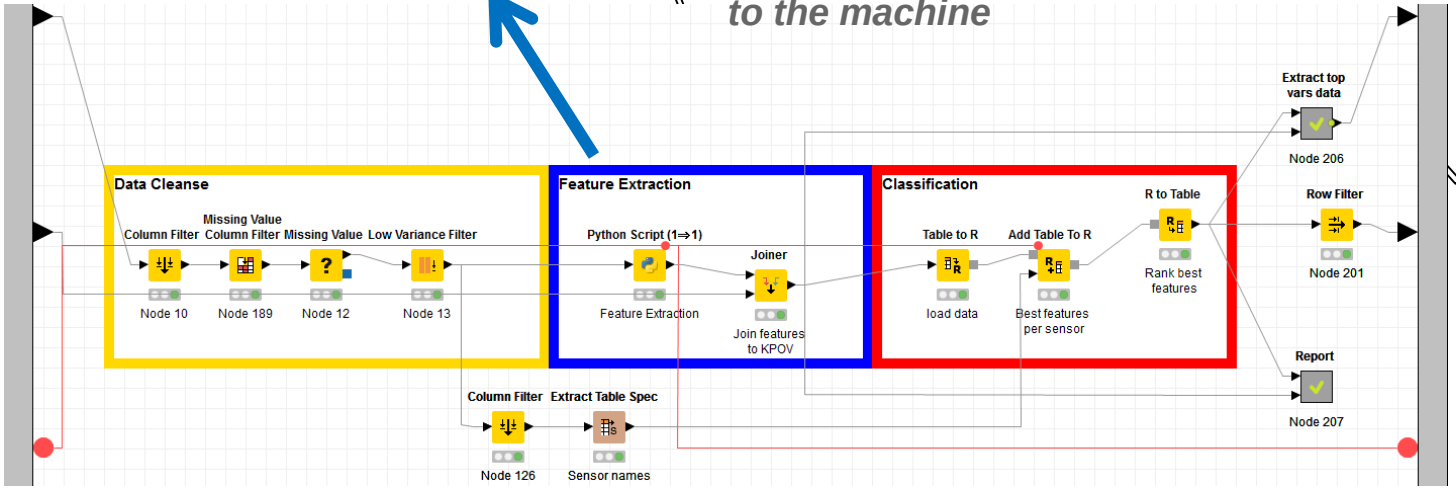
Diagnostics Analysis Meta-node



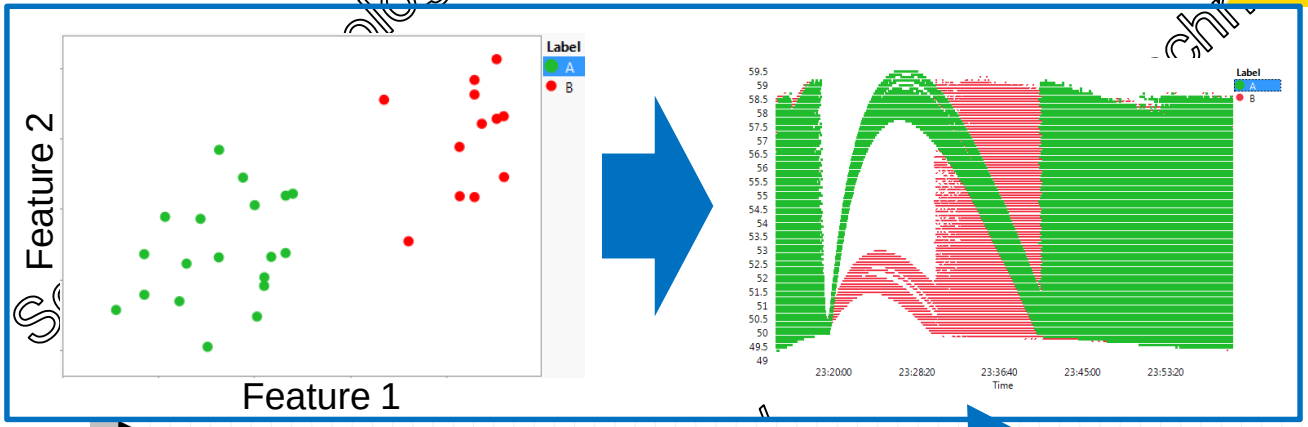
Feature extraction captures the characteristics of a sensor trace via:

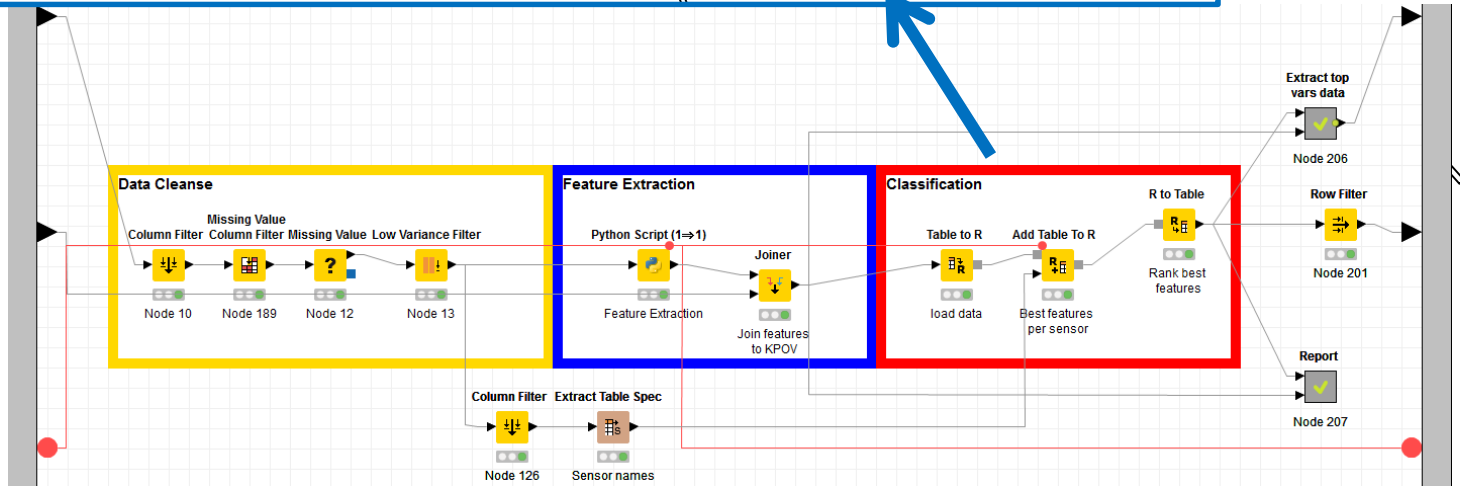
- **Statistical** Features
- **Time Domain** Features
- **Freq Domain** Features

Taking what's visible to the human eye and making it visible to the machine



Diagnostics Analysis Meta-node





Output - Root cause results

Results

Benefits

- Get to root cause of issue in minutes, rather than analysing volumes of data
- Quicker decisions on defective material
- Can hand off complex programming projects to non native programmers

Saving

- Multimillion dollar saving from mitigation of scrap events

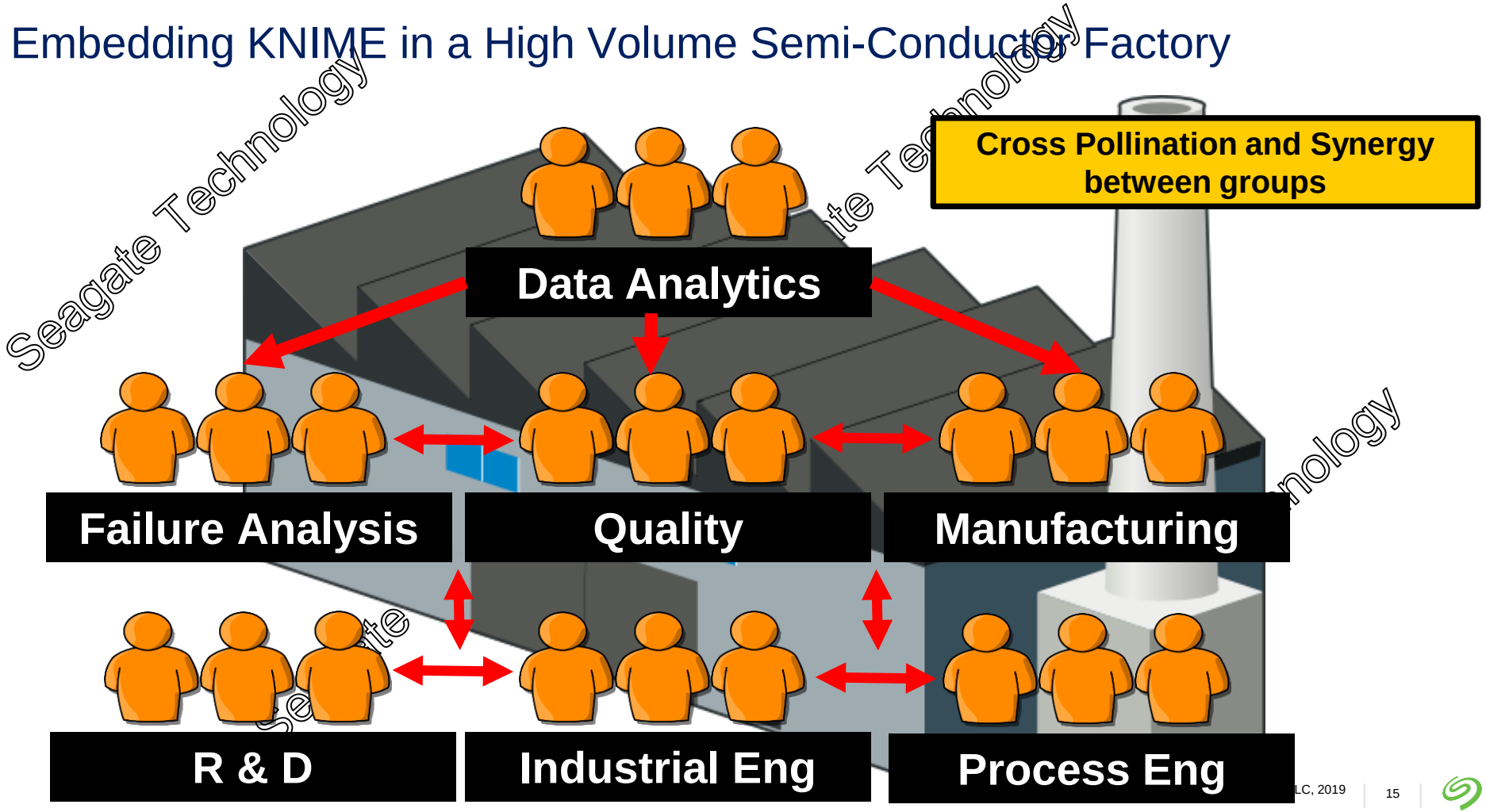
DATA IS POTENTIAL

Embedding KNIME in a Manufacturing Environment

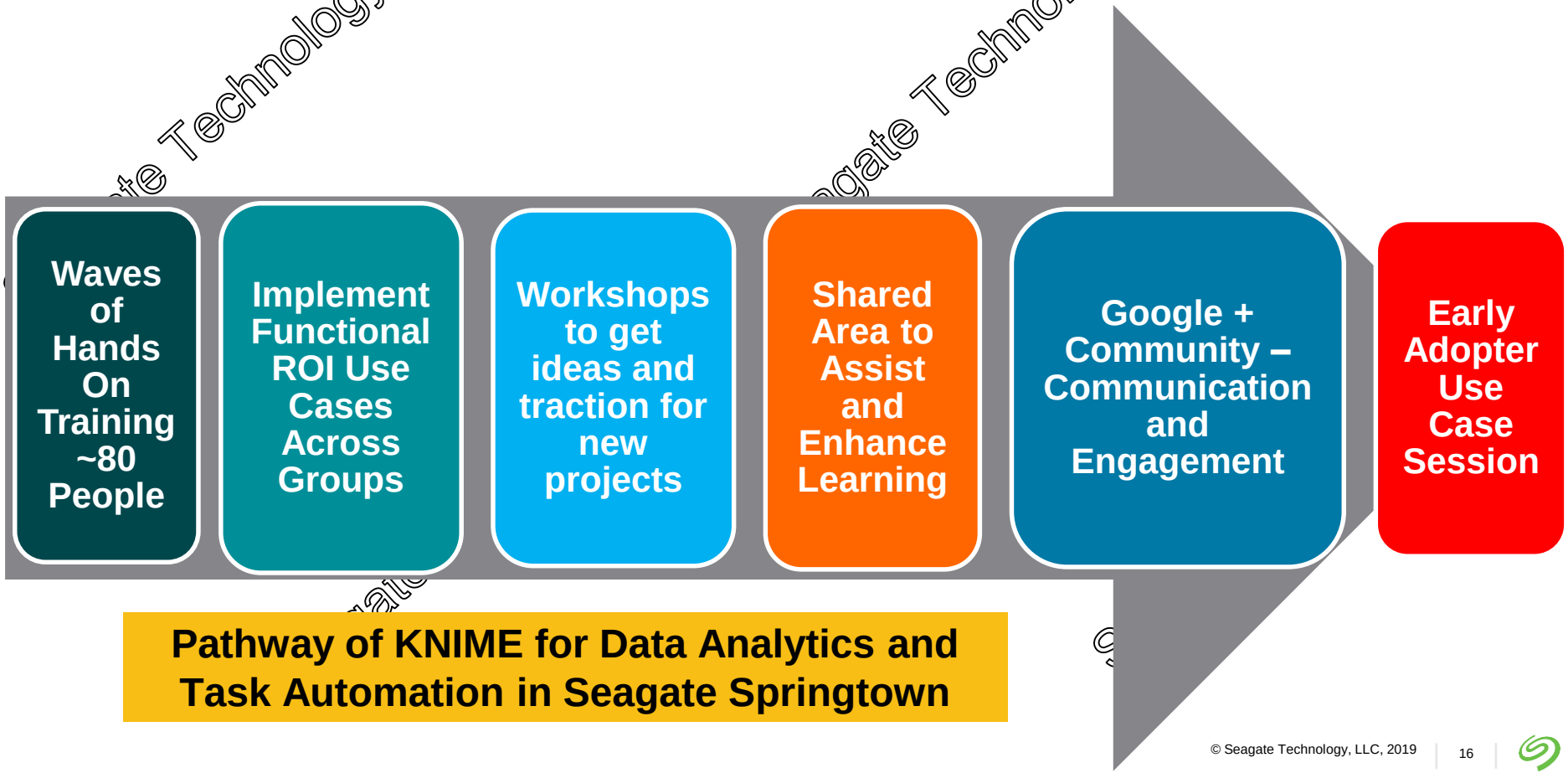
Pathway taken

SEAGATE

Embedding KNIME in a High Volume Semi-Conductor Factory



Embedding KNIME in a High Volume Semi-Conductor Factory



Pathway of KNIME for Data Analytics and Task Automation in Seagate Springtown

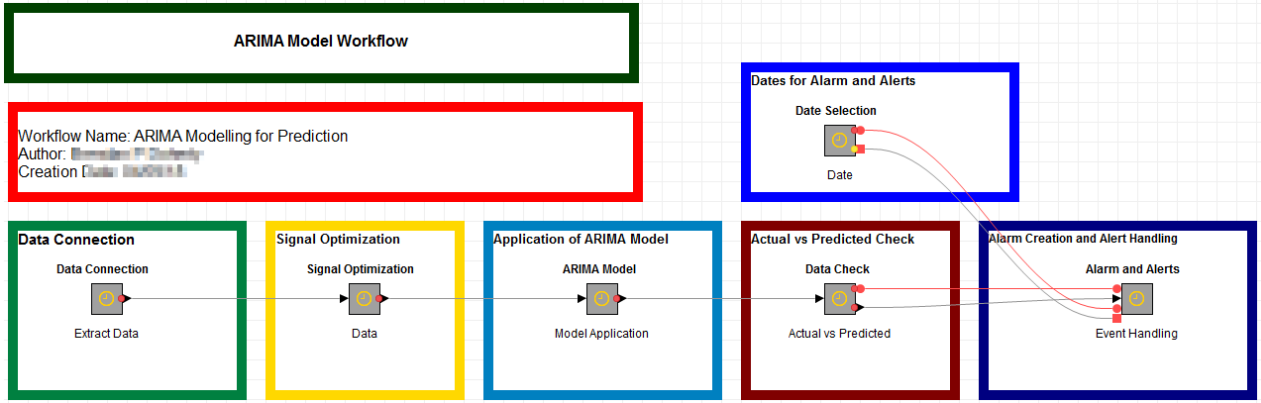
Product Performance Use Case

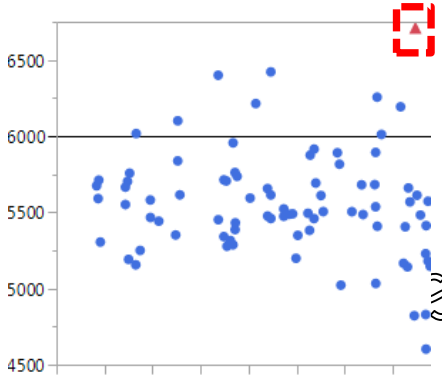
ARIMA Modelling using Product Performance Data

Problem: **Long manufacturing cycle time** for recording heads, **feedback loop** on how the product performs can be in the order of **months**.

Benefit: Predicted trends of product performance provide quicker feedback. **ARIMA** modelling used to highlight **single anomalous data points** months **before** the actual physical product arrives at the point of measurement.

Saving: Cost and **resource** benefit saving from being able to highlight under performing products much earlier in the manufacturing cycle.

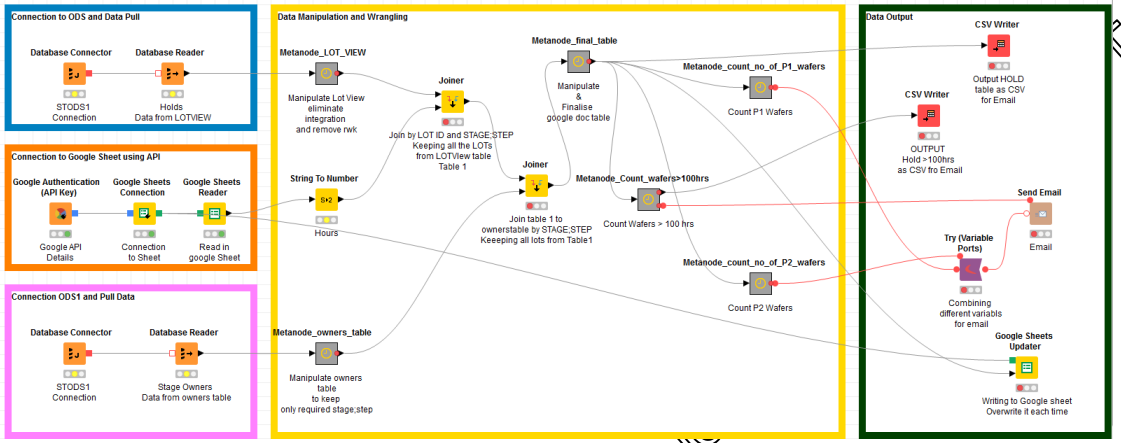


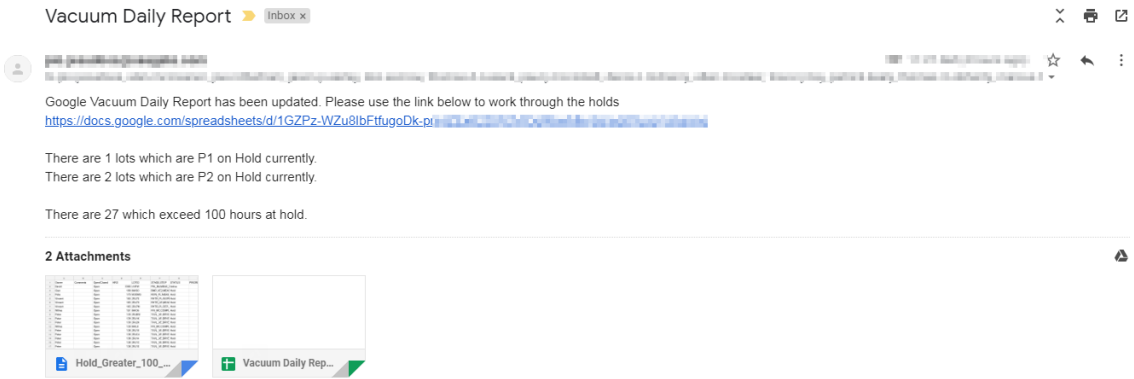


Example of Single Point Anomaly Detected

Process Engineering Use Case

Engineering Holds Report Automation





Problem: ~3 hours per day spent by Process Engineering compiling a WIP (work-in-process) hold report. Very **time consuming process**. Report was **not standardised** and **not always available** first thing in the morning when required.

Benefits: Automated process **saves daily manual engineering input**. Workflow runs in seconds every morning before engineers arrive at work allowing them to **tackle WIP in a hold state faster** and through a **standardised format**.

Saving: The use of KNIME to **automate the process** has **standardised the format** and **saved ~90 hours** of engineering time per month allowing them to concentrate on other tool issues and projects. Improved **cycle time** through engineers **tackling WIP on hold faster**.

Industrial Engineering Use Case

Cycle Time Improvement through Data Wrangling

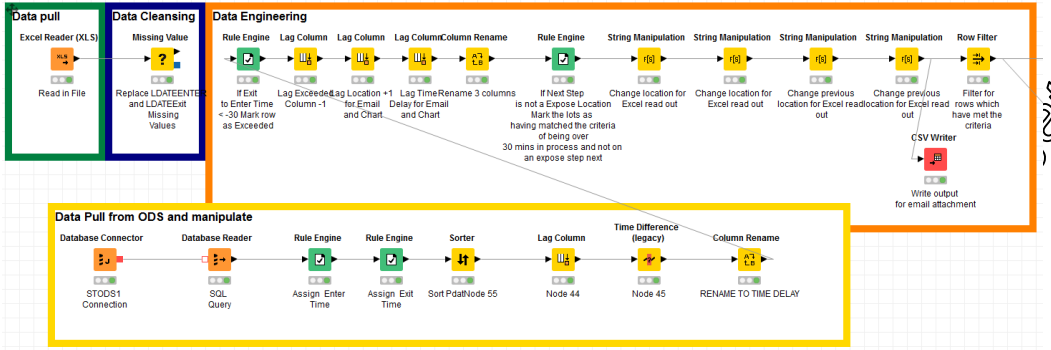


Photo Lots Exceed 30 Minutes Intransit Time where next step is not expose Inbox x

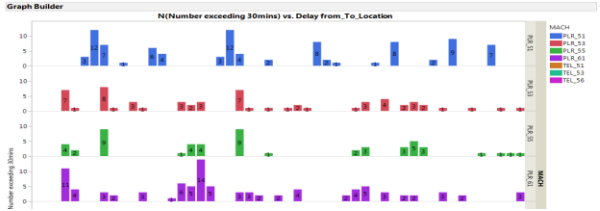
[redacted]

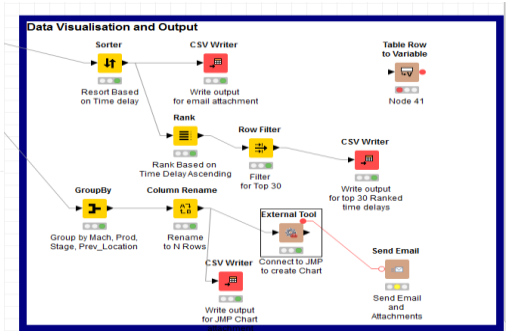
The lots in the attached file exceeded 30 mins processing and were not on an Expose step next.

Please look further into these lots.

The toped ranked lot was [redacted]

Regards





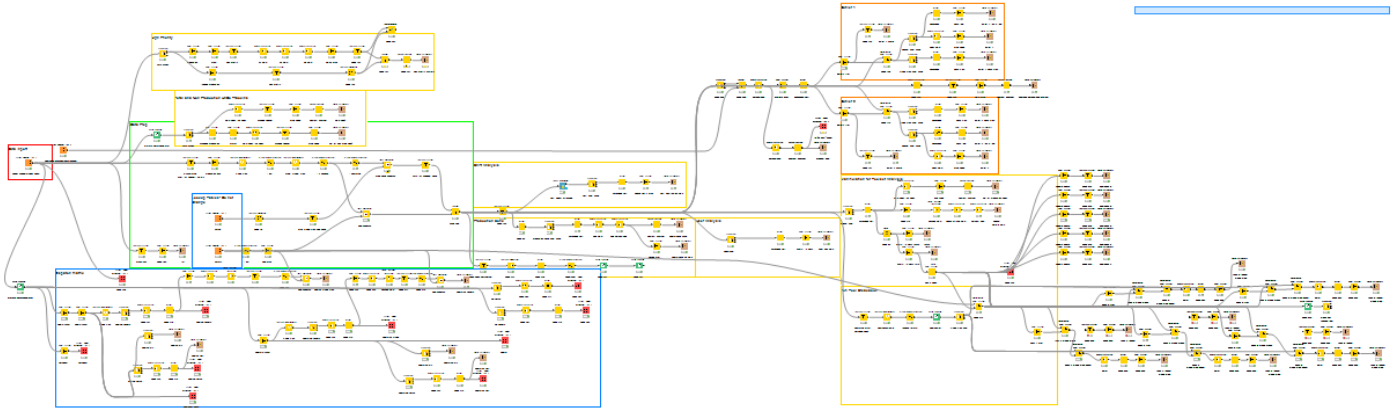
Problem: Very **difficult to identify** wafers which **exceeded** particular attribute **criteria** using either JMP or SQL for users in Industrial Engineering group.

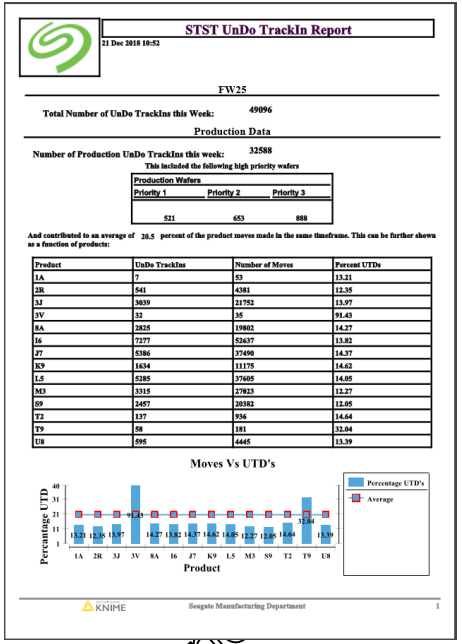
Benefit: Makes **data cleansing and data manipulation easier for non-programmers**. **Integration** with **JMP** to provide familiar charts which are easily interpreted by engineers.

Saving: Improved cycle times through the identification of tools and locations when lots are spending an increased time and the next processing step is not located close by.

Manufacturing Use Case

Undo Track In Analysis Automation and Reporting





STST UnDo TrackIn Report

21 Dec 2018 10:52

FW25

Total Number of UnDo TrackIns this Week: 49096

Production Data

Number of Production UnDo TrackIns this week: 32588

This included the following high priority wafers

Production Wafers		
Priority 1	Priority 2	Priority 3
521	653	888

And contributed to an average of 20.5 percent of the product moves made in the same timeframe. This can be further shown as a function of products:

Product	UnDo TrackIns	Number of Moves	Percent UTDs
1A	7	53	13.21
2R	541	4381	12.35
3J	3039	21752	13.97
3V	32	35	91.43
8A	2825	19802	14.27
I6	7277	52637	13.82
J7	5386	37490	14.37
K9	1634	11175	14.62
L5	5285	37605	14.05
M3	3315	27023	12.27
S9	2457	20382	12.05
T2	137	936	14.64
T9	58	181	32.04
U8	595	4445	13.39

Moves Vs UTD's

KNIME Seagate Manufacturing Department 1

Problem: Tool sets have **higher cycle times** due to products being scanned into them erroneously and then scanned out again by operators without the tool processing the product

Benefit: Project by an **intern** after some training on KNIME was able to **provide ROI** in terms of highlighting tools with higher incidents of issue occurring, in turn **improving cycle time**. Intern was able to provide a real **ROI to the business and their own skill set** after **short amount of training time**.

Saving: Improved cycle times and **ROI in a short space** of time from **new employees**.

Sample Use Cases

Department	Use Case	Project Theme	Benefit
Quality	ARIMA Modelling on time series predicted data trends to highlight anolous point where product quality does not meet criteria	Data Modelling	Product Quality
Process Engineering	Auotmated identification and notification of products which surpassesd defined transit times and criteria	Data Engineering	Improved cycle time
Process Engineering	Automation of complicated manual data wrangling tasks to analyse and visualize data over 2 year time span from a Hadoop database	Data Engineering, Visualisation and Automation	Product Quality
Process Engineering	Automation of Process Engineering WIP on hold process document for using Google Sheets and Oracle	Data Engineering, Automation	Increased engineering time for other tasks / Improved cycle time
Process Engineering	Prediction of processing metrology times for Process Eng Toolsets to reduce the amount of Scrap	Data Modelling	Reduction in Scrap Material
Industrial Engineering	Analysis and visualization of Industrial Engineering data based around products tracked and then untracked into tools in order to improve cycle times	Data Engineering, Automation	Improved cycle time
Industrial Engineering	Automation and modelling of manufacturing tranistion data to identify when wafers are exceeding time limits for intransit between processing stages. Integration with JMP for visualization	Data Engineering, Visualisation and Automation	Improved cycle time
R & D	Automation of R & D cycletime scorecard. Use of Google Sheets and Oracle databases for Cycle Time Reporting. Wide range of reports to analyse and improve cycle time for R & D wafers.	Automation	Improved cycle time / Product Quality
R & D	Automation of the shift handover document creation using data from Oracle databases and Google Sheets	Automation	Improved cycle time
Yield	Visulisation of Yield Data using Heat Maps and Integration with Tableau	Data Engineering, Visualisation and Automation	Product Quality

Summary

- **Rapid** adoption by **non native programmers**
- Can solve **complex** business problems **much more easily**
- Allows users to **quickly** provide a **return on investment** from training
- **Cross pollination** and **collaboration** from users helped generate new projects