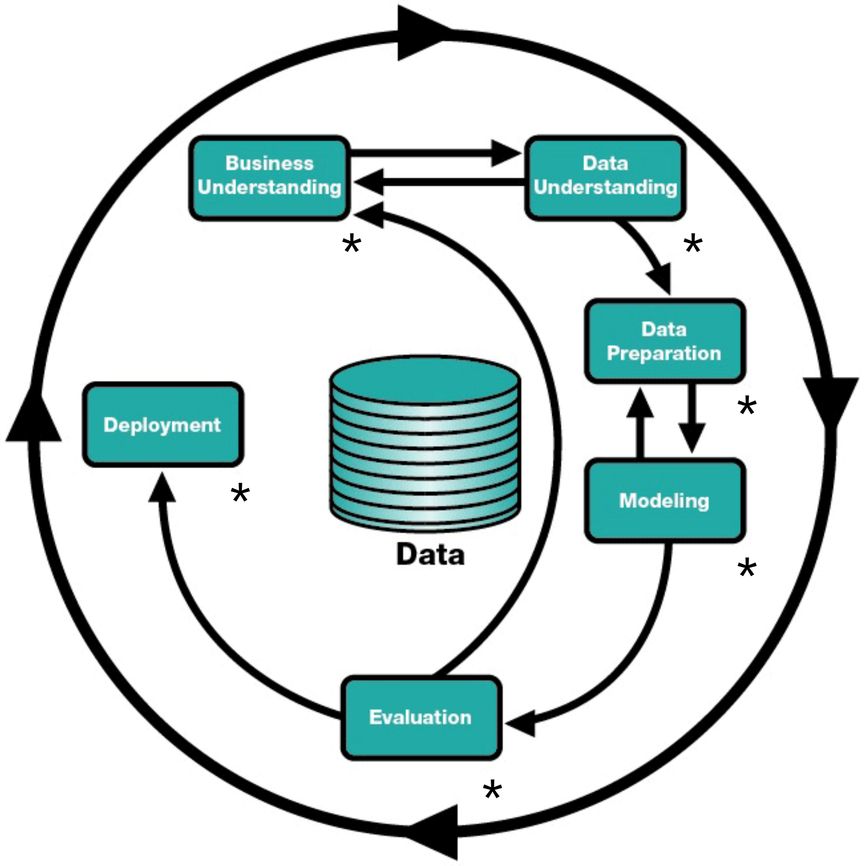
Customer

SAP Predictive Maintenance and Service, on-premise edition and Data Science

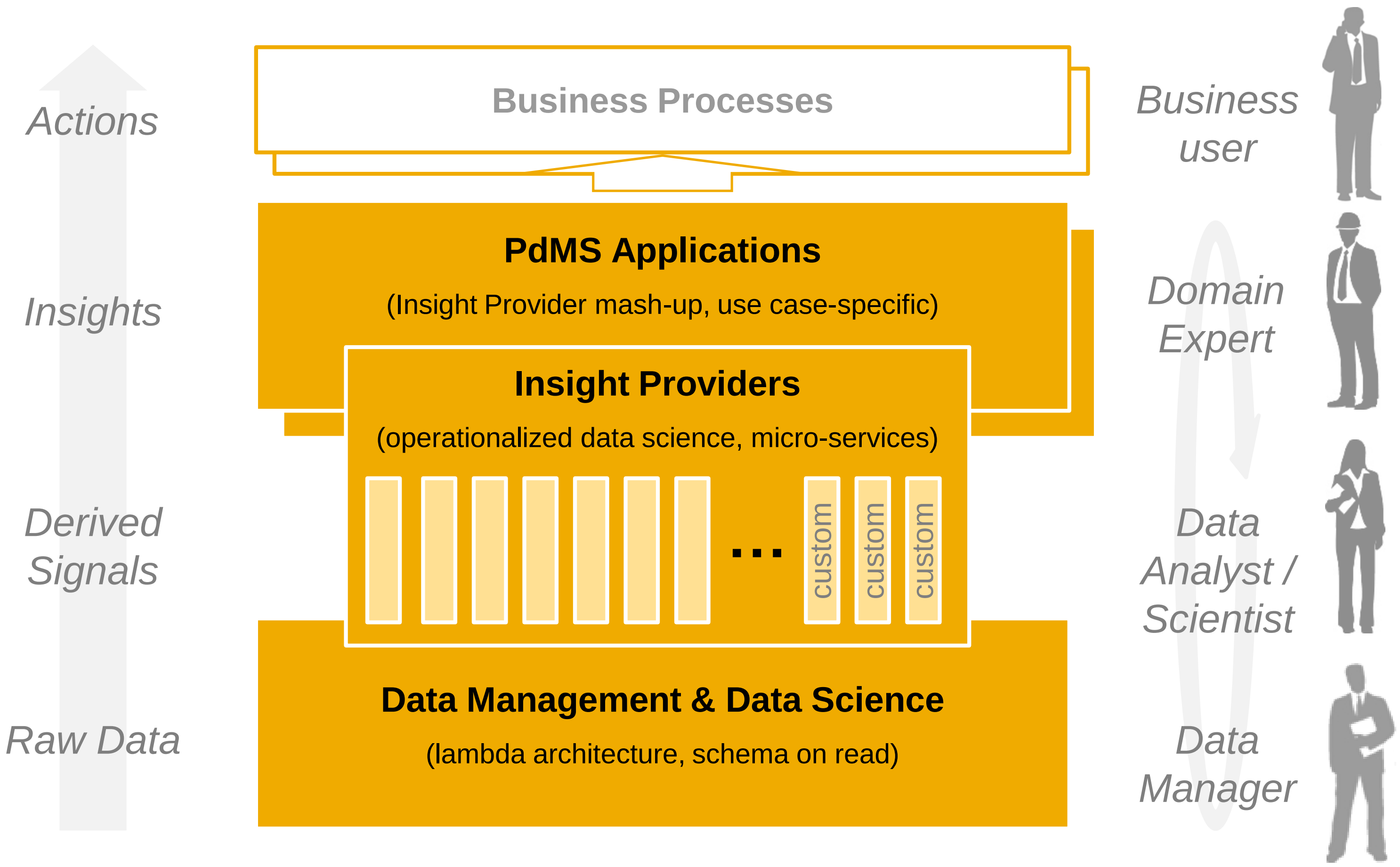
Predictive Maintenance is a process, not just an algorithm…
Domain expertise is as important as data science, if not more…



CRISP-DM Cross Industry Standard Process for Data Mining

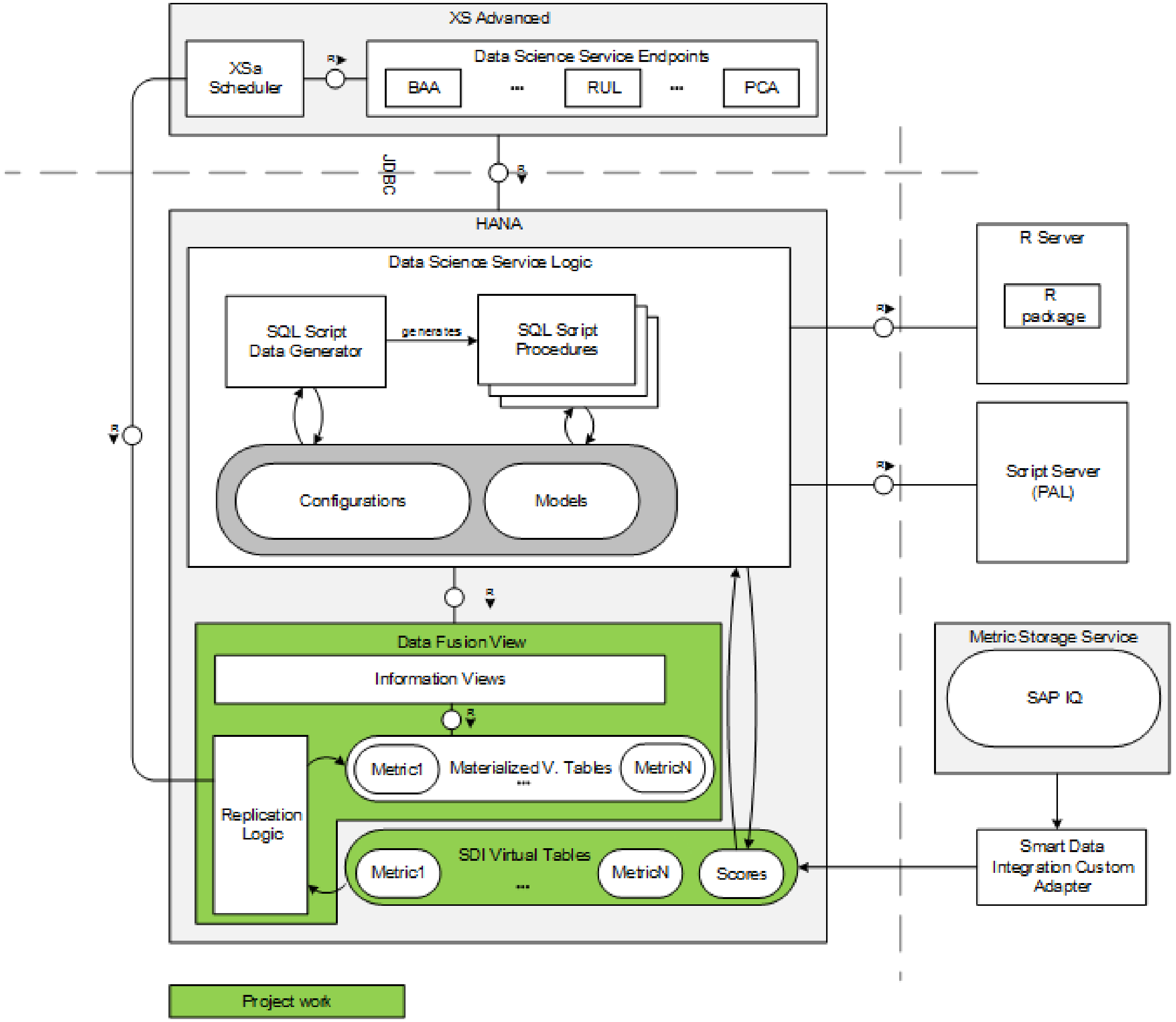
* Involvement of domain expert necessary

Data Science – Predictive Maintenance Architecture

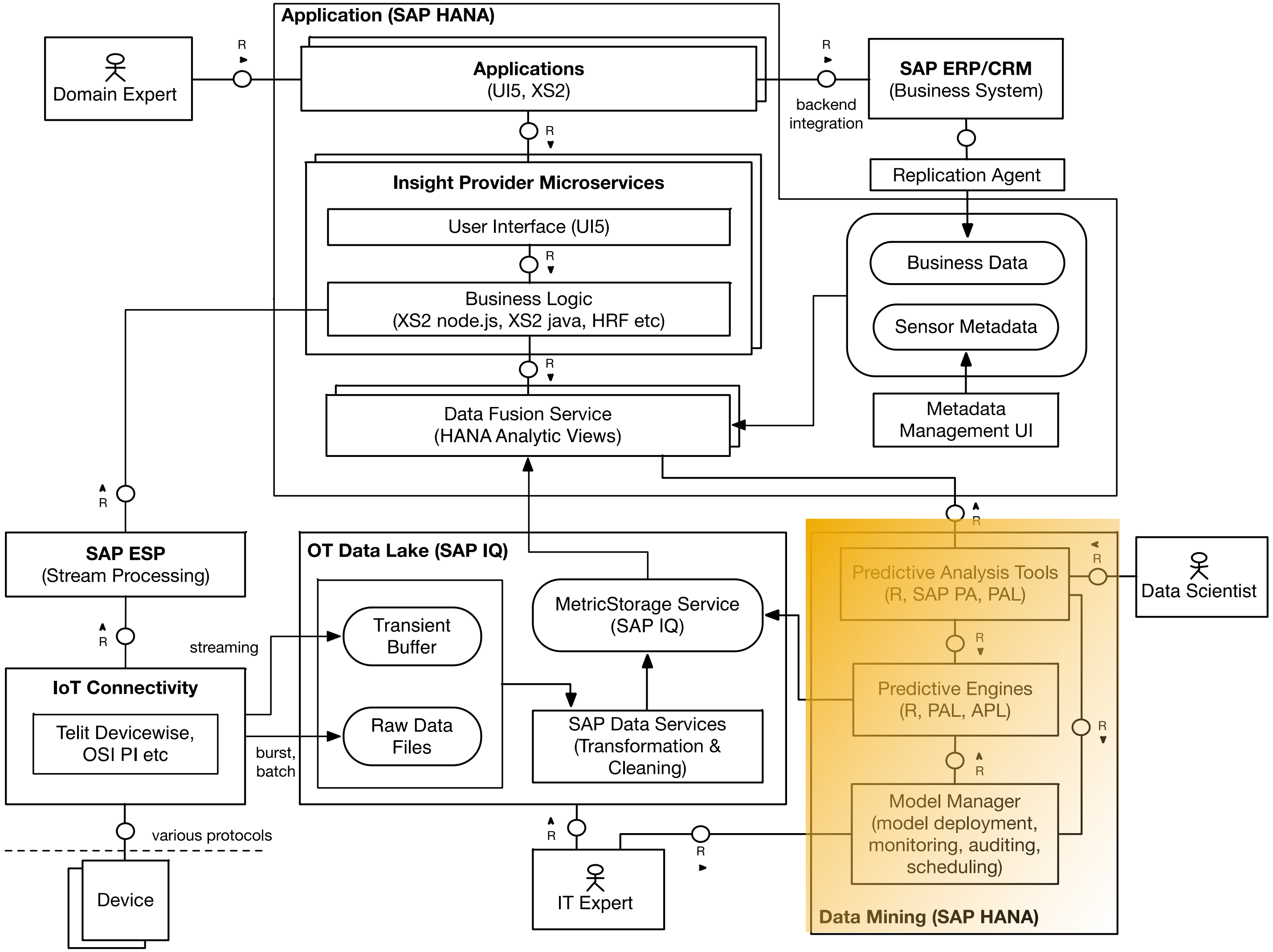


Operationalize Data Science by embedding it in applications and use it to trigger follow-up actions.
Data Science algorithms from many sources – R, PAL, APL… plus customer's own.

Data Science – Predictive Maintenance Architecture



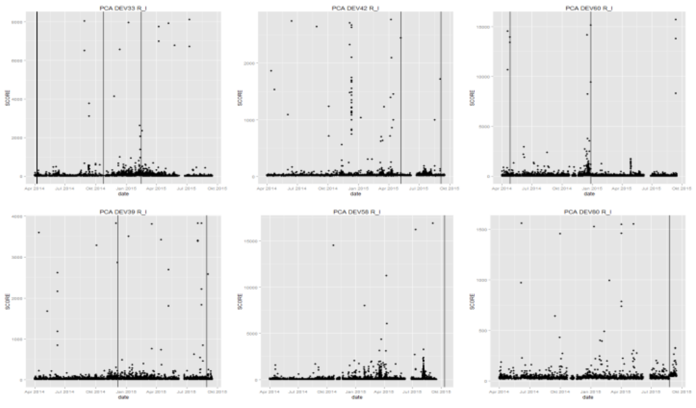
Data Science – Predictive Maintenance Architecture



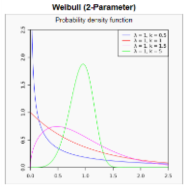
Data Science – Predictive Maintenance Analyses

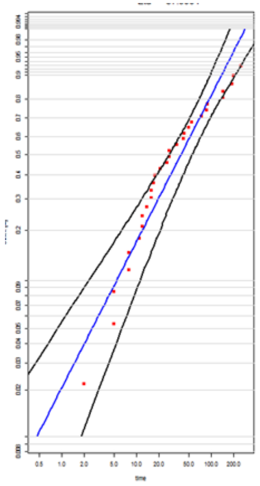
- **Defect Pattern Identification**
  - Statistical analysis, text clustering, association analysis, and decision trees.
  - Visualization of big data with parallel coordinates and multi-dimensional scaling.
- **Systems Trending and Alert Management**
  - Detect outliers and anomalies in the data with supervised and unsupervised machine learning
  - Text analysis and text mining to classify scheduled vs. unscheduled maintenance events
- **Machine Health Prediction**
  - Historic machine data are used to predict breakdowns via decision trees
  - Energy consumption pattern profiles are calculated with k-means clustering
  - Domain expert knowledge was modelled in SAP HANA with decision tables
- **Vehicle Health Prediction**
  - Used association rule mining and regression tree learning to correlate production rework and customer satisfaction data
  - SAP HANA/R data mining and data visualization capabilities applied to surveys and production data sets
- **Emerging Issues**
  - Analysing telematics data & relating them to equipment's service and warranty data using text mining, association analysis
- **Predictive Quality Assurance**
  - Visual detection of cracks (image processing techniques)
  - Heat image comparison of “areas of interest” of sample images of material with issues to current material (Euclidean vector distance calculation)
- **Maintenance Prioritization**
  - Forecast production and probability of failure calculation at asset level.
  - Prioritize maintenance activities based on actual and forecasted KPIs
- **Health Prediction for aircraft components**
  - 5 aircrafts over 5 years, 400 sensors each aircraft – 44.1 Billion sensor readings
  - Correlated with maintenance history (notifications), weather data & geo locations
  - Text Analysis to understand maintenance activities
  - Use of Statistical Process Control and Symbolic Aggregate Approximation for anomaly detection
- **Bad Actor Analytics**
  - Weibull life time analyses
  - Use classification techniques to identify rotating equipment likely to fail based on past patterns
- **Root Cause Analysis for Quality Issues**
  - Find causal relationships between claims and production settings from machine readings
  - Improve on Statistical Process Control usage
- **Maximize Machine Efficiency in Production**
  - Augmenting (human) expert rules with (machine) rule mining (regression trees)
  - Approximating machine state to circumvent “rare event problem” (anomaly detection)
  - De-clutter sensor data for root cause analysis (trend analysis)
- **Asset Health Prediction**
  - Optimize testing and crew efficiency based on limited resources
  - Optimize capital investment for URD Cables
  - Machine health prediction from historic data and outages

Data Science – Predictive Maintenance Algorithms



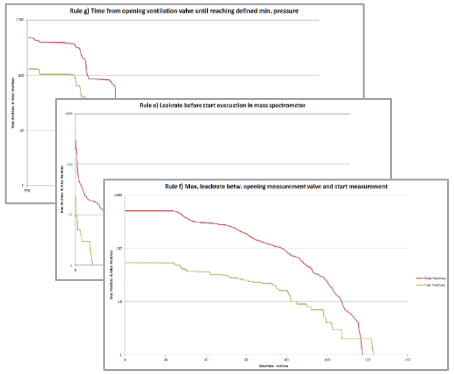
Principal Component Analysis of Switches



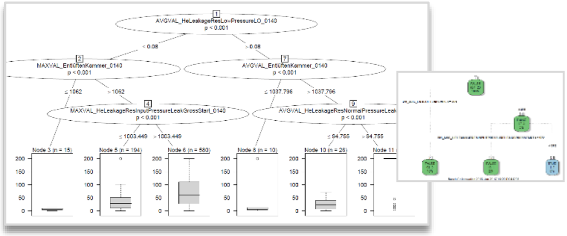


Weibull Remaining Useful Life Estimation

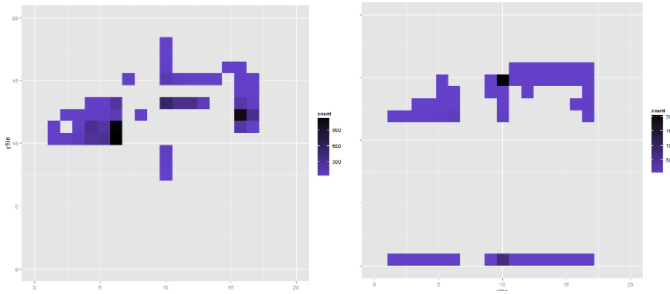
$$f(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0, \\ 0 & x < 0, \end{cases}$$



Expert Rule Validation

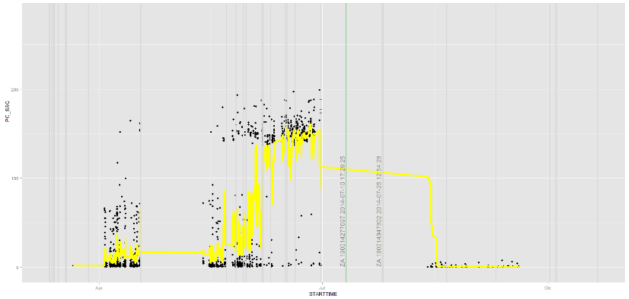


Automatic Rule Extraction with Decision Trees



$$W(P,Q) \;= min\left\{\sum_{x\in M,\, y\in M} dist(x,y)\cdot T(x,y)\,\middle|\, T: M\times M \to [0,1]\ with\ \sum_{y\in M} T(x,y) = P(x),\ \sum_{x\in M} T(x,y) = Q(y)\right\}$$

Wasserstein Metric for battery performance analysis



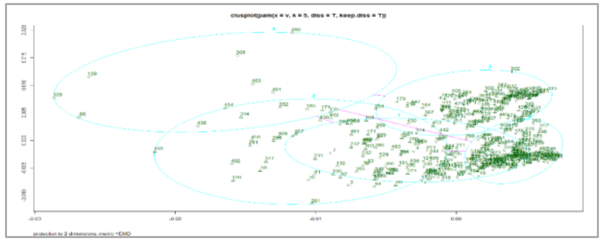
Anomaly Detection with Principal Component Analysis scores

$$\mathrm{Var}[a^T X] = \frac{1}{n}\sum_{i=1}^{n}\left\{a^T\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)\right\}^2 = a^T V_{XX} a.$$

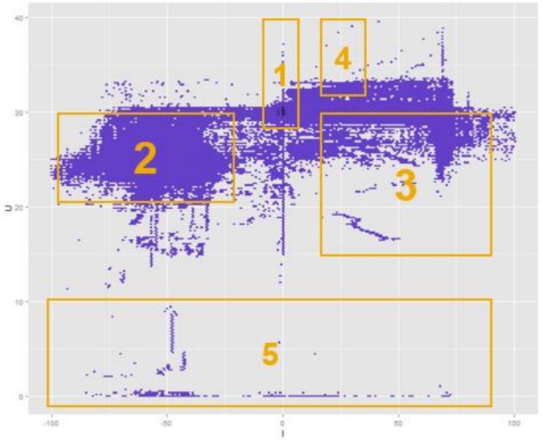
where

$$V_{XX} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)^T$$

PCA



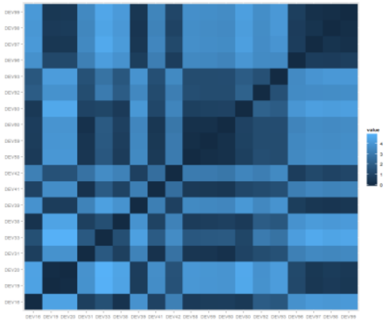
K-Mediod cluster analysis to partition the population into classes of similar devices

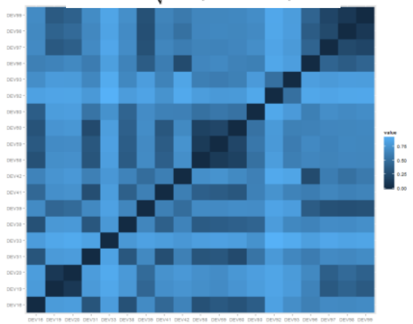


Battery behavioural groupings

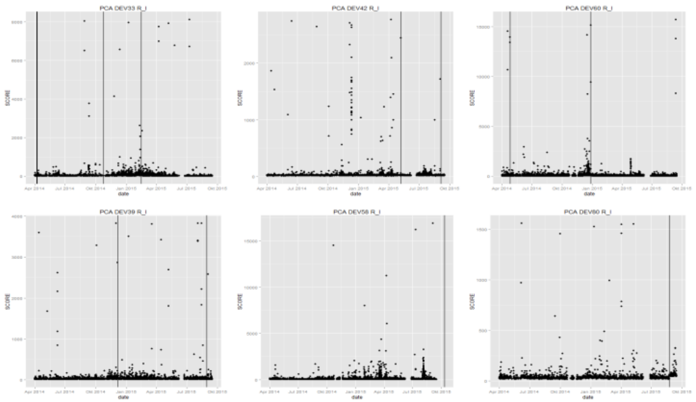
$$W_p(\mu,\nu) := \left(\inf_{\gamma\in\Gamma(\mu,\nu)} \int_{M\times M} d(x,y)^p \,\mathrm{d}\gamma(x,y)\right)^{1/p}$$

$$H(P,Q) := \sqrt{\frac{1}{2}\int\left(\sqrt{\frac{dP}{d\lambda}} - \sqrt{\frac{dQ}{d\lambda}}\right)^2 d\lambda}$$

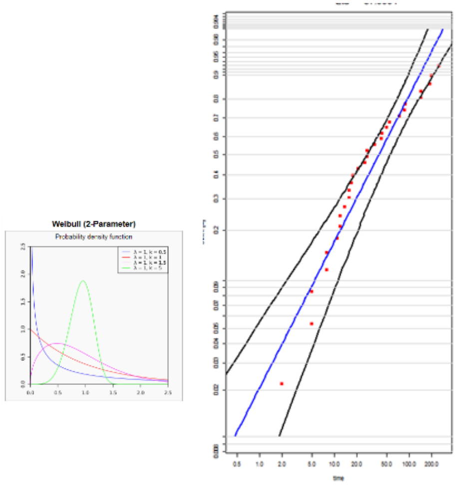




Many Data Science Algorithms are used in Predictive Maintenance

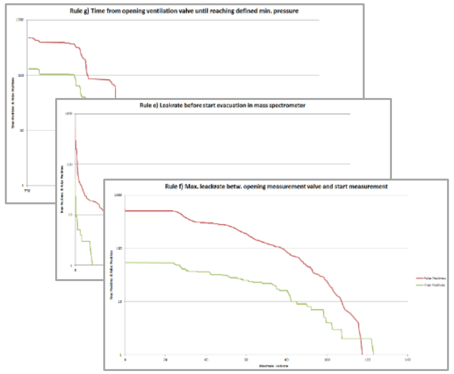


Principal Component Analysis of Switches

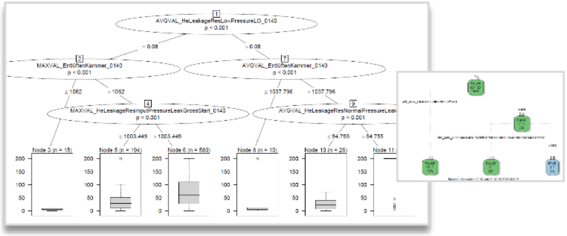


Weibull Remaining Useful Life Estimation

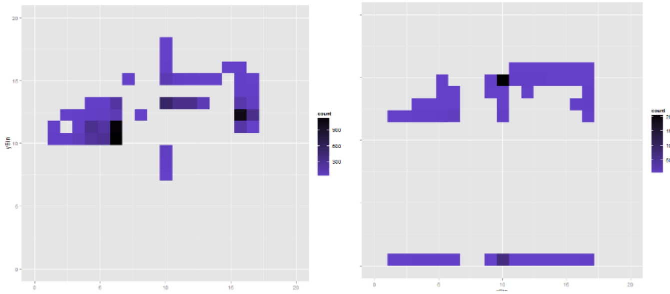
$$f(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0, \\ 0 & x < 0, \end{cases}$$



Expert Rule Validation

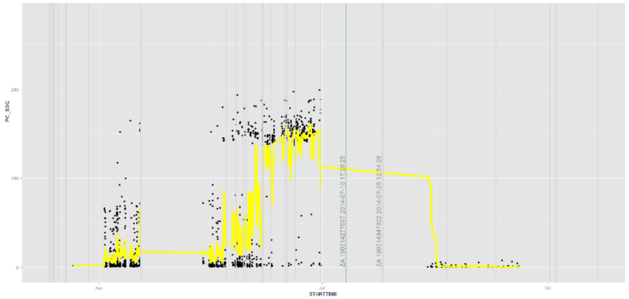


Automatic Rule Extraction with Decision Trees



$$W(P,Q) = \min\left\{\sum_{x\in M,\, y\in M} dist(x,y)\cdot T(x,y) \,\middle|\, T: M\times M \to [0,1] \text{ with } \sum_{y\in M} T(x,y) = P(x),\ \sum_{x\in M} T(x,y) = Q(y)\right\}$$

Wasserstein Metric for battery performance analysis



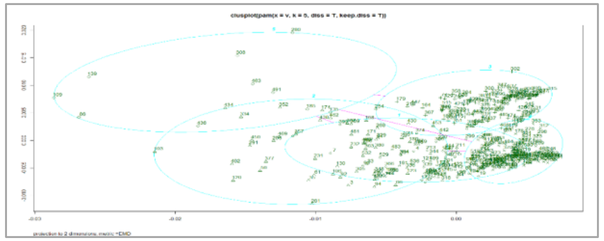
Anomaly Detection with Principal Component Analysis scores

$$\mathrm{Var}[a^T X] = \frac{1}{n}\sum_{i=1}^{n}\left\{a^T\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)\right\}^2 = a^T V_{xx} a.$$

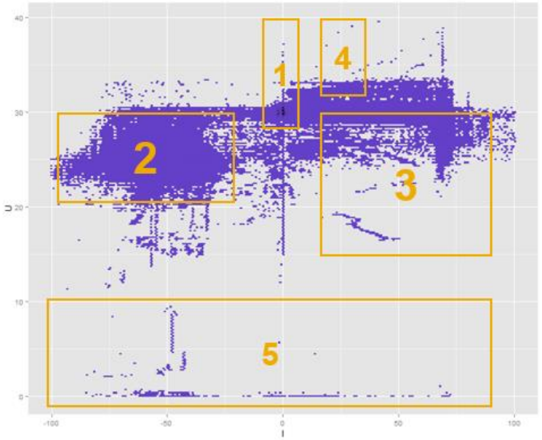
where

$$V_{xx} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)^T$$

PCA



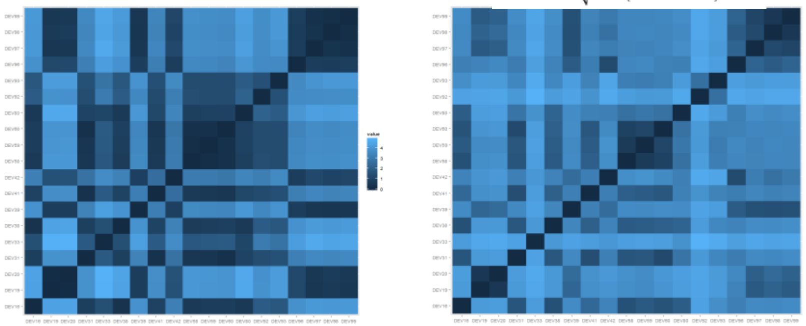
K-Mediod cluster analysis to partition the population into classes of similar devices



Battery behavioural groupings

$$W_p(\mu,\nu) := \left(\inf_{\gamma\in\Gamma(\mu,\nu)} \int_{M\times M} d(x,y)^p \,\mathrm{d}\gamma(x,y)\right)^{1/p}$$

$$H(P,Q) = \sqrt{\frac{1}{2}\int\left(\sqrt{\frac{dP}{d\lambda}} - \sqrt{\frac{dQ}{d\lambda}}\right)^2 d\lambda}$$

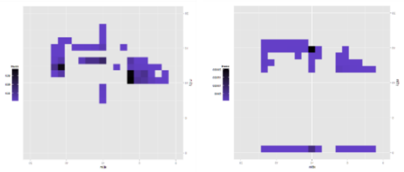


Data Science in Predictive Maintenance (PdMS on Premise)

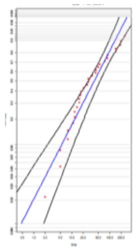
- PdMS on Premise provides three Data Science services ‘out of the box’ that can be applied to customer data.
  - Anomaly Detection, Distance-Based Failure Analysis and Remaining Useful Life Prediction

Anomaly Detection

Distance-Based Failure Analysis



Remaining Useful Life Prediction



Custom Data Science Service

- New Data Science Services (algorithms) can be integrated by embedding the algorithms using R packages that conform to the interface provided.

- PdMS on Premise includes functionality to manage Data Mining Models

Data Science Services – Anomaly Detection

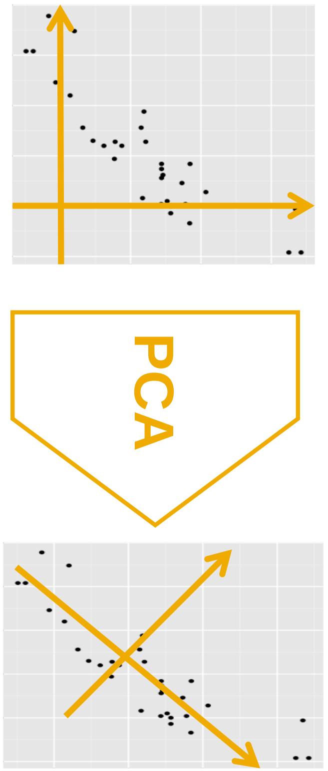
Business Problem: Certain components are crucial for the functioning of a machine or are very expensive, therefore abnormal behavior should be detected and potential problems fixed. Minimizing cost and downtime are important goals as well.

Solution:

1. Data extraction and preparation: Extract the equipment's senor data from the time series storage and prepare it.
2. Learn a model and store it: Apply Principal Component Analysis to the historic sensor data of healthy machines.
3. Apply model to new data and store results (anomaly scores) in the time series storage.
4. Show anomalies in the application and allow to create work orders.

Benefits:

Automatic detection of multivariate anomalies which could lead to failures in components and possibility to take action. Helps to prevent downtimes and minimize maintenance costs.



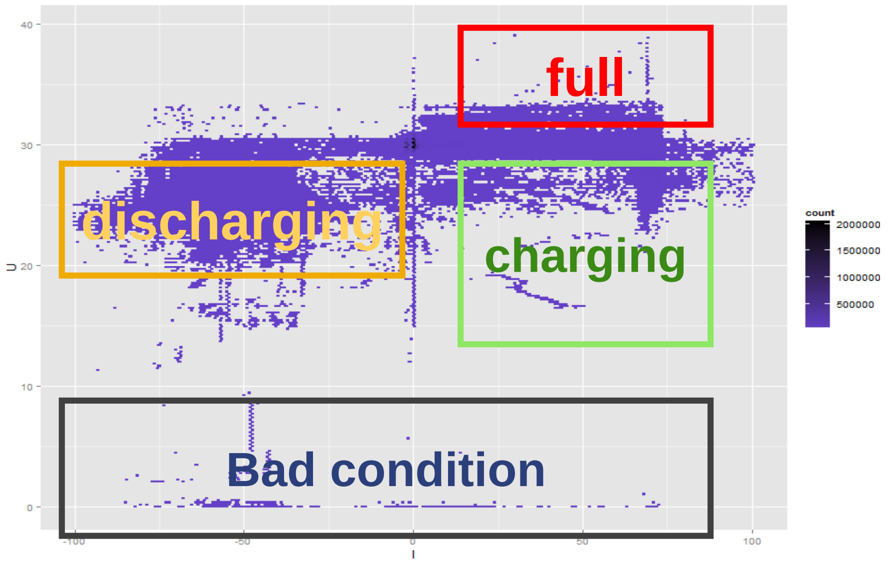
Data Science Services – Distance-Based Failure Analysis

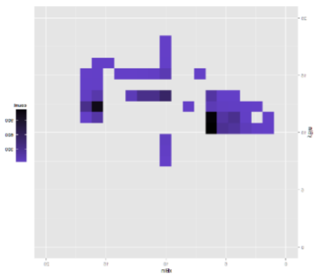
Problem: Some components are crucial assets that can lead to unexpected downtimes when they fail.

Solution:

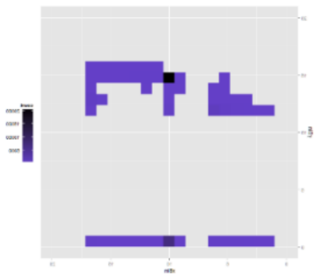
1. Data extraction and preparation: Extract the equipment's senor data from time series storage and prepare it.
2. Learn a model and score new data: Compute the distances of each component to a reference component using earth mover's distance (laz learner type of algorithm) and store distances in time series storage.
3. Show components ranked by distance (in descending order) in the application.

Benefits:
Early identification of malfunctioning components in order to reduce downtime.





Reference component signature



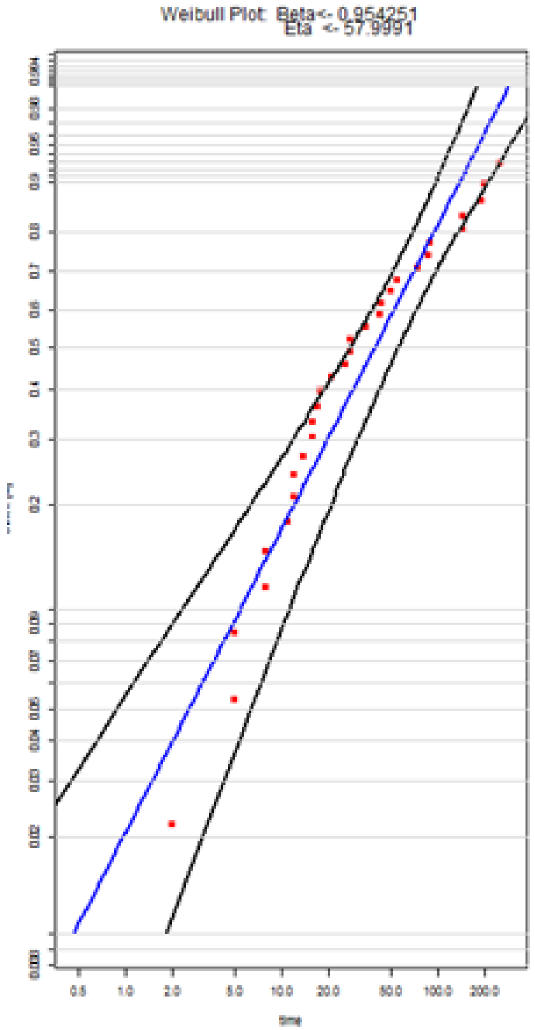
Specific component signature

Data Science Services – Remaining Useful Life Prediction

Problem: The frequent occurrence of failures leads to increased costs for maintenance and downtime.

Solution

1. Data extraction and preparation: Extract repair data (IT data) and compute repair time KPI’s (mean time between repairs, uptime, downtime) for components.
2. Learn a model and store it: Perform Weibull Life Time Analysis on repair data and store model.
3. Score new data: Calculate remaining useful life and probability of failure for each machine/component and store scores in time series storage.
4. Show remaining useful life and probability of failure in application.

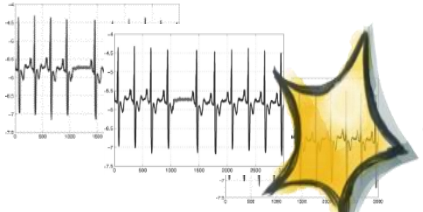


Benefits

Use statistics based approach for estimating lifetime of components. Usage of Weibull Life Time Analysis beneficial if no sensor data available (e.g. new component).

Predictive Maintenance

Discover rules or relationships between attributes and failures



Incident

Extract rules

Find the reason

Root cause analysis

Permanently remove reason if possible

Required:

An incident indication (message or failure criteria)

Attributes with potential relation to failure (sensor data, configurations, countries, product lines, …)

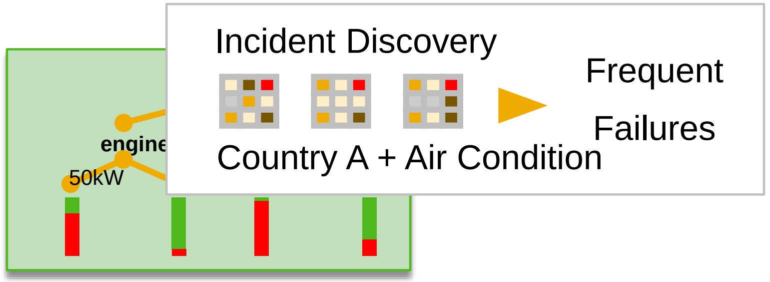
Usage:

Serves as starting point for expert review

Product/process improvement based on analysis results

Algorithms:

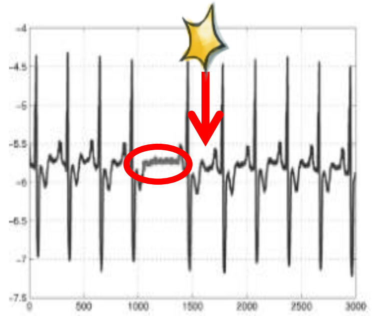
Decision Tree, Association Rule Mining



A	X
A B	X
A	X
C	X
A	X
B	X

A -> X

Predictive Maintenance
Detect anomalies that led to upcoming component failures



Anomaly Detection

Automatically find early deviations from normal

Analyze and repair if necessary

Required:

An incident indication (message or failure criteria)

Attributes with potential relation to failure (sensor data, configurations, countries, product lines, …)

Usage:

Early failure indicator

Product/process improvement and understanding

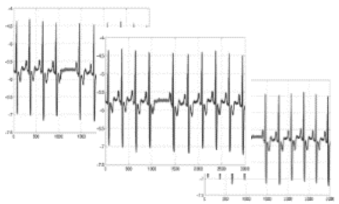
Algorithms:

Univariate methods (Box Plot Rule, Statistical Control Charts, Student’s Test, Grub’s Test, Likelihood Ratio Test)

Multivariate methods (PCA, One-Class Support Vector Machines, Self-Organizing Maps, Neural Networks)

Predictive Maintenance

Create statistical models based on past sensor data and failures



Classification

Learn a prediction Model from past sensor data and failures

Repair if probability of failure high

Required:

- A classification of past data (failure / no failure)
- Attributes with potential relation to failure (sensor data, configurations, countries, product lines, …)
- Sufficient number of failures that occurred in the past that can be related to the sensor data

Usage:

- Optimize maintenance to reduce costs and downtime
- Optimize spare parts planning

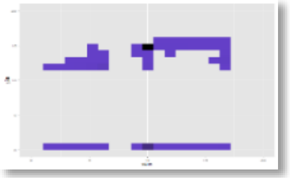
Algorithms:

- Decision Trees, Logistic Regression, Support Vector Machines, Neural Networks, Binary Classification…

Example: From battery data to battery health to action for Trenitalia

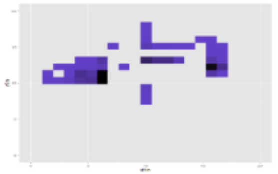
Collect Real Time Telemetry Data from Batteries

Real Time Data



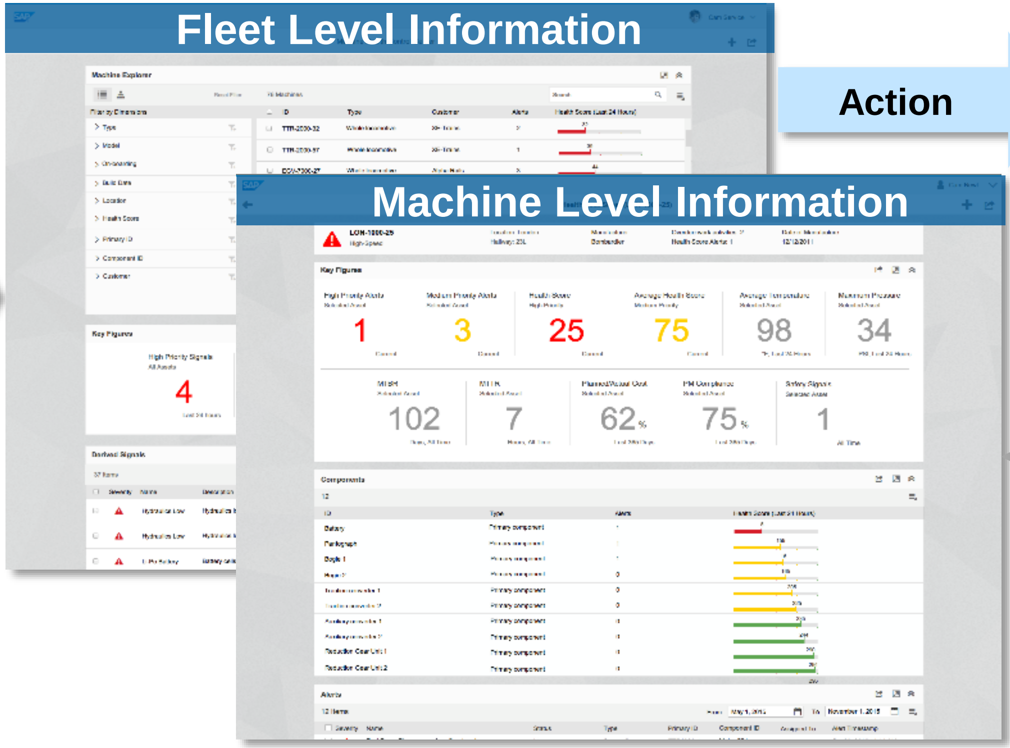
=

Ideal State



Compare and analyze real time battery performance to ideal state

Insights



PdMS Application shows insights about health score of machines and business information

Action

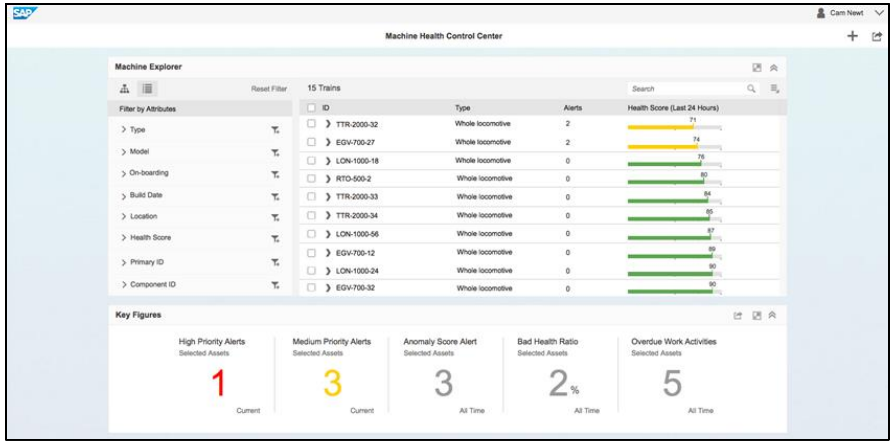
ERP + MRS

Maintenance Schedules & Financial Information

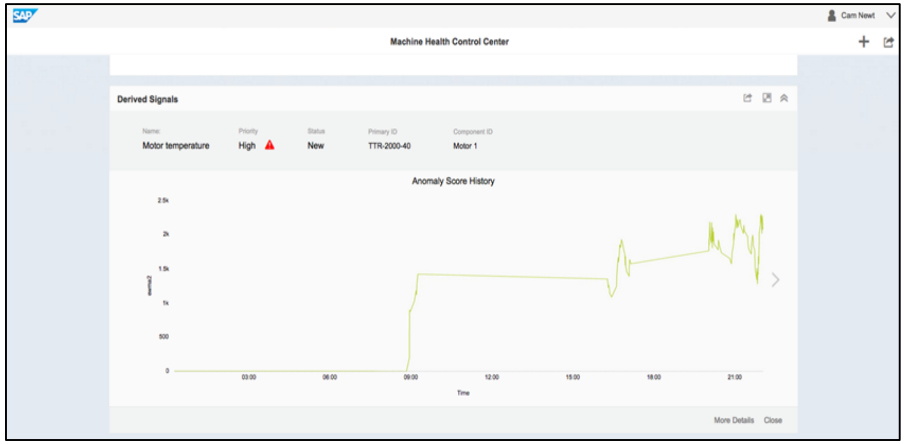
Outcome

60% maintenance cost reduction for batteries

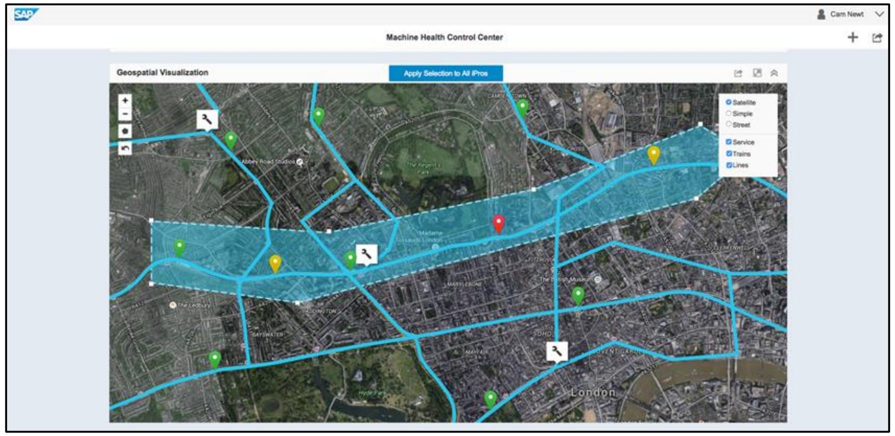
Data Science – Predictive Maintenance Application



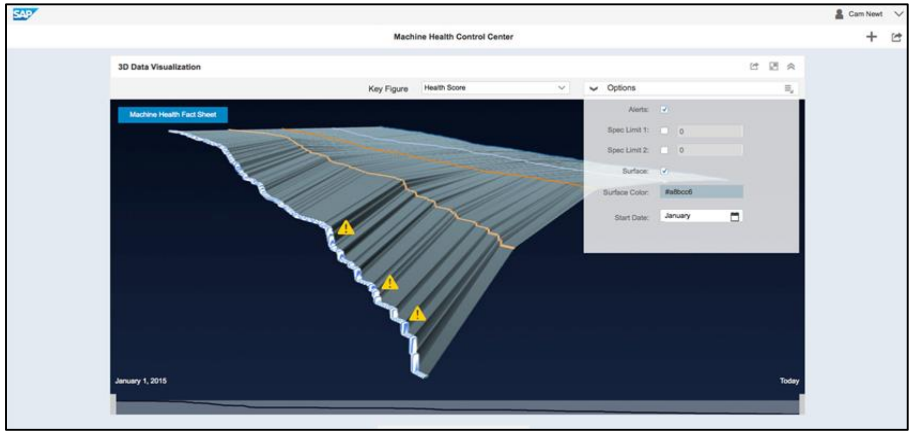
Machine Health Control Center - Machine Explorer



Machine Health Control Center - Anomaly Score History

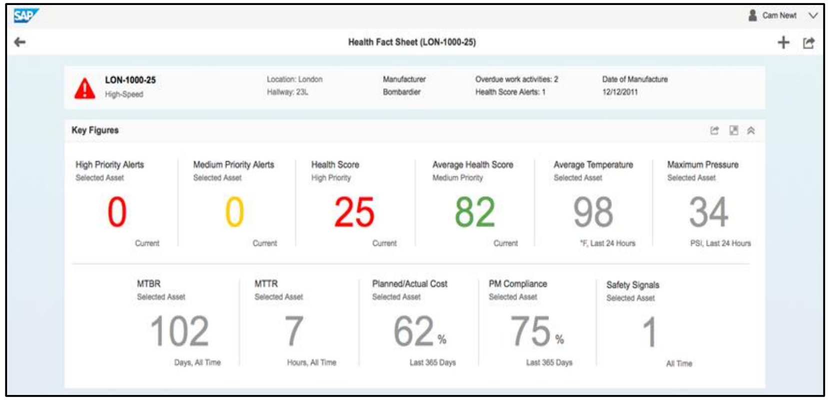


Machine Health Control Center - Geospatial Visualization

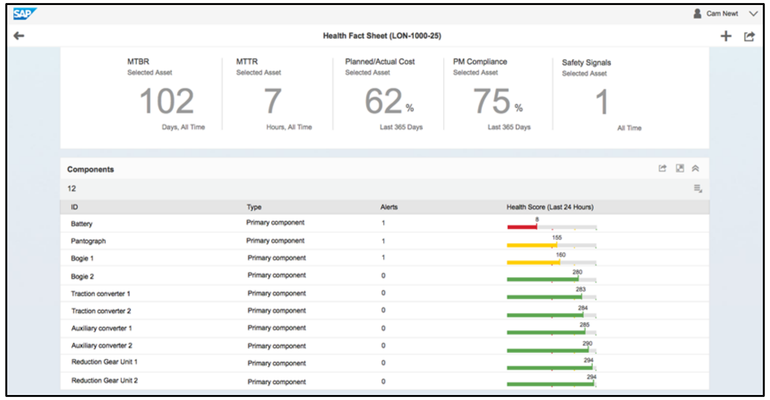


Machine Health Control Center - 3D Data Visualization

Data Science – Predictive Maintenance Application

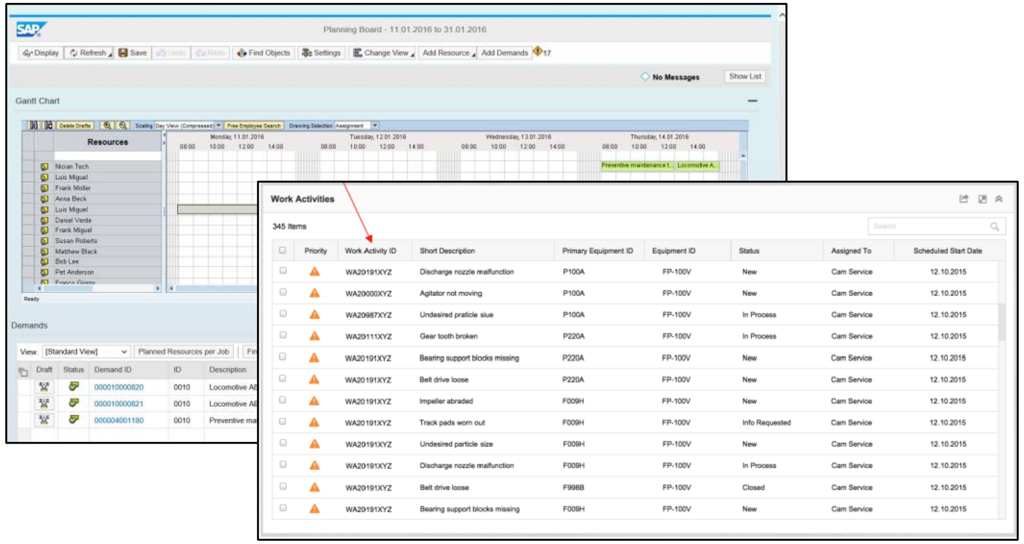


Health Fact Sheet



Health Fact Sheet by Components

Risk Matrix and Survival Curve



MRS Planning board with re scheduled maintenance