

지역통계 데이터 베이스 구축방안⁺

이화영

동신대학교 통계학과

이희춘

상지대학교 통계학과

홍기학

동신대학교 통계학과

A Study on Data Base of Region Statistics⁺

Hwa-Young Lee

Dept. of Statistics, Dongshin University

Hee-Choon Lee

Dept. of Statistics, Sangji University

Ki-Hak Hong

Dept. of Statistics, Dongshin University

Abstract

This study suggests a data base scheme of region statistics whose demand has been rapidly increased as the local self-governing body is introduced in Korea. A program for the region statistics management(registration, reference, modification, deletion) is developed and it can be used by personal computer users.

1. 서론

한 국가 또는 어느 일정지역에서 합리적인 경제, 사회발전계획의 수립은 물론 민간기관과 학계에서도 각종 정책수립 및 업무수행 또는 연구사업수행에는 반드시 그에 관련된 분

⁺ 본 연구는 1991년도 광주은행 학술연구비 지원에 의한 연구 결과임

야의 통계정보가 바탕이 되어야 한다. 지금까지 우리 나라의 국가정책에 관한 의사결정은 중앙집권적이었기 때문에 대부분의 법정통계가 전국 단위로 수집되었다. 또 제1차 사회, 경제발전계획이 시작된 1962년 부터 우리 나라의 국가정책은 경제성장을 위주로 하였기 때문에 발표된 통계 역시 전국단위의 경제통계가 주를 이루었으며 양적, 질적으로 괄목할만한 성장을 이룩한 반면에 지역단위의 통계는 상대적인 무관심속에 등한시 되어왔다. 그러나 지방자치제가 도입된 현시점에서는 지역의 경제, 사회적 발전을 위하여 지역 분석 및 각종 개발 계획의 수립, 집행, 관리, 평가를 목적으로 하는 지역통계에 대한 수요는 명확하게 계속적으로 증가되고 있는 실정이다 [국토개발 연구원, 1986]. 따라서 국가의 통계도 전국단위의 통계편중에서 지역통계를 중시하는 방향으로 전환되어야 할 것이다. 그럼에도 불구하고 아직도 실효성 있는 지역통계는 생산되지 못하고 있다. 그 원인을 우리나라 통계의 제도적인 측면과 종류 및 내용적인 측면 두 가지로 구분하여 설명하고 있는데 [노근호, 1990] 가장 큰 이유는 발표되는 통계가 주로 전국 및 시, 도 단위로 한정되어 있기 때문이라 하겠다. 또한 그렇게 발표되는 통계자료나마 지역통계의 공급체계상의 문제점으로 결과보고서의 배포가 제한되어 있어 [김선기, 임석희, 1990] 필요한 자료를 구하려면 통계정보 이용자가 작성기관을 일일이 파악하여 자료를 문의하거나 직접 방문하여야 하기 때문에 많은 시간과 노력이 소요되는 어려움이 있다.

현대사회는 정보사회라 불리며 점점 복잡해지는 사회에서 개개인들 마저도 시간적, 공간적 문제를 해결하기 위하여 홈 बैं킹 같은 수단을 동원하고 있는데, 지역의 발전을 위하여 정책을 수립하고 운영하며 지역에 관련된 연구를 하는 사람들이 지역에 관한 통계자료를 구하기 위하여 많은 낭비를 하는 데는 큰 문제가 있다고 하겠다. 이와같이 지역에 관련된 통계자료의 활용에 따르는 현실성을 반영하면 이미 발표된 지역관련 통계일지라도 과거부터 현재까지 시계열적으로 수집, 수록하여 지방의회, 기업가와 연구자들이 필요할 때 쉽게 효율적으로 활용할 수 있도록 지역통계의 데이터 베이스 구축은 매우 시급하다고 하겠다.

2. 지역 통계 현황

지역통계는 정부기관(중앙정부, 지방자치단체), 지정 통계기관 및 기타 관련기관에서 분산, 작성되고 있다. 그 중에서도 중앙정부에서 작성되어 발표되는 지역통계가 신뢰성 있는 것으로 알려지고 있다. 따라서 중앙정부에서 실시하는 통계조사의 종류 및 결과발표의 지역단위를 중점적으로 살펴본다. 통계조사 및 발표는 주로 통계청의 통계조사국에서 담당하는데 1991년 9월 현재 정부의 승인하에 실시되는 통계조사는 30종류이며[경제기획원 조사통계국, 1989] 이들의 발표되는 행정단위를 살펴보면 대부분이 전국 및 시, 도 단위로만 공표되고, 읍, 면, 동의 소 단위 행정구역까지 발표되는 것은 세 종류(인구총조사, 주택총조사, 총사업체 통계조사) 뿐이다. 특히 재정수지 및 지역소득에 관한 통계는 아예 공표되지 않고 있다. 한편 공표자료의 배부 또한 서울 위주로 매우 제한되어 지방의 정책입안자, 기업가 또는 연구자들이 쉽게 활용 하는데 많은 어려움이 있다.

지방자치의 실시와 더불어 통계청에서는 지역별 통계자료의 수요가 차츰 증대되고 다

양화 되리라고 전망하여 이에 대처하고, 우리 나라 지역에 관한 유용한 통계자료를 중앙 및 지방자치단체의 각종 정책수립과 지역 관련 연구기관등에 제공하기 위하여 1988년부터 매년 '지역통계년보'를 발간하고 있으며, 해마다 이에 수록되는 통계의 범위는 확대되고 있음을 밝히고 있다. 그러나 1990년 지역통계년보에 수록된 지역통계의 종류를 분야 및 발표되는 행정단위는 소 행정단위 읍, 면, 동에 해당하는 통계는 전혀 없으며, 구, 시, 군의 행정단위까지 수록되는 통계마저도 전체의 10%에도 미치지 못한다 [경제기획원 조사통계국, 1990 a]. 또한 통계청에서는 1976년부터 여러기관에서 작성되는 통계정보를 컴퓨터에 체계적으로 입력하여 필요한 이용자들이 온라인으로 적시에 편리하게 이용할 수 있는 통계정보시스템(KOSIS, Korea Statistical Information System)을 구축하여 이용자들이 사용할 수 있도록 하였다 [경제기획원 조사통계국, 1988; 통계청, 1991]. KOSIS를 이용하여 검색할 수 있는 통계는 총 275,256계열의 통계자료로 30년의 시계열을 유지하고 있다고 밝히고 있다. 그러나 KOSIS 역시 전국을 단위로 발표된 통계의 수록이 76%로 가장 많고, 시, 도 단위의 통계수록은 22%를 차지하는 반면에 구, 시, 군 행정단위 별 통계정보의 수록은 2%에 불과하고, 읍, 면, 동 단위별 통계는 전혀 없다. 따라서 소 행정단위의 지역개발계획 및 연구의 기초자료로 활용될 수 있으리라는 기대를 하기가 어렵다.

광주, 전남지역에 있는 통계기구는 통계청 전남통계사무소와 광주시와 전라남도의 통계담당관실로 이원화 되어 있으며 상호 업무의 협조는 이루어지지 않고 있다. 통계청 전남 통계사무소에서는 지장자치제가 도입된 후에도 이전의 역할과 동일하게 여전히 경상조사 19가지와 특별조사 7가지로 지역경제와 관련된 통계조사도 상당수 되지만 조사된 자료를 제도적으로 이 지역에서 보관할 수 없어서 해당 지역에서 이용하기에는 불가능한 현실이다 [이화영, 이희춘, 1991]. 따라서 지역에서 실시하는 조사를 활용하여 지역통계를 발전시킬 수 있는 방안이 마련 되어야 한다. 이와 관련하여 통계청의 기획국장은 1991년 한국 통계학회에서 다음과 같이 건의 하였다 [김일현, 1991]. 지방의 통계사무소에서 조사하여 중앙에 보고하는 광공업동태조사, 경제활동인구조사 이외의 주요 경상조사의 지역통계화를 추진하여 지역별로 생산, 고용, 물가등 각 부문별 통계를 균형적으로 작성할 수 있게 하며, 지방자치단체에서 실시하는 상주인구조사의 문제점을 개선하고 농수산부문의 지역 단위 통계를 확대하는 등 지역개발정책의 기초자료를 생산할 수 있도록 한다는 것이다.

광주, 전남지역 지방자치단체에서 실시되는 통계조사 역시 지방자치제의 도입전과 바뀐 것은 전혀 없다 [광주직할시, 1986, 1990; 전라남도, 1990]. 다른 지역의 지방자치단체에서는 자체 통계조사도 실시되고 있는 점(서울시의 서울시 인구동태조사, 서울시민 소득추계, 서울시 인구가동조사와 부산시의 차량교통량조사, 충청남도의 농어가 소득조사, 경상북도의 기업경영과 경기전망, [김선기, 임석희, 1990])을 감안할 때 광주와 전남지역에서도 지역특성에 맞고 계획수립에 기초자료가 될만한 자체조사의 개발이 시급하게 이루어져야 할 것이다.

3. 데이터 베이스 구축

계획수립의 기초단계에 필요되는 지역관련 통계가 발표되는 간행물은 아주 다양하다

[노봉남 외, 1988; 송인성, 1987]. 통계정보 이용자들이 간행물을 손쉽게 구하지 못하는 어려움을 해소하기 위하여 이미 발표된 통계자료는 가능한 연도별, 그리고 구, 시, 군과 읍, 면, 동 단위별로 수집하고, 간행물로 발표되지 못한 자료는 해당기관에 의뢰하여 수집하도록 한다. 이와 같이 수집된 많은 통계정보는 정보의 신속한 처리와 이용의 과학화를 필요로 하며 최근 수년간 컴퓨터의 빠른 보급으로 수집된 정보를 축적하고 필요한 것을 쉽게 찾도록, 그리고 용도에 따라 재구성이 가능하도록 자료를 화일로 만드는 절차가 절실히 요구되고 있다. 이와 같이 데이터 베이스가 이루어지면 이용자들이 개인용 컴퓨터를 이용하여 간편하고 신속하게 필요한 통계자료의 검색은 물론 장기적인 보관과 관리가 용이하게 된다. 따라서 본 연구에서는 이같은 목적에 부합될 수 있도록 지역통계관리 프로그램을 개발하였다. 데이터 베이스는 작성, 축적, 제공의 3과정으로 구분하고 있는데 첫째로 조사단위 지역과 통계항목의 코드화가 이루어져야 한다고 지적되었다 [김선기, 임석희, 1990].

한편 본 연구에서는 우선 우리나라 15개 시, 도중에서 광주직할시와 전라남도에 속하는 구, 시, 군 및 읍, 면, 동에 관한 지역통계로 제한되었으며, 전국 수준의 통계와 비교를 위하여 대한민국이 포함되었다.

먼저 연구대상지역인 광주직할시와 전라남도에서 최초의 지방자치단체인 시와 군이 그 자치행정을 수행할 때 필요로 하는 행정단위인 읍, 면, 동을 지역통계의 최소단위로 하였다. 수집된 자료의 등록을 간편하게 하기 위하여 이들 지역에 속하는 구, 시, 군 및 읍, 면, 동은 통계청의 '행정구역분류표' [경제기획원 조사통계국, 1990 b]에 준하여 고유번호가 부여되었다. 즉, 각 시, 도에 두자리, 구, 시, 군에 두자리, 읍, 면, 동에 두자리 등 여섯 숫자로 구성하였다 ([이화영, 이회춘, 1991] 부록참조).

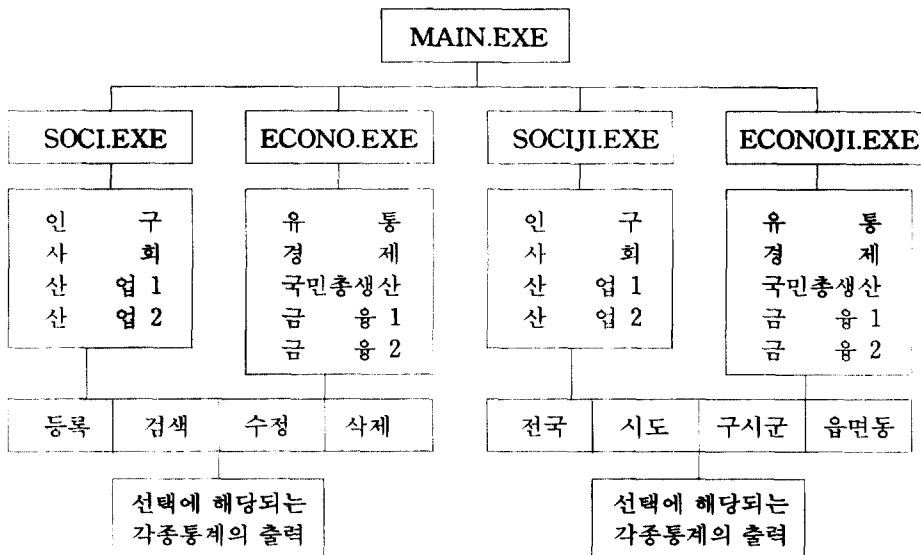
시, 도에서 계획수립에 필요한 분야별 통계자료를 국토개발연구원에서는 지역공간, 지역주민, 지역경제, 기본수요, 유통시설, 복지시설, 주민의식, 행정조직등 8가지 분야로 구분하고, 각 분야별 기본자료에 대하여 제시하였다 [국토개발연구원, 1981]. 그러나 이것이 엄청나게 많은 통계종류에 관한 자료를 한번에 전부 수집 및 수록한다는 것은 매우 어렵다. 따라서 통계청에서 통계조사 및 발표를 담당하는 통계조사국에 소속된 과를 기준으로 하여 인구, 사회, 산업1, 산업2, 유통, 경제, 금융으로 대분하여, 각 분야에서 가장 시급하다고 간주되는 항목을 다음과 같이 선정하였다.

분	야	항	목
인	구	-	가구수
		-	성별, 연령별(5세) 인구수
사	회	-	경제활동인구, 비 경제활동인구
		-	산업별 사업체수 및 근로자수
산	업 1	-	종업원 규모별 사업체수
		-	종업원 규모별 근로자수
산	업 2	-	광공업 사업체수
		-	종업원수
		-	생산액
		-	부가가치

분 야	항 목
유 통	- 도, 소매, 음식, 숙박업별 사업체수 - 도, 소매, 음식, 숙박업별 종업원수 및 판매액 - 수입, 수출
경 제	- 세입(원천별 일반회계 및 특별회계) - 세출(기능별 일반회계 및 특별회계)
국민총생산	- GNP/GRP
금 용 1	- 점포수 - 총금융저축 - 예금은행, 예금, 대출금
금 용 2	- 제2금융권 예금, 대출금 - 지방은행별 예금, 대출금

선정된 항목 중 인구분야는 읍, 면, 동단위까지 자료의 수집이 가능하였으나, 국민총생산은 시, 도 단위까지, 그리고 나머지 분야에서는 구, 시, 군 단위까지 자료수록이 가능하였다.

컴퓨터를 이용한 통계 데이터 베이스의 구축에는 (1) 자료의 등록, 추가 갱신 (2) 자료의 검색, 가공, 편집 (3) 자료의 해석, 모델분석의 세가지 영역을 필요로 하며 이를 위한 시스템의 구성도가 제시되었다 [김선기, 임석희, 1990]. 본 연구에서는 제시된 세가지 영역중에서 첫번째와 두번째의 영역에 중점을 두어 프로그래머인 Quick BASIC V. 4.5를 사용하여 프로그램을 개발하였다. 프로그램 작성의 효율을 높이기 위하여 선정된 분야 중 인구, 사회, 산업1, 산업2는 사회 분야로, 나머지는 경제 분야로 구분하였다. 프로그램은 크게 등록, 검색, 수정, 삭제와 지역별 검색(메뉴방식검색)으로 나누었으며 이를 위한 프로그램의 구성도는 그림1과 같다.



< 그림 1 > 지역통계관리 프로그램 구성도

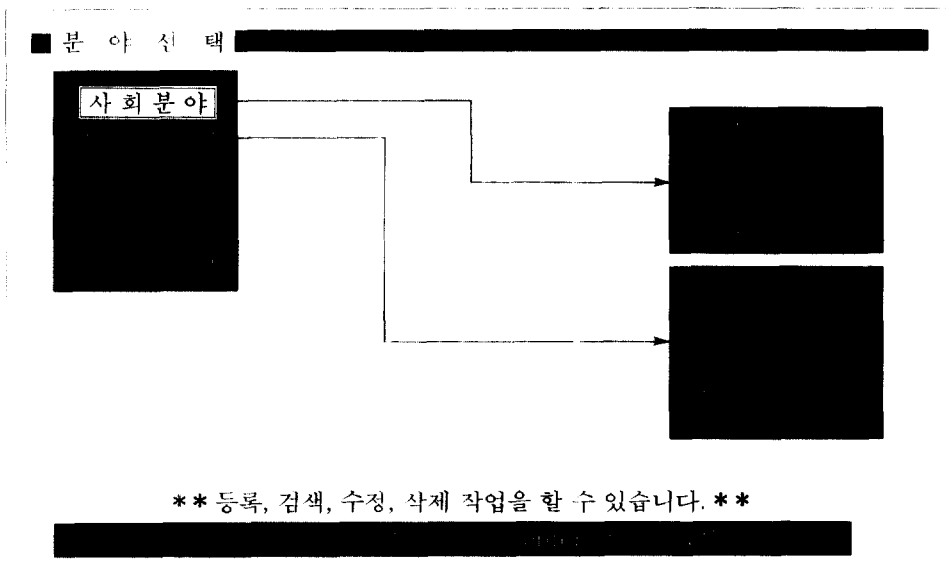
각 화일의 내용은 다음과 같다.

- 1) MAIN.EXE : 주 메뉴를 실행하는 프로그램
- 2) SOCI.EXE : 사회분야(인구, 사회, 산업1, 산업2) 통계를 등록, 검색, 수정, 삭제할 수 있는 실행 프로그램
- 3) ECONO.EXE : 경제분야(유통, 경제, 국민총생산, 금융1, 금융2) 통계를 등록, 검색, 수정, 삭제할 수 있는 실행 프로그램
- 4) SOCOJI.EXE : 사회분야 통계를 지역별 메뉴방식으로 검색 및 출력할 수 있는 실행 프로그램
- 5) ECONOJI.EXE : 경제분야 통계를 지역별 메뉴방식으로 검색 및 출력할 수 있는 실행 프로그램

4. 사용방법

작성된 프로그램을 사용하려면 우선 프로그램을 컴퓨터의 하드디스크에 설치해야 한다 (설치 절차는 [이화영, 이회춘, 1991]의 부록에 자세하게 설명되어 있으며, 이용을 희망하는 사람은 광주은행 조사부에 요청하면 된다). 프로그램이 설치되면 화면 1과 같은 초기 화면이 나타난다. 여기에서는 개발된 프로그램이 할 수 있는 기능인 등록, 검색, 삭제, 수정 및 지역별검색중에서 등록에 관하여 간략하게 설명한다. 나머지 기능에 대한 사용법은 등록과 유사하며, 자세한 설명은 [이화영, 이회춘, 1991]의 부록을 참조하면 된다.

〔화면 1〕 실행의 초기화면



자료의 등록은 두 가지 방법으로 할 수 있다. 첫째로 이 프로그램에서 분야마다 자료를

읽어 들이는 입력방식에 따라 자료화일을 만들어 자료를 등록할 수 있다. 이 경우에는 앞에서 각 지역에 부여된 코드를 사용하여 입력할 수도 있으나 매번 해당지역의 코드를 정확하게 기재하여야 하고 이 코드를 다시 해당지역명으로 변환시키는 과정을 거쳐야 하는 불편성이 있다. 둘째로는 메뉴방식으로 등록하는 것이다. 이는 설치된 프로그램의 초기화면에서(화면 1 참조) 사회 분야를 선택하여 '리턴'키를 누르면 화면 2가 나타나고 여기에서 등록하고자 하는 통계를 화살표키를 이용하여 선택한 뒤 '리턴'키를 친다. 예를 들어 인구를 선택했다면 화면 3이 나타나며 커서가 있는 밑줄 그어진 부분에 해당 자료를 입력한다. 한 항목의 입력이 끝나면 '탭'키를 치면 다음 항목으로 이동하게 된다. 자료의 입력이 끝나면 '맞습니까?(Y/n)' 라는 물음이 화면에 나타나고 ([이화영, 이희춘, 1991]의 부록화면 4 참조) 이때 입력된 자료의 내용이 맞으면 'Y'를 쳐서 자료의 등록을 할 수 있고, 만일에 틀렸을 경우에는 'n'을 치면 화면 3이 다시 나타나서 자료를 다시 입력할 수 있다.

〈화면 2〉

〈화면 3〉

년 령	남	여	년 령	남	여
0~4	_____	_____	5~9	_____	_____
10~14	_____	_____	15~19	_____	_____
20~24	_____	_____	25~29	_____	_____
30~34	_____	_____	35~39	_____	_____
40~44	_____	_____	45~49	_____	_____
50~54	_____	_____	55~59	_____	_____
60~64	_____	_____	65~69	_____	_____
70~74	_____	_____	75~79	_____	_____
80~84	_____	_____	85~	_____	_____

5. 통계적 분석

지역통계의 데이터 베이스 구축 가능성을 시사해주기 위하여 선정된 항목의 등록, 검색, 수정, 삭제에 대해 본 연구과정에서 개발된 프로그램은 김선기, 임석희(1990)이 제시한 시스템의 (1)과 (2)인 자료이용 하부 시스템에 해당된다. 여기에서는 제시된 시스템의 하부에 해당하는 자료분석으로 기존의 통계패키지(BMDP, SAS, SPSS, STATGRAPHICS 등)를 이용하여 검색된 자료의 통계적 분석의 가능성을 설명한다. 이 분야는 사용자가 통계패키지에 대한 사전 지식이 있어야 하며 이를 이용하려면 우선 분석될 자료가 준비되어야 한다. 분석하고자 검색된 자료가 많지 않으면 패키지 안에서 자료를 입력(Inline data)할 수도 있지만 검색된 자료의 양이 많으면 자료를 매번 다시 입력 한다는 것은 인력과 시간의 낭비이다. 따라서 본 연구에서 개발된 프로그램에서는 자료의 검색과정중 두 번째 방법인 지역찾기에서 검색된 자료를 화면이나 프린터 출력이 아니고 화일출력을 선택하면 검색된 자료가 화일로 저장되어, 이 화일을 통계패키지 사용시 외부 화일로 불러들여 분석하는 방법을 제시하였다(자세한 예제는 [이화영, 이희춘, 1991] 참조).

6. 제안

지역에 관련된 통계가 생산되는 단체는 많으며, 이 기관들이 발표하는 자료를 즉시 입수하여 추가 등록하는 일은 쉽지 않다. 따라서 새로이 발표되는 통계를 신속하게 등록하기 위해서는 지역내의 통계관련 모든 기관의 통계작성 관련자들이 정기적으로 모여 그들이 최근에 발표한 통계정보를 제공할 수 있는 모임이 활성화 되어야 할 것이다. 이보다 더 바람직한 방법은 통계작성 기관들이 온라인으로 연결하여 새로운 통계발표 즉시 그 자료를 송부하여 추가 등록할 수 있는 방법이 모색 되어야 할 것이다.

지역에 관련된 통계가 중앙 및 지방자치단체 및 각종 기관에서 발표되기도 하지만 최근에는 지방자치단체, 언론기관 및 민간단체에서도 통계조사를 실시하여 그 결과를 발표하는 경우가 많다. 이들 대부분이 조사된 사항에 대하여 퍼센트 분포 정도만 발표하고 더 이상의 분석은 하지 않고 자료를 사장 시키고 있다. 그런데 행정가, 기업가 또는 연구자들이 계획 및 연구를 수행하는데 필요한 귀중한 정보가 여러 기관에서 조사하여 수집된 자료를 각자 보관하고 있어 후에 타인이 분석을 원할 때 다시 사용 하는데 상당한 어려움이 따른다. 따라서 대학의 연구소 또는 지역통계의 데이터 베이스를 주관하는 곳에서 지역내에서 실시된 조사에 대한 제목, 조사항목, 조사시기, 조사대상자들을 체계 있게 정리하여 유사한 종류의 조사를 원하는 기관 및 개인들에게 참고가 될 수 있도록 한다. 이와 함께 조사자료를 체계적으로 보관, 관리하여 재 분석을 원하는 개인에게 원자료(raw data)를 대여해 주고 분석목적에 타당한 통계분석기법을 상담하여 줌으로써 많은 비용이 투입된 조사자료가 효율적으로 활용되어 지역관련 통계자료의 중요한 출처가 될 수 있는 방안도 강구 되어야 할 것이다.

참고문헌

- [1] 경제기획원 조사통계국(1985), 「인구 및 주택센서스 보고」.
- [2] _____ (1988), 「통계데이터베이스 이용안내서」.
- [3] _____ (1989), 「한국통계조사현황」.
- [4] _____ (1990 a), 「지역통계연보」 제3호.
- [5] _____ (1990 b), 「행정구역분류표」.
- [6] 광주직할시(1990), 「광주통계연보」.
- [7] _____ (1991), 「91시정주요통계」.
- [8] 국토개발연구원(1981), 「도건설 종합계획 작성요령」.
- [9] _____ (1986), 「지역통계정비방안연구」.
- [10] 김선기, 임석희(1990), 「지방자치제에 대비한 지방통계기반의 정비방안」, 「한국지방행정연구」.
- [11] 김일현(1991), 「정부통계의 발전방향」(초록), 한국통계학회 춘계학술논문발표회.
- [12] 노근호(1990), 「충북지역 통계정비 방안연구」, 「충북경제연구」 창간호, 충북경제연구회.
- [13] 노봉남 외(1988), 「지역기초통계의 현황과 확립방안」, 「전남지역 경제조사」 제9호, 전라남도, 광주은행.
- [14] 송인성(1987), 「지역개발정책의 이론과 수단에 관한 연구」, 고주노강희박사 회갑기념논문집, 국가발전과 공공행정.
- [15] 이화영, 이희춘(1991), 「광주, 전남지역통계 DATA BASE 구축방안」 「학술연구보고서」, 광주은행.
- [16] 전라남도(1990), 「전남통계연보」.
- [17] _____ (1989), 「상주인구조사보고서」.
- [18] 통계청(1991), 「통계정보시스템」 -이용법해설-.