

## 제5장

# 국가승인통계 품질 등급부여방안 및 상대지표 개발 검토

심규호

## 제1절 서론

### 1. 연구의 배경

품질(Quality)이라는 단어의 의미를 한마디로 표현한다는 것은 거의 불가능하다. 품질의 사전적 의미는 다음과 같다. 세계적으로 잘 알려져 있는 Wiktionary에 의하면 첫 번째 의미는 우수한 정도(level of excellence), 두 번째는 사물이나 사람의 차이에 대한 특성 또는 속성, 세 번째는 혼합물에 있어서의 총량에 대한 물질의 비율, 네 번째는 높은 사회적 지위(High social position), 다섯 번째는 사람이 만든 물건이나 시스템이 버그(bugs)나 결점(flaws)에 대해 자유로운 정도를 의미한다고 나와 있다. 우리가 품질이라고 하면 보통 첫 번째나 세 번째 정도의 의미에서 두 번째 의미 정도가 포함되는 정도라 할 수 있을 것이다.

품질이라는 단어를 바로 사용하는 것은 힘들기 때문에 사전적인 의미 외에 품질에 대한 많은 정의가 존재한다. ISO 9000에서는 “요구사항을 만족시키는 고유한 성격의 집합의 일련의 정도”라고 정의하고 있고 Six Sigma에서는 “백만 개당 결함의 수”라고 정의 내리고 있다. Joseph M. Juran은 “사용에 있어서의 적절함”이라고 정의했으며 Noriaki Kano는 “반드시 해야 하는 품질(must-be quality)”와 “매력적인 품질(attractive quality)”의 두 방향의 품질에 대한 모형을 제시하였다. 전자는 사용하기 적합한 정도의 품질을 의미하여 후자는 아직 정해지지 않은 사용자가 만족할 정도의 품질을 의미한다. Robert Pirsig는 “배려의 결과”로 정의했으며 Genichi Taguchi는 “목표값에 근접한 균등성(Uniformity

around a target value)”로 정의했다. 또한 Peter Drucker는 “재화나 서비스에 대한 품질은 제공자가 넣는 것은 아니다. 이것은 고객이 얻고자 하는 것이나 지불하고자 하는 것에 대한 것이다.”로 정의했다. Gerald M. Weinberg는 “사람들에게 가치를 제공하는 것(Value to some person)”이라고 정의했다. 품질의 정의는 시대의 흐름에 따라, 품질이 정의되는 분야에 따라, 정의내리는 사람마다 차이는 있지만 품질의 정의는 크게 두 방향으로 정의할 수 있을 것이다. 제품 자체에 대한 품질과 고객에 대한 품질이다. 제품 자체에 대한 품질은 적합성, 제품의 결함 정도, 균등성 등으로 요약할 수 있으며 고객에 대한 품질은 만족, 지불한 가격에 대비한 성능, 가치 등으로 요약할 수 있을 것이다. 통계청에서는 품질을 다음과 같이 정의하고 있다.

- 질기고 튼튼한 것이 품질
  - : 전통적인 개념으로는 ‘오래 사용할 수 있고 질기고 튼튼한 측면’ 즉, ‘품질의 객관적이고 물리적인 성질’을 주로 강조하여 왔다.
- 이용자의 요구사항을 충족시키는 것이 곧 품질
  - : 최근 산업사회의 발달로 치열한 시장경쟁에서 생존해야 하는 기업경영자들은 ‘이용자의 요구사항을 충족시키는 측면’ 즉, ‘품질의 주관적인 특성’을 강조하고 있다.
- 품질이란 “사용목적에의 적합성(Fitness for use)”이라고 정의
  - : 지난 70여 년 동안 품질관리 분야에 큰 영향력을 끼친 바 있는 J. M. Juran은 품질이란 “사용 목적에의 적합성”(Fitness for use)라고 정의하였다.

국가통계에서의 품질은 통계 자체에 대한 품질이라고 할 수 있겠다. 통계데이터(statistical data)도 통계작성기관이라는 조직의 입장에서는 제품이며, 통계라는 제품도 고객만족 차원에서 관리되어야 한다는 품질관리의 개념이 최근 통계 선진국들에서 중요한 개념으로 자리 잡고 있다. 통계데이터에서의 제일 중요한 개념은 단연 품질이다. 통계데이터를 작성하는 조직의 목적, 인력, 예산, 연구 등 모든 것이 통계 데이터의 품질을 위해 존재한다고 해도 과언이 아닐 것이다. 이러한 통계품질의 개념에는 통계의 정확성(Accuracy), 관련성(Relevance), 시의성(Timeliness), 접근성(Accessibility), 비교성(Comparability), 효율성(Efficiency), 해석성(Interpretability), 일관성(Coherence) 등의 요소가 내포되어 있는데, 이 중 통계의 현실 반영 정도를 나타내는 정확성은 가장 중요한 요소이다. 통계청에서는 다음과 같이 통계품질에 대해 정의내리고 있다.



&lt;

오늘날 통계의 품질은 단순히 통계의 정확성과 신속성만 강조하는 것이 아니라 “통계가 얼마나 이용자에게 이용하기 적합하게 작성되고 제공되고 있는가.”를 나타내는 개념으로 자리 잡고 있다.

이러한 통계품질의 개념에는 통계의 정확성, 관련성, 시의성, 접근성, 비교성 등의 요소가 내포되어 있는데, 이 중 통계의 현실반영 정도를 나타내는 정확성이 가장 중요한 요소로 인식되고 있고 통계가 작성되는 모든 과정과 연관되어 있다. 이 외의 요소들도 최근 고객 지향적 정부행정이 강조됨에 따라 모두 중요시 되고 있다.

&gt;

품질관리란 제품 또는 서비스의 품질을 고객들이 만족할 수 있는 수준으로 유지하기 위한 의식적인 활동이다. 원래 관리(Management)라는 말은 경영의 방침 또는 계획을 수립하고 조직을 구성하여 이를 실행하고 필요한 통제를 가하는 것을 의미한다. 총체적인 품질관리(TQM, Total Quality Management)의 창시자이며 미국 제너럴시스템사의 대표 이사인 A.V.Feigenbaum 박사는 근대적 의미의 품질관리를 “소비자를 완전히 만족시키는 것이 가장 경제적인 수준으로 제품을 생산할 수 있도록 조직의 각 부문의 노력을 품질의 유지와 개선을 위하여 통합시키는 것”으로 정의하고 있다. 그러므로 통계품질관리(Quality Management for Statistics)란 “통계 이용자들에게 최대의 만족감을 주면서 동시에 가장 경제적인 통계를 얻기 위한 모든 수단을 통합한 체계”라고 할 수 있다(통계청 품질관리과 홈페이지). 통계품질관리 활동의 첫 시도는 1985년 캐나다 통계청이 “통계품질 가이드라인(Guideline for statistics quality)”를 처음으로 발간한 것이라고 볼 수 있다.

## 2. 연구의 목적 및 필요성

최근에 국가 차원에서의 통계에 대한 인식이 높아지고 통계청뿐만 아니라 많은 국가기관에서 통계 사용자의 요구를 만족시키기 위한 통계 개발에 박차를 가하고 있다. 통계청은 중앙 통계기관으로서 이러한 통계 개발을 선도하고 관리해 나가려고 노력하고 있다. 이에 따라 2006년에 국가승인통계 106종 품질진단 실시를 시작으로 2007년에는 180종, 2008년에는 170종, 2009년에는 58종의 국가승인통계에 대한 품질진단을 실시하고 있다. 이를 실시하기 위해 매 실시 년도에 각 해당 통계의 전문가로 구성된 품질진단 전문가 팀이 구성된다. 이 품질진단 팀은 각 분야의 통계의 정의, 생산과정,

사용자들의 의견, 제공 통계의 정확성, 정책에서의 사용 정도 등 통계 품질과 관련된 많은 정보를 취합하고 분석하여 품질 개선에 도움이 되는 보고서를 작성하고 있다. 이러한 보고서에는 품질이 좋은 부분은 우수사례로 선정하여 공표하기도 하고 개선해야 할 점은 보고서에서 상세한 개선 방법을 제시하기도 한다. 이 보고서를 기반으로 통계 생산 기관이나 부서에서는 품질을 개선하여 국가통계 전반에 기여하기를 기대하는 것이다.

그러나 품질 관리 담당자들이 느끼는 어려움은 다음의 두 가지이다. 먼저 품질 개선 활동은 열심히 하고 있지만 품질 개선에 대한 명확한 평가를 하기가 어렵다는 것이다. 개선 방법에 대한 것은 보고서에 많이 기재하여 통계 생산 부서에 전달하지만 그 내용처럼 실제적으로 개선을 강제할 법적 권한은 없다. 대부분의 통계 생산 기관은 통계 인력 또는 통계 예산 측면에서 열악한 환경을 가지고 있는 경우가 많다. 또한 품질 진단 자체가 어떠한 비교할 수 있는 점수화나 비교 지수를 개발하여 공표하여야 하는데 현실적인 어려움이 있다. 우선 품질을 정확하게 판단하고 개선하는 데는 두 가지가 선행되어야 한다. 첫째는 품질의 각 성격을 대표하는 지표가 측정 가능해야 한다. 유럽에서는 이미 품질을 측정하는 방법에 대한 연구가 많이 진행되었다. 통계 생산에서부터 공표, 활용 부분 전반에 걸쳐 상세한 측정 기준에 따라 측정하고 점수화를 한다. 또한 이러한 점수화된 결과를 분석하여 품질에 영향을 많이 미치는 지표를 찾아내기도 하였다. 둘째는 이 측정된 지표들이 비교 가능해야 한다. 또한 정확히 측정된 값으로 비교 가능한 종합 지표를 개발하여 진단된 통계들 간의 상호 비교가 가능해야 한다. 상호 비교 가능한 지표들이 개발되면 통계 생산 주체들 간에도 비교와 분석을 통해 품질이 개선되는 효과가 있으리라 기대한다.

본 연구에서는 첫째 품질관리의 현황을 국내와 국외의 현황으로 나누어 비교 분석하고자 한다. 국내의 품질 관리는 어떻게 진행하고 있는지, 문제점은 없는지 알아보고 해외의 품질 관리 현황과 연구 현황에 대해 알아보하고자 한다. 국내 통계에 대한 품질 관리에 있어서는 통계청 품질 관리가 현재는 가장 규모가 크고 많은 발전이 있었다고 생각한다. 따라서 국내 사례는 통계청 품질 진단을 중심으로 알아보하고자 한다. 해외에서는 EUROSTAT과 캐나다 등이 품질진단의 원조격이라 할 수 있고 많은 품질 진단에 대한 연구가 진행되어져 왔다. 해외의 품질 진단 현황에 대해서는 유럽과 캐나다 미국을 중심으로 알아보하고자 한다. 문제의 핵심이 되는 측정 가능한 품질 비교 연구에 대한 부분은 국내에서는 아직은 전무한 상태라 해외 사례에 대해 연구해보고자 한다.



## 제2절 국내 사례

### 1. 통계청 품질 진단

통계청의 품질관리는 품질관리과에서 맡고 있다. 1999년 4월 통계기획국 기획과에 품질평가팀이 신설되었고 2002년 7월 청장직속 부서인 품질관리팀으로 개편, 2005년 품질관리팀에서 품질관리과로 확대 개편되었다. 2006년에 국가승인통계 106종 품질진단 실시를 시작으로 2007년에는 180종, 2008년에는 170종, 2009년에는 58종의 국가승인통계에 대한 품질진단을 실시하고 있다(<표 5-1>).

<표 5-1> 연도별 진단 현황

연도	진단 종류
2002년	12종
2003년	11종
2004년	10종
2005년	대규모 조사통계 6종, 외부 6개 기관 작성 통계 12종
2005년	474종 국가승인 통계 대상
2006년	106종
2007년	108종
2008년	170종
2009년	58종 예정

주요업무로는 통계품질평가 및 기법 연구, 표준 매뉴얼 작성 및 보급, 국가통계품질평가제도 운영, 국가통계품질관리 관련 계획 수립 및 실시를 하고 있다. 그 중에서도 품질진단을 위한 지표개발과 품질진단을 실시하고 품질향상을 위한 권고 및 개선 유도가 중요한 업무라 할 수 있다.

통계청에서는 통계품질 차원을 <표 5-2>의 6개 차원으로 정의하고 있다. 통계청에서는 관련성, 정확성, 시의성, 비교성, 일관성, 접근성을 차원으로 정하였으나 국가마다 조건의 차이가 있다.

〈표 5-2〉 통계청 품질 차원

차원	설명
관련성(Relevance)	이용자의 요구사항을 파악하여 통계자료에 반영하는 정도
정확성(Accuracy)	측정하고자 하는 모집단의 특성이나 크기를 얼마나 정확하게 측정했는가.
시의성(Timeliness) 및 정시성(Punctuality)	시의성은 통계작성기준시점과 결과발표시점 간의 차이에 대한 개념 정시성은 예고된 공표일정을 준수하는가에 대한 개념
비교성(Comparability)	시간적·지리적 또는 포괄하는 영역이 달라도 어떠한 통계가 개념, 분류, 측정도구, 기초자료 등이 동일하게 적용되어 개념적으로 서로 비교가 가능한지를 측정
일관성(coherence)	일관성은 동일한 경제·사회현상에 대해 서로 다른 기초자료나 작성방법에 의해 작성된 다른 통계 또는 소관통계가 얼마나 유사한가를 측정
접근성(accessibility) 및 명확성(clarity)	접근성은 이용자가 통계자료에 얼마나 쉽게 접근할 수 있는지에 대한 정도 명확성은 통계가 어떻게 만들어졌는지에 대한 정보 수준

이러한 품질 차원은 품질 진단을 하는데 있어 중요한 요소가 되며 이러한 차원을 바탕으로 품질 진단을 실시하며 각 차원 별로 개선 사항을 도출하여 결과적으로 품질 향상에 기여하고 있다.

유럽을 비롯한 많은 나라들은 자국의 통계 품질 향상을 위해 오래전부터 품질 관리를 해오고 있다. 통계품질 결정요소의 국가 간 비교는 다음 표와 같다. 대부분의 국가가 공통적인 품질 요소를 사용하고 있으나 각 나라와 기관의 특성에 맞게 차이가 있는 부분이 있다. 통계 생산이 집중형이나 분산 형이냐에 따라 또는 서비스형의 통계가 많으냐에 정책관련 통계가 많으냐에 따라 다른 품질 요소를 가지고 있음을 알 수 있다 (<표 5-3>).



〈표 5-3〉 각각의 품질 차원 비교

IMF	OECD	Eurostat	캐나다	호주	네덜란드	스웨덴	영국	한국
정확성/ 신뢰성	정확성	정확성	정확성	정확성	정확성	정확성	정확성	정확성
	관련성	관련성	관련성	관련성	관련성		관련성	관련성
	시의성/ 정시성	시의성/ 정시성	시의성	시의성	시의성	시의성	시의성	시의성/ 정시성
접근성	접근성	접근성/ 명료성	접근성	접근성		접근성/ 명료성	접근성	접근성/ 편리성
		비교성					비교성	비교성
	일관성	일관성	일관성			일관성/ 비교성	일관성	
	해석성		해석성					
		완결성					완결성	
					효율성			효율성
방법론의 충실성								
서비스성								
무결성								
	신뢰성							
				개편크기 최소화				
				커버리지 포괄성				
					응답 부담성			

통계청은 위와 같은 품질 차원을 기반으로 승인 통계 등에 대한 품질진단을 정기적으로 실시하고 있다. 통계청의 세부적인 품질진단 방법은 <표 5-4>와 같다.

<표 5-4> 통계청 품질 진단 방법

품질관리기반	본격적인 품질진단에 앞서 기관장의 관심, 전략과 방침, 인력자원 관리 및 예산규모, 통계작성 환경을 파악하여 품질관리 기반 진단
이용자 만족도 및 요구사항 반영실태	통계이용자가 통계자료에 대해 얼마나 만족하는지를 조사하여 품질진단에 활용할 필요가 있다. 통계작성기관은 통계의 전문이용자, 일반이용자를 구분하여 리스트를 확보하고 수시로 이용자가 해당통계를 만족스럽게 이용하는지 확인
세부작성절차별 체계	통계작성 과정이 통계작성의 본래 목적을 실현하기 위해 적합하게 이루어지고 있는가를 세부 작성 절차별 진단서를 이용하여 진단 진단작업은 통계작성 담당자가 직접 수행하는 내부진단과 외부전문가가 실시하는 외부진단 두 가지를 실시
수집자료의 정확성	통계자료가 얼마나 정확한가는 수집된 자료가 얼마나 정확한가에 달려 있으며, 이는 조사나 보고 등 자료가 수집되는 시스템의 효율성에 의해 좌우된다. 이 진단 절차는 수집 자료의 정확성과 절차적 오류는 없는지에 대해 진단하는 방법을 설명하는데 이에 대한 점검과정은 통계품질을 결정하는 매우 중요한 과정
통계자료서비스의 충실성	작성과정에서는 오류가 없는 통계일지라도 공표되는 과정에서 오류가 발생한다면 통계품질이 높은지 낮은지 진단할 것도 없이 잘못된 통계가 됨 통계자료 서비스의 충실성 점검을 중요하게 생각하는 이유가 바로 이런 점임

위와 같은 절차에 따라 품질 진단을 실시하고 있으며 관련 통계 영역별로 보고서를 작성하여 배포하고 있으며 품질 진단 워크숍, 품질 진단 세미나 등을 통해 품질 진단에 대한 각 분야의 전문가의 의견을 듣고 있다.





## 2. 국내 품질 연구 사례

통계청 품질관리과에서는 2003년 “통계품질평가지표 개발” 용역을 통해 통계작성절차를 중심으로 하는 평가지표를 개발하였다. 통계품질평가지표 개발에 국내에서 참고할 만한 자료가 부족한 관계로 대부분을 국제기구나 주요 국가들에서 사용하고 있는 평가지표들을 검토하여 사용하였다. 현행 48개의 평가지표를 수정 보완해서 7개 차원별로 81개의 지표를 개발하고 이들 지표 중에서 필수지표로 48개를 선정하였다. 개별 평가지표의 필요성과 평가내용들을 정리하고 절차별로 구분하였다. 또한 전수조사, 표본조사, 그리고 가공통계로 나누어 평가에 필요한 지표를 분류하였다(2003, 류제복).

박성현(2003)은 “통계품질관리 표준매뉴얼 작성 학술연구”를 통해 국가통계작성기관의 자체적인 품질관리를 위한 지침서를 개발하였다. 통계품질(statistics quality)은 통계가 얼마나 사용자들이 이용하기에 적합한가를 나타내는 특성으로, 캐나다가 1985년에 「통계품질 가이드라인」을 발간한 이후에, 선진국 및 국제기관들을 중심으로 높은 관심을 보이고 있다. 통계청에서는 품질관리팀을 설치하여 통계품질 제고에 노력하고 있으며, 품질평가 시에 통계작성단계를 8단계(조사기획, 모집단 및 표본, 조사표 작성, 조사 직원 관리, 조사 실시, 자료처리 및 집계, 자료공표, 자료이용)로 구분하여, 6개의 품질요소(정확성, 시의성, 관련성, 접근성, 비교성, 효율성)와 관련된 48개의 품질평가지표를 개발하여 사용하고 있다. 우리나라에는 통계청 이외에도 농림부, 한국은행, 노동부, 보건복지부 등 134개의 국구통계 작성기관들이 있으며, 이들 기관에서 통계품질을 높이기 위한 노력을 기울이고 있으나, 아직 체계적인 방법으로 통계품질관리를 실시하고 있지는 못하다. 따라서 통계가 국가운영에 매우 중요한 만큼 통계작성과정에서 단계별로 품질을 관리할 수 있도록 상세한 국가적인 통계품질 가이드라인(매뉴얼)을 작성하여 이들 기관들에 제공하는 것은 매우 바람직하다고 하겠다. 이러한 매뉴얼은 통계품질관리의 표준화는 물론 통계작성기관이 자체 품질관리를 실시할 수 있도록 유도하는데도 활용될 수 있을 것이다(2003, 박성현).

이기재(2005)는 “보고통계 품질향상을 위한 품질평가지표개발” 논문에서 최근 들어 관심이 높아지고 있는 행정자료를 활용한 보고통계의 품질에 대해 연구하였다. 보고통계는 우리나라 승인통계 중에서 많은 비중을 차지하고 있으며 자료수집 과정이 조사통계와 현저히 다른 특성이 있다. 따라서 보고통계의 품질을 높이기 위해서는 보고통계의 특성에 맞는 품질관리방안이 필요하다. 그동안 통계품질에 대한 논의는 주로 조사통계를 중심으로 진행되어 왔으며 보고통계에 대해서는 활발하게 이루어지지 않았다. 보고통계는

통계작성을 위한 조사원을 별도로 확보할 필요 없이 기존의 행정조직망을 통해 데이터를 수집한다는 면에서 조사통계와 차이를 보인다. 그리고 대상 집단 전체를 파악함으로써 소지역별 통계를 작성할 수 있는 등 전수조사의 장점을 가지고 있으며 대상응답자에게 별도의 응답부담을 지우지 않아도 된다는 특성도 지닌다. 또한 표본조사가 아니므로 표본설계나 표본이론에 근거한 추정 등의 과정이 불필요하다. 반면에 조사통계와는 달리 통계에 대한 이해가 부족하고 전문성이 떨어지는 행정조직의 담당자들을 통해 자료수집이 이루어지므로 다양한 유형의 오차가 발생할 가능성이 있다. 이와 같이 보고통계는 조사통계와 구분되는 뚜렷한 특성을 지니고 있으므로 보고통계의 특성에 부합되는 품질지표들을 개발하여 보고통계의 품질을 진단하는 것이 바람직하다(2005, 이기재).

## 제3절 해외 사례

### 1. 해외 품질 진단 사례

#### 가. 캐나다

캐나다 통계청의 통계품질관리는 1997년 일반감사 국(Office of the Auditor-General)이 설립됨으로서 본격화되었다. 품질담당은 품질 관련 위원회(Subject matter committee)에서 담당하였으며 일반감사 국의 설립은 통계품질의 정의, 관리절차 등 품질관리에 대한 관심을 불러일으키는데 기여하였다. 캐나다 통계청의 품질관리가 체계화된 계기는 Gordon Brackstone의 “통계 작성기관의 품질관리(Managing data quality in a statistical agency)”라는 보고서의 작성으로서, 캐나다 통계청은 이를 통계품질의 기본 프레임으로 채택하게 되었다.

보고서는 6개의 품질차원과 5개의 품질관리 시스템을 제시하고 있으며, 차원 정의, 차원간의 관계, 품질차원과 품질관리 시스템과의 연관성 등을 기술하였다. 또한 캐나다 통계청은 통계품질 전반을 제고하기 위하여 일반이론, 실무경험 및 지식 등에 기초하여 통계품질가이드라인을 작성, 발간하여 지침서로 활용하고 있다. 통계품질가이드라인(일반지침서)은 기획, 조사표 설계, 자료처리, 품질평가, 결과분석, 자료제공 등 통계작성 단계별 업무수행상 필요한 기본사항과 체크리스트를 수록하고 있다.

#### 나. 영국

영국은 통계조사 기법의 다양화 및 품질에 대한 관심고조에 부응하기 위해 통계품질



관리제도를 2000년부터 도입하여 사용하고 있다. 이러한 제도는 “통계의 신뢰성 구축”이라는 정부백서에 목표로 설정되어 있다. 영국 통계청은 작성 통계의 품질 개선 및 이용자에게 품질 정보 제공 등의 목적으로 5개년 계획(2002~2006)으로 추진하였다. 영국 통계청은 품질 관리 조직으로서 통계청 본부에 품질센터(Quality Centre)를 설치하였다. 이 센터에서는 조사방법론 전문가 및 현지조사 업무 경험자 12명이 매년 품질관리 관련 프로젝트를 선정하여 수행하고 있다. 평가 지표 작성, 품질 평가 등 6개 프로젝트를 진행 중이다.

품질관리의 내용으로서는 품질 표준 및 협약 제정 및 보급, 통계자료 산출을 위한 신 조사방법론 개발, 다목적 조사기법 및 통합시스템 개발, 프로젝트 관리, 위기관리 등 효과적인 리더십 및 관리 체계 구축, 메타 자료 및 통계작성 절차 문서화 시스템 구축(통계조사표 설계, 표준 및 지침 데이터베이스 개발), 품질 모니터 및 개선(조사 진행 평가, 조사 방법 평가, 벤치마킹, 자료 품질 평가 등)이다.

#### 다. 스웨덴

스웨덴은 분산형 통계제도를 도입한 나라이다. 스웨덴 통계청은 통계업무의 효율성을 제고하고 대외 경쟁력을 향상시키고 이용자들의 품질정보 제공 요구에 대응하기 위해 또한 최고 관리자의 품질에 대한 관심 제고 등의 계기로 인해 통합품질관리(TQM) 제도를 도입하게 되었다.

품질관리 조직으로는 통계국장실, 국장급 위원회, 품질 위원회, 가상품질 사무소, 전문가 그룹(6명의 최고 전문가 구성), 통계 기관별 품질 촉진 담당 망 조직(25명)으로 이루어져 있다.

#### 라. 미국

미국은 실업률 통계를 작성하는 노동력 조사(CPS, Current Population Survey)에서 발생할 수 있는 문제점을 연구 분석한 품질 보고서 작성을 시작으로 품질에 대한 연구를 시작하였다. 미국은 70개의 통계 작성 기관, 10개의 부서, 8개의 독립기관에서 통계를 작성하고 있는 분산형 제도를 운영하고 있어 각 부서마다 통계작성 가이드라인을 작성하여 개별 품질관리를 실시하고 있다.

예산을 담당하는 OMB(Office of Management and Budget)에서 “Federal Committee on

statistical methodology"를 설치하여 지난 25년간 연방정부의 통계조사에 대한 방법론적 논의에 대하여 중심적 역할을 담당하며 통계 조정 및 품질 제고를 담당하고 있다. 가장 대표적인 기관은 센서스 국이며, 생산되는 제품의 품질을 확인, 조사 전반에 걸친 계획과 집행, 사전에 위험 방지 등을 목적으로 품질보증/품질통제(QA/QC) 네트워크를 설립, 품질 체크리스트 작성 등 다양한 방법으로 운용하고 있다.

#### 마. EUROSTAT

Eurostat는 유럽연합(EU)에 고품질 통계정보 서비스의 제공을 목적으로 하고 있으며, 주로 국가 간 비교 가능한 통계자료 제공, 표준화된 통계시스템 및 통계프로그램 개발 등을 수행한다. Eurostat는 1953년 석탄 및 철강공동체(Coal and Steel Community)에서 제기되는 문제 해결을 위하여 정식으로 발족되었으며, 1958년 유럽공동체(European Community)의 성립과 함께 유럽위원회(European Commission) 산하 5개 집행위원회(Directorate-General, DG)의 하나로서 활동하였다.

고품질 통계를 위한 법령 제정을 위해 Eurostat는 1997년 회원국과의 통계 협력 제고 및 고품질 통계 서비스와 관련한 제반 업무의 효과적인 수행을 위하여 Treaty of Amsterdam, Statistical Law, The Community Statistical Authority 등 3개 주요법령(three major acts)을 제정하였다. 3개 주요 법령은 통계의 공정성(impartiality), 전문적 독립성(independence), 비밀 보호(statistical confidentiality), 국가정책과 통계관리(작성)의 분리, Eurostat 및 회원국의 통계 책무, 통계작성의 기본요건, 절차, 공식통계의 제공 등을 규정하고 있으며, 통계의 기본원리로서 통계의 신뢰성, 투명성, 지속성, 객관성, 비용 효과성 등에 관하여 규정하고 있다.

통계품질사업으로서 유럽공동체(EU)는 European Statistical System(ESS)을 구축하여, 관련국가 및 UN, OECD, IMF 등과의 통계협력을 추구하고 있다. 통합유로통계시스템(ESS)은 Eurostat가 주재하는 SPC(Statistical Programme Committee)에 의해 추진되고 있으며, 매5년 마다 사업계획(five-year programme)을 수립하여 표준화 된 통계분류, 방법론, 정의, 자료수집, 분석, 제공, 통합된 회계 시스템, 통계자료의 공유(Intrastat), 지표 개발 등의 사업을 추진하고 있다. SPC의 추진 사업 중 통계품질프로그램 개발은 스웨덴의 제안에 따라 1999년 창설한 전문가 그룹(Leadership Expert Group)에 의해 실무적으로 추진되고 있다.



## 2. 측정 가능한 품질 지표 연구

### 가. EUROSTAT

EUROSTAT이 추진한 연구의 목적은 European Statistical System에서 시간에 따라 생산되는 자료의 품질을 측정할 수 있는 제한된 지표의 집합을 가지는 데 있다. 그 지표는 EUROSTAT의 품질 정의에 정의된 품질 차원을 표현하여야 하며 또한 그 지표들을 계산하는 방법론은 잘 구축되어야 하며, 쉽게 설명할 수 있어야 한다. 지표를 정의할 때에는 기준(criteria)을 준수하지 않은 품질에 대한 결과물의 차원에 대해서도 측정가능하게 하기 위해서 모든 지표들이 시험적으로 고려되어야 된다는 것을 알아냈다. 따라서 지표들은 Key Indicators, Supportive Indicators, Indicators for further Indicator로 분류되어야 한다(<표 5-5>). Key Indicators는 기준(criteria)을 무조건 만족시켜야 할 의무가 있다. Supportive Indicators는 데이터 품질을 간접적으로 측정하기 때문에 중요하게 여겨지는 만큼 기준을 만족시켜야 한다. 나머지 세 번째 Indicators들은 심층적인 경험과 조사기관들 간의 토론에 목적이 있다. 서로 다른 주제 영역에 대한 통계의 품질을 모니터링하기 위한 스코어보드를 정의할 때는 지표의 목록이 고려되어야 하며, 지표들이 연관성이 있어야 하고 특정 영역이 선택되도록 구현되어야 한다. 서로 다른 품질의 영역의 올바른 균형을 유지하고 최소한의 지표수를 유지하기 위해 세심한 관심이 요구된다. 어떤 지표에 대한 이상한 값(Strange Values)은 통계 생산 과정(Statistical production process)의 실제 문제에 대한 심층적인 조사로 인해 결론지어진다. 이러한 메타데이터의 정보는 품질 보고서에 항상 담겨 있어야 한다. 사용자들에게 통계의 품질에 대해 알려주기 위해 지표를 사용할 때에는 Qualitative statements, 품질 정보를 설명하는데 필요한 지원, 통계 사용의 주요 효과에 대한 요약을 포함할 것이 요구된다(2005, EUROSTAT).

<표 5-5> 품질 영역과 지표

Quality Component	Indicator	Category
Relevance	R1. User satisfaction index	3
	R2. Rate of available statistics	1
Accuracy	A1. Coefficient of variation	1
	A2. Unit response rate (un-weighted/weighted)	2
	A3. Item response rate (un-weighted/weighted)	2
	A4. Imputation rate and ratio	2
	A5. Over-coverage and misclassification rates	2
	A6. Geographical under-coverage ratio	1
	A7. Average size of revisions	1

Timeliness and Punctuality	T1. Punctuality of time schedule of effective publication	1
	T2. Time lag between the end of reference period and the date of first results	1
	T3. Time lag between the end of reference period and the date of the final results	1
Accessibility and clarity	AC1. Number of publications disseminated and/ or sold	1
	AC2. Number of accesses to databases	1
	AC3. Rate of completeness of metadata information for released statistics.	3
Comparability	C1. Length of comparable time-series	1
	C2. Number of comparable time-series	1
	C3. Rate of differences in concepts and measurement from European norms	3
	C4. Asymmetries for statistics mirror flows	1
Coherence	CH1. Rate of statistics that satisfies the requirements for the main secondary use	3

\* Category : 1 = Key, 2 = Supportive, 3 = Further Experience

## Relevance

### R2. Rate of available statistics

#### Calculation Method

$$\text{rate of available statistics} = \frac{\text{Number of values provided In concrete data set}}{\text{Number of fields for which data has to be provided}}$$

## Accuracy

### A2. Unit response rate

#### Calculation Method

- Un-weight unit response rate = (Number of respondent units used in estimation) / (Number of in-scope units + number of eligibility unknown units)
- Weight unit response rate = (Number of weighted respondent units in estimation) / (Number of weighted in-scope units + number of weighted eligibility unknown units)



#### A4. Imputation rate and ratio

##### Calculation Method

Un-weighted imputation rate = (Number of imputed records for the variable X) /  
(Total Number of records)

Weighted imputation ratio = (Total Weighted quantity for imputed values) /  
(Total weighted quantity for all final values)

제공되는 통계적인 정보에 대한 포괄적인 품질을 평가하기 위한 한 개의 합성 지표를 개발하는 것에 대한 다양한 의견이 존재한다. 이러한 지표가 개발된다면 각각의 기관에 제공되는 통계의 종합적인 품질을 평가하는데 도움이 되며 모든 기관에 제공되는 종합적인 품질 정보를 비교하기 위해 결집하는데 도움이 된다. 기술적인 면에서 보면, 지표가 어떤 한 나라의 포괄적인 품질을 평가하는 것은 가능해 보이며 심지어는 각각의 관계되는 가중 품질 지표를 개발함으로써 쉬워 보이기도 한다. 그러나 개념적이고 실제적인 관점에서 보면 EU와 같이 산재된 곳의 포괄적인 통계 품질을 평가하는 하나의 합성지표를 제안하는 것은 쉽지 않다. 그 이유는 나라 간의 많은 차이점이 존재하고 모든 조사 단계에 있어 미리 정의된 표준이라는 것이 존재하지 않기 때문이다. 또한 어떤 표준이 존재하면 항상 따르지는 않기 때문이다. 그러므로 각국은 그들의 요구에 의존된 각각의 품질 요소를 평가하기 위한 서로 다른 가중치를 적용해야 한다. 그리고 물론 각각의 요소의 평가에는 비용과 부담이 요구된다.

#### 나. 영국통계청

Guidelines for measuring statistical quality(version 3.1)는 통계적인 산출물의 품질을 측정하거나 보고할 때 사용되는 품질 측정값들과 지표의 체크리스트를 제공한다. 이것들은 국가통계(National Statistics)는 아니며 통계적인 산출물의 생산 전반에 품질을 측정하는데 가장 좋은 방법을 표현하고 있다. 이 새로운 버전에는 비밀보호 기법이 추가됨으로 인한 정보 손실에 대한 측정값이 새로 추가되었다. 이 문서의 목적은 Government Statistical Service(GSS) 전반에 걸쳐 품질을 측정하거나 보고하는데 필요한 표준화된 접근법을 제안하는데 있다(2007, Karen).

GSS는 사용자들에게 통계적인 산출물을 수집하고 편집하는데 사용되는 기법에 대한 정보를 제공할 것을 위임받았다. 이러한 위임에 대해서는 National Statistics code of practice에 표현되어 있다.

‘National Statistics를 생산하는 과정과 기법은 사용자가 각각의 목적에 적합하게 접근할 수 있을 정도로 충분히 세밀해야 한다.’

여기에 더해 GSS는 National Statistics Quality Strategy에 자신들의 목적이 통계적 산출물의 품질에 대한 정보를 제공하는 것임을 공표하였다. ‘National Statistics에 대한 품질의 측정은 시스템적으로 보고되어야 하며, 사용자가 사용하고자 하는 그들의 분야에서 알맞은 것인지에 대한 평가를 할 수 있어야 한다.’ 아래 [그림 5-1]은 문서에 포함된 품질 차원이다. 품질 차원별로 측정되는 값들이 잘 정리되어 있다.

Table A.1 Dimensions of quality

Definition	Key components
<b>1. RELEVANCE</b> The degree to which the statistical product meets user needs for both coverage and content.	Any assessment of relevance needs to consider: <ul style="list-style-type: none"> <li>• who are the users of the statistics;</li> <li>• what are their needs; and</li> <li>• how well does the output meet these needs?</li> </ul>
<b>2. ACCURACY</b> The closeness between an estimated result and the (unknown) true value.	Accuracy can be split into sampling error and non-sampling error, where non-sampling error includes: <ul style="list-style-type: none"> <li>• coverage error;</li> <li>• non-response error;</li> <li>• measurement error;</li> <li>• processing error; and</li> <li>• model assumption error.</li> </ul>
<b>3. TIMELINESS AND PUNCTUALITY</b> Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the time lag between the actual and planned dates of publication.	An assessment of timeliness and punctuality should consider the following: <ul style="list-style-type: none"> <li>• production time;</li> <li>• frequency of release; and</li> <li>• punctuality of release.</li> </ul>
<b>4. ACCESSIBILITY AND CLARITY</b> Accessibility is the ease with which users are able to access the data. It also relates to the format(s) in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and accompanying advice.	Specific areas where accessibility and clarity may be addressed include: <ul style="list-style-type: none"> <li>• needs of analysts;</li> <li>• assistance to locate information;</li> <li>• clarity; and</li> <li>• dissemination.</li> </ul>
<b>5. COMPARABILITY</b> The degree to which data can be compared over time and domain.	Comparability should be addressed in terms of comparability over: <ul style="list-style-type: none"> <li>• time;</li> <li>• spatial domains (e.g. sub-national, national, international); and</li> <li>• domain or sub-population (e.g. industrial sector, household type).</li> </ul>
<b>6. COHERENCE</b> The degree to which data that are derived from different sources or methods, but which refer to the same phenomenon, are similar.	Coherence should be addressed in terms of coherence between: <ul style="list-style-type: none"> <li>• data produced at different frequencies;</li> <li>• other statistics in the same socio-economic domain; and</li> <li>• sources and outputs.</li> </ul>

[그림 5-1] 품질 차원

주요 품질 측정값들(KQMs : Key Quality Measures)은 사용자에게 전반적인 품질 결과에 대한 가장 중요한 정보를 제공해주는 지표들에 대한 품질 측정값이다([그림 5-1]). 더 나아가 KQMs는 품질에 대한 성능을 모니터링하고 통계적 결과물의 품질 개선에 대한 관리 정보를 제공하여 준다.





Table A.2 Key Quality Measures

	KEY QUALITY MEASURE	ESS QUALITY DIMENSION	GUIDELINES REFERENCE
1	Where possible, describe how the data relate to the needs of users	Relevance	B1.3
2	Provide a statement of the nationally/internationally agreed definitions and standards used	Comparability	B1.13
3	Unit response rates by sub-groups, weighted and unweighted	Accuracy	B3.4 (Household surveys) B3.5 (Business surveys)
4	Key item response rates	Accuracy	B3.7
5	Total contribution to key estimates from imputed values	Accuracy	B4.7
6	Editing rate (for key items)	Accuracy	B4.11
7	Estimated standard error for key estimates	Accuracy	B5.2 (for key estimates of level) B5.3 (for key estimates of change)
8a	Time lag from the reference date/period to the release of the provisional output	Timeliness and Punctuality	B8.1
8b	Time lag from the reference date/period to the release of the final output	Timeliness and Punctuality	B8.2
9	Estimated mean absolute revision between provisional and final statistics	Accuracy	B8.21
10	Compare estimates with other estimates on the same theme	Coherence	B8.28
11	Identify known gaps between key user needs, in terms of coverage and detail, and current data	Relevance	B8.29

[그림 5-2] 주요 품질 측정값들

[그림 5-3]은 정확성에 대한 주요 품질 측정값의 한 예시이다. 사업체 조사에서의 하위 그룹별 응답률에 대한 측정값이다. 각각의 응답률에 대한 비율을 계산하여 가중값을 계산하고 있다. 이 지표는 무응답 오차가 얼마나 유의한가를 보여주는 지표가 된다. 정교한 계산법으로 품질을 측정하고 있음을 알 수 있다.

[그림 5-4]는 시의성에 대한 항목에서의 품질 측정값의 한 예시이다. 데이터가 얻어진 시간과 통계 조사 결과가 공표된 날 사이의 시차를 측정한 예시이다. 사용자가 얼마나 빨리 조사된 통계를 사용할 수 있는가에 대한 중요성을 알려준다.

Ref.	Notes	Example
B3.5 ①	<p><b>Unit response rate by sub-groups: business surveys.</b></p> <p>The response rate is a measure of the proportion of sampled units who respond to a survey. This indicates to users how significant the non-response bias is likely to be. Where response is high, non-response bias is likely to be less of a problem than when there are high rates of non-response. The rates opposite can be expressed as a percentage figure by multiplying by 100.</p> <p>NB: There may be instances where non-response bias is high even with very high response rates, if there are large differences between responders and non-responders.</p>	<p align="right"><b>(Key Quality Measure)</b></p> <p><b>Overall response rate:</b> Unweighted: <math display="block">\frac{FC + FP + PC + PP}{(FC+FP+PC+PP) + RNU + NR + e(U)}</math>Weighted: <math display="block">\frac{\sum_{i \in R_o} w_i x_i}{\sum_{i \in S} w_i x_i}</math></p> <p><b>Full response rate:</b> Unweighted: <math display="block">\frac{FC}{(FC+FP+PC+PP) + RNU + NR + e(U)}</math>Weighted: <math display="block">\frac{\sum_{i \in R_F} w_i x_i}{\sum_{i \in S} w_i x_i}</math></p> <p><b>Non-response rate:</b> Unweighted: <math display="block">\frac{RNU + NR + e(U)}{(FC+FP+PC+PP) + RNU + NR + e(U)}</math>Weighted: <math display="block">\frac{\sum_{i \in N_r} w_i x_i}{\sum_{i \in S} w_i x_i}</math></p> <p>Where: FC = full period return with complete data; FP = full period return with partial data; PC = part period return with complete data; PP = part period return with partial data; RNU = returned but not used; NR = non-response; U = unknown eligibility; I = ineligible (out of scope); e = estimated proportion of cases of unknown eligibility that are eligible. e can be estimated as: <math display="block">e = \frac{(FC+FP+PC+PP) + RNU + NR}{(FC+FP+PC+PP) + RNU + NR + I}</math></p> <p><math>w_i</math> = weight for unit <math>i</math>; <math>x_i</math> = value of auxiliary variable for unit <math>i</math>; <math>R_o</math> = set of all responders; <math>S</math> = set of all eligible sampled units, <math>R_F</math> = set of full responders; <math>N_r</math> = set of non-responders (note that <math>N_r = S - R_o</math>).</p>

[그림 5-3] 정확성 관련 무응답 품질 인자

Ref.	Notes	Example
B8.1 ①	<p><b>Time lag from the reference date/period to the release of the provisional output.</b></p> <p>This indicates whether provisional outputs are timely with respect to users' needs.</p>	<p align="right"><b>(Key Quality Measure)</b></p> <p>The provisional output is published six weeks after the last day of the reference period.</p>
B8.2 ①	<p><b>Time lag from the reference date/period to the release of the final output.</b></p> <p>This indicates whether final outputs are timely with respect to users' needs.</p>	<p align="right"><b>(Key Quality Measure)</b></p> <p>The final output is published eight weeks after the last day of the reference period.</p>

[그림 5-4] 시의성 중요 품질 인자 예시

### 다. 이탈리아 통계청

유럽과 NSI수준에서의 질적 측정에 대한 품질조사 품질이 평가되어 왔다. 개념적으로는 이러한 질적 측정이 EUROSTAT의 품질component에 관련이 되어 있지만 지표들 간의 품질component 간에는 연관성은 완전히 조사되지 않았다. 이 연구에서는 표준품질 지표들이 품질Component와 조사품질 전반적인 것을 잘 반영하고 있는지를 연구하고자 하였다(2008, Giovanna). 구조 방정식 모형(SEM : Structural Equation Modeling)은 품질 지표와 같이 관측된 변수와 품질 또는 품질 성분과 같은 관측되지 않은 변수사이의 관계를 분석하는 방법론적 접근 방법이다. 여기서 품질 또는 품질 성분은 이론적으로 직접적으로 측정 불가능한 요인을 말한다. 따라서 품질 지표와 같은 변수를 통해 간접적으로



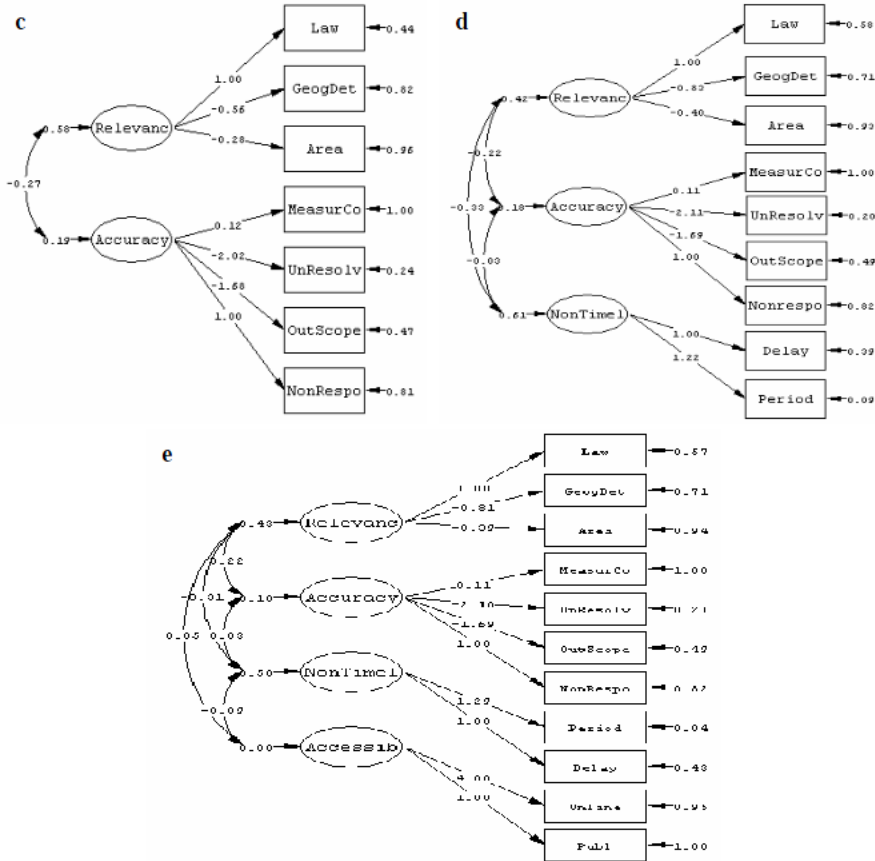
관측하여야 한다. 이 연구에서 품질은 하나 또는 하나 이상의 표준품질 지표의 수준에 영향을 받는 것으로 표현됨을 가정하고 있다. 다른 말로 품질은 하나 또는 그 이상의 품질 지표들 간의 공분산으로부터 나온 측정되지 않은 품질 성분에 의해 표현됨을 가정한다. [그림 5-5]는 분석에 사용된 EUROSTAT의 품질 인자와 이탈리아 통계청의 품질 인자이다.

Table 1. List of latent factors (quality components) and manifest variables (quality indicators) for different data frameworks				
First-order latent components	Manifest variables (Quality indicators)	Eurostat data framework	Istat theoretical data framework	Istat practical data framework
Relevance (and completeness)	User satisfaction index	<input checked="" type="checkbox"/>		
	Rate of available statistics	<input checked="" type="checkbox"/>		
	Law		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Area of interest			<input checked="" type="checkbox"/>
Accuracy	Geographical detail		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	f(variation coefficient)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Frame out-of-scope rate	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Resolved unit rate		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Unit response rate	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	f(item response rate)	<input checked="" type="checkbox"/>		
	Imputation rate	<input checked="" type="checkbox"/>		
	Net imputation rate		<input checked="" type="checkbox"/>	
	Modification rate		<input checked="" type="checkbox"/>	
	Average size of revisions	<input checked="" type="checkbox"/>		
	Measurement error			<input checked="" type="checkbox"/>
Timeliness (and punctuality)	Timeliness	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Punctuality	<input checked="" type="checkbox"/>		
	Periodicity		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Accessibility (and clarity)	Number of publications disseminated	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Number of accesses to DB	<input checked="" type="checkbox"/>		
	Metadata rate of completeness	<input checked="" type="checkbox"/>		
	On-line dissemination		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Coherence and comparability	Length of comp. time series	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	Rate of statistics satisfying requirements for the main secondary use	<input checked="" type="checkbox"/>		
	Coherence provisional vs. final		<input checked="" type="checkbox"/>	
	Coherence with external sources		<input checked="" type="checkbox"/>	

[그림 5-5] SEM 분석을 위한 품질 인자

[그림 5-6]은 SEM 분석 결과를 나타낸다. 위의 분석 결과로서 Relevance와 Accuracy의 품질 관계가 어떻게 이루어지는지 알 수 있다. 분석 결과로서 우리는 품질 지표들이 표현하기 굉장히 복잡함에도 불구하고 품질 영역을 잘 표현하고 있음을 알 수 있다.

Figure 6 Relevance-Accuracy, Relevance-Accuracy-Timeliness and Relevance-Accuracy-Timeliness-Accessibility models



[그림 5-6] SEM 분석 결과 도식

### 제4절 통계청 승인통계 품질 관리

통계청에서 하고 있는 품질관리에 대한 분석을 하기 위해서는 기본적으로 통계청 승인 통계 품질 관리의 절차와 프레임에 대해 알아볼 필요가 있다. 품질관리과에서는 품질관리기반, 이용자 만족도 및 요구사항 반영 실태, 세부 작성절차별 체계, 수집 자료의 정확성, 통계자료 서비스의 충실성의 5가지 단계를 통해 관리를 하고 있다. 각 분야의 전문가를 통해 이러한 작성 단계별로 정보를 수집하여 분석하고 개선 사항을 만들어 반영하여 보고서를 작성하는 것을 목표로 한다. 이러한 단계별 작성 과정을



분석하여 품질에 영향을 미치는 요인들이 정확하게 측정되고 있는지 품질 평가에 있어 누락된 요인들은 없는지를 찾아서 개선하는 것이 연구의 목적이라고 할 수 있겠다. 따라서 이 장에서는 전반적인 품질 관리의 단계별 내용과 성격에 대해 알아보고 다음 장에서는 실제 평가되고 보고된 보고서를 분석하여 세부 개선사항에 대해 중점적으로 알아보고자 한다. 이 절의 내용은 통계청 품질관리 홈페이지 내용을 참고하였음을 밝힌다.

## 가. 품질관리기반

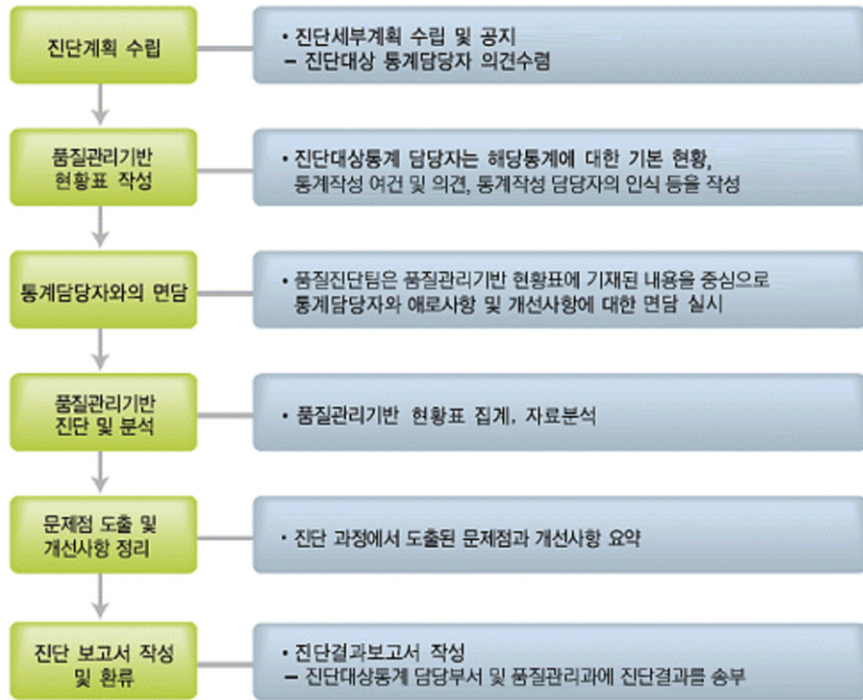
통계품질에 영향을 미치는 요인으로 가장 먼저 통계작성 환경을 들 수 있다. 리더십과 인적자원 관리 등이 이에 해당되며, 이들은 통계를 생산하는 데 필요한 기본 인프라이다. 따라서 본격적인 품질진단에 앞서 기관장의 관심, 전략과 방침, 인력자원 관리 및 예산규모, 통계작성 환경 등을 파악하여 품질관리기반을 진단하는 것이 필요하다. 품질관리기반 진단의 실제 업무는 첫째 세부 진단계획을 수립하고, 둘째 해당통계 담당자에게 품질관리기반 현황표 작성을 의뢰한 후, 셋째 면담을 통해 조사된 내용을 다시 확인하고, 넷째 그 결과를 분석하여 작성환경을 진단하는 과정으로 이루어져 있다. 이 중에서 가장 중요한 과정은 품질관리기반 현황표를 작성하는 일이다. 왜냐하면 통계작성 담당자로부터 절실한 문제점과 의견이 제시되지 않으면 작성환경을 파악하기는 힘들기 때문이다.

### 1) 진단계획 수립

품질진단팀은 진단 시기, 진단에 사용할 현황표, 진단방법 및 절차 등을 포함하여 품질관리기반 진단을 위한 세부계획을 수립하여야 한다. 그리고 이 세부계획을 진단대상 통계담당자에게 공지하여 품질관리기반 현황표 작성에 협조를 받는다.

### 2) 품질관리기반 현황표 작성

품질진단팀은 진단대상통계 담당자에게 품질관리기반 현황표 작성을 요청한다. 현황표는 해당통계에 대한 기본 현황, 통계작성 여건 및 의견, 통계작성 담당자 인식 등의 항목으로 구성되어 있다. 응답자가 쉽게 작성할 수 있도록 작성요령도 함께 배포한다. 품질진단팀은 해당 통계의 작성현황을 사전에 확인하고, 이를 바탕으로 통계담당자가 작성한 품질관리기반 현황표 내용을 검토한다.



[그림 5-7] 품질관리기반 진단 흐름도

### 3) 통계담당자와의 면담

현황표 작성이 완료되면 통계 담당자와의 면담을 요청한다. 면담 시에는 현황표에 기입된 내용을 확인하고 그 근거자료를 수집한다. 특히 조직의 리더십과 방침, 목표, 프로세스에 관한 자료를 수집하여 다른 부문의 진단 시에 활용하여야 한다. 왜냐하면 품질관리기반은 이용자만족도 및 요구사항 반영실태, 세부 작성절차별 체계, 수집자료의 정확성, 통계자료 서비스에 직접적으로 영향을 미치기 때문이다.

### 4) 품질관리기반 진단 및 분석

현황표의 내용 중 해당통계에 대한 제1부 기본현황과 제2부 통계작성 여건 및 제4부 면담 내용들은 집계하여 다른 부문의 진단에 활용한다. 그리고 제3부 통계작성 담당자의 인식 항목들은 결과를 집계하여 5점 척도로 평가한 후 품질관리기반에 대한 평점으로 활용한다.

## 5) 문제점 도출 및 개선사항 정리

진단양식에 제시된 검토의견이나 면접 등을 통해 발견된 문제점을 정리하고 향후에 품질관리기반을 개선하기 위해 필요한 사항을 요약한다.

## 6) 진단보고서 작성 및 환류

통계작성여건 관련 자료, 품질관리기반에 대한 평점 결과, 면담 결과 등을 집계하여 품질관리기반에 대한 진단보고서를 작성하며, 진단대상통계 담당부서의 장에게 보고한 후, 품질개선을 위한 기초자료로 활용하도록 통계담당자에게 환류한다.

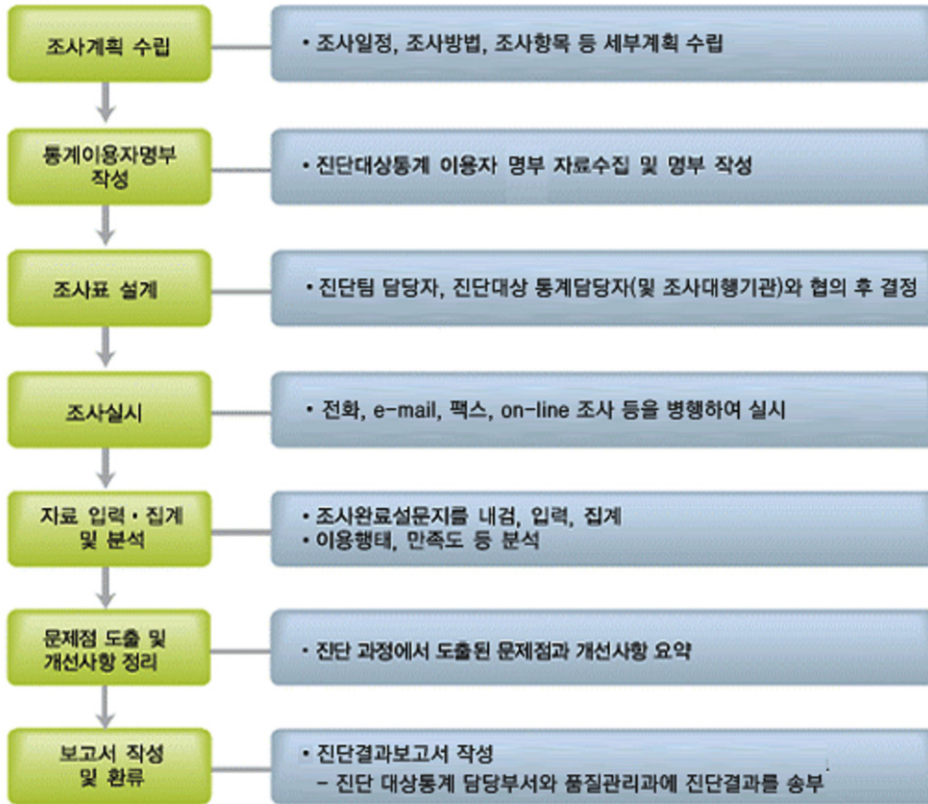
### 나. 이용자 만족도 및 요구사항 반영 실태

통계이용자는 이용하는 통계로부터 기대하는 정보를 충분히 얻기를 원한다. 따라서 통계이용자가 통계자료에 대해 얼마나 만족하는지를 조사하여 품질진단에 이용할 필요가 있다. 통계작성기관은 통계의 전문이용자, 일반이용자를 구분하여 리스트를 확보하고 수시로 이용자가 해당통계를 만족스럽게 이용하는지 확인할 필요가 있다. 이 부분의 진단에서는 통계이용자의 통계에 대한 만족도와 요구사항의 반영정도를 측정하는데 이를 위해서는 해당 통계의 관련 전문가 또는 일반 이용자로 구성된 표적집단면접(Focus Group Interviews)과 통계이용자를 대상으로 하는 전반적인 만족도 조사를 실시한다.

#### 1) 표적집단면접(FGI)

표적집단면접(FGI)은 소집단을 형성하여 특정 주제에 대해서 계획적이면서도 자유롭게 토론하는 형태를 말한다. FGI는 일대일 면접에 비해 정보 획득이 신속하고, 함께 모여 토론함으로써 다른 사람의 아이디어를 힌트로 자신의 생각을 좀 더 확실히 표출하는 시너지 효과가 있어 정성적인 이용자 조사 기법으로 많이 활용된다. 일반이용자(대학생, 대학원생(석.박사), 일반인)를 대상으로 하는 FGI를 통해 이용자의 요구사항, 불만사항 등에 대한 사전 정보를 파악하여 이용자만족도조사 등 이어지는 진단업무에서 어떤 사항을 중점적으로 검토할지에 대한 계획에 활용할 수 있다.

또한 해당 통계에 대한 태도, 인식, 의견 등을 가장 잘 대변해 줄 수 있는 전문이용자(정책고객, 교수, 연구원 등)를 대상으로 하는 FGI를 통해 해당 통계의 문제점과 개선 아이디어를 수집할 수 있다.



[그림 5-8] FGI 흐름도

## 2) 이용자만족도조사

이용자만족도조사는 통계 이용자를 대상으로 실시하는데, 이용자로부터 해당통계에 대한 만족도와 통계에 대한 개선의견 등을 조사한다. 이는 통계생산자가 간과하고 있는 통계이용자들의 요구사항을 통계작성 및 자료제공 과정에 반영하기 위한 것이다. 이 조사 자료로부터 파악되는 통계이용자의 만족도는 이용자적합성(Fitness for Use)에 바탕을 두고 이용자가 직접 진단한 평가점수이다. 아울러 통계이용실태 자료와 개선요구사항 등은 통계개선과제를 발굴하는 기초자료로 활용된다.

이용자만족도조사는 우선 조사를 위한 세부계획을 수립하고 조사대상인 이용자 명부를 작성한 후, 조사표를 설계하여 조사를 실시하고, 조사 자료를 입력·집계하여 통계에 대한 만족도와 다양한 이용실태를 분석하는 것이다.





분류	항목 구성
통계이용실태	1. 통계자료 이용빈도 2. 통계자료 이용목적 3. 통계자료 이용자료 형태 4. 통계자료 이용경로
통계 만족도 (12개 항목)	5-1. 통계자료 공표시기 적절성 5-2. 통계자료 예고 공표 일정 준수 5-3. 통계자료 검색의 용이성 5-4. 통계자료 시계열 비교 편리성 5-5. 통계자료 국가 간 비교 편리성 5-6. 통계자료 이용시 유의사항/관련개념, 용어 정의 제공 5-7. 통계자료의 제공정보 충분성 5-8. 통계자료의 제공정보 다양성 5-9. 통계자료의 신뢰성 5-10. 원시자료 이용의 편리성 5-11. 지출비용에 대한 만족도 5. 통계에 대한 전반적인 만족 수준(전반적 만족도)
재이용 의사	6. 향후 재이용 의사
품질 변화	7. 통계의 품질 변화
개선 의견	8. 개선 의견(개방형)
* 응답자 특성	성별, 연령, 직업, 통계활용비중

[그림 5-9] 이용자만족도조사 흐름도

### 다. 세부 작성절차별 체계

세부 작성절차별 체계 진단은 통계작성 과정이 본래 목적을 실현하기 위해 적합하게 이루어지고 있는가를 세부 작성절차별 진단서를 이용하여 진단하는 것이다. 진단서는 통계 작성과정 중 통계품질에 영향을 미치는 지표들을 5점 척도의 질문들로 구성된 진단도구로서 통계작성과정에 따라 조사통계용, 보고통계용, 가공통계용의 3가지로 구분된다. 진단작업은 통계작성 담당자가 직접 수행하는 내부진단과 외부전문가가 실시하는 외부진단 두 가지를 실시하게 되는데, 외부진단은 진단대상 통계를 주로 이용하는 외부전문가를 진단위원으로 위촉하여 내부진단 결과와 근거자료를 바탕으로 재진단함으로써 이루어진다. 외부진단 시 진단위원은 대상통계의 작성절차에 대해 개선할 사항을 중심으로 품질개선 의견서도 함께 제출한다. 진단팀은 통계담당자 및 외부진단위원에 의한 진단서를 정리하여 보고서를 작성한 후 품질개선의 기초자료로 활용하도록 해당 통계의 담당자에게 환류하여야 한다. [그림 5-10]에서 이러한 업무흐름을 한 눈에 알아 볼 수 있다.



[그림 5-10] 세부 작성절차별 진단 흐름도

통계의 품질을 측정하려면 이를 위한 도구가 있어야 한다. 통계품질지표가 이러한 도구에 해당한다. 통계품질지표는 품질수준을 측정할 목적으로 통계를 작성하는 절차마다 통계품질에 영향을 주는 요소를 발굴·선정하여 구성한다. 일반적으로 통계작성절차를 점검할 수 있는 지표, 기술적으로나 경제적으로 측정에 무리가 없는 지표, 향후 품질개선 방향을 제시할 수 있는 지표 등을 우선적으로 선정한다. 이러한 품질지표는 반드시 측정가능하고 객관적이어야 한다. 품질지표들이 사실에 근거하여 진단되어야 하기 때문에 각 품질지표에 별도의 세부 품질요소들을 제시하였다. 또한 품질요소 항목들의 이행 여부를 표기토록 함으로써 품질진단서가 기본적인 통계작성절차를 준수하는지 여부를 점검하는 체크리스트로 활용될 수 있도록 하였다. 통계품질지표들을 모아 질문 형식으로 구성한 것이 ‘세부 작성절차별 진단서’이다. 진단서에는 선정된 품질지표들이 통계작성 절차별로 배치되어 있으며 통계종류에 따라 조사통계용, 보고통계용 및 가공통계용의 3종류가 있다. 세부 작성절차는 통계작성 기획, 통계 설계, 자료수집, 자료 입력 및 처리, 자료 분석 및 품질평가, 문서화 및 자료제공, 사후관리 등 7개 과정으로 구분한다.



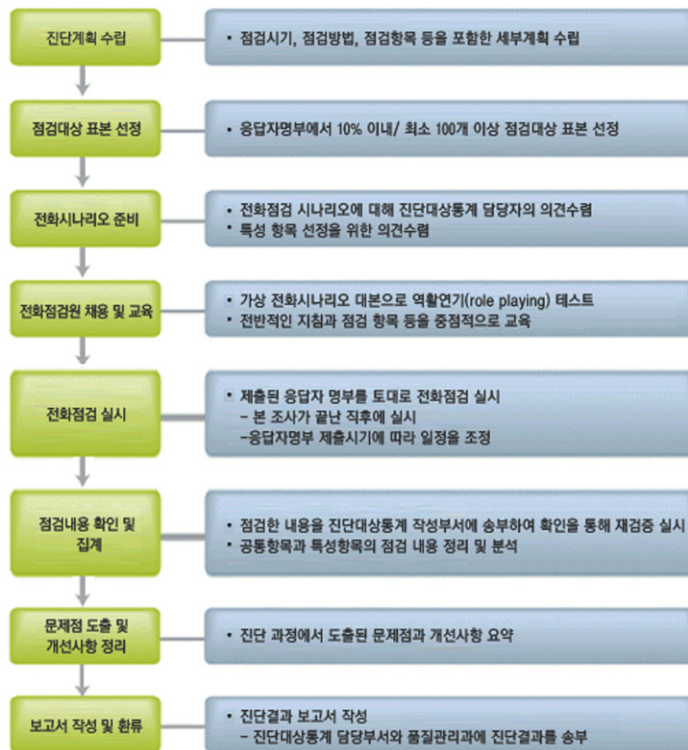
## 라. 수집 자료의 정확성

통계자료가 얼마나 정확한가는 수집된 자료가 얼마나 정확한가에 달려 있으며, 이는 조사나 보고 등 자료가 수집되는 시스템의 효율성에 의해 좌우된다. 이 진단 절차는 수집 자료의 정확성과 절차적 오류는 없는지에 대해 진단하는 방법을 설명하는데 이에 대한 점검과정은 통계품질을 결정하는 매우 중요한 과정이다. 조사통계의 경우에는 현장 조사의 정확성 진단을 하는 방법, 보고통계는 자료수집과정에서 이루어지는 다양한 보고체계의 형태와 최초 보고자에서 나타날 수 있는 자료수집의 오류가능성, 가공통계에서는 통계작성에 투입되는 자료의 선정과정과 최종결과가 산출되기까지의 과정을 분석하여 이에 대한 문제점 및 개선방안을 도출하여 자료수집 과정에서의 품질을 높이는 데 목적이 있다. 만약 진단대상 통계가 조사, 보고, 가공통계로 명확히 구분이 어려운 경우에는 나타난 형태의 자료수집과정을 파악하여 현장업무 수행 내용이 일부라도 누락되어 잘못된 진단 결과가 나오지 않도록 주의한다.

### 1) 조사통계의 현장조사 정확성 진단

조사통계에서는 자료수집이 현장조사를 통해 이루어지므로 수집 자료의 정확성 진단은 현장조사의 정확성 진단으로 고쳐 부를 수 있다. 이 진단은 현장에서 조사에 직접 응한 응답자를 대상으로 전화점검을 실시함으로써 진행된다. 주요 목적은 현장조사의 오류(비표본오차) 유형과 발생 원인에 대한 실증자료를 축적하여 조사품질을 개선시키는데 활용하기 위함인데, 현장조사 조사원의 태도까지도 알 수 있으므로 조사업무에 대한 경고장치(Warning System)로도 기능할 수 있다. 이러한 기능이 작동되면 조사원의 조사수행 성실도가 높아지고 현장조사 단계에서 품질관리에 대한 중요성이 인식되어 지속적인 조사업무 개선을 유도할 수 있다. 조사통계의 경우 자료수집과정은 현장에 조사원이 투입되어 면접, 전화, 인터넷, 이메일 등의 방법을 통해 응답자로부터 조사표를 받아 내용을 검토하고 입력을 하기까지의 일련의 과정을 거친다. 현장조사의 정확성 진단은 보통 전화점검이나 현장방문을 통해 이루어지는데, 전화점검은 진단 대상 조사통계에 대해 현장조사가 완료된 직후 CATI(Computer Assisted Telephone Interviewing, 컴퓨터와 전화를 연결하여 전화점검에 활용할 수 있는 장비 및 프로그램)나 이와 유사한 장비를 활용하여 실제현장에서 적용하는 자료수집방법, 조사여부, 조사원의 태도, 특성항목에 대한 응답일치 여부 등의 조사 실태를 파악하는 작업이다. 이러한 전화점검이 현장조사 실태를 반영하기 어렵다고 판단되는 경우나, 응답자의 조사에 대한 불만 등으로 정확하지 않은 결과가 나타난다고 판단되면 조사현장을 방문하여 현장조사 실태를 직접 확인점검 할 수도 있다. 특히, 전화점검 등을 통한 정확성 진단 시 점검이 10-20% 이루어진 상태에서 조사의 일치 정도를 살펴보고 조사결과와 전화응답 결과가 매우 상이하게 나타나는 경우 진단 대상기관의 담당자에게 진단과정의 문제를 알리고 원인분석을

한 후 업무를 진행시켜야 한다. 이런 문제점을 해결하기 위한 방안으로는 점검용 조사표나 전화점검 시나리오에 문제는 없는지 확인하고 현장 방문점검으로 점검방법 전환 등을 통해 사후의 결과분석 및 진단결과의 환류과정에서 불필요한 논의가 나타나지 않도록 조치하여야 한다. 본 전화점검은 실지조사와 시간차가 발생하여 기억에 의존하는 조사로 정확성 확보에 많은 어려움이 있다. 따라서 진단대상통계의 조사결과는 가능한 한 가장 최근자료를 점검자료로 활용하여야 한다. 조사주기가 2년 이상이고 조사를 실시한 지가 오래되어 응답자 확인이 곤란한 경우는 작성기관 및 품질관리과와 협의하여 현장방문 등의 대안을 마련하여야 한다. 특히 패널조사와 같이 조사대상의 관리가 조사 전체에 영향을 미치는 경우에는 진단대상 기관에서 현장점검 자체에 아주 민감한 반응을 보일 수 있으므로 상세한 상황을 파악하여 품질관리과와 협의한 후 점검토록 한다. 일부 통계에서는 보존기한이 지난 조사표 및 마이크로 데이터의 폐기로 현장점검에 비교할 자료가 없어서 진단을 추진하지 못하는 경우가 발생할 수도 있는데, 그런 경우에도 임의로 판단하여 진단을 중지하지 말고 품질관리과에 즉시 연락하여 대처 방안을 강구토록 한다. 이 업무는 진단계획 수립, 점검대상 표본선정, 전화시나리오 준비, 전화점검원 채용 및 교육, 전화점검 실시, 점검내용 확인 및 집계, 문제점 도출 및 개선사항 정리, 보고서 작성 및 환류 등의 단계로 진행된다.



[그림 5-11] 현장조사 정확성 진단 흐름도

## 2) 보고통계의 자료수집 행태분석

앞에서는 조사통계 수집자료 정확성을 진단하는 업무를 설명하였다. 보고통계는 조사통계의 전화점검 대신 자료수집과정의 행태분석이 매우 중요하다. 자료수집 단계별로 직접 현지에 가서 담당자를 면담하여 자료수집의 방법을 파악하고, 이에 대한 문제점과 개선사항이 있는지 살펴보아야 한다. 특히, 최초 보고자가 자료를 수집하는 과정에서 잘못 조사될 소지가 있는지, 자료를 상급기관에 보고하는 과정에서 오류가 날 가능성이 있는지, 보고양식과 행정 전산망 등을 통해 효율적인 자료 보고가 이루어지고 있는지 등을 살펴보아야 할 것이다. 점검대상은 최소한 10개 이상으로 하되, 시군구별 보고체계의 경우, 한 보고 라인에 있는 광역자치단체(특별·광역시·도)-구시군-읍면동 3곳을 각각 방문한 경우 점검대상 수는 3개로 인정한다. 이 때 점검대상 표본은 임의로 선정하되, 수도권 등 일부지역에만 집중되지 않도록 유의하여야 한다. 점검대상이 적으므로 점검원을 별도 채용하는 것보다는 연구진이 직접 방문하여 점검하는 것이 전체 업무파악과 결과 보고서 작성에 많은 도움을 준다. 보고통계의 현장방문 점검은 통계별 특성에 맞도록 진단 팀에서 자체적으로 현장점검표를 작성하여 현장방문 시 업무의 효율성을 높인다. 보고통계의 경우에도 조사현장이 존재하여 개별 자료가 있는 경우에는 조사통계에 준용하여 현장조사의 정확성을 진단하고 추가로 자료수집 행태도 분석하여 개선방안을 도출하도록 한다.

## 3) 가공통계 투입자료 선정의 적정성 분석

가공통계의 경우 자료수집과정은 통계작성을 위해 투입되는 자료의 최초 선정 과정, 수집에서 입력, 중간결과의 산출 및 최종 통계의 산출까지의 과정을 살펴보아야 한다. 물론 가공통계의 유형이 다양하여 진단업무량의 차이가 있을 수 있으나 하나의 처리단계가 결국 최종 통계작성 결과의 정확성으로 나타나고 통계의 신뢰성에 직결되므로 유의하여 분석해야 할 것이다. 가공통계의 품질은 최초 투입자료의 선정이 매우 중요하다. 따라서 현재 투입하고 있는 자료의 입력과 중간산출물, 최종결과까지의 과정 외에도 통계작성에 투입되는 자료의 선정과정을 결정하기까지의 과정과 절차상에 문제는 없는지를 확인해야 한다. 가공통계는 투입자료 및 통계작성 방법에 따라 투입자료편집형(자료 분류·집계형-한국관광통계, 관련통계 재편집·집중형-한국의 사회지표), 투입자료 단계별 가공형(추계형-주택보급률, 추정형-경기종합지수), 복합형(종합국민경제통계-국민계정), 내부자료 단계별 가공형(조사통계)으로 구분할 수 있다. 특히, 일부 가공통계의 경우 투입자료 수집 시 일부 또는 전체를 직접조사에 의해 수집하는 경우에는 조사통계에 준용하여 현장조사의 정확성 진단을 실시하고, 추가로 자료 수집행태를 분석하고 이에 대한 개선방안을 도출하도록 한다.

### 마. 통계자료서비스의 충실성

작성과정에서는 오류가 없는 통계일지라도 공표되는 과정에서 오류가 발생한다면 통계품질이 높은지 낮은지 진단할 것도 없이 잘못된 통계가 되고 만다. 통계자료 서비스의 충실성 점검을 중요하게 생각하는 이유가 바로 이런 점이다. 우리나라에서는 매년 수많은 통계자료들이 통계간행물, 조사보고서, 각종 백서 및 통계DB 등의 형태로 제공되고 있으나 이들에 대한 사전·사후 점검이 취약한 실정이다. 비록 사후 점검이기는 하지만, 통계자료 서비스의 충실성을 진단하는 목적은 두 가지이다. 첫째, 주로 발생하는 오류의 유형과 발생 원인을 파악하여 이러한 오류의 재발 방지방안을 모색하기 위함이다. 둘째, 이용자에게 필요한 기본정보가 통계간행물에 충분히 제공되고 있는지를 점검하여 미흡한 점을 보완하도록 함으로써 통계서비스의 질을 향상시키기 위함이다. 통계자료 서비스의 충실성 진단은 세부 진단계획을 수립하고 통계간행물, 통계DB 등에 대한 오류 및 이용자 편의성 점검을 실시하고 점검내용을 확인·집계하여 진단대상통계 담당자에게 환류하는 업무로 이루어져 있다.



[그림 5-12] 통계자료서비스 충실성 진단 흐름도

## 제5절 통계청 품질 진단 분석

앞에서 설명한 통계청에서 하고 있는 품질 진단을 분석하여 개선 사항을 연구해보고자 한다. 분석 방법은 통계청에서 진단하여 보고하고 있는 품질 진단 보고서를 분석하였다. 분석 대상은 2008년 10월에 발간한 “2008년 정기 통계품질진단 연구용역 최종결과 보고서”이며 영역은 「기업결합동향」, 「문화재관리현황」, 「재래시장실태조사」, 「농림업 생산지수」의 네 개 영역이다. 표 1~4는 각 품질 보고서의 목차이다. 품질진단 개요부터 진단 결과 문제점 및 개선과제, 발전 전략 및 로드맵, 기타 자료 등의 순서로 큰 흐름은 같으나 각각의 진단 보고서는 세부적인 부분에서는 차이가 있음을 알 수 있다. 기업결합 동향과 농림업 생산지수에는 해외사례가 있으나 다른 두 보고서에는 없다. 해외사례라는 것이 품질 진단 영역마다 있는 경우도 있고 없는 경우도 있을 것이지만 보고서의 큰 흐름에 있어서 언급조차 없다는 것은 의아한 부분이다. 또한 품질에 있어서 데이터의 품질이 가장 중요하다고 판단되지만 보고서는 데이터의 품질에 대한 부분보다는 다른 이용자의 측면이나 정책반영 부분 등과 같은 데이터 자체의 품질 외적인 부분에 더 진단이 치중되어 있음을 알 수 있다. 품질 진단 보고서를 분석하는 목적은 첫째, 품질 진단 단계별 사항을 확인하여 서술 식으로 되어 있는 보고서의 진단 결과를 최대한 수식화하여 보고서별 비교가 가능하게 하려고 한다. 두 번째는 보고서 전반에 걸쳐 세부적인 사항에 대해 개선할 점을 제시하고자 한다.

1장에서는 일반적인 품질진단의 개요에 대해 서술하고 있다. 통계작성 개요 및 현황과 같은 일반적인 진단 통계의 정보에서부터 품질 진단을 어떻게 할 것인가에 대해 서술하고 있다. 진단할 통계의 일반 현황이 주 내용은 아니고 보고서에 따라 차이가 있지만 해당 통계에 대해 잘 모르는 사람이 보고 판단하기에는 통계의 중요도라든지 통계의 현황 및 활용도에 대해서는 자세히 서술되어 있지는 않았다. 또한 품질 진단을 하기에 앞서 품질 차원에 대해 정의하는 부분이 있는데 차원이 충분히 진단 목적에 따라 분류되었다고 판단하기는 힘든 부분이 있다. 진단자의 주관적인 부분이 많이 포함이 되어 있고, 보고서에 따라서는 해당 통계의 차원의 정의보다는 일반적인 차원의 정의에 대해 서술한 보고서도 있다. 품질 진단에 있어서 차원의 정의와 분류가 명확해야 진단을 하는데 있어서도 혼동과 누락이 없을 것으로 판단된다.

각 품질 관리 단계별로 점수화에 대한 분석을 해보고자 한다. 품질관리 기반을 측정

하는 문제에 있어서 점수화가 가능한 부분은 “제공시기의 정확성 여부”, “인력”, “예산”, “하드웨어”, “소프트웨어”, “이용자 만족도 조사” 등이 있다. 제공시기의 정확성 부분에서는 제공시기의 기간이 정해져 있으면 100, 제공시기의 오차 부분만큼의 점수화를 할 수 있을 것이다. 인력 부분은 전문 인력의 수로 점수화 할 수 있는데 통계 전문가가 1명 이상이면 100, 통계 전문가가 아니지만 100% 전담인력이 있으면 90 이런 식으로 점수화가 가능할 것이다. 또한 통계 교육 이수 정도를 점수화할 수 있는데 3년 이내 전문 통계 교육을 받았으면 100, 5년 이내 통계 교육을 이수 했으면 80 이렇게 점수화 할 수 있다. 예산 부분은 전년대비 예산 증감률을 가지고 점수화를 한다. 작년대비 5% 이상의 예산 증가가 있다면 100, 3% 미만 50, 예산 감소 20 이렇게 점수화를 한다. 하드웨어는 통계 처리 시스템의 보유 여부로 판단한다. 통계 처리 시스템을 자체 보유하면서 자체 운영하면 100, 자체 보유하면서 외부 운영이면 80, 외부 보유하면서 자체 운영은 60 등으로 점수화한다. 소프트웨어 사용 부분은 전문 통계 처리 소프트웨어 사용으로 판단한다. 자료 처리에 있어서 통계 처리 소프트웨어의 사용은 중요하다. 사무용 소프트웨어인 엑셀로도 어느 정도의 통계 처리는 가능하나 그 정확성에 대해서는 장담할 수 없다. 따라서 전문 소프트웨어를 하나라도 사용한다면 100, 엑셀에서 고급 통계 기능을 사용한다면 80 등으로 산정한다. 이용자 만족도 부분에 대해서는 신뢰도를 표시해야 하고 전체 표본 대상에 대한 표시가 필요하다. 또한 이러한 전체 표본 수 대비 이용자 만족도 조사에 참여한 사람의 수에 대한 점수화가 필요하다.

통계이용실태 부분에서는 “이용 빈도”, “통계자료 이용 목적”, “재사용 의사”가 점수화가 가능할 것으로 보인다. 통계 이용자의 통계 이용 빈도는 연간 사용횟수와 인원에 대한 비율로 점수화가 가능하다. 연간 사용횟수와 인원을 곱해서 연간 총 이용자 횟수와 나누면 비율이 계산된다. 이것을 점수로 사용한다. 통계 이용 목적은 해당 통계의 이용 항목이 있는데 이 항목의 중요도를 설정하여 중요도와 이용 인원을 비율로 계산하여 점수로 사용한다. 재사용 의사도 마찬가지로 이용할 의사 항목의 중요도를 설정하고 이용 인원을 곱하여 비율로 계산하여 점수화하여 사용할 수 있다.

품질 진단에 있어서 수집 자료의 정확성은 가장 중요한 부분이라고 할 수 있다. 이 부분에서는 주로 입력 시스템과 입력된 자료의 정확성 판단에 대한 점수화가 필요하다. 대부분의 진단 보고서를 보면 공통적으로 입력 시스템이 전산화되어 있는 부분은 믿고 넘어가는 경우가 있다. 하지만 입력 시스템이 공통적인 입력 오류를 가지고 있다면 입력





하는 데이터 모두를 신뢰할 수 없을 정도로 치명적이다. 따라서 입력 시스템을 맹신하기 보다는 더욱 철저한 진단이 필요할 것으로 판단된다. 소프트웨어의 진단은 직접 입력해 보는 것이 가장 좋다. 통계 작성 기관에서 조사된 데이터를 직접 받아서 최소 10번 정도는 직접 입력해보고 오류는 없는지 오류 가능성은 없는지 꼼꼼히 진단해 보아야 할 것이다. 이러한 작업이 어렵다면 전산 전문가의 도움을 받아서라도 진행해야 할 사항임은 분명하다. 또 다른 진단 사항은 입력 전산 시스템이 자체 보유, 운영 중이라면 보안에 취약할 가능성이 있다. 만일에 하나 보안상의 문제 발생으로 인하여 데이터에 문제가 생길 수 있다. 이러한 부분은 보안 전문가나 보안 전문 업체에 의뢰하여 분석하고 결과대로 수정할 필요가 있다.

입력 시스템 진단 다음에 중요한 것이 입력된 값을 확인하는 작업일 것이다. 입력 시스템에서 잘 입력된 자료라 하더라도 입력 당시에 확인되지 못하거나 잘못된 값인데 그냥 입력된 값일 수 있다. 보통 조사된 값은 모두 같은 값이 아니고 일정한 변동을 가진 값이기 때문에 잘못 입력된 값을 눈으로 확인할 수는 없다. 또한 입력된 값의 개수가 많은 경우 또한 눈으로 일일이 확인하기란 어려운 일이다. 이러한 경우 자료의 평균이나 분산, 사분위수, 이상점 등의 통계 보고서를 작성하여 확인하는 것이 한 방법이라고 할 수 있을 것이다. 그것이 정말 잘못된 값이 아니더라도 잘못된 값의 후보를 찾아내어 실제 값과 비교하는 방법은 사용할 수 있을 것이다. 식스시그마 소프트웨어나 일반 통계 패키지의 품질관리 부분의 분석 기능을 사용하면 어느 정도 식별 가능한 이상한 값의 후보군을 찾아낼 수 있다. 다만 이러한 분석 기능은 우리가 품질 진단을 하는데 있어 시간이 많이 걸리거나 찾는 기능이 잘 완비된 소프트웨어를 찾기는 어려울 수도 있다. 따라서 품질 진단에서 입력 값의 진단을 수행할 수 있는 별도의 품질 진단 소프트웨어를 제작할 것을 제안한다. 정확성을 진단해야 하는 자료를 입력하면 기초적인 통계 값 출력 이외에 품질에 영향을 미칠 수 있는 값의 후보군을 찾아서 경고를 해줄 수 있는 소프트웨어가 필요하다고 판단된다. 이렇게 하면 데이터 품질에 가장 큰 영향을 미치는 정확성을 어렵지 않게 수행할 수 있을 것이라고 판단된다. 장기적인 품질에 영향을 미치는 요인으로는 응답 부담이 있을 수 있다. 응답자는 응답 대내외 환경, 응답문항수, 방문횟수 등에 응답부담을 느끼게 된다. 응답 부담은 응답 품질에 안 좋은 영향을 미친다. 피로도가 높아지고 응답 자체에는 신경을 쓰지 않게 된다. 따라서 이러한 응답 부담을 줄여주는 것이 높은 품질을 얻기 위한 노력이라고 할 수 있다. 이러한 요인에 대해서는 먼저 조사 문항의 개수로 점수화가 가능할 것이다. 응답 문항이 필요에 의해 어쩔 수 없이 많을 수는 있으나 비

정상적으로 많거나 할 경우 이를 줄이는 노력을 통해 품질에 힘써야 할 것이다. 두 번째로는 평균 방문 횟수가 있을 것이다. 평균 방문횟수가 비정상적으로 높을 경우 응답자가 의도적으로 응답을 회피하는 건지 아니면 조사 방법에 문제가 있는지 확인할 필요가 있다. 이러한 방문횟수를 줄이는 것 또한 품질을 높이기 위한 기초 노력이 될 것이다.

수집된 자료를 사용자에게 제공하는 것은 사용자들이 자료의 정확성과는 별개로 통계의 품질을 인식하게 되는 중요한 수단이다. 과거에는 인쇄된 매체로 제공하던 것을 최근에는 정보통신의 발달로 인터넷이나 이메일 CD나 DVD, USB 매체로 제공하는 것이 일반화되었다. 매체의 중요도에 따른 순위를 매기기는 어렵지만 매체 제공의 종류의 개수로 점수화가 가능할 것이다. 종류의 개수가 많아지면 이용자들이 그만큼 편하게 자료를 활용할 수 있는 척도가 될 수 있기 때문이다. 예를 들어 4가지 이상의 종류를 제공하면 100, 3가지 이상은 80 등과 같이 점수화가 가능할 것이다. 제공 매체에 담긴 자료의 내용도 중요할 것이다. 일반적으로 조사된 데이터를 모두 공개하지는 않는다. 데이터에 민감한 자료도 있을 수 있고 개인정보와 사업체 정보 등이 수록될 수도 있기 때문이다. 하지만 이러한 한계점 속에서도 조사된 자료는 사용자에게 가능하면 최대한 제공하는 것이 통계 조사의 목적이 될 것이다. 자료가 불필요하게 제한적으로 제공되면 사용자는 반대로 통계의 품질이 좋지 않다고 생각할 수 있기 때문이다. 따라서 제공 자료의 범위에 대한 점수화가 필요하다. 100% 모든 데이터를 제공하면 100 등과 같이 자료 제공 %에 따른 점수화가 가능하다. 영국 통계청의 사례에서처럼 이러한 통계 자료 제공에는 비밀보호 수준이 동반되어야 한다. 또한 통계가 조사된 시점부터 공표된 시점에 대한 시간 간격을 측정하여 얼마나 사용자에게 빨리 그 통계를 사용할 수 있게 해주었는가에 대한 점수화도 필요하다.



〈표 5-6〉 품질 진단 단계별 개선 사항

진단 영역	항목	내용	예시
품질 관리 기반	제공 시기의 정확성 여부	제공 시기 확정 여부에 따른 점수화	월 일 확정 게시 시 100
	통계 전문 인력 여부	통계 전문 인력이 전담하고 있는가에 대한 점수화	통계 전문 인력이 업무 전담 : 100 비전문 인력 업무 전담 100% : 100
	통계교육 이수 여부	통계 전문 교육 이수 여부에 따른 점수화	1년 이내 전문 교육 이수 : 100 2년 이내 전문 교육 이수 : 80

진단 영역	항목	내용	예시
	통계 예산	통계 관련 예산 증액 여부	전년 대비 5% 이상 증가 : 100 전년 대비 3% 미만 증가 : 80
	하드웨어	통계 처리(입력, 분석) 시스템 보유 여부	자체 보유, 자체 운용 : 100 자체 보유, 외부 운용 : 80
	소프트웨어	통계 처리 시 전문 통계 소프트웨어 사용 여부	전문 소프트웨어 사용 : 100 엑셀 통계 기능 사용 : 80
	이용자 만족도 조사	이용자 만족도 조사 시 표본 비율 점수화	만족도 조사 참여 응답자 비율
	만족도 조사 조사 방법	조사 방법 제공 종류에 따른 점수화	4가지 이상 : 100 3가지 이상 : 80
통계 이용 실태	통계 이용 빈도	연간 이용 빈도에 따른 점수화	연간 사용 횟수 x 이용 인원 비율
	통계 자료 이용 목적	이용 목적의 중요도에 따른 점수화	이용 목적의 중요도 가중치 x 이용 인원
	이용자 만족도	통계 이용자의 만족도 점수	종합만족도(이미 존재)
	재사용 의사	재사용 의사의 중요도에 따른 점수화	재사용 의사 중요도 가중치 x 응답 인원
수집 자료의 정확성	입력 시스템 점검	입력 시스템 점검 여부에 따른 점수화	보안+정기 점검 + 업데이트 : 100 보안 + 정기 점검 : 80
	입력된 데이터	품질 분석 결과에 따른 점수화	별도의 프로그램을 이용하여 점수화
	입력 데이터 내검	데이터 내검 방법에 대한 점수화	5단계 이상 내검 : 100 4단계 이상 내검 : 80
	내검 수준	내검 정도에 따른 점수화	수정된 자료의 수 / 조사된 데이터의 총수
	무응답 대체	무응답 대체 정도에 따른 점수화	대체된 데이터의 총수 / 모든 최종 데이터의 수
	조사문항수	조사문항 개수에 따른 품질 점수화	10개 미만 : 100 10개 이상 20개 미만 : 80
	평균 방문 횟수	응답 부담에 대한 품질 점수화	5회 미만 : 100 5회 이상 10회 미만 : 90
수집된 자료 제공	제공 매체 종류	제공 매체 종류에 의한 점수화	4가지 이상 매체 제공 : 100 3가지 이상 매체 제공 : 80
자료 공표	자료 공표 범위	자료 공표 범위에 따른 점수화	자료 공표 %에 따른 점수화 셀 단위 카운트
	비밀보호 수준	비밀보호 기법 적용 여부에 대한 점수화	감춰진 셀의 비율
	공표시기	통계 조사가 이루어진 시기와 공표시기와의 시간차	통계마다의 차이가 있지만 적절한 시기를 고려해서 점수화

<표 5-6>은 각 단계별 점수화에 대한 항목이다. 보고서를 참조하여 최대한 측정 가능한 지표를 작성하였다. 일부 항목들은 공통되지 않고 상세한 정보가 없어 지금의 보고서 단계에서는 측정이 불가능하였다. <표 5-7>에서 <표 5-10>은 앞의 점수화 방안에 의해 4개의 품질 진단을 점수화하고 비교하였다.

<표 5-7> 농림업 생산지수 점수화

진단 영역	항목	내용	점수	
품질 관리 기반	제공시기의 정확성 여부	매년 8월경 제공	80	
	통계 전문 인력 여부	전문 인력 미비에 부분 참여	60	
	통계교육 이수 여부	3년 이내 8회 40일 이수	60	
	통계 예산	별도 예산 X	0	
	하드웨어	농업사업통합정보시스템 Agrix	100	
	소프트웨어	농업사업통합정보시스템 Agrix	100	
	이용자 만족도 조사	66/257(제공 받은 통계 사용자 리스트)	25	
	만족도 조사 조사 방법	전화/e-mail	40	
	통계이용실태	통계 이용 빈도	연 2회 이상 비율	60.61
		통계 자료 이용 목적	정책진단/연구 학습	84.85
이용자 만족도		3.51(보고서 참조) / 5	70.2	
제사용 의사		그렇다 응답 이상	86.36	
수집자료의 정확성	입력 시스템 점검	자료 없음	X	
	입력된 데이터	자료 없음	X	
	입력 데이터 내검	자료 없음	X	
	내검 수준	자료 없음	X	
	무응답 대체	자료 없음	X	
	조사문항수	자료 없음	X	
	평균 방문 횟수	자료 없음	X	
수집된 자료 제공	제공 매체 종류	PDF, WEB	60	
	자료 공표	자료 없음	X	
	비밀보호 수준	자료 없음	X	
	공표시기 차이	자료 없음	X	
총점	827.02			
평균	63.61			



〈표 5-8〉 문화재 관리현황 점수화

진단 영역	항목	내용	점수
품질 관리 기반	제공시기의 정확성 여부	익년 1월	80
	통계 전문 인력 여부	전문 인력 미비에 부분 참여	60
	통계교육 이수 여부	없음	0
	통계 예산	360% 증가	100
	하드웨어	X 측정 불가	X
	소프트웨어	문화재 홈페이지에 등록	60
	이용자 만족도 조사	77/1000(제공 받은 통계 사용자 리스트)	7
통계이용실태	만족도 조사 조사 방법	전화/e-mail/Web	60
	통계 이용 빈도	연 2회 이상 비율	71
	통계 자료 이용 목적	정책진단/연구 학습	40
	이용자 만족도	3.55(보고서 참조) / 5	71
수집자료의 정확성	재사용 의사	그렇다 응답 이상	89.6
	입력 시스템 점검	자료 없음	X
	입력된 데이터	자료 없음	X
	입력 데이터 내검	자료 없음	X
	내검 수준	자료 없음	X
	무응답 대체	자료 없음	X
	조사문항수	자료 없음	X
수집된 자료 제공	평균 방문 횟수	자료 없음	X
	제공 매체 종류	발간물, WEB	60
	자료 공표	자료 없음	X
자료 공표	비밀보호 수준	자료 없음	X
	공표시기 차이	1개월	80
총점	778.6		
평균	59.89		

〈표 5-9〉 기업결합동향 점수화

진단 영역	항목	내용	점수
품질 관리 기반	제공시기의 정확성 여부	매년 1/4분기 경	60
	통계 전문 인력 여부	전문 인력 미비에 부분 참여	40
	통계교육 이수 여부	없음	0
	통계 예산	예산 편성 없음	0
	하드웨어	X 측정 불가	X
	소프트웨어	엑셀 사용	60
	이용자 만족도 조사	31/ 총 사용자 없음	60
	만족도 조사 조사 방법	전화/e-mail/방문	60
	통계이용실태	통계 이용 빈도	연 2회 이상 비율
통계 자료 이용 목적		정책진단/연구 학습	32.3
이용자 만족도		3.33(보고서 참조) / 5	66
제사용 의사		그렇다 응답 이상	58
수집자료의 정확성		입력 시스템 점검	자료 없음
	입력된 데이터	자료 없음	X
	입력 데이터 내검	자료 없음	X
	내검 수준	자료 없음	X
	무응답 대체	자료 없음	X
	조사문항수	자료 없음	X
	평균 방문 횟수	자료 없음	X
	수집된 자료 제공	제공 매체 종류	발간물, WEB
자료 공표		자료 공표 범위	자료 없음
	비밀보호 수준	자료 없음	X
	공표시기 차이	1개월	X
총점	518.9		
평균	43.24		



〈표 5-10〉 재래시장 실태조사 점수화

진단 영역	항목	내용	점수
품질 관리 기반	제공시기의 정확성 여부	조사년도 익년 3월	80
	통계 전문 인력 여부	전문 인력 확보 및 전담	100
	통계교육 이수 여부	없음	0
	통계 예산	6억4천만에서 10억으로 증액	58.8
	하드웨어	통계청 서버 사용	60
	소프트웨어	SPSS, 엑셀	100
	이용자 만족도 조사	68/ 총 사용자 없음	60
	만족도 조사 조사 방법	Web	40
	통계이용실태	통계 이용 빈도	연 2회 이상 비율
통계 자료 이용 목적		정책진단/연구 학습	36.8
이용자 만족도		3.47(보고서 참조) / 5	69.4
재사용 의사		그렇다 응답 이상	0
수집자료의 정확성		입력 시스템 점검	자료 없음
	입력된 데이터	자료 없음	X
	입력 데이터 내검	자료 없음	X
	내검 수준	자료 없음	X
	무응답 대체	자료 없음	X
	조사문항수	자료 없음	X
	평균 방문 횟수	자료 없음	X
	수집된 자료 제공	제공 매체 종류	발간물, WEB
자료 공표		자료 공표 범위	자료 없음
	비밀보호 수준	자료 없음	X
	공표시기 차이	3개월	60
총점	772.1		
평균	55.15		

앞에서 각 품질 진단에 대해 점수화하고 비교한 것을 다시 전체비교를 해보면 <표 5-11>과 같다.

<표 5-11> 품질 진단 별 점수화 비교표

진단명	총점	점수
농림업 생산지수	827.02	63.62
문화재 관리현황	778.6	59.89
기업 결합 활동	518.9	43.24
재래시장 실태조사	772.1	55.15

전반적으로 낮은 점수를 받았으나 현재 상태에서는 이 점수는 큰 의미를 갖지는 않는다. 정보가 없는 항목은 제외를 했지만 점수화를 위해 선정했던 항목에 대한 자세한 정보가 없는 상태에서 점수화를 하기엔 힘든 부분이 있었다. 이렇게 점수화를 하면서 알게 된 것은 점수가 큰 의미는 없지만 전반적으로 개별 품질 진단으로 자체의 품질에 대한 어느 정도의 평가가 가능하다는 것이다. 어느 품질 진단 보고서에는 있는 내용인데 다른 보고서에는 없는 항목도 있다. 세부적으로 들어가 보면 측정해야 하는 항목임에도 불구하고 보고서에는 자세히 서술되어 있지 않거나 누락된 경우가 있었다. 즉 점수가 낮은 이유가 품질 진단의 자체 점수가 낮아서 일수도 있고 자세한 정보를 기록하지 않은 보고서의 품질을 대변할 수도 있다는 것이다. 향후에는 이러한 품질 진단 보고서를 정형화하여 누락되거나 상세 정보가 없는 경우를 제거하여 품질 진단 자체에 대한 평가만 가능하게 하여야 할 것이라고 생각된다.

이러한 평가를 제대로 하게 하려면 일단 평가를 하려는 항목에 대해 필수적으로 보고서에 기재하여야 하고 상세한 항목을 연구하여 보고서 작성 항목에 넣어야 할 것이다. 대부분의 보고서의 상세 내용이 온라인이나 제출된 문서의 정보를 그대로 가져오거나 절대적으로 의존하는 경우가 많은데 이렇게 진단을 하게 되면 상세한 정보를 듣지 못해 잘못된 진단을 내릴 수 있는 가능성을 가지게 된다. 상세한 정보, 정확한 정보를 알 필요가 있을 경우엔 담당자나 기타 방법을 통해 알아내서 정확하게 기록하게 할 필요가 있다고 판단된다. <표 5-12>에서 <표 5-15>는 각 품질 진단 별 보고서의 목차이다. 보고서의 목차를 보면 정형화의 필요성과 어떠한 품질의 목적에 집중하여야 할지에 대한 방향성을 볼 수 있다고 판단된다.





<표 5-12> 기업결합동향

기업결합동향

목 차	
제 1 장 「기업결합동향」 품질진단 개요	1
제 1 절 통계 개요	1
제 2 절 품질진단 개요	3
제 2 장 「기업결합동향」 품질진단 결과	8
제 1 절 기업결합현황 통계의 부문별 품질진단 결과	8
1. 품질관리 기반	8
2. 이용자 만족도 및 요구사항 반영실태	13
3. 세부작성절차별 오차관리체계	35
4. 수집자료의 정확성 점검	44
5. 통계자료 서비스 점검	48
6. 우수사례 검토	54
제 2 절 기업결합현황 통계품질진단 결과 종합	55
제 3 장 해외사례 연구	61
제 1 절 일본의 기업결합추이	61
제 2 절 영국의 기업결합추이	63
제 4 장 문제점 및 개선과제	66
제 1 절 문제점	66
제 2 절 개선과제	67
제 5 장 「기업결합동향」 발전전략 및 로드맵	70
제 1 절 발전 전략과 로드맵	70
제 2 절 정책적 시사점 및 발전 방향	72
참고문헌	74
<부록> 진단에 사용된 진단서, 조사표 양식 등	75

〈표 5-13〉 문화재 관리 현황

문화재관리현황	
목 차	
제 1 장 「문화재관리현황」 품질진단 개요	1
제 1 절 품질진단의 필요성 및 목적	1
제 2 절 품질진단 단계별 진단방법	12
제 3 절 품질진단 절차	19
제 4 절 진단대상 통계 현황	22
제 2 장 「문화재관리현황」 품질진단 결과	25
제 1 절 품질기반진단 결과	25
제 2 절 이용자 만족도 및 요구사항 반영 실태 결과	32
제 3 절 세부 작성절차별 오차관리체계 진단 결과	76
제 4 절 수집자료의 정확성 진단 결과	89
제 5 절 통계자료 서비스의 충실성 진단 결과	107
제 6 절 진단결과 종합평가 및 발전방안	118
제 3 장 문제점 및 개선과제	135
제 4 장 「문화재관리현황」 발전전략 및 로드맵	141
제 1 절 발전전략	141
제 2 절 로드맵	143
참고문헌	144
<부록 1> 진단에 사용된 진단서, 조사표 양식 등	145



<표 5-14> 재래시장 실태 조사

재래시장실태조사

제 1장 「재래시장 실태조사」 품질진단 연구 개요	1
제1절 재래시장 실태조사 통계개요	6
1. 재래시장 실태조사의 목적과 필요성	2. 작성기관
3. 재래시장 실태조사의 연혁	4. 재래시장 실태조사의 통계종류 및 승인번호
5. 조사 항목	6. 작성주기, 대상 기간·시점, 작성기간
7. 작성방법 및 체계	8. 결과공표 주기
제2절 재래시장 실태조사 품질진단 개요	9
1. 재래시장 실태조사 통계의 품질진단 방법	2. 재래시장 실태조사 통계의 품질진단 절차와 내용
제2장 재래시장 실태조사 부문별 통계품질진단 결과	37
제1절 품질관리기반 진단 결과	37
1. 통계작성 환경 품질진단결과	2. 전문가 그룹 면접
제2절 이용자 만족도 및 요구사항 반영 실태	46
1. 조사개요	2. 주요 조사결과
제3절 세부 작성절차별 오차관리 체계 진단	75
1. 진단결과	2. 진단지표별 분석
3. 개선검토의견	
제4절 수집자료의 정확성	85
제5절 통계자료서비스의 충실성	89
제 3 장 문제점 및 개선 과제	92
제1절 차원별 품질요약	92
1. 관련성	2. 정확성
3. 시의성 및 정시성	4. 비교성/일관성
5. 접근성 및 명확성	6. 정확성
제2절 문제점 및 개선과제	97
1. 품질관리기반	2. 통계이용실태 및 이용자 만족도 진단
3. 세부작성절차별 오차관리 체계진단	4. 수집자료의 정확성
5. 통계자료 서비스의 충실성	
제3절 우수사례	103
1. 우수사례 유형	2. 우수사례 요약표
제 4 장 종합 제언	106
제 5 장 재래시장 실태조사 발전전략 및 로드맵	111
제1절 발전전략	111
제2절 로드맵	112
참고문헌	113
부록	114

## 〈표 5-15〉 농림업 생산지수

농림업 생산지수	
제 1 장 「농림업생산지수 통계」 품질진단 개요	1
제 1 절 품질진단의 필요성과 효과	1
1. 품질진단 배경 및 필요성	1
2. 품질진단의 목적	2
3. 기대효과	2
제 2 절 통계 개요	3
1. 기본현황 2. 통계 생산절차 3. 공표자료	9
제 3 절 품질진단 개요	11
1. 통계품질관리기반 진단 2. 통계활용 실태 및 이용자만족도 진단	
3. 전문가그룹면접(FGI) 4. 세부 작성절차별 오차관리체계 진단	19
5. 수집자료의 정확성 진단 6. 통계자료 서비스의 충실성 진단	22
제 2 장 「농림업생산지수 통계」 품질진단 결과	25
제 1 절 통계품질관리기반 품질진단 결과	25
1. 통계작성 여건 2. 조직관리 실태 및 통계작성 담당자의 인식	26
제 2 절 통계활용 실태 및 이용자만족도 품질진단 결과	28
1. 응답자 특성 2. 이용실태 3. 이용자 만족도 4. 만족도의 측정	40
제 3 절 전문가그룹면접(FGI) 결과	44
1. 면접결과 정리 2. 개선요구사항	
제 4 절 세부 작성절차별 오차관리체계 품질진단 결과	46
1. 작성절차별 진단결과 2. 품질차원별 진단결과	60
제 5 절 수집 자료의 정확성 품질진단 결과	66
1. 통계작성에 투입되는 자료의 품질	66
2. 보고양식 작성 및 제출	66
제 6 절 통계자료 서비스의 충실성 품질진단 결과	67
1. 발간물 오류점검 진단결과 2. 이용자 편의사항 점검 진단결과	
제 3 장 해외사례 연구	71
제 4 장 진단내용 종합 및 개선과제	73
제 1 절 진단부문별 결과 종합	73
제 2 절 품질차원별 결과 종합	78
제 3 절 우수사례	85
제 4 절 종합개선과제 및 개선방안	88
제 5 장 「농림업생산지수 통계」 발전전략 및 로드맵	92
참고문헌	94
부록	95



## 제6절 결 론

본 연구에서는 통계의 광의의 품질이란 무엇인가에 대해 알아보고 품질 진단이란 무엇인가에 대해 정의하였다. 또한 통계청이 품질 진단을 하는데 사용하는 품질의 정의에 대해 알아보고 각 품질 진단의 각 단계에 대해 알아보았다. 품질 진단이 어디서 출발하였는지 역사에 대해 알아보고 각국이 품질 진단을 어떻게 시작하여 실시하고 있는지 알아보았다. 그리고 해외에서는 품질 진단을 개선하기 위한 연구로 어떠한 것을 실시하여 개선에 반영하였는지 알아보고 국내의 연구 성과에 대해 알아보았다. 최근의 통계청 품질 진단의 방향과 방법, 절차 등에 대해 알아보았으며 2008년에 실시된 품질 진단의 몇 가지 실제 사례를 살펴보았다.

국가 통계가 아닌 일반 제조업 등의 산업에서의 품질이란 제품의 성능이나 제품의 결함 등의 수치 즉 통계를 가지고 측정하며 품질 진단을 통한 품질 향상 노력을 한다. 그러나 통계에 있어서의 품질이란 정의하기도 어렵고 측정하기도 어렵다는 것을 해외 연구 사례나 국내 사례에서 쉽게 알 수 있다. 통계청에서 실시하는 품질 진단도 가장 중요하다고 생각되는 6개의 차원에 대해서 할 뿐이지 정말 제대로 세밀하게 품질 진단을 한다면 그보다는 많은 차원에 대해 품질을 정의하고 각각의 차원에 대한 항목별로 품질 진단을 하여야 할 것이다. 그러나 한정된 시간에 한정된 자원을 가지고 많은 수의 통계를 심층적으로 하는 것은 현실적으로 불가능할 것이다.

품질 진단이 어렵고 힘든 작업이라고 해서 통계청이나 통계생산 기관이 통계의 품질 향상에 소홀할 수는 없을 것이다. 본 연구에서 찾아낸 어려움은 첫째, 각각의 통계에 대해 품질 진단을 개별적으로 하다 보니 진단마다의 차이가 많이 발생하고 있음을 알 수 있다. 결론적으로 보면 시간과 인력을 많이 들여서 하는 품질 진단에 대한 품질을 보장 못하는 결과가 얻어지게 되는 것이다. 품질 진단을 하는 개별 개체가 생각이 다르고 방향이 달라서 차이가 발생하는 것 같다. 다시 말하면 품질 진단에 개별 진단 주체의 주관적인 생각이 너무 많이 반영되는 것과 같은 문제점이 발생하는 것이 보인다. 제조업의 생산과 관련된 품질관리를 예를 들면 구두를 생산하는데 구두를 100개를 뽑아서 품질을 측정한다 했을 때 측정하는 진단 주체에게 품질이 보증되지 않는 또는 주관적으로 생각 가능한 측정 도구를 지급했다고 한다면 품질에 대한 측정값 자체가 알 수 없는 변동에 의해 미확인된 품질 인자에 의한 품질 성격이 결정될 것이다. 따라서 품질 진단을 하는 측정도구들 매뉴얼이나 설문지 등을 더 세밀하고 측정 가능한 진단 주체들끼리도 혼란의 여지가 없게 설계할 필요가 있을 것이다.

두 번째로는 품질에 있어서 수집 자료의 정확성이 가장 중요한 차원이라고 많은 기존 연구에서 밝혔음에도 불구하고 정확성에 대한 측정을 대부분 입력 시스템의 전산화나 눈에 띄는 이상치가 없으면 그냥 넘어가는 경우가 종종 있다는 것이다. 입력 시스템이 전산화되어 있는 것과 수집 자료의 정확성과는 크게 관련이 없다. 입력이 정확하다는 한 가지 장점이 있는 것이지 조사된 자료 자체가 정확하다는 것과는 별개의 것일 것이다. 따라서 해외 연구 자료에 의하면 정확성을 측정하기 위해 더 많은 측정 항목을 개발하고 있는 실정이다. 이러한 자료의 정확성은 사실 진단 주체들이 일일이 분석하고 자료의 정확성을 판단하기에는 너무 많은 시간과 인적 자원 및 전문 인력이 필요하다. 따라서 진단 주체에 이러한 몫을 남길 것이 아니라 품질 진단을 위한 별도의 자료의 정확성 판단 품질 진단 소프트웨어를 장기적으로 설계하여 개발할 필요가 있다. 이러한 소프트웨어는 기초통계 이외에 사업체나 가구냐에 따른 자료의 성격에 따라 경제 자료인 경우 간단한 시계열 분석 등을 포함해야 한다. 이러한 소프트웨어를 사용하여 진단 주체들은 많은 시간을 할애하지 않아도 수집 자료를 간단하게 DB 형태나 자료 형태 그대로 입력하여 품질에 대한 분석을 가능하게 하여야 한다.

세 번째는 품질 진단이 통계 품질 자체보다는 정책사용 가능성이나 이용자 측면에 대한 부분이 더 많다는 것이다. 입력 자료의 정확성 부분이 다른 부분보다는 더 중요하다는 것은 아무도 이의를 제기하기는 어려울 것이다. 자료 자체의 품질이 보장되지 않으면 정책에 잘못 사용될 가능성, 이용자들이 잘못된 사용으로 인해 혼란을 야기할 가능성이 너무나도 많다. 그러나 품질 진단 매뉴얼에는 수집 자료의 정확성을 세밀하게 분석하고 확인할 수 있는 설계가 부족한 것 같다. 해외 사례에서 보았듯이 자료의 기초 통계는 물론 자료의 수정, 응답자의 비율, 자료의 대체 등 품질에 영향을 미칠 수 있는 많은 자료의 수집부터 가공의 단계를 측정하고 수치화할 필요가 있다. 본 연구에서 제시한 몇 가지 수치는 공통적인 것이지만 통계의 특성상 예를 들면 행정 통계인지 가구통계인지 경제 통계인지 등의 성격에 따라 세밀한 측정 도구를 설계하여 제시해야 할 것이다.



## 참고문헌

- 품질관리과(2008), “「기업결합동향」 2008년 정기 통계품질진단 연구용역 최종결과보고서”, 통계청.
- 품질관리과(2008), “「농림업생산지수」 2008년 정기 통계품질진단 연구용역 최종결과보고서”, 통계청.
- 품질관리과(2008), “「문화재 관리현황」 2008년 정기 통계품질진단 연구용역 최종결과보고서”, 통계청.
- 품질관리과(2008), “「재래시장 실태조사」 2008년 정기 통계품질진단 연구용역 최종결과보고서”, 통계청.
- 류제복(2003), “통계품질평가지표개발”, 통계청.
- 박성현(2003), “통계품질관리 표준매뉴얼 작성 학술연구”, 통계청.
- 품질관리팀(2004), “외국의 품질관리”, 통계청.
- 이기재(2005), “보고통계 품질향상을 위한 품질평가지표 개발”, 조사연구 7권 1호.
- Giovanna Brancato, Giorgia Simeoni(2008), "Modelling Survey Quality by Structural Equation Models", Istat.
- EUROSTAT(2005), "Standard quality indicators", Quality in Statistics.
- Karen Dunnell(2007), "Guidelines for Measuring statistical Quality", Office for National Statistics.
- Erwin Seyfried(2003), "A limited set of coherent quality indicators proposed by the technical working group on quality", Office for National Statistics. Final report of the European Forum on Quality in VET.