

Big Data 기반 제조품질분석

2015. 11. 06

신계영 수석컨설턴트
MES개발팀
솔루션그룹



강사 소개

주요 경력

- > 삼성 SDS, ICTO 사업부 MES 개발팀, Data Scientist (2012/01 ~)
- > 삼성전자, 생산기술연구소, Researcher (~ 2011/12)
- > Virginia Tech, Simulation & Optimization Lab, Research Scientist (~2011/03)
- > Virginia Tech, ISE(산업공학), Ph.D. (2010)
- Random Vector Generation on Large Discrete Spaces

전문 분야

- > Big Data Analytics Modeling on MES
- > Modeling, Data Mining, Statistics
- > Simulation & Optimization Methodologies

주요 업무

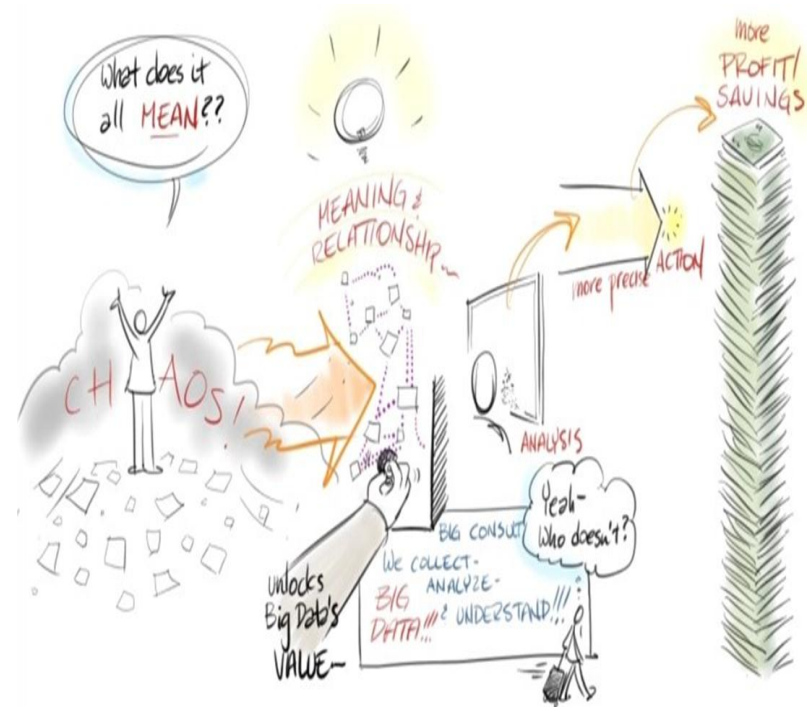
- > 솔루션 기획/PM, 新 사업 기획, 지적재산권 확보
- > 사업 발굴, POC, 과제 제안, 과제 수주
- > PI, Data Scientist, PM

I. FAQs about BIG DATA

II. 제조 품질분석 & Brightics

III. Brightics 적용 사례

Agenda



1. FAQs about Big Data

과제 제안
솔루션 소개
Delivery 할 때,
가장 많이
듣는 질문들

BEST 5

도대체 BIG DATA가 뭐야?

1

얼마나 커야 BIG DATA야?

2

기존의 통계분석, MINING이랑 뭐가 달라?

3

BIG DATA 도입하면 뭐가 좋아지는데?

4

우리가 BIG DATA 도입할 만한 수준은 되는
거야?

5

1. FAQs about Big Data (1/5)

1

도대체 BIG DATA 가 뭐야?

IBM 4Vs



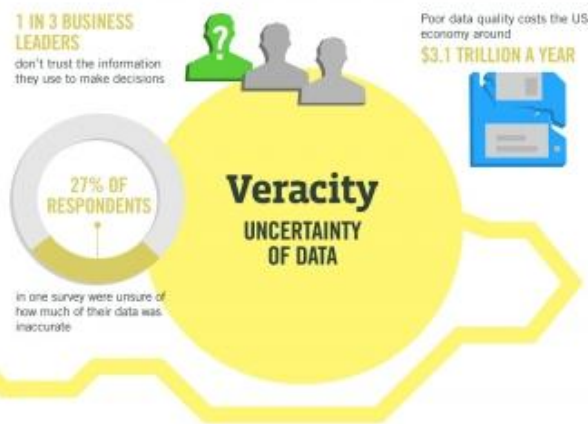
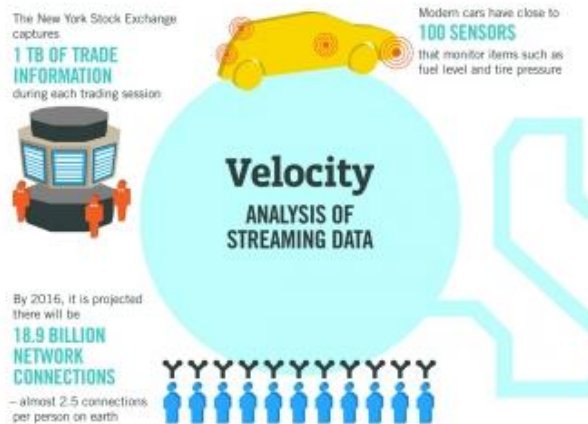
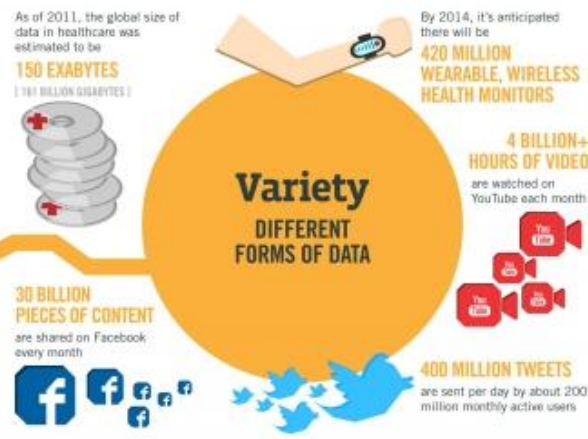
The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**.

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015, **4.4 MILLION IT JOBS** will be created globally to support big data, with 1.9 million in the United States



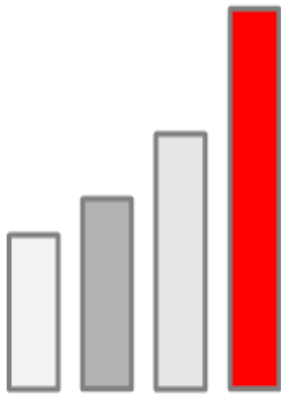
Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTEC, Q&A



1 도대체 BIG DATA 가 뭐야?

Oracle 4Vs

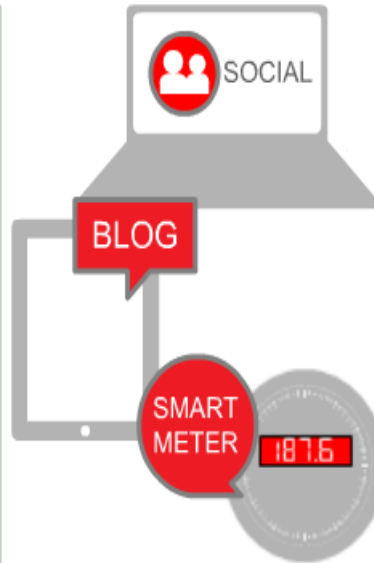
What Makes it Big Data?



VOLUME



VELOCITY



VARIETY

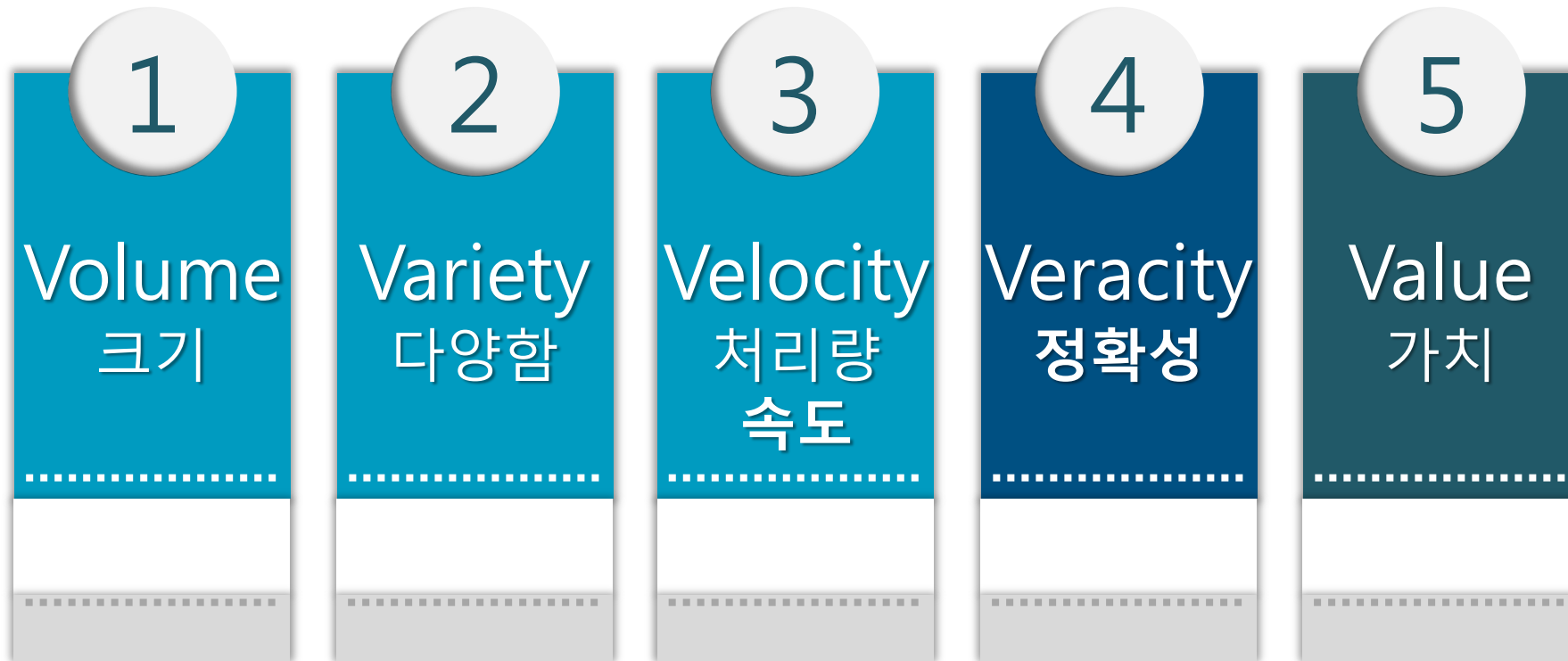


VALUE

1. FAQs about Big Data (1/5)

1 도대체 BIG DATA 가 뭐야?

SAS 5Vs



1. FAQs about Big Data (1/5)

1 도대체 BIG DATA 가 뭐야? SDS 전동수 ceo



1. FAQs about Big Data (2/5)

2 얼마나 커야 BIG DATA 야?



BIG DATA is ANY attribute that challenges CONSTRAINTS of a SYSTEM CAPABILITY or BUSINESS NEED..

1. FAQs about Big Data (2/5)

3 기존의 통계분석, MINING 이랑 뭐가 달라?












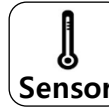
1. FAQs about Big Data (2/5)

3 기존의 통계분석, MINING 이랑 뭐가 달라?



1. FAQs about Big Data (3/5)

3 기존의 통계분석, MINING 이랑 뭐가 달라?

	데이터 유형	연산처리	알고리즘
통계분석 Data Mining	<p>> Structured Data</p> <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  Finance </div> <div style="text-align: center;">  Weblog </div> <div style="text-align: center;">  Point of Sales </div> </div>	Serial (Memory)	Data Mining Statistics
Big Data Analytics	<p>> Un-Structured Data</p> <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  Facebook </div> <div style="text-align: center;">  SNS </div> <div style="text-align: center;">  Voice </div> <div style="text-align: center;">  Image </div> </div>	Parallel (Hadoop)	Text Mining Speech Mining Image Mining
	<p>> Semi-Structured Data</p> <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  XML </div> <div style="text-align: center;">  JSON </div> <div style="text-align: center;">  Sensor </div> </div>		Time Series Analytics

1. FAQs about Big Data (5/5)

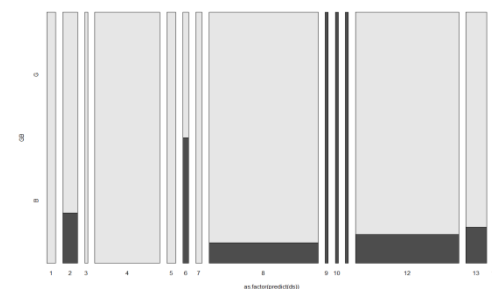
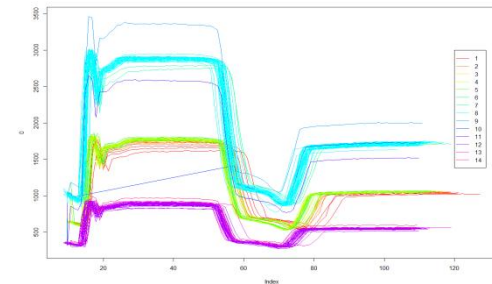
4

BIG DATA 도입하면 뭐가 좋아지는데?

ROI 증명 가능?



PoC 통해서 가능성 증명



1. FAQs about Big Data (4/5)

5 우리가 BIG DATA 도입할 만한 수준은 되는 거야?

우리 現 수준은?



진단 Tool Kit STAP¹⁾

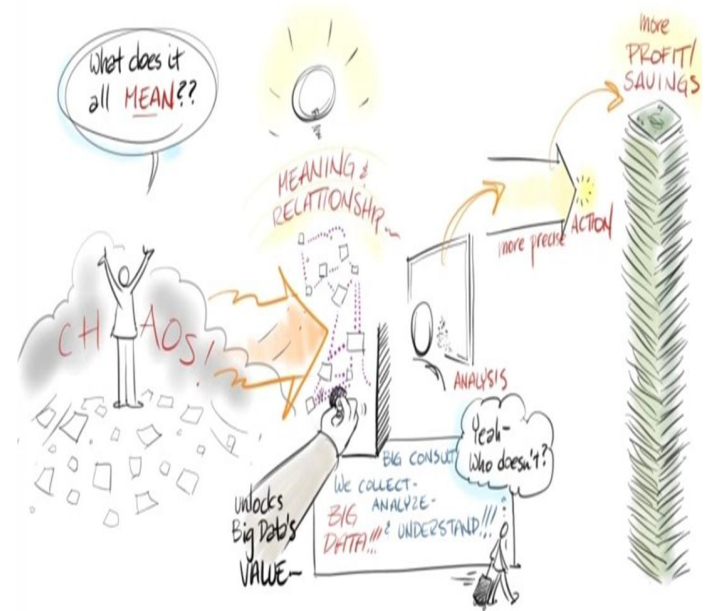
Definition	선진 수준의 정량화된 평가척도 기반으로 제조 품질 부문의 Process/Methodology/Infra&Architecture에 대한 Speedy한 진단을 통해 지속적 개선 관점의 다면적 수준 진단 Tool		
Objective	『 환경 변화에 신속 대응 가능한 제조 품질역량 지원 체계 유지 』 위한 중점 개선 과제 도출		
	업무 Process Gap 및 수준 진단을 위해	분석 개념 및 Flow 에 대한 진단을 위해	IT 혁신 지원 수준 진단을 위해
Construction	Process 운영수준 진단	분석 수준 진단	Infra 지원 수준 진단
		Data Sampling Clustering Analytic	Application Data Model Technology Deployment



1) STAP (Short Term Assessment Program)

- 고객 수준을 단기간 내에 진단 가능하도록 고안된 컨설팅 Tool Kit
- 설비/공정/제품 품질수준 및 Data 확보/활용 수준 등 진단

[Ⅱ. 제조 품질분석 & Brightics]



2. 제조 품질 분석 & Brightics

- 제조 품질 분석은 [설비, 공정, 제품] 품질 이상을 [감지, 분류, 분석], 나아가 [예측/제어]하는 일련의 모든 분석 활동

MES
Manufacturing
Execution
System

Scheduling/Dispatching (MSS)

Operation Control (MOS)

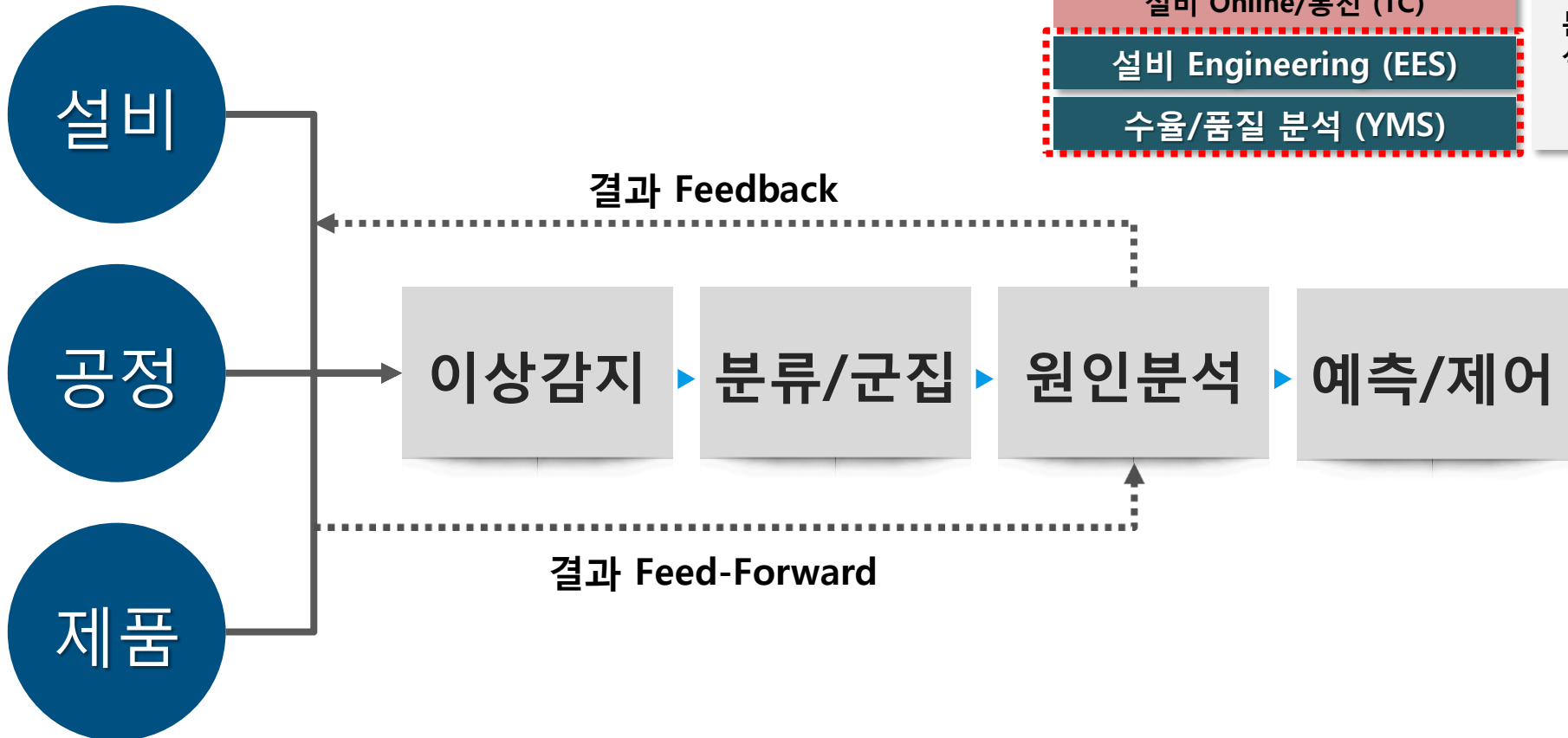
최적 생산/반송 제어 (MCS)

설비 Online/통신 (TC)

설비 Engineering (EES)

수율/품질 분석 (YMS)

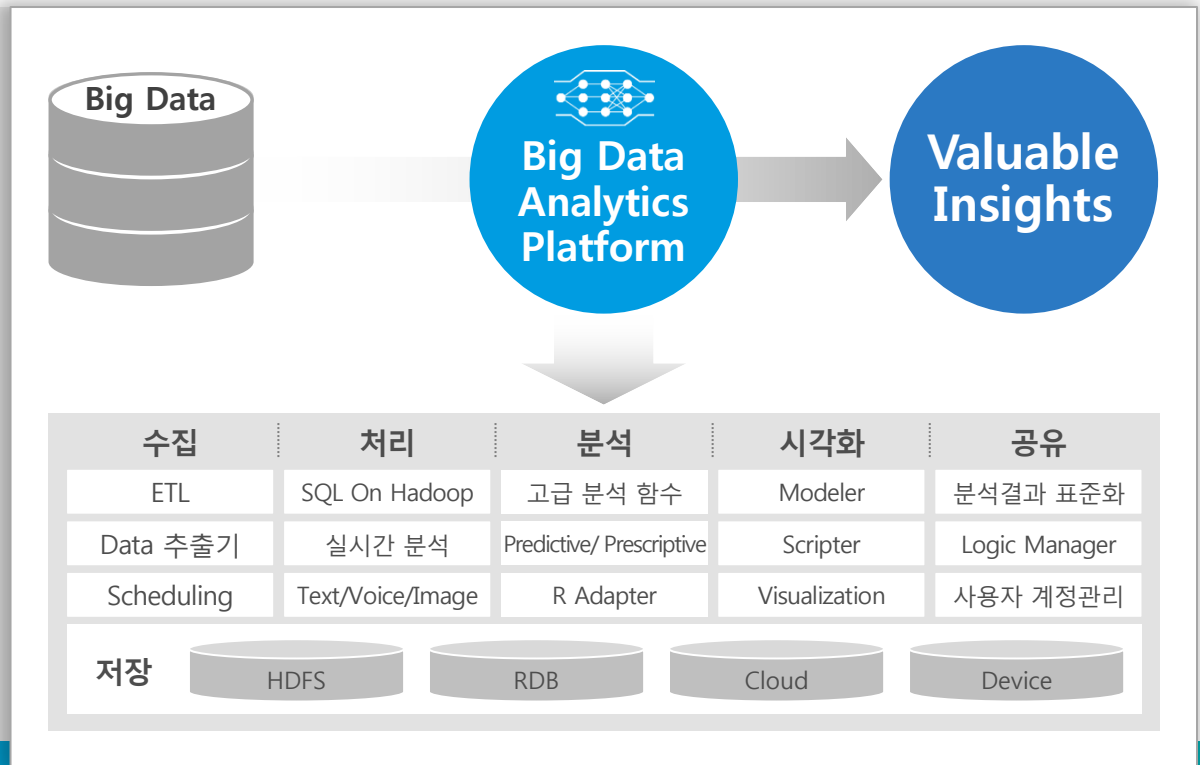
생산정보분석



2. 제조 품질 분석 & Brightics

- Brightics는 Hadoop 환경에서, 고속 병렬처리가 가능한 다양한 통계 Library를 바탕으로, 검증된 사례 기반의 Built-in Analytic Workflow를 제공하는 Big Data 분석 솔루션

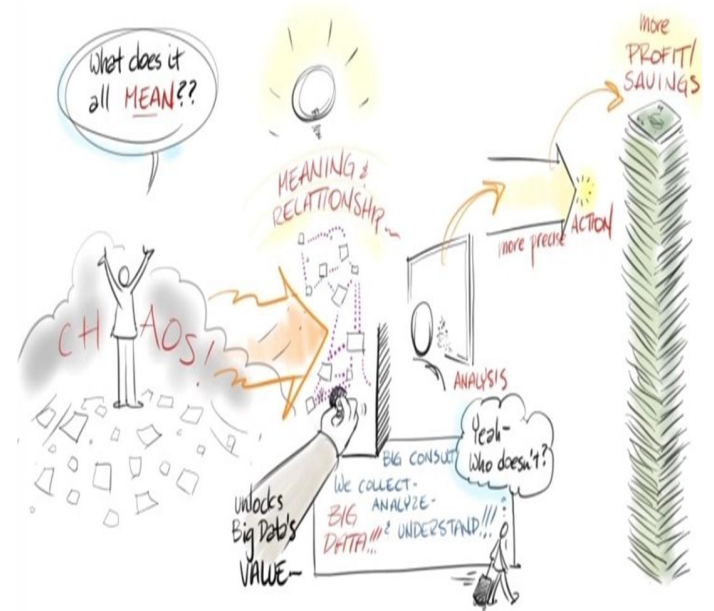
- Hadoop 기반으로 경제적인 Big Data 분석 시스템 구축 가능
- In-Memory 병렬 연산 기반의 다양한 통계 함수 제공
- 분석 모델 설계와 공유가 용이한 Workflow 기반의 모델링 Tool 지원
- 사용자 Interaction이 가능한 다양한 종류의 Chart 제공으로 분석 시각화 용이



Value >

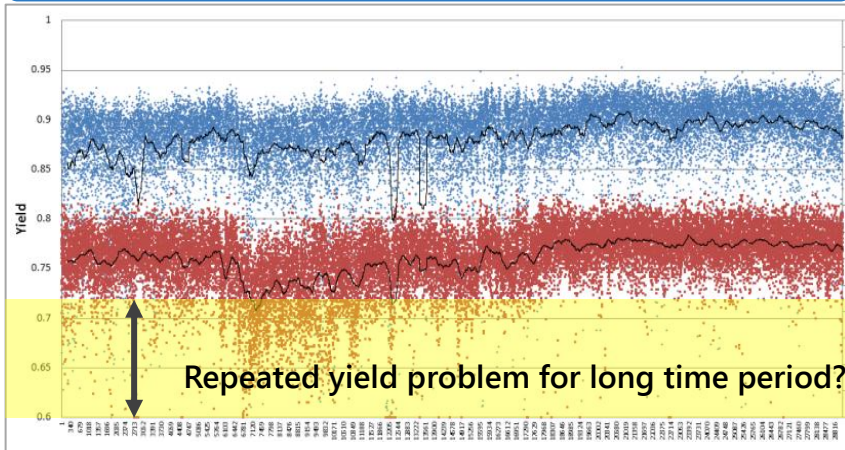
In-memory 기반의 대용량 고속 병렬 처리 기술을 활용하여 분석 함수를 제공함으로써 경제성, 효율성, 호환성, 확장성 제공

[Ⅲ. Brightics 적용 사례]

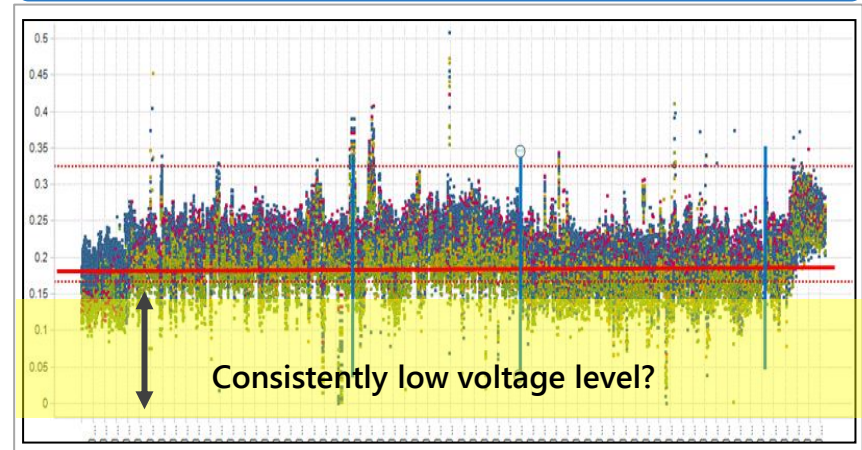


3. Brightics 적용사례 - 제조 현장의 품질 분석 문제는?

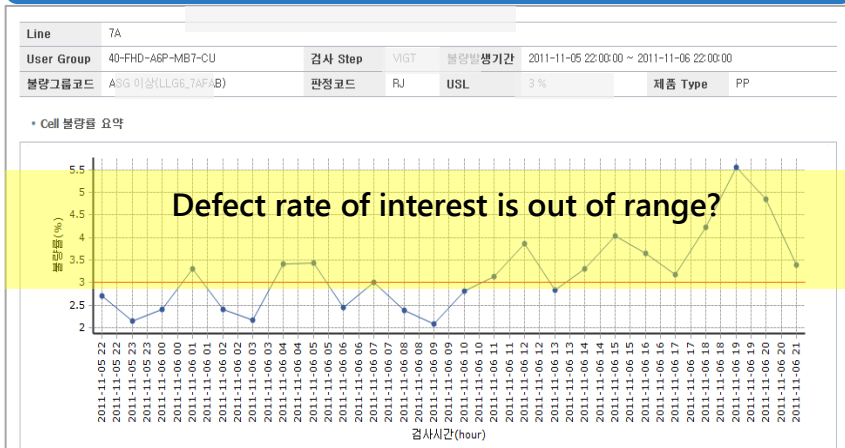
Repeated Yield Drop



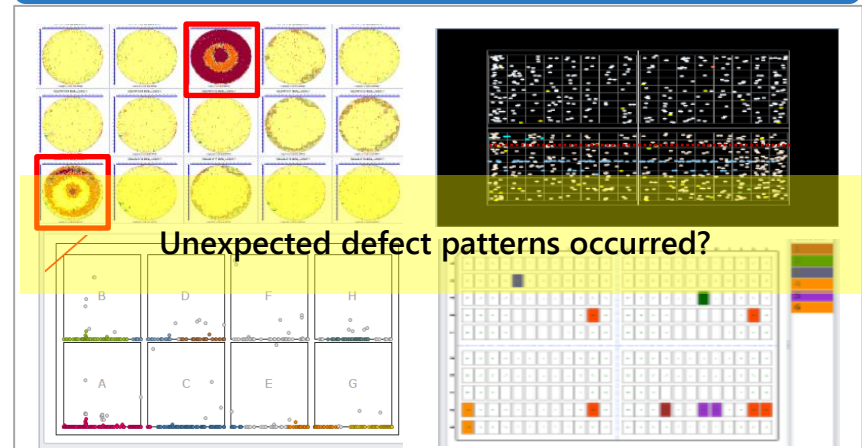
Abnormal Trend in Metrologic data



Important Defect Suddenly Increased



Particular Defect Map Pattern

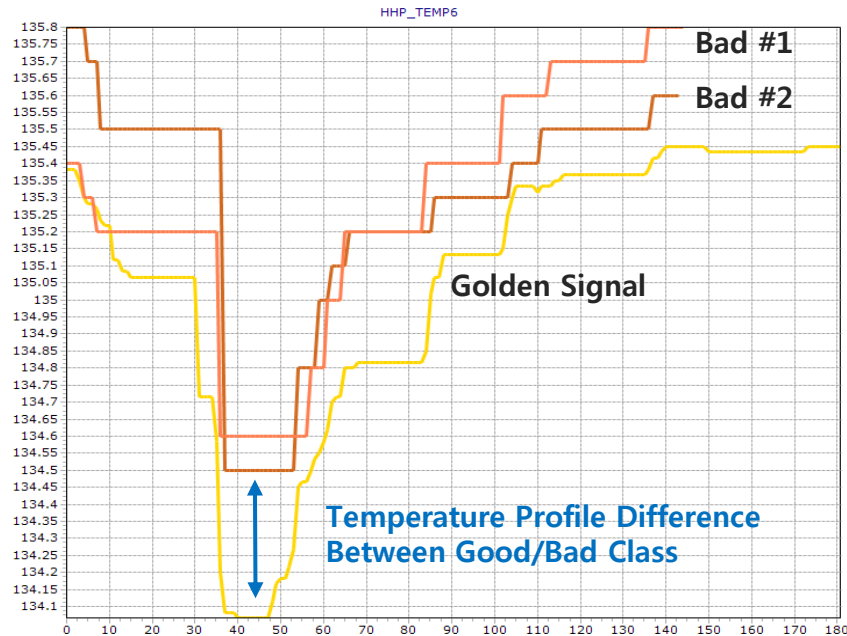


3. Brightics 적용사례 - 대용량 설비 센서 데이터 패턴 해석(1/2)

- 초 단위 설비 센서 패턴 차이에 따른, 품질 영향도 미세 분석 통해, 고질 불량 원인 설명력 향상

Tool Sensor Profiles

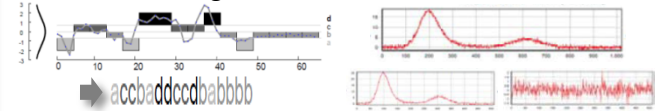
Temperature Sensor data collected from tools



Time Series Sensor Data Analytics

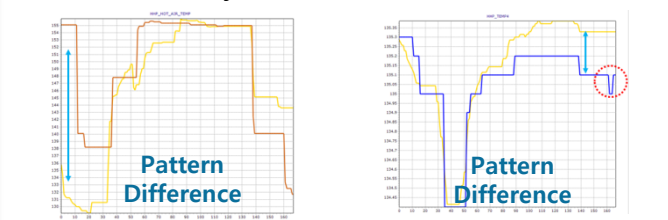
1 Big Data Preprocessing

▪ Noise Cancelling & Dimension Reduction



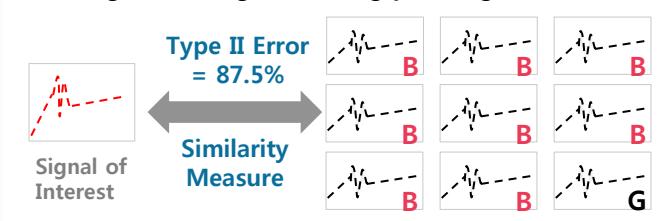
2 Root Cause Analysis

▪ Root Cause Analysis via Sensor Data



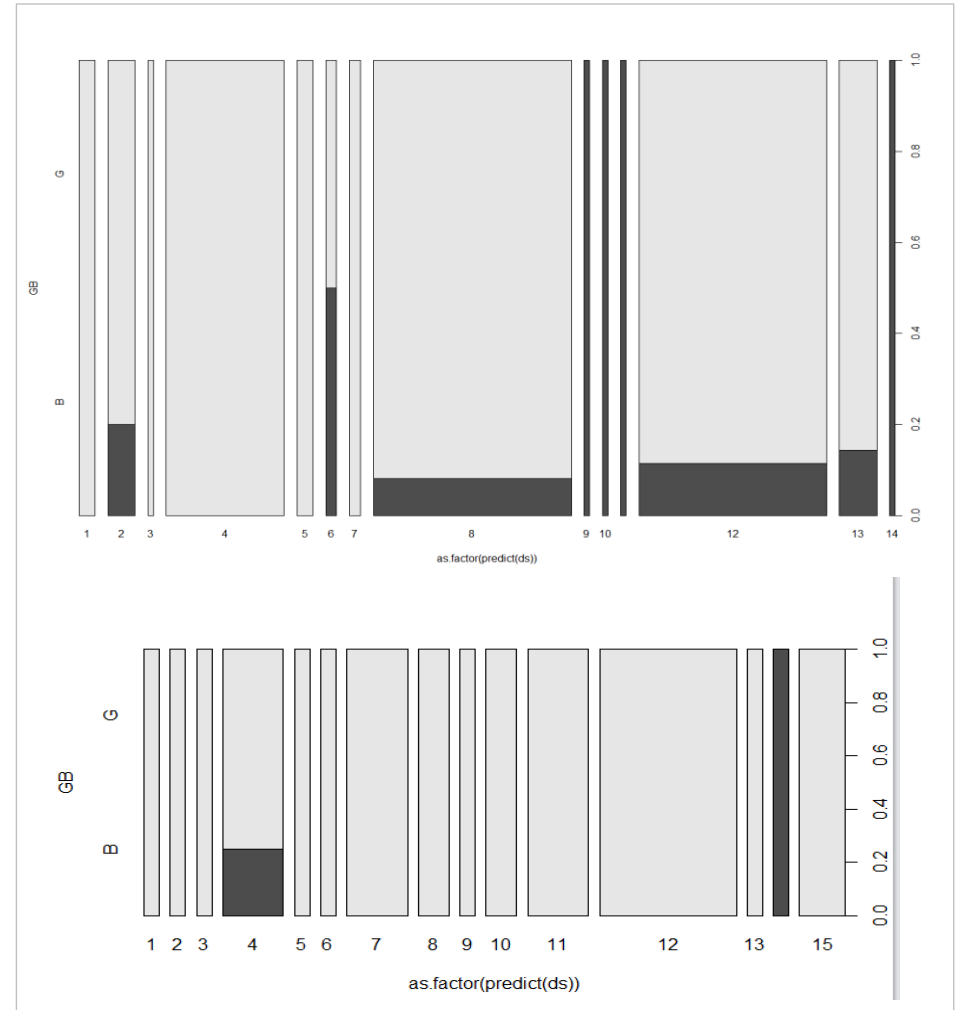
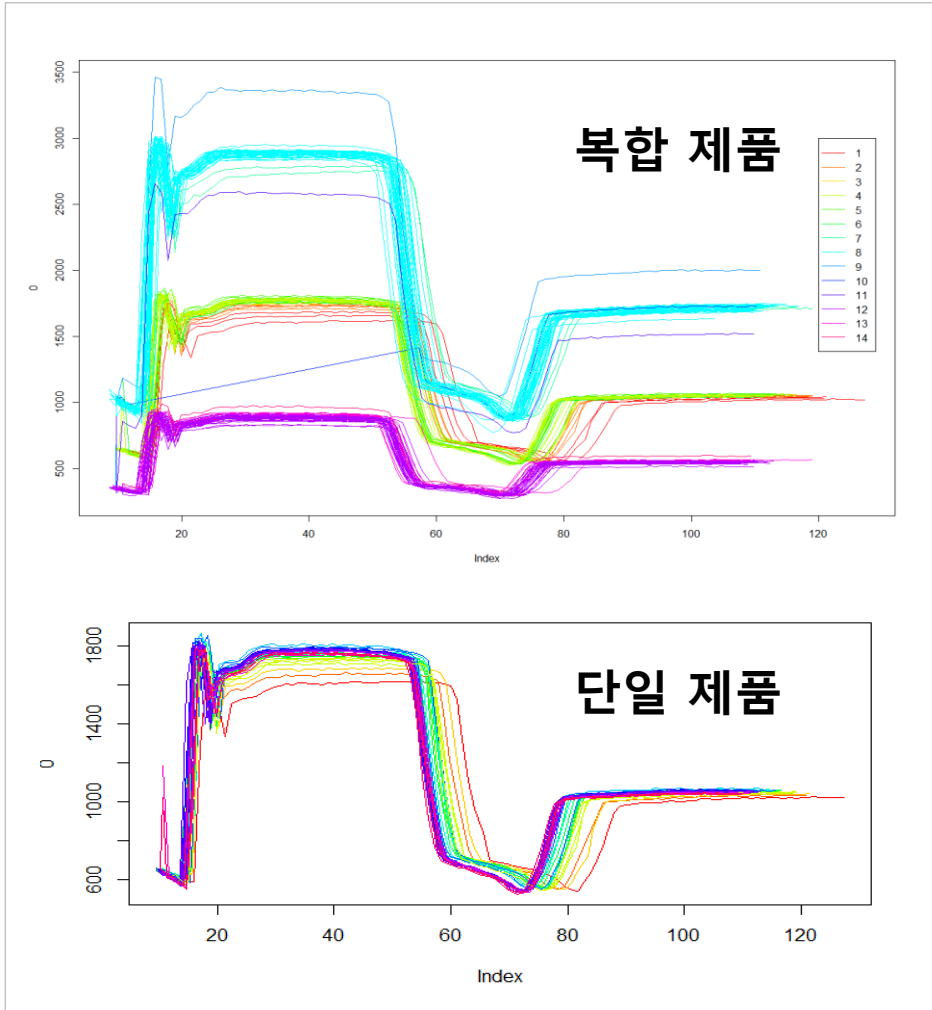
3 Validation

▪ Finding Similar Signals among past (big) data



3. Brightics 적용사례 - 대용량 설비 센서 데이터 패턴 해석(2/2)

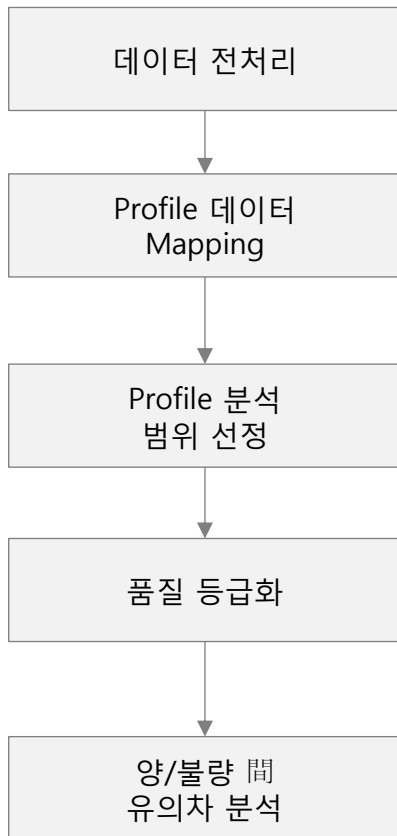
- 센서 패턴 별로 Cluster 분류되었고, 불량률이 높은 제품이 속한 군집을 쉽게 분류 가능함



3. Brightics 적용사례 - 설비 전압 신호 Profile의 품질 연계 해석 (1/3)

- 품질 판정 결과 분포 기준, 5단계 미세 등급화 後, 품질 등급 別 전압 Profile 유의차 분석

분석 절차



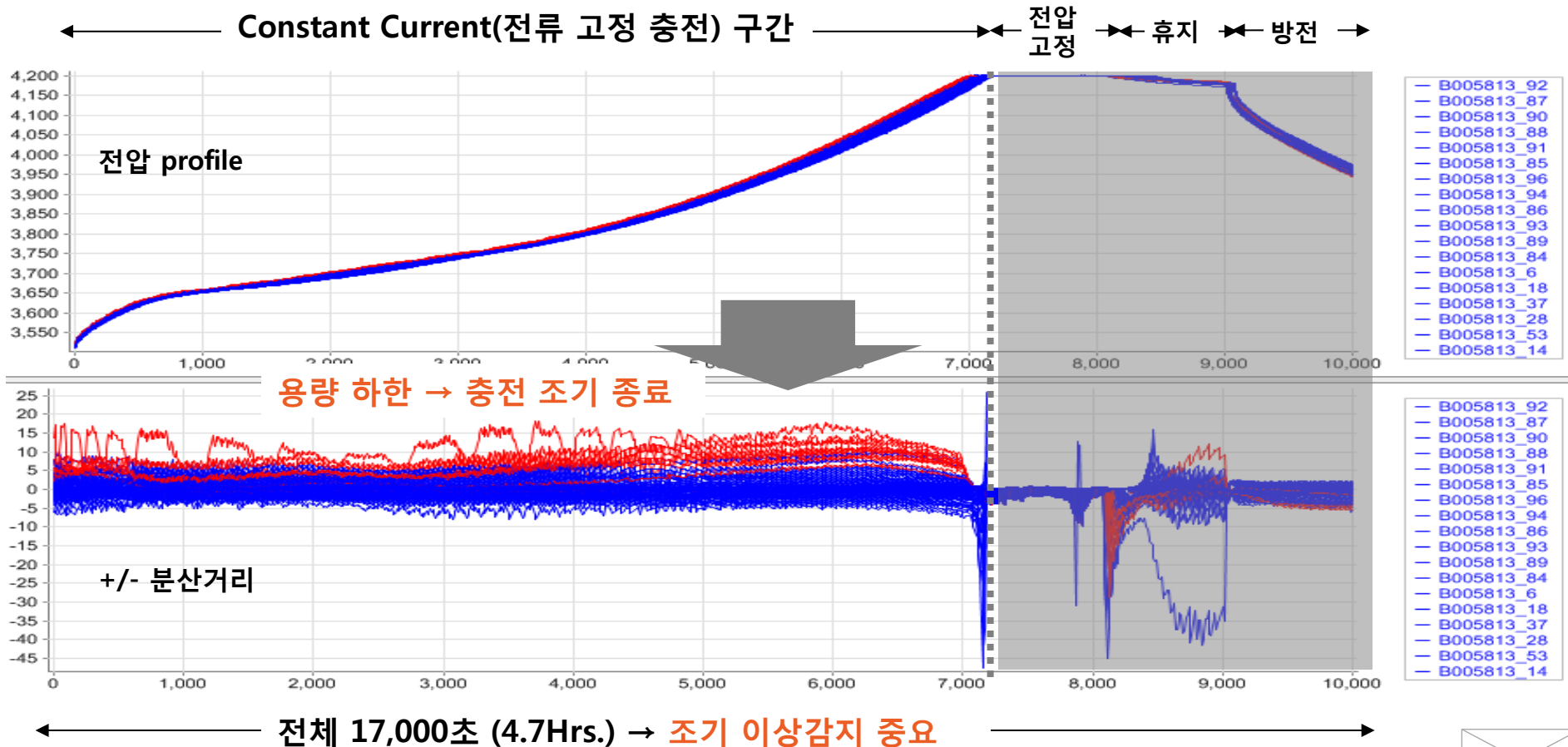
Analytics Reference



1. 품질 판정 결과 기준으로 등급 분할 (예: 용량 상/하한 기준, 4등급까지 양품, 5등급 불량)
2. 양품선정 : 1~4등급 (5등급 제외)으로 나누고, 산포 추정 및 용량 상/하한의 중간을 평균으로, 가정하여 Percentile로 등분
3. 품질 등급화 및 등급 別 경향성 분석을 위한 목적임

3. Brightics 적용사례 - 설비 전압 신호 Profile의 품질 연계 해석 (2/3)

- 전압 Profile 데이터 Transformation을 통한, 양/불 패턴 유의 차 가시성 확보



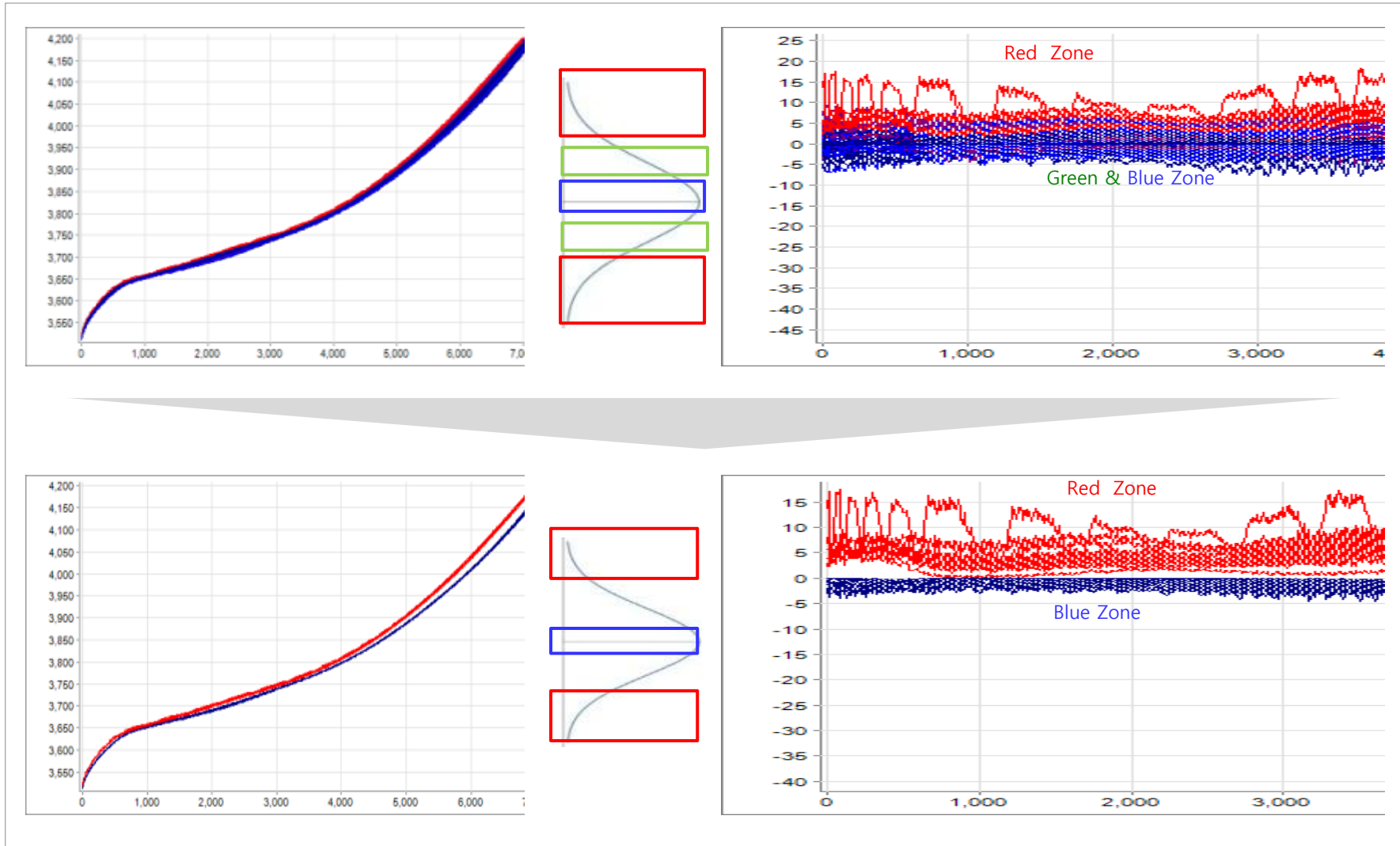
- 시점 別 Normalization : 각 시점 別 평균, 표준편차 계산 후 각 시점 데이터를 정규화
- 시점 別 +/-분산 계산 : 각 시점 別 Z^2 값 계산(분산 의미), 평균 상/하 위치에 따라 +/- 부호 반영

$$Z_t = \text{[Box with X]}$$

$$V_t^2 = \text{[Box with X]}$$

3. Brightics 적용사례 - 설비 전압 신호 Profile의 품질 연계 해석 (3/3)

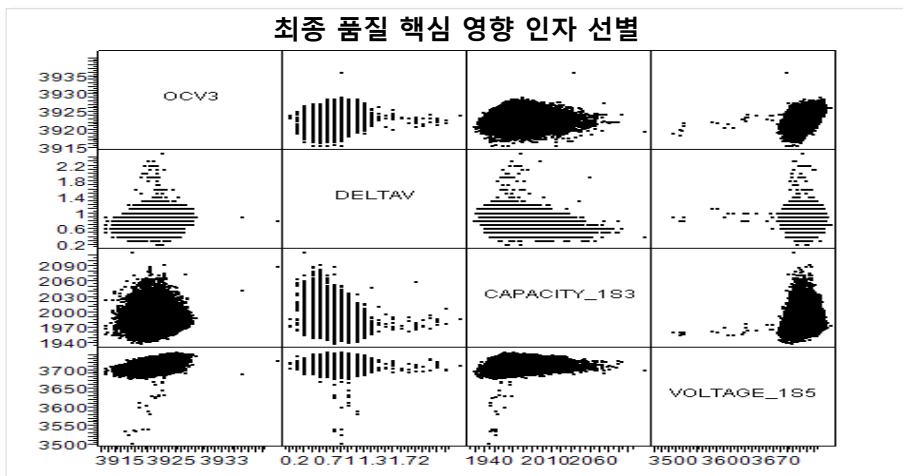
- 양품과 불량 신호間 경계를 명확히 구별하여, 불량 신호 패턴 학습 후 실시간 감지에 활용



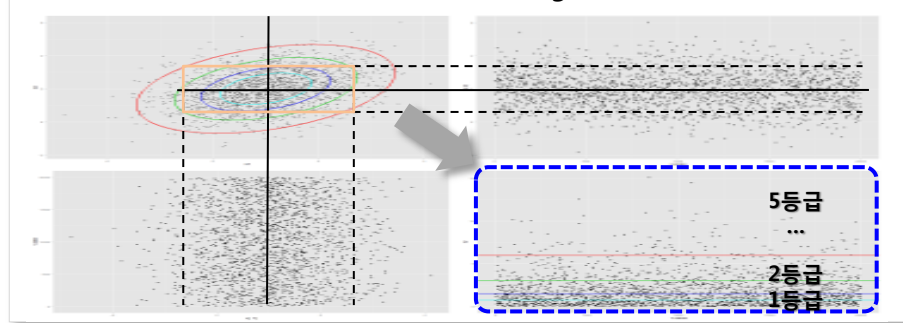
3. Brightics 적용사례 - 양품 등급 분석

- 불량 위주의 감시 및 원인 분석 방식에서 '양품 Trend 관측 및 불량 사전 예측' 체계로

양품 품질 미세 등급화

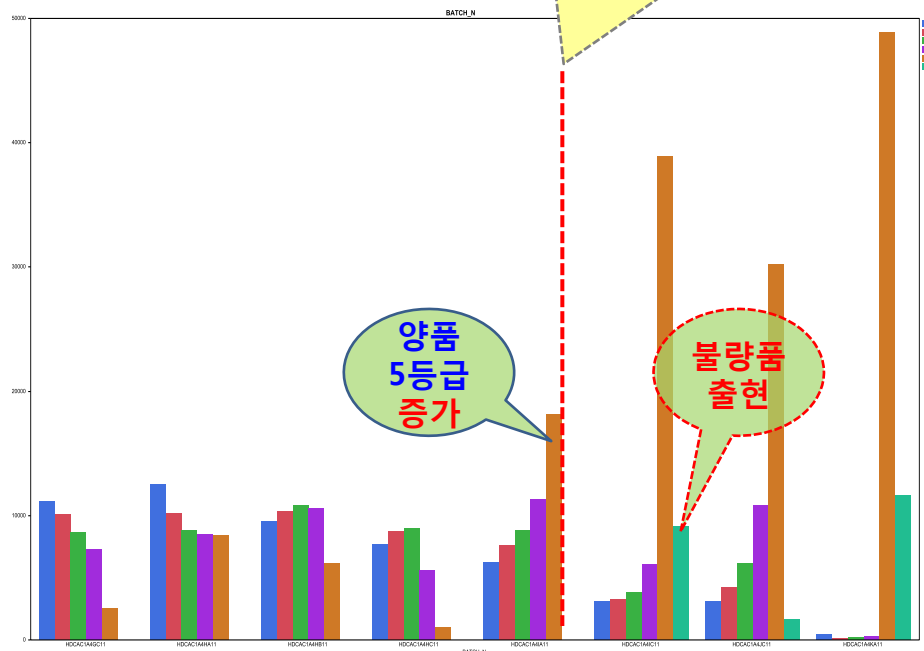


양품 품질 미세 등급화 (Hotelling's T2 Statistic)



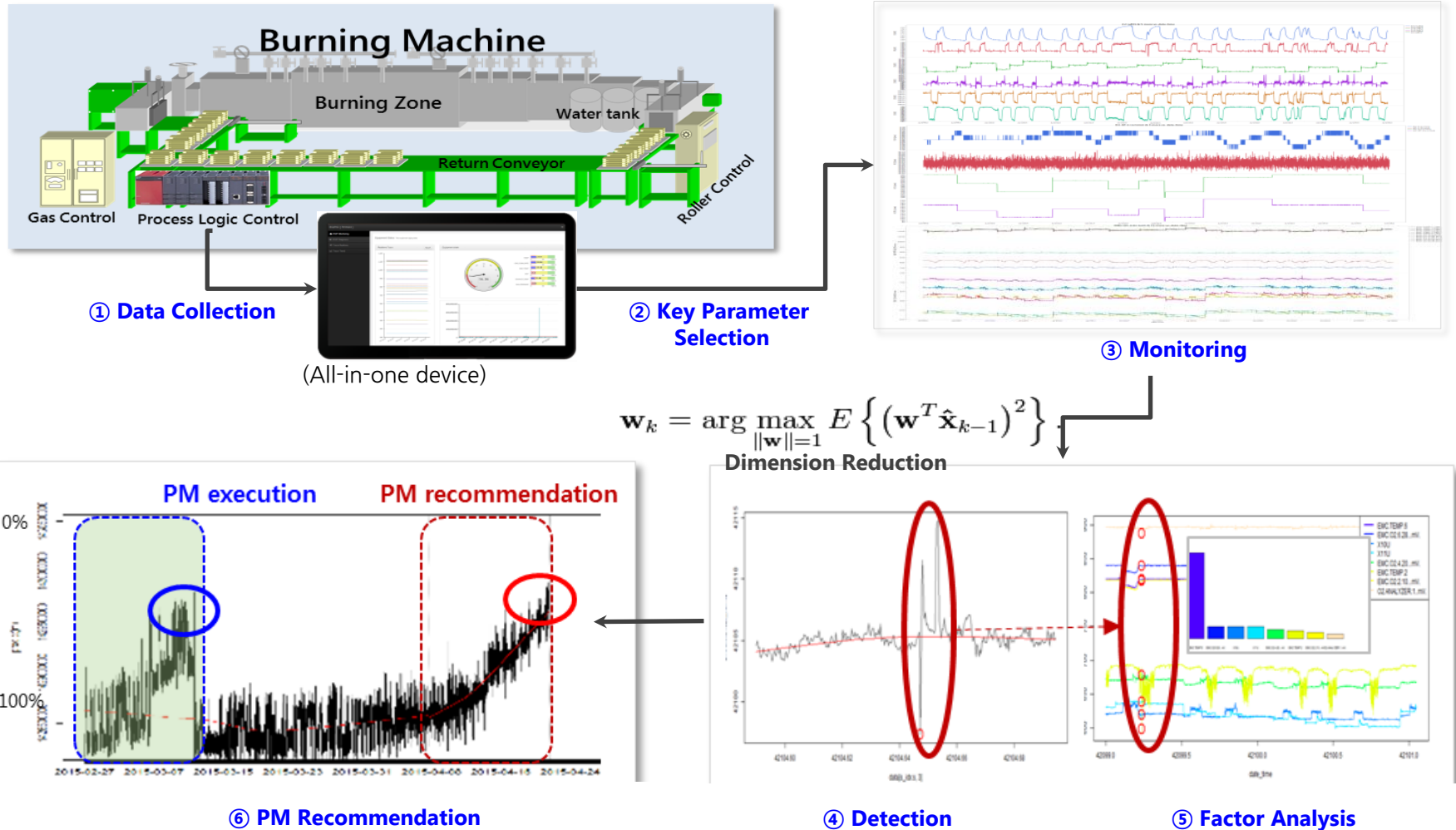
양품 Trend 감시 → 불량 징후 사전 예측

"5등급 양품 다발 시점 인지를 통해, 이후 Batch부터 불량 제품 증가 예측 가능"



3. Brightics 적용사례 - 설비 예지 보전

- 전체 불량 대상 품질 등급화 및 시장 불량 연계분석 결과



Thank you

신계영 (kay.shin@samsung.com)



Copyright © 2015 Samsung SDS Co., Ltd. All rights reserved