

# Analysis of the Effectiveness of Big Data-Based Six Sigma Methodology: Focus on DX SS

Kim Jung Hyuk<sup>†</sup> · Kim Yoon Ki<sup>††</sup>

## ABSTRACT

Over recent years, 6 Sigma has become a key methodology in manufacturing for quality improvement and cost reduction. However, challenges have arisen due to the difficulty in analyzing large-scale data generated by smart factories and its traditional, formal application. To address these limitations, a big data-based 6 Sigma approach has been developed, integrating the strengths of 6 Sigma and big data analysis, including statistical verification, mathematical optimization, interpretability, and machine learning. Despite its potential, the practical impact of this big data-based 6 Sigma on manufacturing processes and management performance has not been adequately verified, leading to its limited reliability and underutilization in practice. This study investigates the efficiency impact of DX SS, a big data-based 6 Sigma, on manufacturing processes, and identifies key success policies for its effective introduction and implementation in enterprises. The study highlights the importance of involving all executives and employees and researching key success policies, as demonstrated by cases where methodology implementation failed due to incorrect policies. This research aims to assist manufacturing companies in achieving successful outcomes by actively adopting and utilizing the methodologies presented.

Keywords : Big Data, 6 Sigma, Effectiveness Analysis, Key Success Policy

# 빅데이터 기반 6시그마 방법론의 유효성 분석: DX SS를 중심으로

김 정 혁\* · 김 윤 기\*\*

## 요 약

지난 수년간 6시그마는 제조업의 주요 혁신 방법론으로, 품질개선과 경비 절감을 위해 사용되었다. 그러나 스마트공장 확산으로 인한 초 단위 데이터 생성 등, 방대한 양의 데이터를 분석하기 어려운 문제와, 오랫동안 정착된 형식적 사용으로 인해, 6시그마의 한계가 지적되었다. 6시그마의 한계를 극복하기 위해, 최근에 빅데이터 기반 6시그마 기법이 연구되고 있다. 빅데이터 기반 6시그마는, 6시그마의 강점인 통계적 검증, 수학적 최적화, 높은 해석력과, 빅데이터 분석의 강점인 기계학습을 모두 활용할 수 있다. 그러나, 최근 연구된 빅데이터 기반 6시그마 기법이 제조공정 및 경영 성과에 미치는 영향에 대한 검증은 미비하다. 이러한 이유로 실무에서는, 빅데이터 기반 6시그마 기법에 대한 신뢰성이 높지 않아 제대로 활용하지 못하고 있다. 본 연구에서는, 빅데이터 기반 6시그마인 DX SS의 유효성 분석을 통해 제조공정의 효율성에 미치는 영향을 알아본다. 또한 기업에서 이 기법을 성공적으로 도입 및 정착시키기 위한 핵심 성공 정책을 도출한다. 추가적으로, 성공 정책에 대한 연구 없이 전 임직원의 참여가 수반되지 못한 잘못된 정책으로 방법론이 중단된 사례는, 핵심 성공 정책 연구에 대한 중요성을 보여준다. 제조기업들이 본 연구에서 제시하는 방법론을 적극 도입하고 사용하여 성공적인 성과를 거둘 수 있도록 본 연구가 도움이 되기를 기대한다.

키워드 : 빅데이터, 6시그마, 유효성 분석, 핵심 성공 정책

## 1. 서 론

기업의 주된 목표는 이윤 창출이다. 소비자 요구사항이 증가하고, 기업 경쟁이 치열해지면서 국내외 기업들은 경영혁신

방법을 지속적으로 진화시키며 경쟁 우위를 확보하기 위해 노력해 왔다. 특히 제조업 기업들은 품질관리와 경비 절감을 달성하기 위해 1970년대에 대량 생산에서 다품종 소량 생산으로 전환과 함께 TQM(Total Quality Management, 총체적 품질관리)을 도입하였다. 이어서 1987년 모토로라에서 개발된 6시그마 방법론(품질개선 경영혁신 방법론)은 Allied Signal, GE 등을 거쳐 2000년 전후로 전 세계로 확산되었으며, 현재 세계적 기업의 40% 이상 기업이 6시그마 경영을 추진하고 있다[1]. 국내에서도 1996년 LG 전자, 삼성 SDI 등 대기업을 중

\* 이 논문은 고려사이버대학교 학술연구비(2년)에 의하여 연구되었음.

† 준 회 원 : 고려사이버대학교 융합정보대학원 석사과정

†† 정 회 원 : 고려사이버대학교 전임교수

Manuscript Received : October 6, 2023

First Revision : December 5, 2023

Accepted : December 11, 2023

\* Corresponding Author : Kim Yoon Ki(ykkim77@cuk.edu)

심으로 도입되어, 2000년대 중 후반에는 중소기업을 포함하여 모든 산업 부문 및 대학으로까지 확산하였다. 모토로라에서는 6시그마를 도입 후 5년간 제조 공정 부적합품률을 150배 이상 감소시켰으며, 이에 따라 제조 부문과 비제조 부문에서 총 34억 달러의 비용을 절감하기도 하였다[2]. GE에서는 6시그마 방법론을 적용하여 도입 초기에 그 효과가 7억 불, 다음 해에는 12억 불에 달하는 것으로 파악하였으며, 국내 13개 대기업도 3조 원 이상의 비용을 절감하고 있다고 하였다[3]. 하지만, 1996년 국내에 6시그마 도입 이후 약 27년이 지난 현재 상황은 조금 다르다. 많은 기업이 6시그마를 도입하여 많은 재무적, 비재무적 효과를 얻어냈지만, 그렇지 못한 기업도 많다. 선행연구에 따르면, 우리나라 기업의 경우 국내 1,000대 기업들 가운데 약 70%가 6시그마를 포함한 경영혁신 활동을 추진하였으나, 최고경영자 층과 조직구성원들의 경영혁신에 대한 인식 및 이해 부족으로 추진업체 중 약 60%의 기업은 성공하지 못한 것으로 나타났다[3]. 과거에는 6시그마의 실패 사례는 최고경영자의 의지 부족, 추진인력 및 역량 부족, 통계 사용으로 접근이 어렵다는 점, 지나친 문서작업과 DMAIC 방법론 절차에 대한 시간 낭비/잔업 증가 등이 주요 요인으로 언급되었다[4]. 이러한 문제들은 성공 정책과 성과에 대한 연구를 통해 어느 정도 보완할 수 있었지만, 현재는 6시그마를 형식적으로 사용하여 효율성이 떨어지는 사례를 흔히 발견할 수 있고, 새로 도입하는 기업들을 찾아보기 어렵고, 연구 활동 또한 매우 적어졌다. 이는 스마트 공장의 확산으로 초 단위로 생성되는 대량의 데이터를 6시그마 기법으로 분석하는데 한계가 있다고 판단한 것이다.

Fig. 1은 6시그마와 빅데이터 기반 6시그마의 데이터 사용 영역을 비교한 것이다. 6시그마는 사전에 엔지니어의 경험과 지식 도메인 내에서 문제의 잠재인자가 있을 것으로 판단되는 데이터 영역만 활용한다. 모집단에서 표본을 샘플링하여 일부 데이터를 추출하여 모수를 추정하므로, 선택한 데이터 영역에 불량유발인자가 없으면 품질 개선이 되지 않는다. 그러나 빅데이터 분석 기법은 전수 데이터를 활용하므로, 모든 불량유발인자를 찾아낼 수 있으며 6시그마를 결합하여 주요 인자를

최적화하는 실험을 활용하면, 더욱 효과적인 품질개선 효과를 기대할 수 있다[5].

LG에서는 TMS(Total Monitoring System)을 통해 Clean Room, GAS, Chemical, 열원, 공조, 대기, DI, 정수 공정 등의 설비 모니터링 관제를 위한 System을 구축하고 해당 공정의 압력/유량/품질 농도 등을 계측하기 위한 Sensor류를 배관에 설치하고, PLC Data를 중앙 Server에 송신받아 초단위로 Data를 생성받는다. 공정마다 측정 Sensor의 개수가 상이하지만, 일반적으로 공정별로 수백개에서 수천개의 Sensor가 설치되어 있다. 1개의 Sensor에서 생성되는 Data를 분 단위 사용한다고 가정시, 60분\*24시간 = 1,440개의 Data가 생성된다. 100개의 Sensor를 1분 단위로 1일치 Data만 활용하려고 해도 144,000개의 Raw Data 분석이 필요하다. 기존 6시그마의 경우 Raw Data의 양이 1,000개만 넘어가도 미니탭과, JMP와 같은 통계프로그램을 주로 활용하므로 처리속도가 느려져서 분석이 지연되지만, 빅데이터 기반 6시그마의 경우 주력 통계프로그램으로 R과 Python을 사용하므로 일반적인 PC 사양으로도 1억개 이상의 Raw Data를 처리 가능하다. 이처럼 빅데이터 기반 6시그마는 전수 데이터를 활용하여 모든 불량유발 인자 분석이 가능하므로, 효과적인 품질 개선효과를 기대할 수 있다.

최근에는 기존 6시그마의 한계를 극복하기 위해 빅데이터 기법을 추가한 빅데이터 기반 6시그마 방법론에 대한 연구가 활성화되고 있다. 최근 10년간 총 7건의 빅데이터 기반 6시그마 방법론의 연구가 있었으며, 그 중 1건은 6시그마 방법론에 데이터마이닝과 BSC(Balanced Scorecard), BPM(Business Process Management)을 결합한 방법론이었고, 빅데이터 분석기법과 6시그마 방법론을 결합한 방법론은 6건이었다. 이 중 국내 연구 논문은 6건, 해외 연구 논문은 1건으로, 빅데이터 기반 6시그마 방법론 연구는 국내연구 비중이 전체의 86%를 차지할 정도로 높았다[5-11]. 비록 연구논문의 수가 많지 않음에도 불구하고, 기존의 통계 기법을 활용한 6시그마 방법론이 일부 데이터를 활용하는 한계에 직면함에 따라, 6시그마와 빅데이터 분석 기법과의 결합을 통한 신규 방법론의 개발은 혁신적인 시도로 평가될 수 있다. 이 방법론은 전수 데이터 분석을 하여 모든 불량유발인자를 찾아낼 수 있으므로 데이터 분석 효율을 향상시키고, 품질 개선 및 비용 절감에 상당한 기여를 하고 있다. 이러한 효과성에 기인하여, 학계에서는 이 분야에 대한 연구가 지속적으로 수행되고 있다.

빅데이터 기반 6시그마는 기존 6시그마 방법론에 빅데이터 분석에서의 데이터 전처리, 군집/분류/예측 기계학습, 모델링, 하이퍼 파라미터 최적화 등의 빅데이터 분석기법을 추가한 방법론이다. 6시그마의 통계적 검·추정과 실험계획법을 통한 수학적 최적화를 통해 높은 해석력을 제공 받는 동시에, 빅데이터 분석기법의 기계 학습을 통한 높은 예측력을 활용할 수 있다는 점이 강점인 경영혁신 방법론이다[12].

LG 전자는 2020년부터 업계 최초로 LG만의 빅데이터 기반

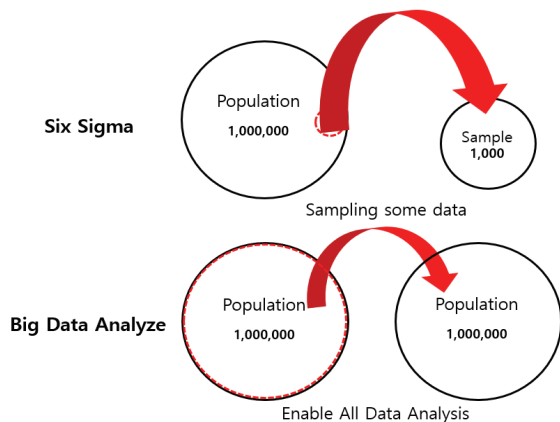


Fig. 1. Comparison of Data Usage Areas to Improve Bad Data

6시그마 기법인 DX SS(Digital Transformation Six Sigma)방법론을 개발하여 사용 중이며, 2022년부터는 LG 디스플레이를 비롯한 LG 계열사에 확대되어 LG만의 고유한 혁신 방법으로 발전시켜 오고 있다.

하지만 아직 빅데이터 기반 6시그마의 유효성을 검증하기 위한 연구는 매우 부족한 실정이며, 대부분 개인의 실무프로젝트를 진행 한 사례였다. 빅데이터 기반 6시그마 방법론을 전사적으로 도입하여 이를 기업 차원에서 실행한 사례는 LG가 처음이며, LG 역시 도입한 지 얼마 되지 않아 검증과 확립을 위한 실증연구가 필요한 상황이다.

또한 방법론이 성공적으로 정착되기 위해서는 임직원들의 전사원적인 참여가 필요하며, 성공 정책을 잘 따라야지만 성공률을 높일 수 있다. 이를 지키지 않고, 무리한 도입으로 제조기업 중 과반 이상이 6시그마를 중단한 사례는, 성공 정책의 연구에 대한 중요성을 보여준다[4, 30]. 이들 핵심 성공 정책에 대한 연구는 과거부터 방법론의 유효성 분석에서 필수로 다루어져 왔다.

따라서 본 연구자는 빅데이터 기반 6시그마 방법론인 LG DX SS의 효과를 2장에서 객관적으로 입증하고, 빅데이터 분석이 필요한 제조 기업들이 빅데이터 기반 6시그마를 성공적으로 도입하고 활용하여 성과를 창출할 수 있도록, 핵심 성공 정책을 3장에서 문헌 연구를 통해 제시하였다.

## 2. 6시그마와 빅데이터 기반 6시그마의 비교

### 2.1 6시그마 방법론

6시그마 방법론은 제조기업에서 품질수준을 높이고 비용 절감을 하기 위해, 공정이나 프로세스의 불량유발인자를 찾아내고 최적화하는 경영혁신 방법론이다. 6시그마 프로젝트는  $y=f(x)$  관계를 통계적으로 검·추정하고, 실험계획법 등을 이용하여 최적의 해결책을 도출하는 과정으로 진행된다. 6시그마 방법론에는 제조 분야에서 사용되는 DMAIC 실험방법론과, 비제조 분야(간접/사무 부문)에서 사용되는 DIDOV, DMADOV 등의 설계방법론으로 구분된다[9, 13].

Table 1은 6시그마 방법론을 구분한 것이다[9, 13]. 데이터를 확보하여 통계적/정량적 분석을 진행하는 DMAIC 개선 방

Table 1. 6 Sigma Methodology Classification

Division	Areas of application	6 Sigma Application Methodology
Improvement	manufacturing sector	DMAIC :Define-Measure-Analyze-Improve-Control
Design	non-manufacturing sector (office work/ indirect)	DIDOV :Define-Identify-Design-Optimize-Verify DMADOV :Define-Measure-Analyze-Design-Optimize-Verify

법론은 주로 제조 분야에서 활용되며, 일반적으로 데이터를 활용한 실험을 통해 치명인자를 최적화하여 개선하는 통계적 방법론이다. 6시그마 과제 대부분은 데이터를 확보하여 통계적 분석기법을 최대한 많이 활용하는 DMAIC 방법론으로 수행된다.

DFSS(Design For Six Sigma) 방법론은 DIDOV, DMADOV 방법론으로 나뉘지며, 주로 비제조 분야(사무/간접)에서 많이 사용되는 설계방법론이다. 설계 방법론은 통계적 기법은 최소로 사용되되 브레인 스토밍 Tool 등의 정성적 기법을 대부분 사용하므로, 데이터를 확보하여 통계적, 정량적으로 분석하는 DMAIC 방법론보다 활용도가 낮다.

Table 2는 6시그마 DMAIC 방법론의 절차와 통계적 기법에 대해 나타냈다[9, 13]. DMAIC 방법론은 현장 문제를 이론 문제로 전환하여 문제를 정의해 보고, 이론적으로 해결책을 찾으면 다시 현장에 적용하여 개선하는 형태로 진행된다.

Table 2. 6 Sigma DMAIC Methodology Procedures and Techniques

Process	Step	Technique
Define	1. Problem Analysis and Project Selection	- Project registration form
	2. Big Y, Small Y Selection	- Multivariate analysis(PCA, FA, Clustering, etc) - Statistical Hypothesis Test (Correlation/regression Analysis, etc) - SIPOC, FDPM
Measure	3. Measurement System Analysis and Identify current levels	- t test, Gage R&R - Process Capability Analysis(Capability Sixpack) - Control Chart, Run Chart
	4. Big Y, Small Y Identifying the relationship	- Multivariate analysis(PCA, FA, Clustering, etc) - Statistical Hypothesis Test (Correlation/regression Analysis, etc)
	5. Derivation of latent factors	- Statistical Hypothesis Test (Correlation/regression Analysis, etc) - Graph Analysis(Boxplot/Pareto chart)
Analyze	6. Potential factor analysis and critical factor selection	- Statistical Hypothesis Test (t test, proportions test, ANOVA, Chi-Square Test, Correlation/Regression Analysis, etc) - DOE(FD, FFD, etc) - Graph Analysis(Boxplot/Pareto)do) - Interval estimation(CI)
Improve	7. Development of improvement plans	- Brainstorming, TRIZ
	8. Process Optimization	- DOE(SAM, RSM, Taguchi, etc)
Control	9. Confirmation of improvement results	- Process Capability Analysis - t test
	10. Establishment of control plan	- Control Chart ( $\bar{x}$ , $R$ , $S$ , $I$ -mR, np, p, c, u, etc)
	11. Execute the control plan	- Best Practice

Define 과정에서는 먼저 조직의 현업과제에서 핵심적인 문제점이 무엇인지 찾아보고, 공정/프로세스 분석 및 고객 인터뷰 등을 통해 문제를 구체화하고, KPI(Key Performance Indicator)에 영향을 주는 핵심 CTQ(Critical To Quality)를 선정한다. CTQ 선정 시 다변량 분석, 통계적 가설검정 등의 통계적 기법과 SIPOC(Supplier-Input-Process-Output-Customer), FDP(Functional Deployment Process Map) 등 프로세스 분석 기법 등이 활용된다.

Measure 과정에서는 측정시스템 분석을 통해 데이터의 신뢰성을 확보하고, 공정능력분석을 통해 현재 CTQ의 품질수준과 목표 시그마 수준을 설정한 후, 잠재인자를 도출한다. 이 단계에서 t Test, Gage R&R, 프로세스 능력분석 및 관리도, Run Chart 등의 통계적 기법이 활용된다.

Analyze 과정에서는 CTQ와 잠재인자 간 영향성을 통계적 가설 검·추정 및 FD(Fractional Design), FFD(Factorial Fractional Design) 등의 DOE(Design of Experiment)를 통해 확인하며, 잠재인자 중에 가장 영향력이 큰 치명 인자를 도출한다.

Improve 과정에서는 SAM(Steepest Ascent Method) 또는 RSM(Response Surface Methodology), Taguchi 등의 DOE 또는 브레인스토밍 Tool을 통해 치명 인자를 최적화하여 현장에 적용하고, 개선된 CTQ의 품질수준과 KPI 성과를 확인하여 성과를 검증한다.

Control 과정에서는 관리도를 활용하여, 개선공정의 변화 또는 이상유무를 지속 모니터링하고 확인한다[9, 13].

2.2 빅데이터 기반 6시그마 방법론

빅데이터 기반 6시그마 방법론은 기계학습(Machine Learning)

등의 빅데이터 분석기법과 6시그마 방법의 통계 기법을 결합한 방법론이며, 2011년부터 현재까지 연구가 활발하게 진행되고 있다. 빅데이터 기반 6시그마 방법론 연구자별로 사용한 빅데이터 기법과 도구, 과제성과를 Table 3에 요약하였다.

Table 3의 빅데이터 기반 6시그마 방법론별 기법과 성과 연구 사례를 살펴보면, 주로 빅데이터 분석기법의 전처리, 모델링, 평가, 예측 등의 방법을 사용하였으며, 설비 관련 파라미터의 분석에서 주요 인자를 추출하는 데 활용되었다. 주요 인자를 찾은 후, 6시그마의 통계기법으로 최적화 실험을 수행하는 것이 공통점이다. 이 방법론을 통해 연 13.96억 원 비용 절감 효과와 공정효율 및 불량 개선 등 품질향상에도 상당한 효과가 있음을 확인할 수 있었다. 또한 기업의 경영 이슈 등의 사결정 사항에도 해당 방법론이 유효함을 객관적으로 입증하였다.

2.3 LG DX SS 방법론

LG 전자는 1996년 국내 최초로 6시그마를 도입하여 경영 혁신을 추진하였고, 연평균 25%의 성장률과 품질 개선, 비용 절감 등의 성과를 달성하였다[14]. LG는 27년 이상 6시그마를 LG의 문화와 특성에 맞는 고유한 혁신방법론으로 발전시켜 왔다. 최근에는 빅데이터 환경에 적응하기 위해 DX(Digital Transformation), 디지털 전환을 강조하는 분위기가 형성되었으며, 빅데이터 분석과 6시그마가 결합한 DX SS(Digital Transformation Six Sigma) 방법론을 2020년 LG전자에 도입하였고, 2022년에는 LG Display 등 다른 계열사에도 확대하였다[12].

Table 4는 LG DX SS 방법론의 프로세스 및 기법에 대해 나타냈다[12]. DX SS 방법론은 6시그마의 DMAIC 방법론을 빅데이터 분석과 6시그마가 결합한 New DIDOV 방법론으로

Table 3. Big Data-Based 6 Sigma Methodology-specific Techniques and Project Performance

Methodology	Techniques	Project Performance
New DMAIC[5]	- Data Pre-processing - Feature Selection - ML Model Creating and evaluating (RF, Regression analysis, 5-Fold Validation, etc): Use Step DMAIC All Steps	'L's Company solar cell process efficiency max improved by 0.23%
New DMAIC[6-7]	- ML Model Creating and evaluating(RF, CART, NN): Use Step Analyze. Improve	Improving process thermal efficiency due to the 'P' steel company's craze, reducing GAS amount used by 1.396 billion won per year
PSI(Prepare-Search-Improve)[8]	- Data Pre-processing - ML Model Creating and evaluating(Regression, CART, NN, K nearest neighbor, SVM, SVR, GPR, Clustering, Analysis of ssoication rules, etc): Use Stap Prepare, Search	A's Company display B defect detection process is applied, 11% improvement in defect detection power, 87% improvement in defect occurrence rate
DPELR(Define-PreProcess-Explore-Learn-Reinforce)[9,10]	- Data Pre-processing - ML Model Creating and evaluating(GLM, LDA, QDA, DT, NNet, SVM, etc): Use Step PreProcess, Explore, Learn, Reinforce	Perform projects that require big data utilization due to quality issues and corporate management issues
HyDAPI (DMAIC&CRISP-DM combined NEW DMAIC)[11]	- Data Pre-processing - ML Model Creating and evaluating(Evaluation Tool not specifically presented): Use Step Define, Measure, Analyze, Improve	When evaluating the existing methodology, 6 Sigma and CRISP-DM have the highest evaluation scores, suggesting a methodology that combines the two methodologies.

재정의하였다. DMAIC 방법론을 New DIDOV 방법론으로 표현한 이유는, 빅데이터 분석기법이 추가되면서 기존 6시그마 프로세스인 DMAIC의 기존 프로세스 개선 절차보다 설계적 개념이 중요시되었기 때문이다. Define 과정, Optimize 과정, Validation 과정은, 기존 6시그마 DMAIC의 Define 과정, Improve 과정, Control 과정과 동일하며, 6시그마 기법과 동일한 통계적 기법을 적용한다. 여기에 Identify 과정, Design 과정에서 빅데이터 분석기법을 추가하여 집중적으로 다룬다.

Identify 과정에서는 데이터 전처리, 정규화, 밸런싱, Feature Selection 기법을 사용하여 Data Set을 확정한다.

Design 단계에서는 여러 회귀/분류/군집 모델들을 활용하여 Model 성능을 비교 평가하고, 최적 성능을 가진 모델을 선정하여 Feature Importance를 통해 중요 변수를 추출하고, 하이퍼 파라미터 최적화를 통해 성능을 향상한다.

Fig. 2는 Six Sigma 방법론과 LG DX SS 방법론의 절차를 비교하였다. 빅데이터 기반 6시그마 방법론인 LG DX SS 방법론으로 진화하면서 6시그마 DMAIC 방법론에서 변경된 부분은, Measure와 Analyze 과정이다.

Measure 과정에서 측정시스템 분석(MSA, Measurement Systems Analysis)과 데이터 수집, 공정능력분석 및 잠재인자를 도출하는 단계로 구성되어 있으며, Analyze 과정에서는 통계적 가설검정 및 실험계획법을 통해 잠재인자 중 가장 영향력이 큰 치명 인자를 도출한다. LG DX SS는 이 두 과정 대신 Identify 과정과 Design 과정이 추가되었다. 이 두 과정에서는 빅데이터 분석기법만을 활용하며, 전수 데이터를 활용한 데이터 수집, 전처리 및 기계학습 초기 모델링, 최적 모델링을 통해 성능이 뛰어난 회귀/분류/예측 모델을 생성한다. 6시그마에서는 불량유발인자가 있을 것으로 추정되는 일부 데이터 영

Table 4. Process and Techniques of LG DX SS Methodology

Process	Step	Technique
Define	1. Confirmation of improvement opportunities	*
	2. CTQ Selection	- Multivariate analysis (PCA, FA, Clustering, etc) - Statistical Hypothesis Test (Correlation/regression Analysis, etc) - SIPOC, FDP - Process Capability Analysis (Capability Sixpack)
Identify	3. Dataset Confirmation	- Data Pre-processing, Normalization, balancing (Y discrete), Feature Selection
Design	4. Initial Modeling	- Modeling Comparative evaluation (Clustering Model K-Means, etc 8 Model, Regression Model CART, etc 16 Model, Classification Model NNet, etc 14 Model)
	5. Optimal Modeling	- Prediction/classification Performance Assessment(R Square, MSE, Accuracy/specificity/sensitivity/precision, ROC curve, etc) - Feature Importance - Optimizing Hyperparameters
Optimize	6. Derive optimal conditions	- Simulation DOE (FD, FFD, SAM, RSM, Taguchi, etc)
Validation	7. Actual Validation	- Pilot Test - Process Capability Analysis - t test
	8. Process Capability Analysis	- Multivariate chart

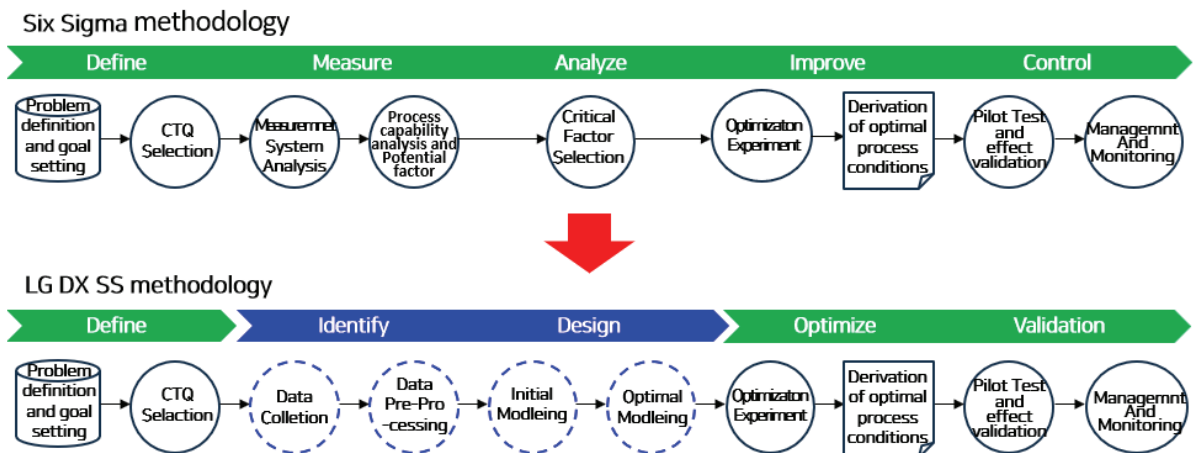


Fig. 2. Comparison of 6 Sigma and LG DX SS Methodology Processes Procedures

역만 사용하는 반면, LG DX SS는 전수 데이터를 사용하므로, 불량 개선 및 최적 공정 조건을 도출하는데 6시그마 방법론보다 효과가 탁월하다. LG DX SS 과정별 절차에 대해 상세히 기술하면 다음과 같다.

Define 과정에서는 기존 6시그마 DMAIC 방법론의 Define 과정과 동일하게 문제/이슈를 정의하고, 기업 경영목표와 고객 요구사항과 연계된 Project Y를 선정한다. SMART 관점(Specific, Measurable, Attainable, Relevant, Time Bound)에서 Big Y와 Little Y를 전개하고, 현재 수준과 목표 수준의 Gap 차이를 확인한다. 문제 현상과 메커니즘 분석을 통해 과제 범위와 관련된 변수들을 확인하고, CTQ와 연관지표를 선정한다.

Identify 과정은 빅데이터 수집 및 전처리를 하는 단계로, 프로젝트의 시간과 노력이 가장 많이 소요된다. 빅데이터는 정형화되지 않은 데이터가 사용되는 경우도 있어, 이를 분석 가능하게 하는 전처리 과정이 필수이다. 데이터 전처리란, 데이터를 사용하기에 앞서 원자료의 정보손실이나 왜곡이 없는 범위에서 데이터를 분석할 수 있는 형태로 만드는 것을 말하며, 파이썬, R 등의 프로그램을 이용하여 데이터 전처리를 수행한다. 이 단계에서 데이터의 품질이 평가되는데, 데이터 품질은 양적, 질적 변수들을 모두 포함하는 데이터의 전반적인 상태를 의미하며, 데이터가 분석 목적에 적합한 정도가 높을수록, 데이터의 품질은 높아진다. 데이터 전처리 방법은 데이터 정제, 통합, 변환, 축소 등의 방법으로 구분된다. 필수과정은 아니지만, 데이터의 스케일을 조정하여 분석 모델링에 적합한 형태로 만들어, 분석의 효율성과 정확도를 높이는 데이터 정규화 과정도 선택적으로 수행할 수 있다. 데이터 전처리 과정이 완료되면, Y 변수가 암환자를 분류하는 양성/음성과 같은 이산형이고, 데이터의 개수가 불균형 상황일 때 데이터 밸런싱 작업을 선택적으로 수행할 수 있고, 원본 데이터에서 불필요한 변수들을 제거하는 Feature Selection 과정을 거친다.

Design 과정에서는 Data Set의 주어진 변수로 초기 기계학습, 딥러닝 모델을 만들고, 여러 모델의 성능을 비교하고 평가하여, 가장 적합한 모델을 선택한다. 빅데이터를 활용하는 기계학습 방법에서는, 전체 데이터를 학습 데이터와 테스트 데이터로 나누어 모델의 성능을 확인한다. 일반적으로 7:3 또는 8:2의 비율로 나눈다. 학습 데이터로 모델을 학습시키고, 테스트 데이터로 올바르게 학습이 수행되었는지 평가한다. 만약 모델의 성능이 부족하다면, 새로운 방법 또는 기준을 적용하여 다시 학습하고 평가한다. R-Square, MSE, 정확도, 특이도, 민감도, 정밀도, ROC, AUC 등의 주요 평가지표를 사용하여 가장 성능이 뛰어난 모델을 선택하고, 유의변수를 찾기 위한 Feature Importance 과정을 거친 후, 하이퍼 파라미터값을 최적화하여 예측 모델을 평가하고, 개선한다. 모델의 과적합(Over Fitting)을 방지하기 위해서는 최종 분석에 사용되는 변

수의 수를 줄이고, 정규화(Regularization)나 교차 검증(Cross Validation)을 해야 한다.

Optimize 과정은 Design 과정에서 선정한 유의변수로 시뮬레이션 기반의 DOE를 수행하여 유의변수를 최적화하고, 결괏값을 예측한다. 이 단계는 6시그마 DMAIC 방법론의 Improve 과정과 유사하다.

Validation 과정에서는 Pilot Test를 통해 Optimize 과정에서 예측한 결괏값의 재현성을 검증하고, CTQ와 KPI의 개선 효과를 확인한다. 이후 관리계획을 수립하고 실행하여, 개선된 프로세스를 표준화하고, 관리, 모니터링하는 시스템을 구축한다. 이를 통해 프로젝트의 결과가 현장에 정착되고, 지속해 개선성고가 유지될 수 있도록 한다. 이 단계는 6시그마 DMAIC 방법론의 Control 과정과 유사하다.

DX SS 방법론은 기존의 6시그마 방법론의 통계적 분석기법과 빅데이터 분석 기법의 다양한 기계학습 알고리즘을 Modeling하고 비교, 평가하여 최적의 Modeling을 적용하고, 주요 인자를 최적화한다. 전수 데이터를 활용하여 불량유발인자를 찾아내고 개선하므로, 데이터 분석 시간을 많이 줄여주며, 품질향상과 경비 절감에 기여한다[12].

LG의 DX SS(Digital Transformation Six Sigma)의 Belt 체계는 DX MBB(DX Master Black Belt), MBB(Master Black Belt), DX BB(DX Black Belt), GB(Green Belt)로 구성된다. 이 중 GB는 가장 초급단계의 Belt로, 데이터를 사용하여 기초적인 통계분석 방법과 실험방법을 이해하여 간단한 문제를 해결할 수 있는 리더역할 또는 BB 과제의 팀원으로 참여한다.

BB는 GB의 인증과제를 지도하고 육성하며, 다양한 실무과제를 개선하기 위한 리더로 활동하며, MBB나 DX MBB 과제의 팀원으로 과제수행에 참여한다.

MBB는 DX Six Sigma 도입 전에는 최고등급의 Belt였으며, BB와 GB 인증과제를 지도하고 육성하며, 조직 내에 가장 큰 Issue를 해결하기 위한 중요도가 높은 과제를 리더로 수행한다. 기존 MBB의 빅데이터 분석 역량을 강화하기 위해 DX Six Sigma가 도입되었으며, DX MBB는 6시그마 분석역량과 빅데이터 분석역량을 모두 갖춘 문제해결 전문가로 조직에서 가장 핵심적인 역할을 수행한다. MBB와 DX MBB의 차이는 빅데이터 기반 6시그마를 사용할 수 있는지이다[12].

Table 5와 Fig. 3, 4는 LG Display DX MBB 30인의 인증과제성적을 요약한 것이다. Table 5에서는 공정명, KPI, 단위, 개선 전/개선 후 품질수준 및 품질 GAP을 표현하여 DX SS의 적용공정 성과를 나타냈다. Fig. 3에서는 DX SS 적용공정 성과의 품질수준의 범위가 0부터 10까지인 16개의 과제를 나타냈으며, Fig. 4에서는 품질수준의 범위가 10부터 100까지인 14개 과제로 구분하였다. 스케일을 두 그룹으로 나누어, 과제의 개선 전/개선 후의 값이 잘 보이도록 표현하였다.

제조공정에서의 불량률 또는 불량수준 등을 최소화하기 위한 망소특성 과제가 24개이며, 검사판정 적합성이나 설계정



Table. 5. DX SS Applied Process Performance

NO	Process name	KPI	Unit	As-is	To-Be	GAP
1	Improvement '@' defect in customer process	defect rate	%	0.58	0	0.58
2	Improvement of 'B' defect on model 'A' screen	defect rate	%	0.7	0	0.7
3	Improvement of model 'A' gap defect	gap defect rate	%	0.6	0	0.6
4	Improvement of 00 inches of stain	defect rate	%	8.1	0	8.1
5	Improvement of 'C' defect in 00 inches	defect rate	%	3.75	0	3.75
6	Improvement of MNT defect leakage rate of Model 'B'	defect leakage rate	%	1.21	0.72	0.49
7	Improvement of 'D' defect in Model 'C'	defect rate	%	0.41	0.0014	0.4086
8	Improvement of MNT defect level on model 'C'	defect level	Level	1.85	0.65	1.2
9	Improvement of 'E' defect in high-resolution models	defect rate	%	0.2	0	0.2
10	3 connections on 00 inch MNT models(Improvement of Damage)	defect rate	%	1.37	0.86	0.51
11	Curved MNT 'F' defect improvement of model 'D'	defect rate	%	0.81	0.51	0.3
12	MNT 'G' defect improvement on model 'E'	defect rate	%	8.5	5.3	3.2
13	Improvement of defective leakage of model 'F'	inspec-tion defect rate	%	6.1	0.8	5.3
14	Improvement of 00 inch 'G' defect	defect rate	%	0.4	0	0.4
15	Improvement of 00 inch 'H' defect	defect rate	%	6.4	0	6.4
16	Improve accuracy of inspection robots	Number of defects	number of times	1.1	1	0.1
17	Improve the acceptance rate of the GB theoretical test in the organization	the rate of accept-ance rate	%	15	100	85
18	Stain improvement of membrane applied model 'G'	stain defect rate	%	100	3.3	96.7
19	Oxidation improvement challenges for model 'H'	defect rate	DPU	17.9	2.8	15.1
20	Securing the quality of Model 'I'	defect rate	%	15.4	0	15.4
21	Securing automatic inspection detection power of inspection equipment	inspec-tion detection power	%	39.9	73.1	33.2
22	Stain Improvement of High Shield Application Model	stain occur-ence rate	%	50	0	50
23	Improvement of foreign substance defects in Model 'J'	defect rate	%	30	6.6	23.4
24	Improvement of 'K' quality of model 'K'	success rate	%	81.9	84.6	2.7
25	Establishment of a system for predicting remaining capacity	predic-tion error late	%	100	1	99
26	Automate optimal tolerance design based on AI technology	design consis-tency	%	96	99.8	3.8
27	Development of floor number and area prediction models	predic-tive consis-tency	%	86	94	8
28	Improve mobility failure rates in 00 inches	Mobility defect rate	%	30	2	28
29	Field 'k' Defective Abnormal Detection System Built	predic-tive consis-tency	%	0	70	70
30	Optimizing an image inspection pattern	Inspec-tion pattern time	sec	23.4	21	2.4

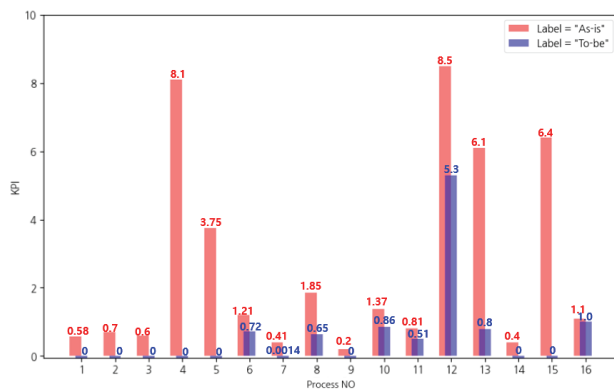


Fig. 3. DX SS Application Process Performance [NO. 1-16]

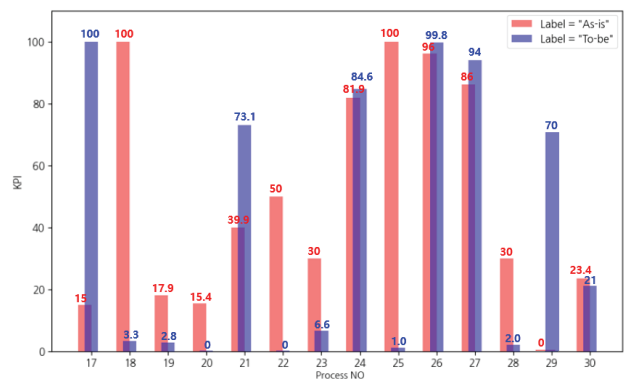


Fig. 4. DX SS Application Process Performance [NO. 17-30]

합성, 예측 정확성 등을 최대화시키기 위한 망대특성 과제가 6개로 구분된다. 적용 공정별로는 기존 공정에 적용한 과제 25개와, 신규 공정(개발 및 대체물질/모델 변경 적용)에 적용한 과제 5개로 분류된다.

LCD나, OLED 생산공정에서 발생할 수 있는 불량은, 대표적으로 이물 불량, 굽힘 불량, Encap 불량, 증착 Align 불량, Pattern 불량, 이물 불량, LTPS 특성 불량, Pixel 수축 불량, Dark Spot 불량, 미 발광 Pixel 불량 등이 있는데, 실제로는 디스플레이 Size, Model 별로 제조공법에 차이가 있으므로, 이외에도 수많은 불량 유형이 있다[15]. 그동안 기존 6시그마 방법론은 일부 데이터 영역만 분석했지만, 빅데이터 기반 6시그마를 활용한 DX MBB 30인의 과제에서는 전수 데이터를 활용하여 불량유발인자를 찾아내고 최적화하여, DX 과제에서 개선한 불량 유형만 15종에 달할 정도로, 여러 불량유형을 개선하였다. 일부 과제에서는 검사 시간과 제조공정에 필요한 유기물 잔량을 최소화하여 생산시간과 제조 비용을 절감하였다.

과제 수행성과를 종합해 보면 품질개선 전/후 Gap 수치가 평균 18.83으로 크게 개선되었으며, 신규공정(개발 및 대체물질/모델 변경 적용)에 적용된 과제에서는 Gap 수치가 평균 53으로 높은 품질개선 효과를 보였다. 빅데이터 기반 6시그마 방법론 사용 시, 불량을 감소 등 품질개선에 상당한 효과가 있음을 과제 수행 성과를 통해 확인할 수 있다.

2.4 6시그마 방법론과 빅데이터 기반 6시그마 방법론 비교

6시그마 방법론과 빅데이터 기반 6시그마 방법론은 빅데이터 분석기법이 추가된 점이 가장 큰 차이점이다. Table 6은 두 방법론의 차이점이 잘 나타날 수 있도록 요약하였다[5-12].

6시그마의 전처리 과정은 데이터의 분포 적합도로 검증하며, 일반적으로 정규성검정 또는 관리도로 이상치와 결측치를 추출하고, 제거하거나, 대체 값을 적용한다. 이에 반해 빅데이터 기반 6시그마는 전처리 과정이 굉장히 체계적이다. 데이터 정제/통합/축소/변환 등을 통해 이상치와 결측치를 추출, 제

거, 대체하며, 불필요한 변수를 제거하고, 유의변수만 분석하는 것이 모델의 성능을 향상하는 데 도움이 되기 때문에, Feature Selection 등을 통해 예측과 관련이 높은 변수들을 선택하고, 적정 수로 축소하는 과정도 거친다. 추가로 필수는 아니지만, 데이터의 스케일을 조정하여 분석 모델링에 적합한 형태로 만들어 분석의 효율성과 정확도를 높이는 데이터 정규화 과정도 수행할 수 있다.

6시그마는 일부 잠재인자와 치명 인자만 개선 대상으로 한다. 빅데이터 기반 6시그마는 모든 인자를 대상으로 하고 있어, 불량유발인자를 모두 찾아내 개선할 수 있다는 점이 장점이다.

6시그마 방법론은 사전에 엔지니어의 경험과 지식 도메인 내에서 문제의 잠재인자가 있을 것으로 판단되는 데이터 영역만 사용하므로, 해당 영역에 불량유발인자가 없으면, 품질이 개선되지 않을 가능성이 존재한다. 반면 빅데이터 기반 6시그마 방법론은 전수 데이터를 대상으로 분석을 진행하므로, 모든 불량인자를 찾아낼 수 있으며, 실험계획법을 진행하여 유의변수에 대한 최적화까지 가능해 높은 품질개선 효과를 기대할 수 있다.

6시그마는 미니탭, JMP 등의 통계 프로그램을 이용하여 데이터 분석을 진행하는데, 분석하는 인자의 수가 100개 이상만 되더라도 분석이 어려울 정도로 처리속도가 느려지고, 결과장표마저 복잡해지는 단점이 있다. 반면에 빅데이터 기반 6시그마의 경우, 파이썬과 R 등의 프로그램을 사용으로 데이터 분석을 하므로, 모든 인자의 분석이 가능하며, 결과값도 요약된 형태로 확인할 수 있다.

6시그마는 상관분석/회귀분석/통계적 가설 검·추정/DOE 등을 통해 명확한 수식으로  $y=f(x)$  관계를 정의할 수 있다는 점에서 높은 해석력을 가지지만, 빅데이터 분석은 인공지능망과 같은 주요 기계학습 알고리즘이 추적이 불가능한 블랙박스(Black Box) 영역이 있다는 점이 약점이다. 빅데이터 기반 6시그마 방법론은 모델링하여 성능을 비교, 평가, 예측하는 절차에서는 기계학습 알고리즘이 사용되지만, 유의변수에

Table 6. Comparison of 6 Sigma Methodology with Big Data-based 6 Sigma Methodology

Division	6 sigma methodology	big data-based 6 sigma methodology
Data Collection	Partial Data(Sample)	Total Data(Population)
Pre-proce-ssing process	Detection of outliers with normality tests or control charts	Pre-processing process is organized to account for more than 80% of the data analysis process(data purification/integration/reduction/transformation, feature selection, etc), data normalization
Subject to improvement	Potential factor, Critical factor	Total factor
Characteristic	High interpretation power through clear definition of analysis group and application of statistical model	Machine learning and application of statistical models provide high predictive accuracy and interpretability
Use Program	Minitab, JMP, etc	Minitab, JMP, Python, R, etc
Analysis Tool	Correlation/regression Analysis, Statistical Hypothesis Test, DOE(FD, FFD, SAM, RSM Taguchi, etc)	Machine Learning(K-Means, CART, SVM, Nnet etc 27+ utilization models, Feature Importance, Correlation/regression Analysis, Statistical Hypothesis Test, DOE



대한 최적화는 설명이 가능한 실험계획법을 통해 진행되므로, 높은 예측 정확도와 해석력을 동시에 가질 수 있다는 점이 장점이다.

2.5 빅데이터 기반 6시그마 방법론과 DX SS 방법론 비교

빅데이터 기반 6시그마의 국내외 연구자 7명의 논문과 LG DX SS 방법론을 비교해 보면, 개선영역에 대한 전수 데이터를 수집하여, 데이터 전처리, 모델링, 예측하고 유의변수를 찾아내 DOE를 통해 최적화하는 과정은 동일하였다. 하지만 알

고리즘 측면에서는 LG DX SS 방법론이 알고리즘개수가 더 많으며, 체계적으로 구분하고 있음을 알 수 있다[5-12].

Table 7은 LG DX SS 방법론의 Machine Learning 비교평가 Model을 구분하였으며, 44개 알고리즘으로 구성되어 있다. Clustering Model은 7개, Regression Model은 16개, Classification Model은 14개, Association Rule Model 4개, Time Series Model 3개로 구분된다[12]. (Clustering / Regression / Classification Model 간 중복으로 사용되는 모델은 \* 표시함)

Table 7. Comparative Evaluation Model Classification of LG DX SS Methodology

Division	Analysis algorithm	Description
Clustering	Hierarchical	Cluster analysis that integrates and separates data based on how hierarchical decomposition is formed
	Clustering Using Representatives	Clustering using representative values by performing an analysis of the object's "association"
	K-Means	Algorithm to group given data into k clusters
	Self-Organizing maps	Neural network-based cluster analysis
	DBSCAN	Cluster formation, connecting neighborhood radius and minimum neighborhood-based data density points
	Gaussian mixture models	Cluster analysis assuming that the distribution of data consists of the weighted sum of k normal distributions
	Spectral Clustering	Reduce dimensions by utilizing the eigenvalues of the distance matrix between each point, and cluster analysis based on this
Regression	Multivariate Linear Mode	Assuming that the relationship between the dependent variable and two or more independent variables is a linear relationship, and the model is fitted using the least squares method.
	Principle Component Regression	Principal component analysis is a way to reduce high-dimensional data that is likely to be associated with each other to low-dimensional data
	Partial Least Square	It is similar to PCR, but there is a difference in considering the correlation with the dependent variable when it is reduced to a lower dimension
	LASSO	LASSO regression, like MLR, utilizes the least squares method, but imposes constraints on the sum of the absolute values of the regression coefficients
	Ridge	Similar to LASSO, but imposes constraints on the sum of regression coefficient squared values
	Elastic Net	A model containing LASSO and Ridge that constraints the sum of absolute values of regression coefficients and the sum of squares weighted sum of squares
	Non-Linear Regression	NLR assumes that the relationship between dependent and independent variables is nonlinear
	*K Nearest Neighbor	k-NN is a nonparametric method used for classification or regression
	*Logistic Regression	LR is a special case of GLM, which expresses the class into which individual observations are classified as a probability using a logistic function.
	*Neural Network	A neural network is a combination of synapses, a model in which artificial neurons (nodes) that form a network have problem-solving capabilities through learning
	*Deep Learning	Deep learning is machine learning that attempts a high level of abstraction through a combination of several nonlinear transformation techniques
	*SVM	The SVM algorithm finds the most wide of the two classes distinguishable boundaries
	*CART	Create a decision tree by splitting subsets of datasets to create two child nodes over and over again using all independent variables
	*Bagging	Representative algorithms incorporating multiple models designed to improve stability and accuracy
	*Boosting	An algorithm that combines models with weak predictive accuracy into strong models
	*Random Forest	Similar to the Bagging method, but characterized by random selection of variables when fitting a model

Table 7. (Continued)

Division	Analysis algorithm	Description
Classification	GLM	How to make the relationship between the dependent variable and the independent variable into a linear relationship after changing the dependent variable through the link function
	Naive Bayes	Assume independence between independent variables and calculate and classify the posterior probability to be included in each class applied with Bayes Theorem
	Bayesian Network	Use a set of random variables and a directional acyclic graph to express a set of variables as conditional independence
	*K Nearest Neighbor	k-NN is a nonparametric method used for classification or regression
	*Logistic Regression	LR is a special case of GLM, which expresses the class into which individual observations are classified as a probability using a logistic function.
	*Neural Network	A neural network is a combination of synapses, a model in which artificial neurons (nodes) that form a network have problem-solving capabilities through learning
	*Deep Learning	Deep learning is machine learning that attempts a high level of abstraction through a combination of several nonlinear transformation techniques
	Linear Discriminant Analysis	LDA assumes the normality and equal variance of the distribution of independent variables of each class, and then finds a linear line that is a hyperplane that can divide the class well
	Quadratic Discriminant Analysis	Unlike LDA, QDA assumes that each class has a different covariance matrix and appears in the form of a quadratic function
	*SVM	The SVM algorithm finds the most wide of the two classes distinguishable boundaries
	*CART	Create a decision tree by splitting subsets of datasets to create two child nodes over and over again using all independent variables
	*Bagging	Representative algorithms incorporating multiple models designed to improve stability and accuracy
	*Boosting	An algorithm that combines models with weak predictive accuracy into strong models
*Random Forest	Similar to the Bagging method, but characterized by random selection of variables when fitting a model	
Association Rule	Collaborative Filtering	Analyzing users' past behaviors or preferences to provide customized services
	Contents-Based Filtering	Recommend new items based on the characteristics of items that users have shown interest in in the past
	Knowledge-based Filtering	Provides recommendations based on the user's explicit requirements or preferences
	Hybrid Filtering	Combine several techniques such as collaborative filtering, contents-based filtering, and knowledge-based filtering to build a hybrid system that combines each advantage
Time Series	Auto Regression	Prediction of current time series values by a linear regression combination of past observations of time series data
	Moving Average	Calculate the average over a specific period of time from time series data to reduce short-term volatility and identify long-term trends
	Auto Regressive Integrated Moving Average	Model complex patterns of time series data by combining three components: automatic regression (AR), integration (I), and moving average (MA)

Table 8에서는 빅데이터 기반 6시그마 방법론 연구자 7인의 평가모델 개수와, LG DX SS 방법론의 비교평가 모델 개수 및 변수 중요도 분석 사용 유무를 비교하였다. Model 44개 중 중복되는 모델의 수를 제외하면, 34개의 모델이 사용되었으며, 다른 빅데이터 기반 6시그마 방법론에서 사용되고 있는 3개~11개의 알고리즘 보다 모델 평가개수가 많고, 체계적이다. DX SS 방법론은 초기 Modeling 과정에서 다수의 군집/회귀/분류 알고리즘간의 성능을 비교 평가하여, 가장 성능이 뛰어

난 알고리즘을 최종 선택하여 모델링하고, 하이퍼 파라미터 최적화를 진행하므로, 분석에 평가할 모델이 많을수록 최적 모델을 선정하는 데 유리하다. 또한 다양한 모델 비교를 통해, 어떤 알고리즘이 적합한지 비교할 수 있으며, 훈련 데이터에 과적합 될 가능성도 방지할 수 있고, 성능이 더 좋은 모델을 선택할 수 있다. 여러 모델을 비교 평가하는 것으로 신뢰성이 확보되는 강점도 있다.

LG DX SS 방법론은 Model 비교평가 이후, 최종 선정된 모

Table 8. Comparison of the Number of Evaluation Models &amp; Feature Importance Analysis by Big Data-based 6 Sigma Methodology

Methodology Reference	Use Model	Number of evaluation models	Use Model Feature Importance analysis
New DMAIC [5]	Logistic Regression, Lasso Regression, Random Forest, SVM, GLM, Neural Network	6	0
New DMAIC [6,7]	Polynomial Regression, CART, Neural Network	3	0
PSI(Prepare-Search-Improve) [8]	Logistic Regression, CART, Neural Network, K Nearest Neighbor, SVM, Polynomial Regression, SVR, GPR, K-Means, Hierarchical, A-Priori	11	0
DPELR(Define-PreProcess-Explore-Learn-Reinforce) [9,10]	GLM, LDA, QDA, DT, NNet, SVM.L, SVM.P, SVM.R	8	0
HyDAPI(DMAIC&CRISP-DM combined NEWDMAIC) [11]	Data Pre-processing, ML Model Creating and evaluating(Evaluation Tool not specifically presented): Use Step Define, Measure, Analyze, Improve	-	-

델로 변수 중요도를 분석하여, 유의변수를 추출한다. 변수 중요도는, 종속변수에 미치는 독립변수의 영향을 나타내는 지표이다. Feature Importance를 사용하는 이유는, 데이터의 차원을 축소하고, 모델의 복잡도를 줄여 모델의 성능을 향상하고 과적합을 방지하기 위함이다.

따라서, 모델이 어떤 기준으로 변수의 중요도를 계산하는지 이해하고, 설명할 수 있어야 한다. 랜덤포레스트의 경우, MSE는 Mean Squared Error로 회귀모델의 오차를 측정하는 지표로 작을수록 좋다. incMSE가 크다는 것은, 해당 변수가 제거되면 MSE가 많이 증가한다는 것으로, 변수의 중요도가 높다고 해석할 수 있다. incNodePurity는 노드 순도의 증가량을 의미하는 것으로, 의사결정나무에서 노드를 분할할 때 사용하는 기준으로, 높을수록 좋다. 따라서 incNodePurity 값이 크다는 것은, 해당 변수가 노드를 분할하는데 많은 기여를 하는 것으로, 변수의 중요도가 높다고 해석할 수 있다.

SVM Model은 변수 중요도를 판단하는 기준으로, SVM 계수(coefficient)를 사용하는데, SVM 계수는 각 특성이 결정 경계에 미치는 영향력을 나타내며, 계수의 크기가 클수록 중요한 특성을 가진다[12].

Feature Importance는 Model의 성능 향상을 위한 필수과제이다. LG의 DX SS 방법론은 다양한 알고리즘의 Feature Importance 통계량을 활용하지만, Table 8의 다른 빅데이터 기반 6시그마 방법론에서는, 변수 중요도 분석을 특정 통계량에 의존하고 있어, 변수 중요도 판단의 다양성이 부족하다.

### 3. 빅데이터 기반 6시그마 방법론의 핵심 성공 정책

#### 3.1 빅데이터 기반 6시그마 방법론의 핵심 성공 정책 연구목적

앞서 최근 빅데이터 기반 6시그마 방법론의 7건의 연구와 LG DX SS 프로젝트 30건의 사례를 통해, 빅데이터 기반 6시그마가 품질 향상과 비용 절감에 효과적임을 객관적으로 입증하였다. 이러한 방법론의 유효성에도 불구하고, 기업에 도입하여 성공률을 높이기 위해서는, 방법론의 성공 정책을 잘 따라야 하며, 전사적 차원에서 임직원의 노력이 수반되어야 한

다. 기업에서 사용하는 경영혁신방법론이 아무리 효과적인 방법론이라 하더라도, 사용하는 임직원이 제대로 활용하지 못하면 성과를 기대하기 어렵고, 방법론이 지속되기 어렵다. 1998년부터 2022년까지 연구된 6시그마 관련 논문 342개를 분석한 결과, 이 중 67개의 논문이 핵심 성공 정책에 대한 연구였다. 이는 전체 6시그마 논문의 20%의 비중을 차지한다.

선행연구사례에 따르면, 6시그마의 주요 실패 원인으로는, 6시그마 운동을 실시한 CEO가 퇴임하면서 중단된 사례, 사전 분석과 준비 없이 무리하게 도입한 사례, 전사원적 참여를 끌어내지 못한 사례, 통계 접근이 너무 어렵다는 점 등이 있다. 또한 프로젝트 선정 시 쉬운 것만 선택하는 문제, 프로젝트 추진 절차와 기법 등에 대해 잘못 이해하고 진행하거나, 준비시간이 부족한 사례, 데이터의 허구성 등이 대표적이었다. 잘못된 6시그마 적용으로 인해, 6시그마를 도입한 제조기업 중 과반 이상이 6시그마를 중단한 사례도 있었다[4, 30]. 이는 방법론에 대한 성공 정책에 대한 연구 없이 무리한 도입 탓에 실패한 것이 주요 원인이다.

기업에서 방법론이 성공하기 위해서는, 방법론의 성공 정책을 전사원이 따라야 하며, 방법론을 수행했을 때, 성과가 효과적으로 나올 것이라는 확신이 있어야 한다. 이에 방법론에 대한 유효성 검증을 위해서는, 방법론이 성공하기 위한 성공 정책에 대한 연구가 필수적이다.

이에 본 연구자는 기업에 방법론이 효과적으로 도입 및 정착되고, 지속될 수 있도록, 방법론이 성공하기 위한 핵심 성공 정책을 제시하고자 한다.

#### 3.2 빅데이터 기반 6시그마 방법론의 핵심 성공 정책

아직 빅데이터 기반 6시그마 방법론의 성공 정책은, 연구된 사례가 없다. 방법론이 성공적으로 수행되고, 정착되어 기업에서 안정적인 품질 개선 및 비용절감 등 성과를 내고, 업무 효율성 향상을 기대하기 위해, 본 연구자는 Table 9에 6시그마의 성공 정책과, 빅데이터 분석의 성공 정책 간의 공통성공 정책을 도출하여, 빅데이터 기반 6시그마 방법론의 성공 정책으로 제시하였다.

Table 9. 6 Sigma, Big Data Analysis Success Policy

Common Policy	6 Sigma Success Policy	Big Data Analysis Success Policy
Top management support [16-29]	<ul style="list-style-type: none"> <li>- CEO's attention and leadership</li> <li>- Reflecting the establishment of management strategies</li> <li>- Strategy understanding and effort</li> <li>- The drive of a dedicated organization</li> <li>- Tolerance for failure</li> <li>- 6 Sigma Promotion Strategy</li> <li>- Linkage with strategy</li> <li>- Operation of the promotion organization</li> <li>- availability of financial resources</li> </ul>	<ul style="list-style-type: none"> <li>- Top management participation</li> <li>- Strategic alignment</li> <li>- Manager's ability</li> <li>- Senior management support</li> <li>- The promise and dedication of top management</li> <li>- Budget</li> </ul>
Process Improvement [19, 22, 25-26, 29]	<ul style="list-style-type: none"> <li>- Improving the Process of Scientific Approach</li> <li>- Integrating and standardizing tasks</li> <li>- Correct Roadmap</li> <li>- Business Process Framework</li> <li>- Standardizing tasks</li> </ul>	<ul style="list-style-type: none"> <li>- Compatibility with organizational processes</li> <li>- A strong business case</li> <li>- Sponsorship and outsourcing</li> </ul>
Education and training [16, 18-23]	<ul style="list-style-type: none"> <li>- Education and training</li> <li>- Education at all levels</li> <li>- Propulsion Leader Training</li> <li>- Human Resources Development through Education and Training</li> <li>- Securing excellent human resources</li> </ul>	<ul style="list-style-type: none"> <li>- Interdisciplinary learning</li> <li>- Methodology education</li> <li>- Education</li> <li>- Reinforce Ability</li> </ul>
Education System [17, 19, 21-26, 28]	<ul style="list-style-type: none"> <li>- Support system</li> <li>- Companywide improvement activities</li> <li>- All employees participate</li> <li>- Systematic system construction</li> <li>- Support Infra</li> <li>- How to operate</li> <li>- Support information system</li> <li>- Linkage of 6 sigma and institutional strategy</li> <li>- Monitoring and Control</li> </ul>	<ul style="list-style-type: none"> <li>- Structure and capabilities for implementation</li> <li>- Personal information management</li> <li>- External pressure management</li> <li>- Effective project management methodology</li> <li>- Continuous monitoring and control</li> <li>- Monitoring and Feedback</li> <li>- Privacy and Security</li> </ul>
Human and skill [17, 20-29]	<ul style="list-style-type: none"> <li>- Level and capability of belt</li> <li>- Use of statistical techniques</li> <li>- Problem-solving ability</li> <li>- Cooperation of black belt and partner</li> <li>- Supplier's participation</li> <li>- Linkage with human resources</li> <li>- Connection with suppliers</li> <li>- Connection with employees</li> <li>- Understanding 6 sigma tools and techniques</li> <li>- Understanding the 6 Sigma Methodology</li> <li>- Role of middle administrator</li> <li>- Project Promoter</li> <li>- Utilization of 6 Sigma Improvement Techniques</li> <li>- Understanding Tools and Technologies</li> </ul>	<ul style="list-style-type: none"> <li>- Overcoming the talent gap</li> <li>- Stakeholder engagement</li> <li>- A skilled skill</li> <li>- Availability of analytics experts</li> <li>- Recognition of project problems</li> <li>- Effective project management technology</li> <li>- A moderate skill</li> <li>- Human Management</li> <li>- A highly trained sponsor</li> <li>- User Group</li> <li>- A business analyst</li> <li>- A data analyst</li> <li>- Data Management Specialist</li> <li>- Project Manager</li> <li>- An appropriate skill</li> <li>- Database management</li> <li>- Data mining application</li> <li>- Data mining skill</li> <li>- Software for Project</li> <li>- Service Type</li> <li>- Interpretation of results</li> <li>- Leveraging the knowledge you discover</li> <li>- An experienced employee</li> <li>- Organizational capability</li> <li>- Technical gap analysis</li> <li>- Technical Tool</li> <li>- Analysis Platform</li> <li>- Search technology</li> <li>- Reporting and visualization skill</li> <li>- Software flexibility/scalability</li> </ul>

Table 9. (Continued)

Common Policy	6 Sigma Success Policy	Big Data Analysis Success Policy
Time [18, 22, 26-27]	<ul style="list-style-type: none"> <li>- Sufficient preparation time</li> <li>- Continuous Push</li> <li>- Preparation time</li> </ul>	<ul style="list-style-type: none"> <li>- Realistic schedule and timing</li> <li>- Time</li> </ul>
Customer and quality-focused thinking [18-24, 26-29]	<ul style="list-style-type: none"> <li>- Reflect customer needs</li> <li>- Understanding Customer Requirements and satisfaction</li> <li>- Quality management</li> <li>- Customer-focused and market-focused</li> <li>- Service Quality and Service</li> <li>- Ambitious culture of quality</li> <li>- Customer Participation</li> <li>- Connection with customers</li> <li>- Customer and quality-focused thinking</li> </ul>	<ul style="list-style-type: none"> <li>- Customer Input</li> <li>- User Participation</li> <li>- Effective quality control</li> <li>- Customer Participation</li> <li>- Customer-focused</li> <li>- System quality</li> </ul>
Communication [19, 21-24]	<ul style="list-style-type: none"> <li>- Cooperation without barriers</li> <li>- Communication Ability</li> <li>- Communication</li> <li>- Communication between team members</li> </ul>	<ul style="list-style-type: none"> <li>- Sharing information</li> <li>- Communication</li> <li>- Strengthen communication among stakeholders</li> <li>- Understanding the process's long-term organizational impact</li> </ul>
Organizational culture [16, 19, 21-25, 29]	<ul style="list-style-type: none"> <li>- Recognize it as an everyday business activity</li> <li>- Process thinking</li> <li>- Ownership and Participation</li> <li>- Cultural change</li> <li>- Substructure and Culture of the Organization</li> <li>- Organizational culture</li> <li>- Active participation</li> <li>- Implementing a Quality Culture</li> </ul>	<ul style="list-style-type: none"> <li>- Cultural Change for Experiment</li> <li>- Review of organizational structure</li> <li>- Emotional maturity</li> </ul>
Performance and rewards [17-23, 25]	<ul style="list-style-type: none"> <li>- Reward and Incentives</li> <li>- Reflecting personnel evaluation</li> <li>- Additional points for promotion</li> <li>- Financial performance</li> <li>- Quantified indicators and expected effects</li> <li>- Evaluation and Rewards</li> <li>- Performance compensation</li> <li>- Remove performance indicator</li> <li>- Benefit</li> </ul>	<ul style="list-style-type: none"> <li>- Reward</li> </ul>
Data [16, 19, 21]	<ul style="list-style-type: none"> <li>- Data-based management</li> <li>- Management by data and facts</li> <li>- Build data</li> <li>- Statistically processing data</li> </ul>	<ul style="list-style-type: none"> <li>- Aim for data</li> <li>- Data availability</li> <li>- Data Quality</li> <li>- High Quality Dataset</li> <li>- Dataset accessibility</li> <li>- Linked data interoperability</li> <li>- Identify critical data requirements</li> <li>- Cleanup and preprocessed datasets</li> <li>- Easily accessible data</li> <li>- Uncomplicated data structure</li> <li>- Management of metadata</li> <li>- Data-based mindset</li> <li>- Representativeness of data</li> </ul>
Project [19, 21-29]	<ul style="list-style-type: none"> <li>- Select the required project</li> <li>- Quality Project Experience</li> <li>- Project Management Technology</li> <li>- Project priority and selection</li> <li>- Project execution quality</li> <li>- Select Team Members and Teamwork</li> <li>- Clear objectives and goal</li> <li>- Team ability to respond to changes</li> <li>- Project team's general expertise</li> <li>- Self-configured Collaboration Team</li> <li>- Project Planning Level</li> <li>- Clear requirements and specification</li> <li>- Structured Project procedures</li> <li>- Structured progress reporting</li> <li>- Effective project manager technology</li> <li>- Project complexity</li> </ul>	<ul style="list-style-type: none"> <li>- Management of expectations and cultural experiment</li> <li>- An agile method</li> <li>- Focus on problem solving</li> <li>- Pre-planning to improve cost performance</li> <li>- An effective milestone</li> <li>- Devotion of project team</li> <li>- Clear objectives and goal</li> <li>- Effective Project Planning</li> <li>- Effective Cost Estimation</li> <li>- Measurement of effective projects</li> <li>- The ability of project team</li> <li>- The purpose of business purposes</li> <li>- Project schedule and planning</li> <li>- Project related costs</li> <li>- Team Skill</li> </ul>

6시그마의 성공 정책과 빅데이터 분석의 성공 정책 분석시 공통된 성공 정책으로, 최고경영진 지원, 프로세스 개선, 교육 훈련, 교육시스템, 인원 및 기술, 시간, 고객 및 품질 중심 사고, 커뮤니케이션, 조직문화, 성과 및 보상, 데이터, 프로젝트의 12가지 항목을 정리하였다.

최고경영진의 지원 항목에서는 최고경영진의 역할과 책임감, 리더십을 강조하였으며, 방법론의 전략적 적용과 최고경영진의 역량, 방법론 지원 예산 등도 성공 정책으로 인식되었다.

프로세스 개선 항목에서는 업무 통합과 표준화, 조직 프로세스의 호환성, 올바른 로드맵 제시 등을 성공 정책으로 인식하였다.

교육훈련 항목에서는 방법론 수행에 꾸준한 교육과 훈련을 지속해야 한다고 보았고, 교육훈련을 통한 우수한 인재의 확보와 구성원들의 역량 강화가 성공 정책이라고 보았다.

교육시스템 항목에서는 방법론 수행이 성공하기 위해 교육 전담 부서(Big Six Sigma 팀)에서 꾸준히 지원할 수 있는 시스템을 갖추어야 하고, 효과적인 프로젝트 수행을 위해 전사원 참여, 전사적 활동, 지원체계와 인프라가 잘 갖추어져야 한다고 보았다. 또 지속적인 모니터링 및 피드백을 통해 구성원들의 교육 진행시 어떤 점을 보완해야 할지에 대한 고민 및 지속적인 대책 수립이 필요하다고 보았다. 또한 데이터 사용에 있어서 개인정보관리와 프라이버시, 보안 등을 고려하였다.

인원 및 기술 항목에서는 수행 인원의 숙련도, 이해관계자와의 관계, 6시그마 기법과 빅데이터 분석기법의 활용 능력 등을 성공 정책으로 판단하였다. 이들 정책은 방법론 수행의 필수적인 요소로서 통계, 데이터마이닝 도구와 기술에 대한 이해를 바탕으로 해야 한다고 인식하였다.

시간 항목에서는 충분한 도입 시간, 준비 시간, 추진 시간, 일정과 타이밍 등을 성공 정책으로 고려하였다. 이들 정책은 방법론 정착과 프로젝트 수행에 영향을 미치는 요소로 시간 관리의 중요성을 강조하였다.

고객 및 품질 중심 사고 항목에서는 고객 참여와 니즈를 반영하는 고객 중심의 사고가 중요하다고 보았으며, 효과적인 품질관리를 위한 의욕적 품질 문화, 시스템 품질, 서비스 품질 관리 등 품질의 중요성을 인식하고, 고객의 VOC를 적극 반영하여 방법론을 수행하는 것이 성공의 핵심이라고 보았다.

커뮤니케이션에서는 구성원들이 원활하게 협력하고 소통할 수 있도록, 장벽 없는 상호협력 및 커뮤니케이션이 방법론 수행의 성공에 필수적이라고 인식하였다. 또한 팀원과 이해관계자 간의 의사소통을 강화하는 것도 중요한 정책이라고 판단하였다.

조직문화에서는 방법론을 경영문화의 일부로 인식하고, 받아들일 수 있는 문화적 변화가 필요하다고 보았다. 모든 구성원이 주인의식을 가지고 적극적으로 방법론 수행에 참여하는 것이 성공의 지름길이라고 인식하였다.

성과 및 보상 항목에서는 방법론 수행의 성공과 실패를 성

과지표에 반영하면 안 된다는 주장과 적극적으로 성과지표에 반영하여 인사평가나 진급가점, 인센티브 등 보상을 해주어야 한다는 주장이 상반되게 나타났다. 연구자의 경험상 방법론을 수행하는 기업에서 기업의 특성과 목표에 맞게 적절한 성과보상 체제를 검토하여 적용하는 것이 바람직하다고 생각한다.

데이터 항목에서는 데이터는 사실에 근거하여 사용해야 한다고 하였으며 데이터의 가용성과 품질, 접근성이 중요하다고 인식하였다. 비정형 데이터를 분석할 수 있는 형태로 변환하고 메타 데이터를 관리하는 등 데이터의 지속적인 관리 역시 방법론 수행의 성공 정책에 속한다고 보았다.

프로젝트 항목에서는 목적과 목표를 명확히 정하고 프로젝트의 중요도에 따라 우선순위를 결정해야 한다고 보았으며, 프로젝트 경험과 팀워크가 뛰어난 팀원, 전문성이 프로젝트 결과에 많은 영향을 준다고 보았다. 사전에 프로젝트 일정과 계획의 기획을 잘하는 것과 프로젝트 관리자의 역량을 향상하는 것이 프로젝트의 수행의 성공 정책이라고 인식하였다.

6시그마와 빅데이터 분석의 공통된 12가지 성공 정책을 빅데이터 기반 6시그마의 방법론의 성공 정책으로 제시하였다. 6시그마와 빅데이터 분석의 성공 정책은 비슷한 항목이 많았지만 일부 차이점도 발견되었다. 6시그마의 경우 최고경영진의 지원, 프로세스 개선, 교육시스템, 고객 및 품질 중심 사고, 조직문화, 성과 및 보상 등의 성공 정책을 더 강조하였고 빅데이터 분석 성공 정책에서는 인원 및 기술, 데이터 관리 등의 성공 정책을 더 강조하였다. 시간, 교육훈련, 커뮤니케이션, 프로젝트 항목은 두 방법론 모두 비슷하게 강조하였다. 이러한 두 방법론의 성공 정책을 효과적으로 조직문화에 잘 적용하고 유지해야 빅데이터 기반 6시그마 방법론의 성공적인 수행과 성과를 기대할 수 있을 것이다.

#### 4. 결 론

빅데이터 기반 6시그마 방법론은 기존 6시그마의 한계를 극복하기 위해 6시그마 기법에 빅데이터 분석기법을 추가한 방법론이다. 이 방법론에 대해 최근 10년간 7건의 국내외 연구가 진행되었으며 스마트 공장의 확산과 자동화 Needs의 증가로 빅데이터 기반 6시그마 방법론에 대한 관심도 높아지고 있다.

그러나 빅데이터 기반 6시그마 방법론과 LG DX SS 방법론이 얼마나 현업에서 성과가 있으며 유효성이 있는 방법론인지에 대해서는 객관적으로 검증되지 않았다. 1장과 2장에서는 기존 6시그마 방법론과 빅데이터 기반 6시그마 방법론에 대한 비교를 통해 빅데이터 기반 6시그마 방법론에 대한 강점을 제시하였다.

빅데이터 기반 6시그마 방법론은 기존 6시그마의 통계적 분석기법은 그대로 활용하되 빅데이터 환경에서 다양한 기계학습 기법들을 적용하고 모델링하고, 비교/평가하여 최적의 모델을 선택하고, 변수 중요도 분석 및 하이퍼 파라미터 최적

화 등을 통해 모델 성능을 최적화한 후, 유의변수에 대한 실험 최적화까지 진행한다. 전수 데이터를 활용하여 불량유발인자를 찾아내므로 데이터 분석시간을 단축하고, 품질 향상 및 경비 절감이 가능한 점이 강점이다.

최근 10년간 진행된 빅데이터 기반 6시그마 방법론 연구자 7명의 과제성과(연 13.96억 비용절감, 예측정확도 90% 확보, 불량 발생률 87% 개선 등)와, LG DX SS 인증과제 DX MBB 30인의 성과(품질개선 GAP 수치 평균 18.83)를 통해 빅데이터 기반 6시그마 방법론이 품질 향상과 경비 절감에 유효하다는 것을 객관적으로 입증하였다.

또한 3장에서는 빅데이터 기반 6시그마 방법론의 성공을 위한 핵심 성공 정책(Key Success Policy)을 제시하였다. 빅데이터 기반 6시그마 방법론의 핵심 성공 정책은, 최고경영진 지원, 프로세스 개선, 교육훈련, 교육시스템, 인원 및 기술, 시간, 고객 및 품질 중심 사고, 커뮤니케이션, 조직문화, 성과 및 보상, 데이터, 프로젝트 등의 12가지로 구분할 수 있었다. 이들 핵심 성공 정책을 전 임직원이 따르고 올바르게 시행되었을 때 빅데이터 기반 6시그마 방법론의 성공률이 높아질 것이며 해당 방법론의 도입을 시도하는 많은 기업이 안정적으로 방법론을 정착하고 지속할 수 있을 것이다.

본 연구의 한계는 다음과 같다. 첫째, 기존 연구 문헌과 LG Display에서 전문가 30인이 수행한 일부 프로젝트 사례를 통해 빅데이터 기반 6시그마 방법론의 실증 분석을 하였으므로 더 많은 실제 사례를 통한 유효성 검증이 필요하다.

둘째, 실제 빅데이터 기반 6시그마 방법론을 전사적으로 수행하고 있는 기업이 LG Display를 포함한 계열사에서만 수행되고 있어 더 많은 기업에서 해당 방법론에 대한 성공정책에 대한 연구가 진행되고, 제조 외 비제조 부문(간접/사무)에 대한 성과도 검증하여 실증분석이 이루어져야 한다고 생각한다.

본 연구 목적은 빅데이터 기반 6시그마 방법론인 LG DX SS의 유효성을 검증하는 것이다. 이 방법론은 전수 데이터를 활용하여 불량유발인자를 파악하고, 빅데이터 기법과 6시그마의 통계기법 등을 통해 유의변수에 대한 최적화까지 진행하여 효과적인 품질개선, 경비 절감 달성이 가능하다. 본 연구에서는 이 방법론의 강점과 효과, 방법론이 성공하기 위한 핵심 성공 정책(KSP)을 제시하였다. 이를 통해 앞으로 스마트공장 확산 및 데이터가 넘쳐나는 빅데이터 환경에서 국내외 많은 제조기업이 빅데이터 기반 6시그마 방법론을 성공적으로 도입하고 안정적으로 정착 및 지속시켜 많은 성과를 얻기를 기대한다.

## References

- [1] Y. K. Park, C. H. Yoon, and Y. H. Ryu, "Current situation, issue and development direction of six sigma innovation in korea," *Journal of the Korean Institute of Industrial Engineers*, Vol.32, No.4, pp.253-267, 2006.
- [2] S. Jin, "A study on the effects of the success factors of six sigma management on quality performance - Focusing on Kyongbuk area company -, " M.S. dissertation, Hanbat National University atDeajeon, Korea, pp.8-10, 2014.
- [3] W. S. Joo, "A case study on the success factors of the 6 sigma innovation in the automotive parts sector," M.A. dissertation, Korea University of Technology & Education at Chungcheong-do, Korea, 2015.
- [4] Y. J. Ahn, "The future of Six Sigma," *Journal of the Society of Service Management*, Vol.15, No.1, pp.157-172, 2014.
- [5] D. K. Go, "A study on the Methodology for Application of 6 Sigma Based on Big Data - On the case of Process Optimization for Solar Cell -, " M.S. dissertation, Hanyang University at Seoul, Korea, pp.1-46, 2018.
- [6] J. M. Go and G. S. Jang, "A six sigma methodology based on data mining, BSC, and BPM," *Industrial Innovation Research*, Vol.27, No.1, pp.57-83, 2011.
- [7] G. S. Jang, "A six sigma methodology using data mining: A case study of "P" steel manufacturing company," *Information System Research*, Vol.20, No.3, pp.1-21, 2011.
- [8] H. T. Kim, "A study on the methodology of quality innovation using big data-based data mining - On the case of display product's defect improvement -, " M.A. dissertation, Seokyeong University at Seoul, Korea, 2019.
- [9] S. G. Ko, "New six sigma breakthrough using big data," *Creation and Innovation*, Vol.12, No.4, pp.115-133, 2019.
- [10] Y. J. Woo, "Manufacturing quality improvement methodologies for big data application," M.S. dissertation, Gachon University at Gyeonggi, Korea, 2019.
- [11] Rose Clency, "Data-driven quality improvement approach to reducing waste in manufacturing," *The TQM Journal*, Vol.35, No.1, pp.51-69, 2021.
- [12] LG Display, "6 Sigma MBB Repair Process (DX SS Methodology)," in Gyeonggi, Korea: LG Display, inc, pp.4-135, 2022.
- [13] I. S. An, "A study on the success factors of six sigma and their implementation performance," M.A. dissertation, Paichai University at Daejeon Korea, 2011.
- [14] C. S. Choi, J. S. Le, and W. Y. Yun, "Present and future of six sigma: LG electronics case," *Journal of the Korean Institute of Industrial Engineers*, Vol.32, No.4, pp.338-346, 2006.
- [15] J. H. Yang and J. G. Yun, "AMOLED defective and reliability," *Journal of the Korean Society for Information Display*, Vol.8, No.5, pp.11-21, 2007.
- [16] W. S. Joo and D. S. Ha, "The strategy of the product quality optimization through the six sigma," *Management Information System Review*, No.4, pp.557-570, 2000.



- [17] D. S. Shin, "An empirical study on the major factors of implementing six sigma successfully through blackbelts," Ph.D. dissertation, Dankook University at Gyeonggi, Korea, 2001.
- [18] H. S. Kim, "A study on the effects of the success factors on business performance for six sigma," M.S. dissertation, Nam Seoul University at Chungcheong-do, Korea, 2008.
- [19] H. N. Choi, "A study of the impact of six sigma's success factor on service quality improvement: Focused on s tele-com company," M.S. dissertation, Dankook at Gyeonggi, Korea, 2009.
- [20] H. H. Park, "A study on the effects of the success factors of six sigma on job satisfaction and quality performance," M.S. dissertation, Ulsan University at Ulsan, Korea, 2013.
- [21] Y. K. Lee, "The effects of six sigma success factors on innovation and business performance - The intervening effect of business innovation capability -," Ph.D. dissertation, Hanyang University at Seoul, Korea, 2015.
- [22] W. J. An, "A study on the effects of the success factors of six sigma on change management activity and business performance," M.A. dissertation, Ulsan University at Ulsan, Korea, 2017.
- [23] M. M. Marco, L. R. Jorg, G. A. Carlos, Diego Tlapa, B. L. Yolanda, and S. L. Hector Antonio, "Measuring critical success factors for six sigma in higher education institutions: Development and validation of a surveying instrument," *IEEE ACCESS*, No.8, pp.1813-1821, 2020.
- [24] M. Dursun, N. Goker and Hakan Mutlu, "Evaluation of project management methodologies success factors using fuzzy cognitive map method: Waterfall, agile, and lean six sigma cases," *International Journal of Intelligent Systems and Applications in Engineering*, No.10, pp.35-42, 2022.
- [25] Bruno Muniz Felix, Elaine Tavares and Ney Wagner Freitas Cavalcante, "Critical success factors for Big Data adoption in the virtual retail: Magazine Luiza case study," *REVISTA BRASILEIRA DE GESTAO DE NEGOCIOS Review of Business Management*, No.20, pp.112-126, 2017.
- [26] A. Wisnuwardhana, A. N. Hidayanto, N. F. A. Budi, I. C. Hapsari, and A. Haidaroh, "Systematic literature review: Critical success factor in the application of data maning", *The 8th Engineering International Conference 2019, Journal of Physics: Conference Series*, pp.1-8, 2023.
- [27] J. S. Sim, "Critical Success Factors in Data Maning Projects", Ph.D. dissertation, University of North TEXAS at Denton TEXAS, USA, 2003.
- [28] Jayanthi Ranjan and Vishal Bhatnagar, "Critical Success Factors For Implementing CRM Using Data Mining", *Journal of Knowledge Management Practice*, Vol.9, No.3, pp.1-3, 2008.
- [29] A. S. Zaher Ali, Rosni Abdullah and Mohd Heikal Husin, "Critical Success Factors for Big Data : A Systematic Literature Review," *IEEE ACCESS*, No.8, pp.118940-118955, 2020.
- [30] N. Y. Baek, "A Study on the Failure of Six Sigma Implementation," M.S. dissertation, Kyungnam University at Kyungnam, Korea, 2008.



### 김정혁

<https://orcid.org/0009-0002-4632-7642>

e-mail : kaiwenz12@cuk.edu

2012년 서울사이버대학교 법무행정학부  
(학사)

2022년 ~ 현 재 고려사이버대학교  
융합정보대학원 석사과정

2008년 ~ 현 재 LG 디스플레이 환경기술담당 교육관리

관심분야 : Six Sigma & Big Data Methodology



### 김윤기

<https://orcid.org/0000-0001-9348-702X>

e-mail : ykkim77@cuk.edu

2012년 서경대학교 컴퓨터공학과(학사)

2019년 고려대학교 전산학과(석·박사)

2019년 ~ 2021년 SK 주식회사 C&C

빅데이터·AI 플랫폼 기술연구개발

2021년 ~ 현 재 고려사이버대학교 전임교수

관심분야 : 분산 병렬 시스템, AI 빅데이터 플랫폼, HPC