

「2008 국제공동연구」 국외출장 결과 보고서

매크로 데이터의
비밀보호방법에 관한 연구

2008. 10.

통계개발원

목 차

□ 매크로데이터의 비밀보호방법 국제공동연구결과(요약)	1
I. 출장 개요	1
II. 연구 결과	4
가. 연구 목적	4
나. 연구 내용	4
다. 향후 계획	11
□ 연구보고서	12
제 1절 서 론	18
제 2절 집계표 데이터의 비밀보호 방법	20
제 3절 산업구조통계의 비밀보호 현황	49
제 4절 실증분석	66
제 5절 결 론	81
< 부 록 >	84
A. 종사자규모별 주요지표 보조셀감추기 방법 적용 결과	84
B. 출하액규모별 주요지표 보조셀감추기 방법 적용 결과	101

매크로 데이터의 비밀보호방법 국제공동연구 결과(요약)

I. 출장개요

가. 출장목적

- 통계이용자들의 자료제공 요구 시 개인정보의 비밀보호에 대한 중요성이 대두되어 통계자료의 비밀보호 기법의 습득과 활용이 요구되고 있음
- 이에 따라 통계개발원에서는 국제공동연구사업의 일환으로 '06년부터 3개년 계획으로 「통계자료 비밀보호 기법」에 관한 국제공동연구를 추진하였음
- '06~'07년 기간 중에는 마이크로 데이터(이산형, 연속형)의 비밀보호 기법연구를 수행하였으며, '08년에는 매크로(집계표) 데이터의 비밀보호에 대한 해외전문가와 국제공동연구 수행을 통해,
 - 매크로데이터의 통계적 비밀보호 방법 관련 선진기법의 습득
 - 동 선진기법을 사업체단위 통계조사 결과에 적용함으로써 개인정보 보호 및 통계자료의 확대 제공을 위한 통계정보의 유용성 확보
 - 비밀보호와 관련된 지속적인 연구 및 실무적용을 위함

나. 출장기간

- 2008. 7. 12 ~ 9. 5 (8주간)

다. 출장자 및 출장지역

소 속	직 급	이 름	출장지역
경제통계실	통계사무관	정 동 욱	미 국
연구기획실	통계주무관	김 경 미	(워싱턴 D.C.)

※ 방문기관 : 국립보건통계센터(National Center for Health Statistics)

라. 주요 수행사항

- 매크로데이터의 비밀보호 기법에 대한 기초연구
 - 선행연구 결과에 대한 문헌연구
- 비밀 보호된 매크로 데이터의 파일작성(SAS 프로그램 이용)
 - 대상자료 : 사업체 단위 통계조사
 - 결과자료 : 매크로데이터의 비밀보호기법 적용파일 작성
- 연구보고서 작성 및 연구결과 발표 등

마. 소요예산

- 소요예산 : 29,000천원 내외

바. 기대효과

- 통계청에서 실시하는 각종 통계조사의 집계표 데이터 제공시 응답자의 비밀보호를 위한 효과적인 적용 기대
- 향후 매크로데이터의 비밀보호 기법에 대한 지속적인 공동 연구의 발판 마련

사. 세부일정

	수 행 사 항	접촉인사	비고
7.12(토)	인천 → 워싱턴D.C.	-	이동
1주차 (7.14~20)	- 기초문헌연구 · 매크로자료의 비밀보호기법 개념 · 주요 기법의 적용사례 검토	Lawrence H. COX (부서장)	-
2주차 (7.21~7.27)	- 자료분석 I · 분석대상 자료의 특성 파악 (SAS프로그램을 이용한 분석)	"	SAS 프로그래밍
3주차 (7.28~8.3)	- 자료분석 II · 분석대상에 적합한 보호기법 탐색	"	"
4주차 (8.4~8.10)	- 자료분석 III · 보호기법별 시뮬레이션 수행	"	"
5주차 (8.11~8.17)	- 자료분석 IV · 보호기법별 시뮬레이션 수행(계속)	"	"
6주차 (8.18~8.24)	- 자료분석 V · 시뮬레이션 결과에 대한 유의성 분석	"	"
7주차 (8.25~8.31)	- 자료분석 VI · 시뮬레이션 결과에 대한 유의성 분석(계속) · 통계조사에 적합한 보호기법 도출	"	"
8주차 (9.1~9.3)	- 연구결과 정리 · 연구결과 보고서 작성	"	-
9.4(목) ~9.5(금)	워싱턴D.C. → 인천	-	이동

II. 연구결과

가. 연구목적

- 자료의 개별정보에 대한 비밀보호강화와 제공범위 확대 증가
 - 통계수요측면에서 보면, 다양한 분석을 위한 세부자료와 신규통계작성에 대한 수요가 증가하고 있음
 - 반면, 통계작성기관과 응답자는 개인 및 사업체의 비밀 보호 강화에 대한 요구 증대
- 두 제약점을 해결하는 대안마련 시급
 - 각종 정보의 DB화 자료결합 등으로 개별정보 노출위험이 증가하고 있으며, 정보의 원천인 응답자의 비밀보호 준수를 최우선으로 실현하기 위한 대안 마련 모색
- 제공범위 확대와 응답자의 비밀보호를 동시에 만족시킬 수 있는 비밀보호 기법 마련을 위한 연구 필요

나. 연구내용

- 대상자료 : 자료제공 형태중 집계표(매크로) 데이터
- 집계표 데이터형태
 - 집계표는 크게 두 가지 데이터 형태로 구성되어 있음
 - 빈도데이터(Count Data) : 개수 자체에 의미가 있는 자료
 - 크기데이터(Magnitude Data) : 개수 보다는 양적인 부분에 의미가 있는 자료
 - 한국과 미국의 집계표 형태 비교
 - 한국 : 표두(Column) 항목에 빈도데이터와 크기데이터 혼합형

- 미국 : 빈도데이터와 크기데이터를 각각 구분하여 집계표를 구성

○ 집계표 데이터의 비밀보호 방법

- 노출위험이 있는 민감한 셀 찾는 방법

- 빈도데이터 : 기준(Base)값 이하인 셀을 노출위험이 있는 셀로 정의

- 크기데이터

① (n,k) dominance rule

셀을 구성하는 사업체별 조사결과가 셀 합계의 k% 이상을 차지하는 경우

② p-percent rule

셀을 구성하는 사업체별 조사결과중 두 번째로 큰 값 (X2)이 가장 큰 값(X1)을 p% 이내에서 추정 가능한 경우

③ p/q- ambiguity rule

p% rule을 강화한 방법으로 개별정보중 일부를 알고 있을 가능성이 있을 때 적용하는 방법

- 민감한 셀 : $S_{pq}(X) = x_1 - \left(\frac{q}{p}\right) \times x_{3+} > 0$

- 집계표 데이터 형태에 따른 비밀보호 방법

빈도데이터	크기데이터
반올림(Rounding) 셀 감추기(Cell suppression) 자료변조(data perturbation) 분류구조변경(data collapsing) 기타(swapping, switching 등)	셀 감추기(Cell suppression)

① 반올림(Rounding)

- 전통적인 반올림
: 기준(Base)을 정하고 0과 기준값(B)으로 변환하는 방법
- 제어된 반올림
: 모든 셀값이 기준(Base)의 배수가 되도록 가감을 반복하는 방법 (주변합과 총합을 바꾸지 않으면서 반올림)
- Zero Restriction 50/50 Rounding
: 0과 기준값(B)으로 변환하는 과정에서 난수를 발생하여 확률적인 개념에서 반올림을 하는 방법

② 셀감추기(Cell Suppression)

- 노출위험이 있는 셀만 감추는 방법(SODC)
: 기준(Base) 이하의 셀만 감추기 하는 방법
※ 집계표내의 선형관련식으로 감추어진 셀을 찾을 수 있음
- 보조셀감추기(CCS)
: 직접적인 노출이 있는 셀을 감추기 위해 주변의 셀을 보조적으로 감추기 하는 방법
※ 집계표내의 선형관련식으로 감추어진 셀을 찾을 수 없는 장점이 있음

③ 기타

- CTA(Controlled Tabular Adjustment)
- 분류구조방법(Data Collapsing) 등

○ 비밀보호된 집계표 데이터 작성 (별첨보고서 참조)

- 분석대상 자료 : 2006년 기준 광업 및 제조업 통계조사
중분류 27. 제1차 금속산업

<분석대상의 주요지표>

(단위 : 개, 명, 10억원)

	사업체수	월평균 종사자수	출하액	부가가치	유형자산 연말잔액
제조업 전체	119,181	2,910,935	909,067	326,844	310,176
중분류27	3,173	117,684	88,721	25,282	30,990
구성비	2.7%	4.0%	9.8%	7.7%	10.0%

- 적용기법 : 보조셀 감추기(Complementary Cell Suppression)

$$\text{목적함수(1): } V(a) = \min \sum a_{ij} z_{ij}$$

$$\text{목적함수(2): } V(a) = \min \sum z_{ij}$$

- (1) (2)중 제한되는 정보량을 최소화하는 목적함수(1)을 적용
- 최적경로 탐색은 3가지 접근법중 네트워크기법 적용
 - i) 결합기법(combinational technique)
 - ii) 선형기법(linear programming)
 - iii) 수학적 네트워크기법(mathematical network)

- 세부 적용단계

- Step 1 : PDC(Primary Disclosure Cell)을 정의함
- Step 2 : <표1>과 같이 임의의 PDC(Primary disclosure cell) a_{81} 을 출발점으로 가능한 모든 경로(Circuit)를 찾은 후, 각 경로에 해당하는 셀 값의 합을 구함
- Step 3 : 경로 중 「0」 값을 포함하고 있는 경우, 이를 찾아 제외시킴. 왜냐하면 「0」 값은 보고서를 통해 이용자에게 제공해야 하는 정보로서 보조감추기셀(Complementary Suppression Cell: CSC)로 이용할 수 없기 때문임
- Step 4 : Step1, Step2에서 남은 경로 중 감추어진 셀 값이 최소인 경로를 선택함 <표1>
- Step 5 : 모든 PDC에 대한 보조감추기 경로를 찾아 총합을 최소화하는 셀을 선택함<표2>
- Step 6 : 동일한 방법으로 PDC a_{71}, a_{53}, a_{63} 에 대한 최소경로 탐색결과는 <표3>과 같음

<표2> 보조셀감추기(CCS) 결과 (세분류)

산업분류 종사자규모	D2721	D2722	D2729
계	217	467	64
5 ~ 9	72	147	22
10 ~ 19	79	137	16
20 ~ 49	47	101	21
50 ~ 99	CSC(11)	45	CSC(3)
100 ~ 199	CSC(4)	21	PDC(1)
200 ~ 299	0	CSC(7)	PDC(1)
300 ~ 499	PDC(2)	CSC(5)	0
500명 이상	PDC(2)	CSC(4)	0

<표3> 보조감추기 셀 찾기(a_{71}, a_{53}, a_{63})

PDC	Circuit			감추어진 셀값의 합	찾기
(7,1)=2	(7,2)	(8,2)	(8,1)	13	최소경로
(5,3)=1	(5,1)	(4,1)	(4,3)	19	최소경로
(6,3)=1	(6,2)	(5,2)	(5,3)	30	최소경로

- 타당성 검증

- 감추어진 셀의 수와 정보량을 통해 노출위험이 있는 셀만 감추기하는 방법과 보조셀 감추기 방법을 비교 <표4, 5>
- CCS방법이 SODC 적용결과 보다 감추어진 셀과 정보량이 크나, 이는 SODC방법에 의해 충분한 비밀보호가 이루어지지 않았기 때문임

<표4> CCS와 SODC방법 적용결과 비교 (감추어진 셀의 수)

산업분류		D272	D2721	D2722	D2729
원자료	셀수(A)	24	8	8	8
CCS방법	감추어진 셀 수(B)	10	4	3	3
	$C=(B/A) \times 100$	41.7	50.0	37.5	37.5
SODC방법	감추어진 셀수(D)	4	2	0	2
	$E=(B/A) \times 100$	16.7	25.0	0.0	25.0
Difference	C - E	25.0	25.0	37.5	12.5

<표5> CCS와 SODC방법 적용결과 비교 (감추어진 정보량)

산업분류		D272	D2721	D2722	D2729
원자료	정보량(H)	748	217	467	64
CCS방법	감추어진 정보량(I)	40	19	16	5
	$J=(I/H) \times 100$	5.4	8.8	3.4	7.8
SODC방법	감추어진 정보량(K)	6	4	0	2
	$L=(K/H) \times 100$	0.8	1.8	0.0	3.1
Difference	J - L	4.6	7.0	3.4	4.7

- 선형관련식에 의한 재계산 여부를 확인하여 효과적인 비밀보호 여부 체크

$$\begin{aligned}
 a_{41} + a_{43} &= 14 \\
 a_{51} + a_{53} &= 5 \\
 a_{62} + a_{63} &= 8 \\
 a_{71} + a_{72} &= 7 \\
 a_{81} + a_{82} &= 6 \\
 a_{41} + a_{51} + a_{71} + a_{81} &= 19 \\
 a_{62} + a_{72} + a_{82} &= 16 \\
 a_{43} + a_{53} + a_{63} &= 5 \\
 a_{ij} &\geq 1
 \end{aligned}$$

<표2>의 결과를 예시로 위와 같은 선형관련식을 세우고 선형방정식을 풀면 유일한 해를 찾을 수 없기 때문에 성공적인 비밀보호가 이루어진 것으로 검증

※ 자세한 분석내용은 “연구보고서” 참조

다. 향후계획

- 광업 및 제조업 통계조사에 대한 비밀보호 방법의 추가 연구수행
 - 집계표 구성형태를 데이터유형(빈도 및 크기데이터)에 따라 구분하는 집계표 설계방법 연구
 - 데이터유형에 따른 다양한 민감한 셀 찾는 방법 적용 및 기법적용에 따른 정보제공 효율성에 관한 연구 수행
- 산업구조통계조사에 대한 비밀보호 방법 확대 적용방안 마련
 - 통계청에서 작성중인 산업구조통계조사에 대한 새로운 비밀보호기법 확대 적용방안 연구

연구보고서

매크로 데이터의 비밀보호방법에 대한 고찰

통계개발원

목 차

제1절 서론	18
제2절 집계표 데이터의 비밀보호 방법	20
1. 집계표 자료형태	20
2. 민감한 셀 찾는 방법	22
가. 빈도 데이터	22
나. 크기 데이터	22
1) (n, k) - dominance rule	23
2) p - percent rule	24
3) p/q - ambiguity rule	25
3. 비밀보호방법	27
가. 반올림 방법	28
1) 전통적인 반올림 방법	28
2) 제어된 반올림 방법	29
3) zero restriction 50/50 Rounding	33
나. 셀감추기 방법	35
1) 개요	35
2) 셀감추기 방법론	35
3) 보조셀감추기 방법론	37
가) 일반적인 CCS 방법론	40

나) 네트워크 이론을 통한 CCS 방법론	42
다. CIA 방법	46
1) 개요	46
2) CIA 방법론	47
제3절 산업구조통계의 비밀보호 현황	49
1. 광업 및 제조업 통계조사 작성 개요	51
2. 비밀보호 방법	52
3. 현행 SODC 방법 적용에 따른 한계점	64
제4절 실증분석	66
1. 분석대상	66
2. 비밀보호 현황 및 한계점	67
3. 보조셀감추기 방법의 적용	68
4. 타당성 검증	77
5. 실증분석 결과 요약	80
제5절 결 론	81
< 부 록 >	84
A. 종사자규모별 주요지표 보조셀감추기 방법 적용결과	84
B. 출하액규모별 주요지표 보조셀감추기 방법 적용결과	101

표 목 차

<표 2- 1> 집계표 예시(한국)	21
<표 2- 2> 집계표 예시(미국)	21
<표 2- 3> 빈도데이터의 민감한 셀 찾기(예제)	22
<표 2- 4> (3,80) - dominance rule(예제)	23
<표 2- 5> 20 - percent rule(예제)	24
<표 2- 6> 20/50 - ambiguity rule(예제)	26
<표 2- 7> 자료형태에 따른 전통적인 비밀보호 기법(Cox, 2008)	28
<표 2- 8> 전통적 반올림(예제)	28
<표 2- 9> (4×5) Original Data	30
<표 2-10> (5×6) Original Data	36
<표 2-11> HCCS방법 적용결과	38
<표 2-12> (3×3) Original Data	40
<표 2-13> (3×3) CCS방법 적용결과	41
<표 2-14> CCS방법에 따른 정보제한 총량	42
<표 2-15> Network에 의한 최소경로 탐색(a_{33})	44
<표 2-16> Network에 의한 최소경로 탐색(a_{23})	45
<표 2-17> 원자료 집계결과	48
<표 2-18> 셀감추기 적용결과	48
<표 2-19> CTA 방법 적용결과	48

<표 3- 1> 집계표 예시	50
<표 3- 2> 광업 및 제조업 통계조사 항목	51
<표 3- 3> 코크스, 석유정제품 및 핵연료 제조업(D23)	54
<표 3- 4> 셀값찾기 Step 1	58
<표 3- 5> 셀값찾기 Step 2	59
<표 3- 6> 셀값찾기 Step 3	59
<표 3- 7> 셀값찾기 Step 4	60
<표 3- 8> 셀값찾기 Step 5	60
<표 3- 9> 선형관련식에 의한 감추어진 셀 「 X 」 값 찾기 최종결과	61
<표 3- 10> 감추어진 셀 「 X 」 값 찾기(D2713)	65
<표 4- 1> 분석대상의 주요지표	67
<표 4- 2> 소분류별 사업체수(D272)	69
<표 4- 3> 세분류별 사업체수(D2721)	71
<표 4- 4> 보조셀감추기(CCS) 결과 (소분류)	72
<표 4- 5> 보조셀감추기(CCS) 결과 (세분류)	72
<표 4- 6> 보조셀감추기(CCS) 결과 (세세분류)	75
<표 4- 7> CCS와 SODC방법 적용결과 비교(감추어진 셀의 수)	77
<표 4- 8> CCS와 SODC방법 적용결과 비교(감추어진 정보량)	78

그림 목차

[그림 2-1] <표2-12>에 대한 Network Table	43
[그림 3-1] 코크스, 석유정제품 및 핵연료 제조업 분류체계	57
[그림 4-1] 제1차 금속산업 분류체계(중분류 27)	67

제 1절 서론

우리나라 산업구조는 대내외 환경변화에 따라 전반적으로 다양한 측면에서의 구조적인 변화기에 진입하고 있다. 통계수요 측면에서 보면, 종전의 개별 업종 중심의 접근(Sectoral Approach)에서 경제주체간의 협력과 보완관계를 중시하는 클러스터 접근(Cluster Approach)이 요구되는 구조적인 전환기를 맞고 있으며, 소득수준 향상으로 교육, 의료, 지식기반서비스업 등의 국민경제의 새로운 동력으로 제기되면서 서비스업에 대한 관심이 커지기 시작하면서 새로운 통계 수요를 창출하고 있다.

자료제공 측면에서 보면, 통계법 개정(2007.10.)과 아울러 통계이용자들의 자료요구에 대응하기 위해 응답자의 개별정보 노출을 제한하면서 활용성을 증대시킬 방안모색의 필요성이 대두되고 있다. 그러나 자료수집 시 절대적 개인 및 사업체의 비밀보호라는 서약에 따라 수집되고 있지만, 각종 정보의 DB화로 통계적 자료결합이나 분석과정에서 개인 또는 사업체의 비밀보호 자료가 공개될 위험은 더욱 커져가고 있다. 따라서 세분화된 정보까지 요구하는 이용자와 응답자의 개별정보 비밀보호라는 제약을 동시에 만족하기 위한 방법에 대한 연구의 필요성이 대두되고 있는 시기이다. 이와 같은 현안을 해결하기 위해 각 국은 여러 방법을 모색하고 있는데, 미국, 캐나다, 호주 등에서는 노출위험이 있는 자료에 대해 다양한 방법론을 적용하여 비밀보호된 자료를 제공하고 있으며, UNECE/Eurostat에서는 “통계자료비밀보호 국제공동회의(Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality)”를 개최하여 각 국의 연구결과 및 적용사례에 대한 토의를 활발하게 진행하고 있다.

통계청 통계개발원에서도 통계자료의 비밀보호에 대한 문제를 인식하고 국제공동연구사업의 일환으로 2006년부터 3개년간 「통계적 비밀보호기법(Masking)」의 연구를 추진하고 있는데, '06~'07년 기간 중에는 마이크로데이터(이산형데이터 및 연속형데이터)의 비밀보호 기법연구를 수행하였다. 이어 실시되는 본 연구는 미국 국립보건통계센터(National Center for Health Statistics :NCHS)에서 동 분야에 대한 선행연구를 수행한 Dr. Lawrence Henry COX의 연구를 토대로 매크로데이터(빈도데이터 및 크기데이터)에 관한 비밀보호 방법에 관한 연구이다.

본 연구에서 다루는 내용은 자료제공 형태 중 집계표(매크로데이터) 자료에 대한 비밀보호방법 연구로, 공개범위 제한 및 측정방법과 집계표에서 그 결과의 공개위험 정도를 평가하는 방법 등 노출위험이 있는 민감한 셀의 정보를 보호하여 응답기관 또는 응답자의 개별정보 노출을 제한하는 방법을 연구하는 것이다.

제2절에서는 집계표를 구성하는 자료형태인 빈도데이터(Count Data)와 크기데이터(Magnitude Data)에 따른 전통적인 비밀보호 기법에 대한 방법론과 예제를 소개하였다. 이는 집계표 데이터의 비밀보호방법론에 관해 전반적으로 고찰하였을 뿐만 아니라 통계청에서 적용 가능한 방법을 찾기 위한 기초자료로 활용될 것이다.

제3절에서는 「광업 및 제조업통계조사」의 집계표를 자세히 살펴봄으로써 산업구조통계의 비밀보호 현황과 한계점을 살펴본다. 특히, 노출위험이 있는 셀만 감추는 현행 비밀보호 기법(Suppressing Only the Disclosure Cells: SODC)의 한계점에 대해 선행관련식을 통해 비밀보호의 정도를 재조명해 본다. 이를 통해 연구된 한계점은 실증분석을 통한 새로운 기법 적용의 필요성과 결과를 도출하기 위한 근거로 활용한다.

제4절에서는 광업 및 제조업통계조사 23개 중분류중 매출액, 자산비중 등을 고려하여 적합한 중분류를 선정한 후, 기존의 SODC방식과 새로 적용해본 기법인 보조셀감추기(Complementary Cell Suppression: CCS)방법의 적용 결과를 비교하여 현 상황에서 적합한 비밀보호 기법을 제시하고자 한다.

제 2절 집계표 데이터의 비밀보호방법

1. 집계표 자료형태

집계표를 구성하고 있는 자료는 크게 ①빈도데이터(count data) ②크기데이터(magnitude data)로 구분할 수 있다. 「빈도데이터」는 사업체수와 같이 개수 자체에 의미가 있는 자료형태이며, 「크기데이터」는 매출액, 종사자수 등과 같이 규모에 의미가 있는 형태이다. 자료형태에 따라 노출위험이 있는 셀을 결정하는 방법에 차이가 있다. 따라서 적용하는 비밀보호 방법 또한 상이하기 때문에 집계표를 어떻게 설계하는가는 중요한 과정이다.(Cox, 2008)

한국과 미국은 집계표 구성에서 다소 차이가 있는데 제조업통계조사¹⁾를 통해 아래와 같이 비교해 보고자 한다. 한국의 집계표 형태는 <표2 - 1>과 같이 산업분류, 지역(시도), 종사자규모, 매출액규모 등을 조합하여 표측(Row)항목을 구성하고, 조사항목중 사업체수, 종사자수, 출하액, 유형자산 등 주요항목을 표두(Column)에 배치하고 있다. 즉, 빈도데이터와 크기데이터가 혼합되어 있으며, 비밀보호를 위한 기준변수(Key variable)로서 사업체수가 이용되고 있다.

1) 제조업통계조사 : 한국(광업 및 제조업통계조사), 미국(Annual Survey of Manufacture)

〈표2-1〉 집계표 예시(한국)

산업분류	종사자규모	사업체수	월평균 종사자수	급여액	출하액	...	유형자산 연말잔액
D19 가죽, 가방 및 신발 제조업	5~9명	1,181	7,212	84,543	565,117	...	110,002
	10~19명	435	5,696	90,273	619,753	...	124,855
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	300~499명	2	X	X	X	X	X
	500명이상	3	1,964	56,025	324,478	...	227,235

반면 미국의 제조업통계조사의 경우는 빈도데이터와 크기데이터의 집계표가 구분되어 각각 제공되고 있으며, 노출위험 여부를 결정하는데 상호 영향을 미치지 않도록 설계되어 있다. 이와 같은 집계표 구성 형태의 차이로 인해 비밀보호 방법을 적용함에 있어서도 각각에 적합한 방법이 다를 것이다. 이는 제 3절에서 다시 언급하기로 한다.

〈표2-2-1〉 집계표 예시(미국)

Table 1. Statistics for All Manufacturing by State: 2005
[Data based on the 2005 Annual Survey of Manufactures (ASM). For information on confidentiality protection, sampling error, nonsampling error, and explanation of terms, see Note at end of table. For meaning of abbreviations and symbols, see introductory text. For explanation of terms, see Appendix A.]

Geographic area	All employees		Production workers			Value added (\$1,000)	Total cost of materials (\$1,000)	Total value of shipments ^a (\$1,000)	Relative standard error of estimate ^b (percent) for column—	
	Number ^c	Payroll (\$1,000)	Number ^c	Hours (1,000)	Wages (\$1,000)				A	F
	A	B	C	D	E				F	G
United States	12 169 822	579 690 961	8 238 151	19 969 681	227 890 266	2 204 094 956	2 555 491 508	4 735 282 666	1	1
Alabama	264 159	13 016 229	205 496	424 471	6 679 496	28 372 848	49 884 530	87 680 546	1	1
Alaska	10 810	380 553	9 879	19 806	285 130	1 805 457	4 824 857	6 570 783	1	2
Arizona	158 995	7 654 132	99 810	206 893	3 742 100	23 726 129	20 891 994	43 233 559	2	1
Arkansas	185 459	6 325 482	149 258	324 555	4 524 044	26 991 983	31 425 409	59 180 229	1	1
California	1 297 881	63 922 953	801 594	1 829 756	32 789 986	217 548 932	218 653 894	424 238 357	1	1
Colorado	126 468	6 021 574	85 813	190 368	3 182 880	19 462 751	16 675 852	37 479 999	2	2

〈표2-2-2〉 집계표 예시(미국)

Table 2. Statistics for Industry Groups and Industries: 2005 and Earlier Years
[Data based on the 2005 Annual Survey of Manufactures (ASM). For information on confidentiality protection, sampling error, nonsampling error, and explanation of terms, see Note 2 at end of table. For meaning of abbreviations and symbols, see introductory text. For explanation of terms, see Appendix A.]

NAICS and NAICS-based code ^d	Industry group and industry	All employees		Production workers			Value added (\$1,000)	Total cost of materials (\$1,000)	Total value of shipments ^a (\$1,000)	Relative standard error of estimate ^b (percent) for column—	
		Number ^c	Payroll (\$1,000)	Number ^c	Hours (1,000)	Wages (\$1,000)				A	F
		A	B	C	D	E				F	G
31-33	Manufacturing	2005.. 12 169 822	579 690 961	8 238 151	19 969 681	227 890 266	2 204 094 956	2 555 491 508	4 735 282 666	1	1
		2004.. 13 294 079	659 703 575	9 365 138	19 263 917	322 675 474	2 641 433 991	2 283 143 959	4 289 570 420	1	1
		2003.. 13 627 966	667 802 408	9 706 881	19 462 882	329 480 112	1 923 424 910	2 086 279 624	4 016 287 242	1	1
		2002.. 14 644 385	675 165 127	10 319 528	20 421 721	326 540 663	1 889 290 940	2 022 159 349	3 914 719 162	X	X

2. 민감한 셀 찾는 방법

민감한 셀을 찾는 방법은 자료형태에 따라 다소 차이가 있는데, 다음에서는 빈도데이터와 크기데이터로 구분하여 민감한 셀을 찾는 방법을 살펴보고자 한다.

가. 빈도데이터(Count Data)

빈도데이터(Count Data)는 기준(Base) B값을 정의함으로써 노출위험이 있는 셀을 결정한다. 즉, 임의의 셀 값이 B 이하인 경우 해당 셀을 PDC(Primary Disclosure Cell)로 정의하고 해당 셀의 자료제공을 제한한다. <표2-3>은 기준 B가 2인 경우 노출위험이 있는 셀을 표시한 것이다.

<표2-3> 빈도데이터의 민감한 셀 찾기(예제)

산업분류 \ 종사자규모	D27211	D27212	D27213	D27219
5 ~ 9	9	40	9	14
⋮	⋮	⋮	⋮	⋮
300 ~ 499	0	1	1	0
500명 이상	1	0	1	0

나. 크기데이터(Magnitude Data)

크기데이터(Magnitude data)는 ① (n,k) - dominance rule ② p-percent rule ③ p/q ambiguity rule 등을 적용하여 해당 셀의 노출위험여부를 결정한다. 각각의 방법에 대해 자세히 살펴보면 다음과 같다.

1) (n,k) - dominance rule

이 방법은 임의의 셀에 대한 노출위험 여부를 결정할 때, 해당 셀을 구성하고 있는 각각(사업체별)의 값이 셀 합계의 k% 이상을 차지하는 경우 노출위험이 있는 민감한 셀로 정의하는 방법이다.

민감한 셀을 탐색하는 과정을 보여주기 위해, 우선 임의의 셀을 구성하는 L개의 사업체별 자료를 식(2.1)과 같이 내림차순으로 순차적으로 정렬하고, 셀 합계와 상위 n개 자료의 합을 아래 식(2.2), (2.3)과 같이 정의하자.

$$x_1 \geq x_2 \geq \dots \geq x_n \geq x_{n+1} \geq \dots \geq x_l \quad (2.1)$$

$$V(x) = \sum_{i=1}^l x_i \quad (2.2)$$

$$D(x_n) = \sum_{j=1}^n x_j \quad (2.3)$$

이 때 식(2.4)를 만족하면 임의의 셀은 노출위험이 있는 셀로 정의되어, 해당 셀에 대한 자료제공을 제한한다. 이와 관련한 선형민감도를 표현하면 식(2.5)와 같다.

$$D(x_n) > \left(\frac{k}{100}\right) V(x) \quad (2.4)$$

$$S_{n,k}(X) = \sum_{i=1}^n X_i - \left(X_{(n+1)^+} \cdot \left(\frac{k}{100-k}\right)\right) > 0 \quad (2.5)$$

〈표2-4〉 (3,80) - dominance rule(예제)

x_1	x_2	x_3	x_4	x_5	$V(x)$
100	50	40	30	20	240

〈표2 - 4〉의 예제를 통해 자세히 살펴보자. $n=3$, $k=80$ 으로 정의하면, 위의 (2.2) (2.3)식에 의해 $D(x_3)=190$, $V(x)=240$ 이다. 민감한 셀인지 여부를 판단하기 위해 (2.4)식을 이용해 계산해보면, 아래와 같이 $D(x_3)$ 의 값이 $V(x)$ 의 $k\%$ 값 보다 작으므로 이 셀은 노출위험이 있는 셀이 아님을 알 수 있다.

$$D(x_3) = 190 < \left(\frac{k}{100}\right) V(x) = \left(\frac{80}{100}\right) \times 240 = 192$$

2) p-percent rule

이 방법은 자료에서 두 번째로 큰 값에 해당되는 데이터를 이용해 제일 큰 값에 해당하는 데이터를 얼마나 추정할 수 있는가 하는 정도로 노출위험을 판단하는 방법이다.

민감한 셀을 탐색하는 과정을 살펴보자. 임의의 셀을 구성하는 L개의 사업체별 자료를 내림차순으로 순차적으로 정렬한 결과를 아래의 식(2.6)과 같이 정의하자.

$$x_1 \geq x_2 \geq \dots \geq x_n \geq x_{n+1} \geq \dots \geq x_l \quad (2.6)$$

이 때 x_2 값을 이용하여 x_1 값을 $p\%$ 이내에서 추정할 수 있다면 이 셀을 노출위험이 있는 민감한 셀로 판단한다. 즉, 아래의 식(2.7)을 만족하면 이 셀을 노출위험이 있는 셀로 정의하며, 선형민감도는 식 (2.8)과

같이 표현할 수 있다.

$$V(x) - x_2 < (100 + p)/100 \times x_1 \quad (2.7)$$

$$S_{p\%}(X) = x_1 - \left(\frac{100}{p}\right) \times x_{3+} > 0 \quad (2.8)$$

<표2-5> 20-percent rule(예제)

x_1	x_2	x_3	x_4	x_5	$V(x)$
100	50	40	30	20	240

<표2 - 5>의 예제를 통해 $p = 20$ 인 20-percent rule을 적용해 보자. 노출위험 여부를 알아보기 위해, 식(2.7)을 이용해 계산해 보면 아래와 같이 (2.7)의 부등식을 만족하지 않기 때문에 이 경우 노출위험이 발생하지 않는다고 할 수 있다.

$$240 - 50 > ((100 + 20)/100) \times 100$$

3) p/q-ambiguity rule

이 방법은 p-percent rule을 보완한 방법이다. 먼저 앞서 언급한 p% rule의 수식 (2.7)을 아래와 같이 다시 정리해 보자.

$$V(x) - x_2 < \left(1 + \frac{p}{100}\right) \times x_1$$

$$\Leftrightarrow V(x) - x_2 - x_1 < \left(\frac{p}{100}\right) \times x_1$$

$$\begin{aligned}
&\Leftrightarrow x_{3+} < \left(\frac{p}{100}\right) \times x_1 \\
&\Leftrightarrow \left(\frac{p}{100}\right) \times x_1 - x_{3+} > 0 \\
&\Leftrightarrow x_1 - \left(\frac{100}{p}\right) \times x_{3+} > 0 \tag{2.8}
\end{aligned}$$

만약, 두 번째로 큰 값(x_2)를 알고 있다면 식(2.8)은 다음의 식과 같이 유도될 수 있다.

$$x_{3+} > L = \left(\frac{q}{100}\right) \times x_{3+} \tag{2.9}$$

(2.9)의 식을 (2.8)의 식에 대입하면, 식(2.10)을 도출할 수 있다.

$$x_1 - \left(\frac{q}{p}\right) \times x_{3+} > 0 \tag{2.10}$$

즉, p/q-ambiguity rule은 p% rule을 좀 더 강화한 것으로서 비밀보호된 자료 중 원시자료의 일부 개별정보를 알 수 있는 경우 p/q-ambiguity rule을 이용하여 노출위험이 있는 셀을 찾는다.

이 방법에서 선형민감도는 식(2.11)과 같이 표현할 수 있으며 q=100인 경우는 p% rule의 선형민감도와 동일하게 된다.

$$S_{pq}(X) = x_1 - \left(\frac{q}{p}\right) \times x_{3+} > 0 \tag{2.11}$$

<표2-6> 20/50-ambiguity rule(예제)

x_1	x_2	x_3	x_4	x_5	$V(x)$
100	50	40	30	20	240

<표2 - 6>의 예제를 통해 살펴보면, $V(x) = 240$, $p = 20$, $q = 50$ 조건에서 식(2.10)을 이용해 계산하면 아래와 같이 (2.10)식과 같은 부등식을 만족하지 않아 노출위험이 발생하지 않음을 알 수 있다.

$$100 - \left(\frac{50}{20}\right) \times 90 = -125 < 0$$

이와 같이 민감한 셀을 찾는 방법에 대해 자료형태별로 알아보았다. 민감한 셀이란 노출위험이 있는 셀로서, 집계된 자료를 통해 개인 또는 사업체의 개별정보를 재계산할 수 있는 경우이다. 이렇게 선정된 민감한 셀을 비밀보호하는 방법 역시 빈도데이터(Count data)인지 크기데이터(Magnitude data)인지에 따라 전통적으로 적용하는 방법이 있는데, 다음 장에서는 데이터 형태에 따라 적용될 수 있는 방법과 각각의 방법에 대한 자세한 내용을 소개코자 한다.

3. 비밀보호 방법

비밀보호 방법에는 자료유형에 따라 적용 가능한 기법은 서로 다르다. 빈도데이터(Count data)의 경우, 작은 값을 민감한 셀로 정의하여 비밀보호가 이루어지기 때문에 ①반올림(Rounding) ②셀감추기(Cell suppression) ③자료변조(data perturbation) 또는 ④분류구조변경(data collapsing) 등과 같은 방법이 주로 적용된다.

반면 크기데이터(Magnitude data)의 경우, 셀 값을 결정하는데 영향력이 큰 값을 갖고 있는 개별정보의 제공 여부에 따라 민감한 셀이 결정되기 때문에 주로 셀감추기(Cell suppression)방법이 적용된다.

〈표2-7〉 자료형태에 따른 전통적인 비밀보호 기법(Cox,2008)

빈도데이터	크기데이터
반올림(Rounding) 셀 감추기(Cell suppression) 자료변조(data perturbation) 분류구조변경(data collapsing) 기타(swapping, switching 등)	셀 감추기(Cell suppression)

본 장에서는 비밀보호기법 중 주로 적용되고 있는 반올림, 셀 감추기, CTA방법을 중심으로 적용방법을 상세히 소개하고자 한다.

가. 반올림(Rounding)방법

1) 전통적 반올림방법(Conventional Rounding)

〈표2-8〉 전통적 반올림(예제)

2	6	3	7	6	7	31	0	5	5	5	5	5	25	
6	7	6	5	7	1	32	5	5	5	5	5	0	25	
3	6	5	7	6	7	34	5	5	5	5	5	5	30	
6	7	6	6	7	6	38	5	5	5	5	5	5	30	
1	6	8	2	6	5	28	0	5	5	0	5	5	20	
18	32	28	27	32	26	163	⇒	15	25	25	20	25	20	130

기준(Base) = 5

- ① if $[i, j] \geq 5$ then $[i, j] = 5$
- ② if $[i, j] \leq 2$ then $[i, j] = 0$
- ③ if $[i, j] \geq 3$ and $[i, j] < 5$ then $[i, j] = 5$

전통적인 반올림방법은 흔히 알고 있는 일반적인 반올림방법을 의미한다. 즉, 기준(Base: B)을 10으로 할 경우 0이상 10미만은 0으로 변환하고 10이상은 10으로 변환하는 방법이다. B를 5로 할 경우는 위 <표 2-8>의 예제와 같이 변환될 수 있다.

일반적인 수식으로 표현하면 아래와 같이 나타낼 수 있다.

$$r = \begin{cases} B & , r \geq \frac{B+1}{2} (B: odd) \text{ or } r \geq \frac{B}{2} (B: even) \\ 0 & , otherwise \end{cases}$$

2) 제어된 반올림방법(Controlled Rounding)

제어된 반올림방법은 모든 셀이 기준(Base: B)의 배수가 되도록 더하기(+), 빼기(-)를 반복하는 방법이다. 즉, B가 5인 경우 모든 셀 $[i, j]$ 가 5의 배수가 되도록 더하고 빼기를 반복한다. 이 방법은 특정한 고리를 가지고 같은 수를 더하고 뺌으로써 주변합을 바꾸지 않으면서 자료를 변환하는 방법이다.

〈표 2-9〉 (4×5) Original Data

	Col1	Col2	Col3	Col4	Col5	Row Sums
Row1	37	3	30	6	4	80
Row2	1	16	23	5	15	60
Row3	30	15	8	27	10	90
Row4	7	1	4	7	21	40
Col Sums	75	35	65	45	50	270

위 <표2 - 9>의 예제를 가지고 기준이 5인 제어된반올림방법을 적용하는 과정을 자세히 살펴보자. 먼저, 표에서 5의 배수가 아닌 셀 중 하나를 고른다. 본 예제에서는 첫 번째 열, 두 번째 행의 셀, 즉 셀(1,2)를 선택하기로 한다. 그 다음으로, 첫 번째 선택한 셀에서 시작하여 행과 열이 이어지도록 6개의 셀을 골라 교체고리(alternating cycle)를 만든다. 이 때 선택되는 셀은 마찬가지로 5의 배수가 아닌 셀 중에 고른다. 본 예제에서는 셀(1,2) 셀(1,4) 셀(3,4) 셀(3,3) 셀(4,3) 셀(4,2)로 정하였다. 이는 아래 [제어된반올림방법의 변환과정]의 표 R1과 같다.

이제 정해진 교체고리에 어떤 수 d 를 더하고 빼는데, 두 가지 경우를 생각할 수 있다. 첫 번째 셀을 더하기로 시작하여 더하기 빼기를 반복하는 경우와 빼기로 시작하여 이를 반복하는 경우이다. 첫 번째 셀을 더하기로 시작하는 경우 셀(1,2)에는 d 를 더하고, 셀(1,4)에는 d 를 빼고, 셀(3,4)에는 d 를 더하고, 셀(3,3)에는 d 를 빼고, 셀(4,3)에 d 를 더하고, 셀(4,2)에 d 를 뺀다. 이를 양의 교체고리라고 하자. 반대로 음의 교체고리 즉, 빼기로 시작하는 경우는 셀(1,2) 셀(3,4) 셀(4,3)이 빼는 셀이 되고 셀(1,4) 셀(3,3) 셀(4,2)가 더하는 셀이 된다. 아래의 표에서는 더하는 셀은 격자무늬, 빼는 셀은 빗금무늬로 구분하였다.

교체고리를 만든 후 결정해야 할 것은 더하거나 빼질 수 d 이다. 양

의 교체고리에서 더하거나 빼질 수 d_+ 는 더하는 셀인 셀(1,2) 셀(3,4) 셀(4,3)과 빼는 셀인 셀(1,4) 셀(3,3) 셀 (4,2)을 5의 배수로 만들기까지 최대로 더하거나 뺄 수 있는 수이다. 음의 교체고리에서의 d_- 는 더하는 셀인 셀(1,4) 셀(3,3) 셀 (4,2)과 빼는 셀인 셀(1,2) 셀(3,4) 셀(4,3)을 5의 배수로 만들기까지 최대로 더하거나 뺄 수 있는 수이다. 즉, 예제에서 d_+ 는 1이고 d_- 는 2가 된다.

d 를 정한 후 마지막으로 결정해야할 사항은 양의 교체고리를 이용할 것인지 음의 교체고리를 이용할 것인지 하는 것이다. 즉, 더하기로 시작할 것인지 빼기로 시작할 것인지를 정하는 것이다. 이는 p_+ 와 p_- 를 계산하여 결정한다. 양의 교체고리의 확률을 p_+ 로, 음의 교체고리의 확률을 p_- 로 난수를 발생하여 랜덤하게 교체고리를 선택하는데 p_+ 와 p_- 는 아래와 같이 계산된다. 본 예제는 음의 교체고리를 사용하는 방법을 선택하였다.

$$P_+ = \frac{d_-}{d_+ + d_-} \quad , \quad P_- = \frac{d_+}{d_+ + d_-}$$

[제어된반올림방법의 변환 과정]

<i>Original</i>	Col1	Col2	Col3	Col4	Col5	Row Sums
Row1	37	3	30	6	4	80
Row2	1	16	23	5	15	60
Row3	30	15	8	27	10	90
Row4	7	1	4	7	21	40
Col Sums	75	35	65	45	50	270

↳ **R1**

	Col1	Col2	Col3	Col4	Col5	Row Sums
Row1	37	1	30	8	4	80
Row2	1	16	23	5	15	60
Row3	30	15	10	25	10	90
Row4	7	3	2	7	21	40
Col Sums	75	35	65	45	50	270

↳ **R2**

	Col1	Col2	Col3	Col4	Col5	Row Sums
Row1	37	2	30	8	3	80
Row2	1	15	24	5	15	60
Row3	30	15	10	25	10	90
Row4	7	3	1	7	22	40
Col Sums	75	35	65	45	50	270

↳ **R3**

	Col1	Col2	Col3	Col4	Col5	Row Sums
Row1	38	2	30	8	2	80
Row2	0	15	25	5	15	60
Row3	30	15	10	25	10	90
Row4	7	3	0	7	23	40
Col Sums	75	35	65	45	50	270

<i>R4</i>	Col1	Col2	Col3	Col4	Col5	Row Sums
Row1	35	5	30	5	5	80
Row2	0	15	25	5	15	60
Row3	30	15	10	25	10	90
Row4	10	0	0	10	20	40
Col Sums	75	35	65	45	50	270

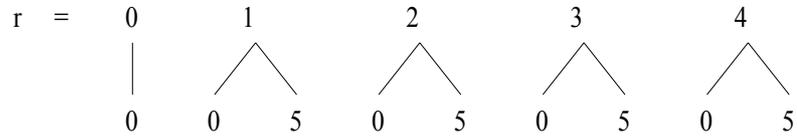
<i>Final</i>	Col1	Col2	Col3	Col4	Col5	Row Sums
Row1	35	5	30	5	5	80
Row2	0	15	25	5	15	60
Row3	30	15	10	25	10	90
Row4	10	0	0	10	20	40
Col Sums	75	35	65	45	50	270

모든 셀의 값이 B의 배수가 되도록, R1에서 R4와 같이 교체고리를 만들어 더하고 빼는 과정을 반복하면 위와 같이 최종적으로 변환된다. 이 방법은 반올림 방법 중 주변합을 바꾸지 않으면서 변환된다는 점에서 큰 의미를 가진다.

3) Zero restriction 50/50 Rounding

이 방법은 기존의 반올림방법과 유사한 방법으로서 확률적인 개념을 적용한 최근의 기법이다. 기존 방법이 반올림된 원래의 값을 추측할 수 있는 반면, 이 방법은 확률적 접근으로 의하여 랜덤하게 반올림한다. 이로 인해 원래의 값을 추측하기가 어렵다. 이 때 반올림할 확률은 50/50을 기준으로 하며, 자세한 방법은 아래와 같다.

예제) 기준(Base :B) = 5



$$\cdot P(R(r)=B) = \frac{B-1}{2B}$$

$$\cdot P(R(r)=0) = \frac{B+1}{2B}$$

예제의 경우 각각의 데이터가 0이 되거나 B가 될 경우를 고려해 볼 때, 0이 될 확률과 B가 될 확률은 위와 같이 계산 되며 평균과 분산은 다음과 같이 계산할 수 있다.

$$\begin{aligned}
 \cdot E(r) &= \sum r \cdot P(r) \\
 &= \frac{1}{B} (0 + \dots + B-1) \\
 &= \frac{1}{B} \cdot \frac{B(B-1)}{2} = \frac{B(B-1)}{2B} = \frac{B-1}{2}
 \end{aligned}$$

$$\therefore \text{assume } r \sim U(0, B-1), \sum_{i=1}^n i = \frac{n(n+1)}{2}$$

$$\begin{aligned}
\cdot V(r) &= E(r^2) - [E(r)]^2 \\
&= \frac{1}{B}(0^2 + 1^2 + \dots + (B-1)^2) - \left(\frac{B-1}{2}\right)^2 \\
&= \frac{1}{B} \cdot \frac{(B-1)B(2B-1)}{6} - \left(\frac{B^2-2B+1}{4}\right) \\
&= \frac{1}{12}(B^2-1)
\end{aligned}$$

$$\therefore \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

나. 셀감추기 방법(Cell Suppression)

1) 개 요

셀감추기(CS)방법은 집계표 자료 중 노출위험이 있는 셀을 보호하기 위한 비밀보호 기법으로 일반적으로 이용되고 있다. 이 방법은 사업체 수와 같은 빈도데이터(count data)와 종사자수, 매출액 등과 같은 크기 데이터(magnitude data) 모두에 적용 가능하기 때문에 그 활용도가 높다. CCS방법을 적용한 미국의 사례를 보면, 1970년대 이후 미센서스국(Census Bureau)을 중심으로 집계표 자료에 대한 비밀보호 기법으로 이 방법이 적용되었는데 특히, 산업구조통계에서는 Economic Census (1994)와 ASM (Annual Survey of Manufacture) 등에 적용되고 있다.

2) 셀감추기(Cell Suppression :CS) 방법론

먼저 일반적인 셀감추기(Cell Suppression :CS) 방법에 대해 살펴보자. 이 방법은 응답자의 비밀과 관련하여 노출위험이 있다고 생각되는

즉, 민감하다고 판단되는 셀을 집계표 공표 시 감추어 보여주지 않는 방법을 말한다. 단지 민감한 셀만을 감춘다고 하여 SODC(Suppressing Only the Disclosure Cells)라고 부르기도 한다.

노출위험이 발생한다고 생각되는 경계점은 각 나라마다 다른데, 사업체수 등의 빈도데이터(count data)를 예로 들면 미국의 센서스국은 5, 미국의 국세청(IRS)는 3, 뉴질랜드는 3, 한국은 2를 그 기준으로 하고 있다. 다시 말해, 한국 통계청의 경우에는 사업체수가 2개 이하에 해당되는 셀은 감추어 보여주지 않고 집계표를 제공하고 있음을 말한다. 이 방법의 단점은 단순히 민감한 셀만을 감춤으로 인해, 집계표 내의 주변합이나 관련된 다른 집계표를 이용하여 감추어진 셀이 계산 가능한 경우가 많이 발생한다는 것이다. 이는 다음의 예제를 통해 쉽게 알 수 있다.

〈표 2-10〉 (5×6) Original Data

	Col1	Col2	Col3	Col4	Col5	Col6	
Row1	2	6	3	7	6	7	31
Row2	6	7	6	5	7	1	32
Row3	3	6	5	7	6	7	34
Row4	6	7	6	6	7	6	38
Row5	1	6	8	2	6	5	28
	18	32	28	27	32	26	163

<표 2 - 10>과 같은 예제에서 노출위험의 경계점을 3으로 하여 민감한 셀에 대해서만 셀감추기 한 경우, 그 주변합만을 이용한 아래의 선형관계식을 통해 그 값을 쉽게 알 수 있다.

- Col 6 : $7 + D_{26} + 7 + 6 + 5 = 26 \Rightarrow D_{26} = 1$
- Col 4 : $7 + 5 + 7 + 6 + D_{54} = 27 \Rightarrow D_{54} = 2$
- Col 3 : $D_{13} + 6 + 5 + 6 + 8 = 28 \Rightarrow D_{13} = 3$
- Row 1 : $D_{11} + 6 + 3 + 7 + 6 + 7 = 31 \Rightarrow D_{11} = 2$
- Row 3 : $D_{31} + 6 + 5 + 7 + 6 + 7 = 34 \Rightarrow D_{31} = 3$
- Row 5 : $D_{51} + 6 + 8 + 2 + 6 + 5 = 28 \Rightarrow D_{51} = 1$

셀감추기방법의 이와 같은 점을 보완한 비밀보호방법이 보조감추기 즉, CCS(Complementary Cell Suppression)방법이다.

3) 보조 셀 감추기(CCS) 방법론

CCS 방법은 먼저 일반적인 셀감추기방법인 CS방법과 동일하게 집계표의 각 셀 중 응답자 비밀과 관련하여 직접적인 노출위험이 있는 PDC(Primary Disclosure Cell)을 정의한다. 그리고 정의된 셀의 값이 관련 집계표의 다른 정보들을 이용하여 계산되지 않도록 하는 CSC(Complementary Suppression Cell)을 찾는다. 즉, PDC가 다른 셀과의 선형관련성으로 재계산되지 않도록 하는 CSC를 찾는다. 이 과정에서 정의되는 PDC와 CSC의 값 모두를 공표하지 않는 방법이 CCS방법이다.

CCS 방법에서의 주요 관건은 PDC(Primary Disclosure Cell)가 노출되지 않도록 하는 CSC(Complementary Suppression Cell)를 찾는 과정이다. 간단히 생각하면 PDC가 존재하는 행과 열에 두개 이상의 셀이 감추어지도록 하면 될 것이다. 이러한 방법을 직관적인 보조셀감추기

(Heuristic Complementary Cell Suppression :HCCS)이라고 한다. <표2-11>의 예제에서 이 방법을 이용하여, 감추어진 PDC가 존재하는 행과 열에 두개 이상의 셀이 감추어지도록 하면 다음과 같다.

<표 2-11> HCCS방법 적용 결과

	Col1	Col2	Col3	Col4	Col5	Col6	
Row1	2	6	3	7	6	7	31
Row2	6	7	6	5	7	4	32
Row3	3	6	5	7	6	7	34
Row4	6	7	6	6	7	6	38
Row5	1	6	8	2	6	5	28
	18	32	28	27	32	26	163

<표 2-11>과 같이 감추기를 한 경우, 감추어진 셀을 계산해 낼 수 없는 것처럼 보이지만 아래의 식을 이용하면 그 값을 찾을 수 있게 된다.

• Equation (1)

$$\begin{aligned} \text{Row2} + \text{Row5} &= (6+7+6+D_{24}+7+D_{26}) \\ &\quad + (D_{51}+6+8+D_{54}+6+D_{56}) \end{aligned}$$

• Equation (2)

$$\begin{aligned} \text{Col4} + \text{Col6} &= (7+D_{24}+7+6+D_{26}) \\ &\quad + (7+D_{26}+7+6+D_{56}) \end{aligned}$$

• Equation (1) - Equation (2) = $D_{51} = 1$

따라서, CSC(Complementary Suppression Cell)를 찾는 과정에 있어서 위와 같이 행과 열에 두개의 셀이 감추어지도록 하는 것이 아니라, 셀들의 선형관련성으로 계산되지 않도록 하는 수학적 접근방법이 필요함을 알 수 있다. 아울러 감추어지는 셀의 수와 정보량을 최소화하는 CSC를 찾아야 함은 물론이다.

이러한 논리적인 접근을 위해 결합기법(combinational technique), 선형기법(linear programming), 수학적 네트워크기법(mathematical network)과 같은 3가지 접근법이 주로 적용된다. 3가지 방법 모두 수리적인 측면에서는 상이하나, 자료제공이 제한되는 셀의 수와 정보량을 최소화하는 목적함수(Objective function)를 정의하고, 이를 이용하여 그 해(解)를 찾아나간다는 점에서는 그 맥락이 같다.

CCS방법의 절차는 「감추어진 셀의 수」와 「감추어진 정보량」중 어떤 것을 최소화 할 것인가에 따라 접근방법이 다소 차이가 있으나, 기본적인 절차는 다음과 같다.

- (1) 기준(Base :B)을 정한다. 즉 PDC(Primary Disclosure Cell)을 구분할 수 있는 경계점을 지정하는 것으로서, 사업체수의 경우 n개 이하 셀 정보를 제한한다면 $B=n$ 이 된다.
- (2) CSC(Complementary Suppression Cell)을 선형모형식 또는 수학적 네트워크 등을 이용하여 찾는다.
- (3) PDC와 CSC 모두를 감추기 한 후, 자료를 제공한다.

가) 일반적인 CCS방법

다음의 예제는 「감추어진 정보량」을 최소화 할 수 있도록 CSC (Complementary Suppression Cell)를 찾는 방법이다. <표2 - 12>의 (3×3)정방행렬의 데이터를 가지고 CCS방법을 적용하는 경우를 생각해 보면 아래와 같은 과정으로 설명할 수 있다.

<표2-12> (3×3)Original data

	Col1	Col2	Col3	합 계
Row1	10	5	10	25
Row2	15	10	1	26
Row3	20	15	1	36
합 계	45	30	12	87

(Step1) 노출위험의 경계점 Base를 b=2로 가정하면, PDC(Primary Disclosure Cell)는 $X_{23} = 1, X_{33} = 1$ 이 된다.

(Step2) $X_{23} = 1, X_{33} = 1$ 정보만 제한한다면, 집계표내 다른 셀과의 선형 관련성으로 인해 비밀보호가 이루어지지 않는다.

$X_{21} + X_{22} + X_{23} = 26 \Leftrightarrow 15 + 10 + X_{23} = 26$ $\therefore X_{23} = 1$

그렇기 때문에 가능한 CSC(Complementary Suppression Cell)의 쌍(pair)을 선형모형식으로 찾으면 $(X_{21}, X_{31}), (X_{22}, X_{32})$ 인데, 공개가 제한되는 정보량이 최소가 되도록 하기 위해서는 (X_{22}, X_{32}) 가 최적의 CSC로 정해진다.

(Step3) PDC : (X_{23}, X_{33})

CSC : (X_{22}, X_{32})

〈표2-13〉 (3×3) CCS방법 적용 결과

	Col1	Col2	Col3	합 계
Row1	10	5	10	25
Row2	15	10	1	26
Row3	20	15	1	36
합 계	45	30	12	87

노출위험이 있는 셀에 대한 CSC의 도출 결과는 <표2 - 13>와 같으며, 동 결과의 효율성(efficiency)은 <표2 - 14>과 같이 측정할 수 있다.

〈표2-14〉 CCS방법에 따른 정보제한총량

(단위 : 개, %)

	PSC ²⁾		CSC ³⁾		합 계	
	개	구성비	개	구성비	개	구성비
- 감추어진 셀의 수	2	22.2	2	22.2	4	44.4
- 감추어진 정보량	2	2.3	25	28.7	27	31.0

즉, PDC(Primary Disclosure Cell) (X_{23}, X_{33})에 대한 노출위험을 보호하기 위해 보조적으로 2개의 CSC(Complementary Suppression Cell) (X_{22}, X_{32})을 추가함으로써, 3×3 table의 노출위험이 있는 셀 값을 보호하였다. 이 과정에서 「감추어진 셀의 수」는 2개(PDC) → 4개(PDC 2개 + CSC 2개)로 증가하였고, 노출위험이 있는 2.3% (PDC)의 정보를 보호하기 위해 28.7% (CSC)의 정보를 추가적으로 감추기 함으로서 「감추어

진 정보량」은 31.0% (PSC+CSC)로 증가하였다.

<표2 - 13> CCS결과에 대한 정도 검증을 위하여, 선형관련성을 이용해 확인해 보면 다음과 같다.

$$\text{Equations) } X_{22} + X_{23} = 11 \text{ (eq1)} \quad X_{32} + X_{33} = 16 \text{ (eq2)}$$

$$X_{22} + X_{32} = 25 \text{ (eq3)} \quad X_{23} + X_{33} = 2 \text{ (eq4)}$$

$$0 \leq X_{33} \leq 2 \text{ (eq5)} \quad 0 \leq X_{23} \leq 2 \text{ (eq6)}$$

☞ X_{ij} : original value i : # of rows j : # of columns

$$\text{solution) } 9 \leq X_{22} \leq 11 \text{ (eq7)} \quad 14 \leq X_{32} \leq 16 \text{ (eq8)}$$

eq7, eq8, eq3으로 가능한 $(X_{22}, X_{32}, X_{23}, X_{33})$ 의 조합은 아래와 같다.

- Feasible solution 1 : (9, 16, 2, 0)
- Feasible solution 2 : (10, 15, 1, 1)
- Feasible solution 3 : (11, 14, 0, 2)

즉, CCS방법을 이용하여 위와 같이 4개의 셀(PDC: 2개, CSC: 2개)을 감추기 한 경우, 선형관련식을 이용하여 감추어진 셀에 대한 unique한 solution을 찾을 수 없기 때문에, <표2 - 13>은 성공적으로 비밀보호가 수행된 것으로 평가할 수 있다.

나) 네트워크(network)이론을 통한 CCS방법

CCS방법을 네트워크(network) 이론으로 접근하면 다음과 같이 일반화 할 수 있다. 목적함수(1)는 「감추어진 정보량」을 최소화하는 것이

며, 목적함수(2)는 「감추어진 셀의 수」를 최소화하는 방법이다.

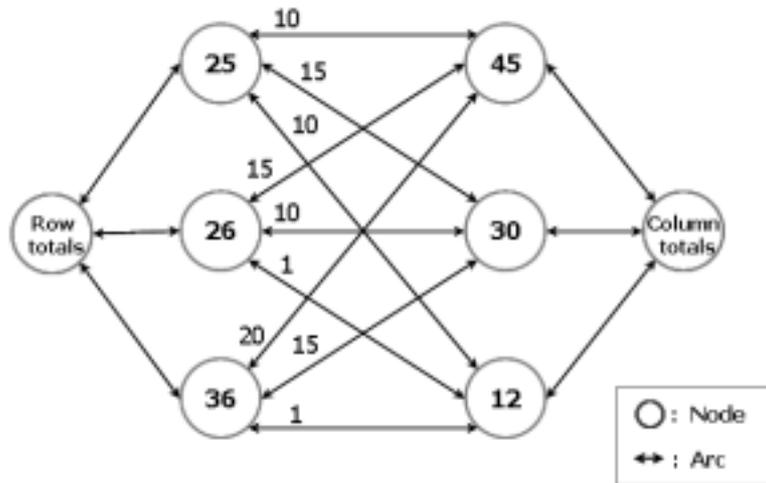
목적함수(1): $V(a) = \min \sum a_{ij} z_{ij}$

목적함수(2): $V(a) = \min \sum z_{ij}$

☞ a_{ij} : 셀값, $z_{ij} = 1$ (감춘셀), $z_{ij} = 0$ (기타 셀)

<표2 - 12>에서 「감추어진 정보량」을 최소화하도록 네트워크테이블(network table)을 작성하면 <그림2 - 1>과 같다.

<그림2-1> 표2-12에 대한 network table



위 <그림2 - 1>은 <표2 - 12>에 대한 네트워크 테이블(network table)을 나타낸 것이다. 네트워크 테이블을 작성하는 방법을 살펴보면, <그림2 - 1>에서 보는바와 같이 노드(Node)의 두 번째 열에 데이터의 행(Row)의 합계를 표시하고 노드의 세 번째 열에는 데이터의 열(Column)의 합계를 표시한다. 또한, 행에서 열으로 보내어지는 값을 아크(Arc ;↔)위에 나타내어 도식화 한다. 이때 행과 열간에는 서로 왕

복할 수 있는 경로를 갖기 때문에 양쪽이 열린 화살표(\leftrightarrow)로 표시한다. 참고로 3차원 이상의 집계표를 network table로 도식화하는 경우, 평면으로 도식화 가능하도록 2차원 집계표로 분리·축소하여 각각 나타내어야 한다.

우선 <표2 - 12>데이터의 기준(Base :B)은 B=2로 가정했기 때문에 <그림2 - 1>에서 PDC(Primary Disclosure Cell)은 a_{23}, a_{33} 임을 알 수 있다. PDC(a_{23}, a_{33}) 중 a_{33} 에서 출발하여 경로를 찾는 방법을 살펴보자. 1개의 PDC 자료를 보호하기 위해서는 최소 3개의 CSC(Complementary Suppression Cell)가 필요한데, 이는 집계표내 다른 셀과의 선형관련성으로 인해 PDC가 재계산되어 비밀보호가 이루어지지 않을 수 있는 경우를 막기 위해서이다. 이와 같이 4step 내에 a_{33} 로 도착할 수 있는 모든 경로를 고려하여 이에 대한 $V(a) = \sum a_{ij}Z_{ij}$ 값의 합을 계산한 후, 목적함수(1)를 만족하는 최소의 경로를 찾는다.

<표2-15> Network에 의한 최소경로 탐색(a_{33})

가능한 경로	감추어질 셀 값	비 고
$(3,3) \rightarrow (2,3) \rightarrow (2,2) \rightarrow (3,2) \rightarrow (3,3)$	27	최소경로
$(3,3) \rightarrow (1,3) \rightarrow (1,2) \rightarrow (3,2) \rightarrow (3,3)$	31	
$(3,3) \rightarrow (1,3) \rightarrow (1,1) \rightarrow (3,1) \rightarrow (3,3)$	41	
$(3,3) \rightarrow (2,3) \rightarrow (2,1) \rightarrow (3,1) \rightarrow (3,3)$	37	

<표2 - 15>의 결과를 보면, PDC a_{33} 에 대한 CSC는 a_{23}, a_{22}, a_{32} 로 정해진다. 이 경로는 「감추어진 정보량」와 「감추어진 셀의 수」를 모두 최소화하는 결과이다.

동일한 방법으로 a_{23} 에 대한 최소경로를 찾으면 <표2 - 16>와 같다.

<표2-16> Network에 의한 최소경로 탐색(a_{23})

가능한 경로	감추어질 셀 값	비 고
(2,3) → (1,3) → (1,2) → (2,2) → (2,3)	26	최소경로
(2,3) → (3,3) → (3,2) → (2,2) → (2,3)	27	
(2,3) → (1,3) → (1,1) → (2,1) → (2,3)	41	
(2,3) → (3,3) → (3,1) → (2,1) → (2,3)	37	

<표2 - 16>의 결과를 보면, PDC a_{23} 에 대한 CSC는 a_{13}, a_{12}, a_{22} 로 정해진다. 그러나 네트워크(network)에 의한 최소경로 탐색결과는 과대감추기(over suppression)를 발생할 수 있다. 이 때문에 네트워크 탐색 후, CSC의 적정성을 선형관련식으로 판단하여 과대감추기된 셀을 제거할 필요가 있다. 이를 통해 최종적으로 「감추어진 정보량」을 최소화하는 CSC를 찾게 된다. 위 예제에 대한 선형관련식은 다음과 같다.

$$\text{Equations) } a_{12} + a_{13} = 15 \text{ (eq1) } \quad a_{22} + a_{23} = 11 \text{ (eq2)}$$

$$a_{32} + a_{33} = 16 \text{ (eq3)}$$

$$0 \leq a_{33} \leq 2 \text{ (eq4) } \quad 0 \leq a_{23} \leq 2 \text{ (eq5)}$$

☞ X_{ij} : original value i : # of rows j : # of columns

위 선형식에서 PDC (a_{23}, a_{33})에 대한 CSC (a_{22}, a_{32}) 만으로도 감추어진 셀의 유일한 해를 찾을 수 없기 때문에, (a_{12}, a_{13})을 CSC로 추가하는 것은 과대감추기(over suppression)가 됨을 알 수 있다. 이와 같은 경우 과대감추기된 셀은 감추기에서 제외한다. 따라서 최종 CCS 결과는

PDC(Primary Disclosure cell) : (a_{23}, a_{33}) , CSC(Complementary Suppression Cell) : (a_{22}, a_{32}) 이다.

셀감추기(CS) 방법은 결합기법(combinational technique), 선형기법(linear programming), 수학적 네트워크기법(mathematical network)과 같은 3가지 접근법이 주로 적용된다. 빈도데이터와 크기데이터 모두에 적용가능한 장점이 있으며, 자료제공이 제한되는 셀의 수와 정보량을 최소화하는 목적함수(Objective function)를 만족시키는 해(解)를 찾아나간다는 점에서는 그 맥락이 같다. 또한, 직관적인 셀감추기(Heuristic Complementary Cell Suppression :HCCS)와 SODC(Suppression Only the Disclosure Cell)방법은 비밀보호된 셀 값이 재계산될 가능성이 높기 때문에 보조셀 감추기(Complementary Cell Suppression)방법을 적용하는 것이 효과적임을 살펴보았다.

다. CTA방법(Controlled Tabular Adjustment)

1) 개 요

CTA방법은 집계표 형태의 데이터에 대한 비밀보호 기법중의 하나로써, 주로 크기데이터(Magnitude data)에 적용된다. 이 방법은 노출위험이 있는 자료를 감추거나 제거하지 않고 셀 값을 조정하는 방법이다. 보조셀감추기(CCS)방법이 노출위험이 있는 셀 값을 제한(감추기, suppression)한다는 것과 다르게 CTA방법은 모든 셀 값을 제공한다는 데에 큰 장점이 있다. 이에 따라 미국에서는 보조셀감추기(CCS)방법의 대안으로 대두되고 있다. 그러나 이 방법은 주변합(marginal total)이 원자료의 집계결과와 달라지기 때문에 우리나라의 경우 이용 상 혼선을 초래할 수 있어 적용 시 사전검토가 필요하다.

2) CTA 방법론

CTA방법이 적용되는 과정을 자세히 살펴보자. 우선 앞에서 소개한 「민감한 셀 찾기」 방법 중 해당 자료형태에 맞는 방법을 이용하여 노출위험이 있는 셀을 정의한다. 그리고 정의된 민감한 셀이 노출위험이 없는 셀이 되기 위한 값(safe cell value)이 무엇인지 찾는다. 이때 노출위험이 없는 셀이 되기 위한 값은 원자료에 임의의 값을 더하거나 빼면서 탐색하며, 선형모형 프로그램을 이용하여 리밸런싱(rebalancing) 과정을 수행한 후 셀 값을 조정한다. 이러한 과정을 모형으로 표현하면 아래와 같다.

$$\min \sum_{i=1}^n (y_i^+ + y_i^-)$$

$$\text{s.t } T(y) = 0 ,$$

$$y_i^- \geq p_i(1 - I_i) ,$$

$$y_i^+ \geq p_i I_i ,$$

$$0 \leq y_i^- ,$$

$$y_i^+ \leq e_i ,$$

$$I_i : \text{binary} \Rightarrow 0 \rightarrow (-), 1 \rightarrow (+)$$

$$\text{where } y = y^+ - y^- ,$$

p_i : 민감한 셀을 보호하기 위한 상하한값,

a : 원자료,

$a + y^+ - y^-$: 조정된 자료

〈표2-17〉 원자료 집계 결과

167	317	1,284	587	4,490	3,981	2,442	1,150	70(21)	14,488
57(1)	1487	172	667	1,006	327	1,683	1,138	46(7)	6,583
616	202	1,899	1,098	2,172	3,825	4,372	300(40)	787	15,271
0	36(10)	0	16(4)	0	0	65	0	140(40)	257
840	2,042	3,355	2,368	7,668	8,133	8,562	2,588	1,043	36,599

노출위험이 있는 셀(7개) : $a_{21}, a_{42}, a_{44}, a_{38}, a_{51}, a_{52}, a_{54}$

괄호() 안의 숫자 : safe cell value

〈표2-18〉 보조셀감추기(CCS)방법 적용 결과

CSC(X)	317	1,284	CSC(X)	4,490	3,981	2,442	1,150	PDC(X)	14,488
PDC(X)	1,487	172	667	1,006	327	1,683	CSC(X)	PDC(X)	6,583
616	CSC(X)	1,899	1,098	2,172	3,825	4,372	PDC(X)	787	15,271
0	PDC(X)	0	PDC(X)	0	0	65	0	PDC(X)	257
840	2,042	3,355	2,368	7,668	8,133	8,562	2,588	1,043	36,599

감추어진 셀 수 : 11개(30%)

감추어진 정보량 : 2,759개(7.5%)

〈표2-19〉 CTA방법 적용 결과

167	317	1,276(-8)	587	4,490	3,981	2,442	1,150	91(+21)	14,501
56(-1)	1,487	172	667	1,006	327	1,679(-4)	1,138	39(-7)	6,571
617(+1)	196(-6)	1899	1,095(-3)	2172	3,825	4,371(-1)	260(-40)	797(+10)	15,232
0	26(-10)	0	12(-4)	0	0	70(+5)	0	180(+40)	288
840	2,026	3,347	2,361	7,668	8,133	8,562	2,548	1,107	36,592

<표2-17>의 예제를 가지고 보조셀감추기(CCS)방법과 CTA방법을 적용해 보기로 하자. 예제의 민감한 셀은 p-percent rule(p=20%)을 적용하여 선정하였으며, 괄호안의 수를 더하거나 뺀을 때 노출위험이 없는 값이 된다. <표2-18>는 보조셀감추기방법을 적용한 결과이며 <표2-19>은 CTA방법을 적용한 결과이다. CTA방법은 위에 나타낸 모형을 바탕으로 행과 열의 균형을 유지하도록 리밸런싱(rebalancing)하면서 노출위험이 있는 셀 값에 일정한 값을 가감하여 민감하지 않은 셀이 되도록 조정하였다.

보조셀감추기 방법을 적용한 경우 <표2-18>와 같이 감추어진 셀의 수가 11개이며 CTA방법을 적용한 경우 <표2-19>과 같이 바뀌어진 셀의 수가 13개이다. 보조셀감추기 방법에서 감추어진 셀의 수에 비해 CTA방법에서 바뀌어지는 셀의 수가 더 많으나, CTA방법의 경우 감추어지는 셀 없이 셀 값 모두를 제공한다는 차이점이 있다. 즉, 이 방법의 경우 제한되는 정보량은 없는 것이다. 그러나 원자료 집계결과와 주변합이 다르기 때문에 우리나라에 이 같은 기법을 적용하기 위해서는 충분한 검토가 필요할 것으로 생각된다.

제 3 절 산업구조통계의 비밀보호 현황

매크로 데이터란 개별자료의 집계표 형태를 의미하는데, 통계조사별로 발간되는 보고서에 수록되는 집계표가 이에 속한다. 이러한 집계표를 구성하고 있는 자료는 크게 ①빈도데이터(count data)와 ②크기데이터(magnitude data)로 구분할 수 있다. 빈도데이터는 사업체수와 같이 개수 자체에 의미가 있는 자료형태이며, 크기데이터는 매출액, 종사자수 등과 같이 그 규모에 의미가 있는 형태이다. 두 데이터는 그 형태에 따라 비밀보호 방법에 있어 적용상 차이가 있다. 앞서 언급한 바와

같이 일반적으로 「빈도데이터」에는 셀 감추기(Cell suppression), 반올림(Rounding), 자료변조(Data perturbation) 등을 「크기데이터」는 셀 감추기(Cell suppression)를 적용하고 있다.(Cox, 2008) 만약 두 데이터 형태가 혼합되어 제공되고 있는 경우는 셀 감추기 방법만이 적용 가능하다. 이 밖의 방법으로는 분류구조를 변경하는 방법으로, 표 크기를 축소하거나 특성을 재분류하여 노출위험이 있는 셀이 나타나지 않도록 조정하는 방법이 있다.

통계청에서 제공되고 있는 산업관련 매크로 데이터의 형태를 보면, 빈도데이터와 크기데이터가 혼합되어 하나의 집계표에 제공된다. 즉, 사업체수를 중심으로 주요 지표들을 산업분류, 종사자규모, 매출액규모 등으로 그룹화하여 주요 항목에 관한 정보를 제공하는데, 그 형태는 <표3 - 1>과 같다.

<표3-1> 집계표 예시

산업분류	종사자규모	사업체수	월평균 종사자수	급여액	출하액	...	유형자산 연말잔액
D19 가죽, 가방 및 신발 제조업	5~9명	1,181	7,212	84,543	565,117	...	110,002
	10~19명	435	5,696	90,273	619,753	...	124,855
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	300~499명	2	X	X	X	X	X
	500명이상	3	1,964	56,025	324,478	...	227,235

현재 통계청에서 작성하고 있는 산업구조통계는 10종에 달한다. 이들에 대한 집계표 제공 시 적용하고 있는 비밀보호방법은 제 2절에서 언급한 방법 중 노출위험이 있는 셀만 감추는 SODC(Suppressing Only the Disclosure Cells) 방법이다. 사업체 수를 기준변수로 셀의 노출위험 여부를 결정하며, 그 기준(base)은 2로 한다. 즉, 해당 셀의 사업체 수가 2이하인 경우는 노출위험이 있는 셀로 결정하고, 노출위험이 있는 셀로 결정되면 동일한 행의 모든 자료는 「X」 표시로 제공하고

본래의 정보는 감추어진다. 다시 말해, 노출위험이 있는 셀만 감추는 SODC(Suppressing Only the Disclosure Cells) 방법으로서는 기준을 2로 적용하고 있다.

본 절에서는 산업구조통계 중 「광업 및 제조업 통계조사」를 대상으로 노출위험이 있는 셀에 대한 비밀보호 현황과 그 문제점을 살펴보고자 한다.

1. 광업 및 제조업 통계조사 작성 개요

「광업 및 제조업 통계조사」는 한국표준산업분류상 「C. 광업」과 「D. 제조업」을 영위하는 사업체 중 종사자수 5인 이상의 사업체를 대상으로 매년 실시되고 있다. 조사결과가 공표되는 형태는 ①산업분류별 ②산업분류별×종사자규모별 ③산업분류별×매출액규모별로 「전국편과 지역편」 「산업편과 품목편」 등으로 구분해 공표된다. 주요지표 도출을 위한 조사항목은 14개 항목으로서 <표3-2>와 같다.

<표3-2> 광업 및 제조업 통계조사 항목

구분	조사항목
일반사항	1)사업체명 및 소재지 2)본사소속 및 타공장 유무 3)본사명 및 소재지 4)경영조직 5)자산과 자본
고용사항	6)종사자수 및 연간급여액
수입사항	7)연간출하액 및 수입액
비용사항	8)연간제조원가 및 판매비와 관리비 9)내국소비세
자산사항	10)유형자산 11)무형자산 및 컴퓨터소프트웨어
기타사항	12)연간제품출하액 및 재고액 내역 13)연간 임가공 수입액의 품목별 내역 14)재고액

<표3 - 2>의 항목 중 주요지표를 중심으로 「광업 및 제조업 통계조사 보고서」 및 국가통계포털(KOSIS)을 통해 제공되고 있는 주요 집계표의 종류는 아래와 같다.

- 1-1) 산업세세분류×종사자규모별 주요지표
- 1-2) 시도×산업중분류×종사자규모별 주요지표
- 2-1) 산업세세분류×출하액규모별 주요지표
- 2-2) 시도×산업중분류×출하액규모별 주요지표
- 3-1) 산업세세분류별 사업체수 및 종사상지위 및 성별 종사자수
- 3-2) 시도×산업중분류별 사업체수 및 종사상지위 및 성별 종사자수
- 4-1) 산업세세분류별 사업체수 및 출하액 및 주요생산비
- 4-2) 시도×산업중분류별 사업체수 및 출하액 및 주요생산비
- 5-1) 산업세세분류별 사업체수 및 재고액
- 5-2) 시도×산업중분류별 사업체수 및 재고액
- 6-1) 산업세세분류별 유형고정자산
- 6-2) 시도×산업중분류별 유형고정자산
- 7-1) 산업세세분류별 부지면적 및 임차료
- 7-2) 시도×산업중분류별 부지면적 및 임차료
- 7-3) 시군구×산업대분류별 부지면적 및 임차료

2. 비밀보호 방법

「광업 및 제조업 통계조사」는 <표3 - 2>와 같이 빈도데이터(count data)와 크기데이터(magnitude data)가 하나의 집계표에 표두 항목으로 제공되고 있다. 제 2절에서 소개한 미국의 ASM(Annual Survey of Manufacture)이 빈도데이터와 크기데이터를 서로 다른 집계표로서 제공되고 있는 점과는 차이가 있는데, 이런 점에서 우리나라의 비밀보호

방법은 미국과 달리 변형된 방법을 적용해야 함을 알 수 있다.

현행 「광업 및 제조업 통계조사」는 노출위험이 있다고 판단되는 셀을 사업체수 2개를 기준으로 정하고 있다. 즉, 사업체수가 2개 이하인 (base=24) 경우 자료제공을 제한하는 방법이다. 예를 들어 산업세세분류별×종사자규모별 집계표에서 해당 셀의 사업체수가 1개이거나 2개인 경우, 사업체수를 제외한 해당 분류 내 셀과 관련된 모든 정보의 자료제공을 제한하고 「X」 표시로 나타낸다. 이는 노출위험이 있는 셀만 감추는 전통적인 SODC(Suppressing Only the Disclosure Cells) 방법을 적용한 경우이다. 그러나 이 방법은 제 2절에서 언급한 바와 같이, 집계표내의 행과 열의 선형관련식으로 계산이 가능하기 때문에 쉽게 「X」 표시된 값을 계산하거나 추측할 수 있다.

이에 본 절에서는 「2006년 기준 광업 및 제조업통계조사」 결과를 통해 통계청의 산업구조 통계관련 비밀보호기법 적용 현황과 그 한계점을 살펴보고자 한다.

「2006년 기준 광업 및 제조업통계조사」 결과 중 한국표준산업분류 코드 D23에 해당하는 「코크스 석유정제품 및 핵연료제조업」의 산업 및 종사자규모별 집계표의 예를 살펴보자. <표3-3>은 이에 해당하는 데이터로써 현재 「2006년 기준 광업 및 제조업통계조사 보고서」와 국가통계포털(KOSIS)에서 제공하고 있는 형태로, 분류 내에 사업체수가 1개이거나 2개인 경우 동일한 행의 모든 정보를 「X」 표시로 감추기하고 있다. 이 과정에서 우선 눈에 띄는 점은 사업체 분포도가 상대적으로 적은 D2321 분류의 경우, 8개 범주의 종사자규모⁵⁾중 「500명 이상」의 정보만이 제공되고 있다는 것을 알 수 있다. 이는 공표를 위한

4) Base=2 : 사업체수가 1개 또는 2개인 경우에 한함(단, 0인 경우는 제외)

5) 종사자규모 : ① 5~9명 ② 10~19명 ③ 20~49명 ④ 50~99명 ⑤ 100~199명

⑥ 200~299 ⑦ 300~499 ⑧ 500명이상

세부 분류단계를 정함에 있어, 산업을 특성을 고려하지 않고 모든 산업에 그 기준을 동일하게 적용한데 기인한 것으로 판단된다. 그러므로 세부분류단계의 공표 기준을 정함에 있어, 산업의 특성을 반영하여 정보량과 실효성을 감안한 기준을 정하는 것에 관한 재검토가 필요할 것이라 생각된다.

〈표3-3〉 코크스,석유정제품 및 핵연료 제조업(D23)

산업 및 종사자규모	사업체수	월평균 종사자수	급여액	출하액	
23	계	118	9,791	674,424	69,840,546
	5 ~ 9	38	257	5,967	217,730
	10 ~ 19	35	469	11,144	211,887
	20 ~ 49	32	966	28,308	630,663
	50 ~ 99	6	406	17,052	392,182
	100 ~ 199	1	X	X	X
	500명 이상	6	7,514	603,092	67,817,139
232	계	117	9,168	642,376	69,730,433
	5 ~ 9	38	257	5,967	217,730
	10 ~ 19	35	469	11,144	211,887
	20 ~ 49	32	966	28,308	630,663
	50 ~ 99	6	406	17,052	392,182
	100 ~ 199	1	X	X	X
	500명 이상	5	6,891	571,044	67,707,026
2321	계	8	7,176	583,927	68,408,790
	5 ~ 9	0	0	0	0
	10 ~ 19	0	0	0	0
	20 ~ 49	1	X	X	X
	50 ~ 99	1	X	X	X
	100 ~ 199	1	X	X	X
	500명 이상	5	6,891	571,044	67,707,026
23210	계	8	7,176	583,927	68,408,790
	5 ~ 9	0	0	0	0
	10 ~ 19	0	0	0	0
	20 ~ 49	1	X	X	X
	50 ~ 99	1	X	X	X
	100 ~ 199	1	X	X	X
	500명 이상	5	6,891	571,044	67,707,026

〈표3-3〉 코크스, 석유정제품 및 핵연료 제조업(D23)(계속)

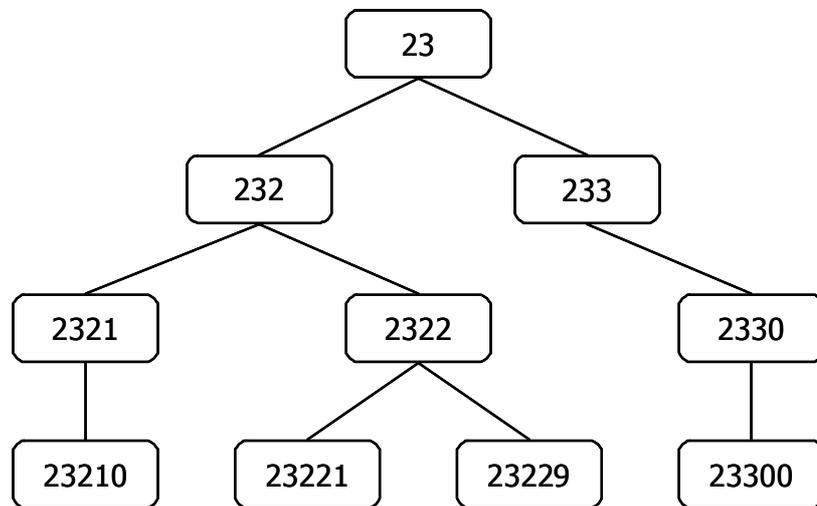
산업 및 종사자규모	사업체수	월평균 종사자수	급여액	출하액
232 계	117	9,168	642,376	69,730,433
5 ~ 9	38	257	5,967	217,730
10 ~ 19	35	469	11,144	211,887
20 ~ 49	32	966	28,308	630,663
50 ~ 99	6	406	17,052	392,182
100 ~ 199	1	X	X	X
500명 이상	5	6,891	571,044	67,707,026
2322 계	109	1,992	58,449	1,321,643
5 ~ 9	38	257	5,967	217,730
10 ~ 19	35	469	11,144	211,887
20 ~ 49	31	922	27,076	568,481
50 ~ 99	5	344	14,262	323,545
100 ~ 199	0	0	0	0
500명 이상	0	0	0	0
23221 계	66	1,210	37,211	1,099,890
5 ~ 9	22	150	3,231	170,564
10 ~ 19	23	315	7,300	177,910
20 ~ 49	18	552	17,452	469,511
50 ~ 99	3	193	9,228	281,905
100 ~ 199	0	0	0	0
500명 이상	0	0	0	0
23229 계	43	782	21,238	221,753
5 ~ 9	16	107	2,736	47,166
10 ~ 19	12	154	3,844	33,977
20 ~ 49	13	370	9,624	98,970
50 ~ 99	2	X	X	X
100 ~ 199	0	0	0	0
500명 이상	0	0	0	0

〈표3-3〉 코크스, 석유정제품 및 핵연료 제조업(D23)(계속)

산업 및 종사자규모		사업체수	월평균 종사자수	급여액	출하액
233	계	1	X	X	X
	5 ~ 9	0	0	0	0
	10 ~ 19	0	0	0	0
	20 ~ 49	0	0	0	0
	50 ~ 99	0	0	0	0
	100 ~ 199	0	0	0	0
	500명 이상	1	X	X	X
2330	계	1	X	X	X
	5 ~ 9	0	0	0	0
	10 ~ 19	0	0	0	0
	20 ~ 49	0	0	0	0
	50 ~ 99	0	0	0	0
	100 ~ 199	0	0	0	0
	500명 이상	1	X	X	X
23300	계	1	X	X	X
	5 ~ 9	0	0	0	0
	10 ~ 19	0	0	0	0
	20 ~ 49	0	0	0	0
	50 ~ 99	0	0	0	0
	100 ~ 199	0	0	0	0
	500명 이상	1	X	X	X

한편, 비밀보호 방법론적인 관점에서 살펴보면, <표2>는 비밀보호를 위해 SODC(Suppressing Only the Disclosure Cells) 방법이 적용되었는데, 이 방법의 보완점을 파악하기 위해 우선 집계표내의 선형관련식을 통해 감추기한 셀 「X」의 값을 계산할 수 있는지 알아보고자 한다. 집계표내의 선형관련식을 만들기 위해 중분류 D23의 분류체계를 보면 <그림3 - 1>과 같다.

<그림3-1> 코크스, 석유정제품 및 핵연료 제조업 분류체계6)



D23 코크스, 석유정제품 및 핵연료	D232 석유정제품 제조업	D2321 원유 정제처리업	D23210 원유 정제처리업
		D2322 석유정제품물 재처리업	D23221 윤활유 및 그리스 제조업 D23229 기타 석유정제품물 재처리업
	D233 핵연료 가공업	D2330 핵연료 가공업	D23300 핵연료 가공업

6) D231(코크스 및 관련제품 제조업)은 조사대상 사업체가 없기 때문에 제외

<그림3 - 1>의 분류체계를 이용해 선형관련식을 만들면 다음과 같다.

$$\begin{aligned}
 D23_i &= D232_i + D233_i \\
 D232_i &= D2321_i + D2322_i & D2321_i &= D23210_i \\
 & & D2322_i &= D23221_i + D23229_i \\
 D233_i &= D2330_i & D2330_i &= D23300_i \\
 i &= 1, 2, \dots, 8 (\text{종사자규모})
 \end{aligned}$$

선형관련식을 이용하여 감추기한 셀 「X」의 값을 하향식(Top down) 접근방식으로 단계별로 계산하면 다음과 같다.

<표3-4> 셀값찾기 Step 1

산업 및 종사자규모	사업체수	월평균 종사자수	급여액	출하액
D23 계	118	9,791	674,424	69,840,546
5 ~ 9명	38	257	5,967	217,730
10 ~ 19명	35	469	11,144	211,887
20 ~ 49명	32	966	28,308	630,663
50 ~ 99명	6	406	17,052	392,182
100 ~ 199명	1	179	8,861	570,945
500명 이상	6	7,514	603,092	67,817,139

노출위험이 있는 「100~199명」 분류에서 「X」로 표시되어 감추어진 월평균종사자수, 급여액 및 출하액 등의 값은 관련식 (3.1)에 의해서 <표3 - 4>와 같이 모든 값을 계산할 수 있다.

$$X_{5j} = X_{+j} - \sum_{i \neq 5} X_{ij} \quad j = 1, 2, \dots \quad (3.1)$$

〈표3-5〉 셀값찾기 Step 2

산업 및 종사자규모		사업체수	월평균 종사자수	급여액	출하액
D23	100 ~ 199명	1	179	8,861	570,945
D232	100 ~ 199명	1	179	8,861	570,945
D233	100 ~ 199명	0	0	0	0

소분류 D232의 「100~199명」 분류에 감추어진 셀의 값은 <표3 - 4>의 step 1에서 선형관련식으로 계산된 값과 관련식 (3.2)를 이용하여 <표3 - 5>와 같이 계산된다.

$$D23_i = D232_i + D233_i \quad (3.2)$$

이와 같은 계산단계를 선형관련식을 이용하여 순차적으로 진행하면, 다음과 같이 중분류 D23에 감추어진 모든 「X」의 값이 <표3 - 9>와 같이 계산된다. 즉, 감추어진 모든 셀의 값을 찾아낼 수 있다.

〈표3-6〉 셀값찾기 Step3

산업 및 종사자규모		사업체수	월평균 종사자수	급여액	출하액
D232	20 ~ 49명	32	966	28,308	630,663
D2321	20 ~ 49명	1	44	1,232	62,182
D2322	20 ~ 49명	31	922	27,076	568,481

선형관련식 : $D232_i = D2321_i + D2322_i$

〈표3-7〉 셀값찾기 Step4

산업 및 종사자규모		사업체수	월평균 종사자수	급여액	출하액
D232	50 ~ 99명	6	406	17,052	392,182
D2321	50 ~ 99명	1	62	2,790	68,637
D2322	50 ~ 99명	5	344	14,262	323,545
D232	100 ~ 199명	1	179	8,861	570,945
D2321	100 ~ 199명	1	179	8,861	570,945
D2322	100 ~ 199명	0	0	0	0

선형관련식 : $D232_i = D2321_i + D2322_i$

〈표3-8〉 셀값찾기 Step5

산업 및 종사자규모		사업체수	월평균 종사자수	급여액	출하액
합 계		118	9,791	674,424	69,840,546
D232	소 계	117	9,168	642,376	69,730,433
D233	소 계	1	623	32,048	110,113
	5 ~ 9명	0	0	0	0
	10 ~ 19명	0	0	0	0
	20 ~ 49명	0	0	0	0
	50 ~ 99명	0	0	0	0
	100 ~ 199명	0	0	0	0
	500명 이상	1	623	32,048	110,113

선형관련식 : $D23_i = D232_i + D233_i$ $D233 = \sum_{i=1} D233_i$

〈표3-9〉 선형관련식에 의한 감추어진 셀 「X」 값 찾기 최종 결과

산업 및 종사자규모	사업체수	월평균 종사자수	급여액	출하액
D23 계	118	9,791	674,424	69,840,546
5 ~ 9	38	257	5,967	217,730
10 ~ 19	35	469	11,144	211,887
20 ~ 49	32	966	28,308	630,663
50 ~ 99	6	406	17,052	392,182
100 ~ 199	1	179	8,861	570,945
500명 이상	6	7,514	603,092	67,817,139
D232 계	117	9,168	642,376	69,730,433
5 ~ 9	38	257	5,967	217,730
10 ~ 19	35	469	11,144	211,887
20 ~ 49	32	966	28,308	630,663
50 ~ 99	6	406	17,052	392,182
100 ~ 199	1	179	8,861	570,945
500명 이상	5	6,891	571,044	67,707,026

〈표3-9〉 선형관련식에 의한 감추어진 셀 「X」 값 찾기 최종 결과(계속)

산업 및 종사자규모	사업체수	월평균 종사자수	급여액	출하액
D2321 계	8	7,176	583,927	68,408,790
5 ~ 9	0	0	0	0
10 ~ 19	0	0	0	0
20 ~ 49	1	44	1,232	62,182
50 ~ 99	1	62	2,790	68,637
100 ~ 199	1	179	8,861	570,945
500명 이상	5	6,891	571,044	67,707,026
D23210 계	8	7,176	583,927	68,408,790
5 ~ 9	0	0	0	0
10 ~ 19	0	0	0	0
20 ~ 49	1	44	1,232	62,182
50 ~ 99	1	62	2,790	68,637
100 ~ 199	1	179	8,861	570,945
500명 이상	5	6,891	571,044	67,707,026
D2322 계	109	1,992	58,449	1,321,643
5 ~ 9	38	257	5,967	217,730
10 ~ 19	35	469	11,144	211,887
20 ~ 49	31	922	27,076	568,481
50 ~ 99	5	344	14,262	323,545
100 ~ 199	0	0	0	0
500명 이상	0	0	0	0
D23221 계	66	1,210	37,211	1,099,890
5 ~ 9	22	150	3,231	170,564
10 ~ 19	23	315	7,300	177,910
20 ~ 49	18	552	17,452	469,511
50 ~ 99	3	193	9,228	281,905
100 ~ 199	0	0	0	0
500명 이상	0	0	0	0

〈표3-9〉 선형관련식에 의한 감추어진 셀 「X」 값 찾기 최종 결과(계속)

산업 및 종사자규모	사업체수	월평균 종사자수	급여액	출하액	
D23229	계	43	782	21,238	221,753
	5 ~ 9	16	107	2,736	47,166
	10 ~ 19	12	154	3,844	33,977
	20 ~ 49	13	370	9,624	98,970
	50 ~ 99	2	151	5,034	41,640
	100 ~ 199	0	0	0	0
	500명 이상	0	0	0	0
D233	계	1	623	32,048	110,113
	5 ~ 9	0	0	0	0
	10 ~ 19	0	0	0	0
	20 ~ 49	0	0	0	0
	50 ~ 99	0	0	0	0
	100 ~ 199	0	0	0	0
	500명 이상	1	623	32,048	110,113
D2330	계	1	623	32,048	110,113
	5 ~ 9	0	0	0	0
	10 ~ 19	0	0	0	0
	20 ~ 49	0	0	0	0
	50 ~ 99	0	0	0	0
	100 ~ 199	0	0	0	0
	500명 이상	1	623	32,048	110,113
D23300	계	1	623	32,048	110,113
	5 ~ 9	0	0	0	0
	10 ~ 19	0	0	0	0
	20 ~ 49	0	0	0	0
	50 ~ 99	0	0	0	0
	100 ~ 199	0	0	0	0
	500명 이상	1	623	32,048	110,113

3. 현행 SODC⁷⁾ 방법 적용에 따른 한계점

통계청에서 실시하고 있는 산업구조통계는 모두 사업체수가 2개 이하인 분류와 관련된 행의 자료를 민감한 셀로 정의하고 보호하는 SODC 방법을 적용하고 있다. 그러나 이 방법은 <표3 - 4>~<표3 - 8>에 제시된 바와 같이, 집계표내의 행과 열의 선형관련식을 이용하여 비밀보호된 셀 값 「X」를 찾을 수 있다.

즉, 중분류 D23의 경우 민감한 셀에 대한 비밀보호가 무의미하다고 볼 수 있으며, 해당 셀에 「X」 표시 대신 실제 집계자료를 제공해 준 것과 다르지 않다. 결과적으로 SODC방법 적용에 따라, 이용자는 「X」 값을 찾기 위해 선형관련식을 풀어야 하는 어려움만 있을 뿐, 그 값을 알아낼 수 있다는 데에 그 한계점을 갖는다.

한편, 중분류 D23과 같이 모든 분류에서 감추어진 셀 「X」 값이 모두 계산되는 것은 아니다. <표8>의 중분류 D27에 속하는 세분류 D2713의 경우를 살펴보자. 먼저 세세분류 D27131의 종사자규모 「200~299명」의 감추어진 셀은 그 합계가 주어져 있으므로 쉽게 계산된다. 다음으로 세분류 D2713의 종사자규모 「300~499명」, 「500명이상」에서 감추어진 셀 값을 계산하기 위해 선형관련식을 만들면 (3.3)과 같다.

$$D2713_i = D271_i - (D2711_i + D2712_i + D2719_i) \quad (3.3)$$

$D2713_i$ 값을 계산하기 위해 필요한 값 중 $D2711_i$ 와 $D2719_i$ 값이 사업체수가 2개 이하로서 감추어진 셀이다. 따라서 하나의 식에 2개의 미지수가 발생하여 선형관련식 (3.3)을 이용하여 감추어진 셀 「X」 값을 계산할 수 없다. 그러나 이와 같은 경우는 우연히 같은 범주의 종사자

7) SODC : Suppressing Only the Disclosure Cells

규모에 해당하는 사업체 수가 2개 이하로 감추어진 경우이며, 이와 같은 우연한 경우를 제외하고는 SODC방법은 노출위험이 있는 셀을 보호하는데 한계점을 가진다.

〈표3-10〉 감추어진 셀 「X」 값 찾기 (D2713)

산업 및 종사자규모	사업체수	월평균 종사자수	급여액	출하액	
D2713	계	695	18,481	465,891	6,791,073
	5 ~ 9명	221	1,511	26,753	290,258
	10 ~ 19명	217	2,917	59,758	715,556
	20 ~ 49명	178	5,543	124,725	1,680,559
	50 ~ 99명	55	3,578	94,269	1,414,480
	100 ~ 199명	17	2,260	69,610	1,372,520
	200 ~ 299명	4	1,125	33,716	335,659
	300 ~ 499명	1	X	X	X
	500명 이상	2	X	X	X
D27131	계	33	1,503	38,322	280,166
	5 ~ 9명	6	44	705	6,969
	10 ~ 19명	9	115	2,398	27,605
	20 ~ 49명	10	314	7,361	47,650
	50 ~ 99명	4	347	8,163	53,909
	100 ~ 199명	3	405	9,403	44,351
	200 ~ 299명	1	278	10,292	99,682
	300 ~ 499명	0	0	0	0
	500명 이상	0	0	0	0
D27132	계	662	16,978	427,569	6,510,907
	5 ~ 9명	215	1,467	26,048	283,289
	10 ~ 19명	208	2,802	57,360	687,951
	20 ~ 49명	168	5,229	117,364	1,632,909
	50 ~ 99명	51	3,231	86,106	1,360,571
	100 ~ 199명	14	1,855	60,207	1,328,169
	200 ~ 299명	3	847	23,424	235,977
	300 ~ 499명	1	X	X	X
	500명 이상	2	X	X	X

이와 같이 현재 통계청에서 제공되고 있는 집계표 자료는 SODC (Suppressing Only the Disclosure Cells)방법 적용에 따른 한계점으로 인하여, 노출위험이 있는 자료에 대한 비밀보호가 이루어지지 않고 있다. 따라서 다음 절에서는 실증분석을 통해 SODC방법을 보완한 「분류 구조변경(Cell Collapsing)방법」과 「보조셀감추기(Complementary Cell Suppression)방법」을 중심으로 집계표 형태의 자료에 대한 비밀보호 방법을 적용해 보고자 한다.

제 4 절 실증분석

1. 분석 대상

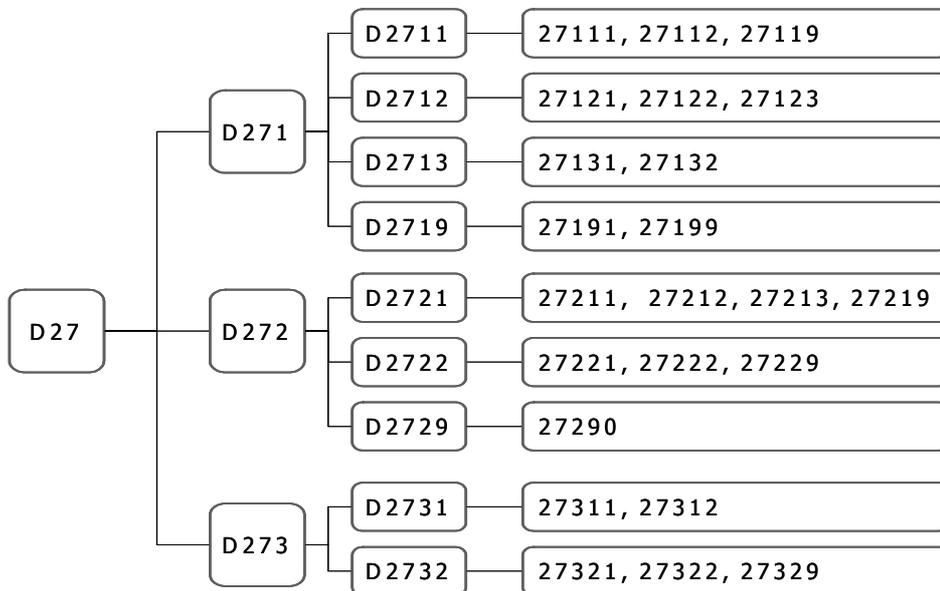
「광업 및 제조업통계조사」는 C.광업 3개, D.제조업 23개 중분류로 구성되어 있다. 분석대상은 이 중 매출액, 유형자산 비중과 소분류내 세분류 및 세세분류별 분포를 기준으로 선정하였다. 분석대상으로 선정한 중분류는 「27. 제1차 금속산업」으로, 분류체계는 <그림4-1>과 같이 소분류 3개, 세분류 9개, 세세분류 23개로 세분류별내 세세분류가 2개 이상으로 적절하게 분포되어 있다. 또한 <표4-1>에서 볼 수 있듯이 출하액과 유형자산 규모는 제조업 전체의 약 10% 정도로 큰 비중을 차지한다. 이에 따라 이 분류를 통해 집계표 형태 자료에 관한 비밀보호 방법의 적용에 관한 실증분석을 해보고자 한다.

〈표4-1〉 분석대상의 주요지표

(단위 : 개, 명, 10억원)

	사업체수	월평균 종사자수	출하액	부가가치	유형자산 연말잔액
제조업 전체	119,181	2,910,935	909,067	326,844	310,176
중분류27	3,173	117,684	88,721	25,282	30,990
구성비	2.7%	4.0%	9.8%	7.7%	10.0%

〈그림4-1〉 제1차 금속산업 분류체계(중분류 27)



2. 비밀보호 현황 및 한계점

다른 산업구조통계조사와 마찬가지로 「광업 및 제조업통계조사」 결과도 노출위험이 있는 셀만 감추는 SODC(Suppressing Only the Disclosure Cells) 방법을 적용하고 있다. 민감한 셀을 결정하는 기준(base)은 2이다. 즉, 사업체수가 1~2개인 경우 <표3-10>의 「300명~499명」 셀과 같이 노출위험이 있는 셀로 정의한 후 해당 셀과 관련된 모든 자료의 제공을 제한한다.

그러나 「광업 및 제조업통계조사」는 물론 산업구조 통계조사에 일반적으로 적용하고 있는 SODC방법은 제 2절에서 언급한 바와 같이 집계표내의 열과 행의 선형관련식으로 비밀보호된 셀 값을 계산 할 수 있기 때문에 적절한 비밀보호가 이루어지지 않는다.

또한, 집계표내에 빈도데이터⁸⁾와 크기데이터⁹⁾가 혼합되어 있기 때문에 비밀보호 여부는 전적으로 빈도데이터의 기준(base)이하 여부에 따라 결정된다. 즉, 사업체수와 기타 크기데이터가 서로 상관(correlate)되어 있기 때문에 사업체수에 따라 노출위험 셀 여부가 결정된다. 이는 크기데이터의 민감한 셀을 결정하는 방법인 (n,k) - dominance rule, p-percent rule 등을 적용하는데 한계점을 갖게 된다.

따라서 현재의 자료제공 형태를 유지하면서 효과적으로 자료의 비밀을 보호하는 방법은 셀 감추기 방법 중 하나인 보조셀감추기(CCS)¹⁰⁾를 적용하는 것이 바람직하다. 본 절에서는 분석대상으로 선정된 중분류 27에 대한 보조셀감추기(CCS) 방법의 적용결과를 실증분석하였다. 중분류 27에 속하는 소분류 272를 중심으로 보조셀감추기방법을 적용한 비밀보호 과정을 자세히 설명하고 있으며, 중분류 27 전체의 적용 결과표는 부록에 수록하였다.

3. 보조셀감추기(CCS)방법의 적용

CCS방법은 제 2절에서 언급한 바와 같이, 노출위험이 있는 셀의 비밀을 보호하기 위해 보조감추기셀(Complementary Supression Cells :CSC)을 찾아 집계표내의 행과 열의 선형관련식으로 계산되지 않도록

8) 빈도데이터 : Count data

9) 크기데이터 : Magnitude data

10) 보조감추기 : Complementary Cells Supression(CCS)

하는 것이다. 이 방법은 열과 행의 총합 및 분포의 변형 등이 없기 때문에 일반적으로 적용되는 방법이다. 그러나 1개의 노출 위험이 있는 민감한 셀(Primary Disclosure Cell :PDC)을 보호하기 위해 일반적으로 3개의 보조감추기셀(CSC)을 추가해야 한다.

이에따라 보조셀감추기(CCS) 방법은 SODC방법에 비해 노출위험이 있는 셀을 완전히 보호할 수 있는 반면, 감추어진 셀 수는 2~3배 이상 증가할 수 있기 때문에 비밀보호 측면과 정보공개 측면이 상충될 수 있다.

본 절에서는 보조셀감추기(CCS) 방법과 분류구조변경(Cell Collapsing)방법을 혼합하여, 소분류 272에 대한 결과를 우선 제시하고 「감추어진 셀 수」와 「감추어진 정보량」으로 적용방법의 효율성을 제고하고자 한다.

〈표4-2〉 소분류별 사업체수(D272)

산업분류 종사자규모	D272	D2721	D2722	D2729
계	748	217	467	64
5 ~ 9	241	72	147	22
10 ~ 19	232	79	137	16
20 ~ 49	169	47	101	21
50 ~ 99	59	11	45	3
100 ~ 199	26	4	21	1
200 ~ 299	8	0	7	1
300 ~ 499	7	2	5	0
500명 이상	6	2	4	0

우선 SODC 방법에 의한 비밀보호 현황을 파악하기 위해 하향식 (Top down) 접근방식으로 살펴보면, 기준(base) 2에서 노출위험이 있는 셀은 세분류 D2721과 D2729에서 각각 2개가 발생한다. 그러나 4개의 셀에 대한 관련정보를 「X」 표시로 제한하더라도 다음과 같은 선형 관련식으로 셀 값을 계산할 수 있다.

$$D2721_i = D272_i - (D2722_i + D2729_i) \quad (4.1)$$

$$D2729_i = D272_i - (D2721_i + D2722_i) \quad (4.2)$$

<표4 - 3>에서 계산된 결과로 세분류 D2721의 비밀보호 정도를 보면, 감추어진 셀 8개중 「50~99명」 행은 관련식(4.3)으로 계산 한 후, (4.4)식에 계산된 값을 대입하면 감추어진 셀 값은 모두 계산된다.

$$D27219_i = \sum_{i=1}^8 D27219_i - \sum_{j \neq i} D27219_j \quad (4.3)$$

$$D27213_i = D2721_i - (D27211_j + D27212_i + D27219_i) \quad (4.4)$$

나머지 6개의 감추어진 셀 값은 우연히 같은 범주의 셀이 민감한 셀로 선정되어 감추어짐에 따라 보조셀감추기(CCS)방법과 비슷한 형태의 조합이 이루어져 계산이 불가능하다. 기준(base) 2이하의 SODC방법에 의한 결과, 노출위험이 있는 모든 셀에 대한 비밀보호가 이루어지지 않았다. 따라서 보조셀감추기 방법으로 추가적인 비밀보호가 필요하며, <표4 - 2>, <표4 - 3>에 대한 소분류 D272의 종사자규모별 집계표의 보조 셀감추기(CCS) 결과는 <표4 - 4>와 같다.

〈표4-3〉 세분류별 사업체수(D2721)

산업분류 중사자규모	D2721	D27211	D27212	D27213	D27219
계	217	22	128	32	35
5 ~ 9	72	9	40	9	14
10 ~ 19	79	5	53	9	12
20 ~ 49	47	5	25	10	7
50 ~ 99	11	0	7	(2)	(2)
100 ~ 199	4	X(2)	X(2)	0	0
200 ~ 299	0	0	0	0	0
300 ~ 499	(2)	0	X(1)	X(1)	0
500명 이상	(2)	X(1)	0	X(1)	0

<그림4 - 1>의 중분류 D27 분류체계를 참조하여, 상위분류의 노출위험 셀을 비밀보호하면서 하위분류를 감추기 하는 하향식(Top down) 접근방식으로 보조셀 감추기를 수행한 결과, <표4 - 4>와 같이 소분류 단계에서는 노출위험이 있는 셀이 없다. 즉, 소분류 단위로 자료를 제공할 경우, 모든 셀 값이 제공 가능함을 알 수 있다.

다음으로 세분류 단계를 살펴보자. 세분류에 관한 보조셀감추기를 위해서는 PDC(Primary disclosure cell)를 중심으로 경로를 구성하여 감추어지는 정보량이 최소인 것을 찾아 보조감추기셀(CSC)로 정해야 한다. 그 결과는 <표4 - 5>와 같으며 그 방법은 다음과 같다.

〈표4-4〉 보조셀감추기(CCS) 결과 (소분류)

산업분류 종사자규모	D271	D272	D273
계	1,805	748	620
5 ~ 9	625	241	162
10 ~ 19	540	232	195
20 ~ 49	418	169	182
50 ~ 99	130	59	59
100 ~ 199	57	26	15
200 ~ 299	11	8	7
300 ~ 499	10	7	0
500명 이상	14	6	0

〈표4-5〉 보조셀감추기(CCS) 결과 (세분류)

산업분류 종사자규모	D2721	D2722	D2729
계	217	467	64
5 ~ 9	72	147	22
10 ~ 19	79	137	16
20 ~ 49	47	101	21
50 ~ 99	CSC(11)	45	CSC(3)
100 ~ 199	CSC(4)	21	PDC(1)
200 ~ 299	0	CSC(7)	PDC(1)
300 ~ 499	PDC(2)	CSC(5)	0
500명 이상	PDC(2)	CSC(4)	0

보조감추기셀(Complementary Suppression Cell :CSC)을 찾는 과정은 셀들의 선형관련성으로 감춘 셀이 계산되지 않도록 하는 수학적 접근방법이 필요하며, 아울러 감추어지는 셀의 수와 정보량을 최소화 하는 보조감추기 셀을 찾아야 한다.

$$\text{목적함수(1): } V(a) = \min \sum a_{ij} z_{ij}$$

$$\text{목적함수(2): } V(a) = \min \sum z_{ij}$$

☞ a_{ij} : 셀값, $z_{ij} = 1$ (감춘셀), $z_{ij} = 0$ (기타 셀)

이러한 논리적인 접근을 위해 결합기법(combination technique), 선형기법(linear programming), 수학적 네트워크기법(mathematical network)과 같은 3가지 접근법이 주로 적용된다. 3가지 방법 모두 수리적인 측면에서는 상이하나, 자료제공이 제한되는 셀의 수와 정보량을 최소화 하는 목적함수(Objective function)를 정의하고, 이를 이용하여 그 해(解)를 찾아나간다는 점에서는 그 맥락이 같다. 이 때 사용되는 목적함수는 위와 같이 감추어지는 정보량을 최소화하는 관점에서의 목적함수(1)과 감추어지는 셀의 수를 최소화하는 관점에서의 목적함수(2)로 나타낼 수 있다. 이러한 목적함수를 이용하여 보조감추기 셀을 찾아간다.

위 내용의 세부과정을 살펴보면 다음과 같은 단계로 요약된다.

Step 1 : <표4 - 5 - 1>과 같이 임의의 PDC(Primary disclosure cell) a_{s1} 을 출발점으로 가능한 모든 경로(Circuit)를 찾은 후, 각 경로에 해당하는 셀 값의 합을 구한다.

Step 2 : 경로 중 「0」 값을 포함하고 있는 경우, 이를 찾아 제외시킨다. 왜냐하면 「0」 값은 보고서를 통해 이용자에게 제공

해야 하는 정보로서 보조감추기 셀로 이용할 수 없기 때문이다.

Step 3 : Step1, Step2에서 남은 경로 중 감추어진 셀 값이 최소인 경로를 선택한다. <표4 - 5 - 1>

Step 4 : 모든 PDC에 대한 보조감추기 경로를 찾아 총합을 최소화 하는 셀을 선택한다.<표4 - 5>

위와 같은 과정으로 a_{81} 에 대한 보조감추기셀은 a_{82}, a_{72}, a_{71} 이 된다.

<표4-5-1> 보조감추기 셀 찾기($a_{81} = 2$)

PDC	Circuit			감추어진 셀값의 합	찾기
(8,1)=2	(8,2)	(7,2)	(7,1)	13	최소경로
	(8,2)	(6,2)	(6,1)	13	X
	(8,2)	(5,2)	(5,1)	31	
	(8,2)	(4,2)	(4,1)	62	
	(8,2)	(3,2)	(3,1)	154	
	(8,2)	(2,2)	(2,1)	222	
	(8,2)	(1,2)	(1,1)	225	
	(8,3)	(7,3)	(7,1)	4	X
	(8,3)	(6,3)	(6,1)	3	X
	(8,3)	(5,3)	(5,1)	8	X
	(8,3)	(4,3)	(4,1)	16	X
	(8,3)	(3,3)	(3,1)	70	X
	(8,3)	(2,3)	(2,1)	97	X
	(8,3)	(1,3)	(1,1)	96	X

동일한 방법으로 PDC a_{71}, a_{53}, a_{63} 에 대한 최소경로 탐색결과는 아래와 같다.

〈표4-5-2〉 보조감추기 셀 찾기(a_{71}, a_{53}, a_{63})

PDC	Circuit			감추어진 셀값의 합	찾기
(7,1)=2	(7,2)	(8,2)	(8,1)	13	최소경로
(5,3)=1	(5,1)	(4,1)	(4,3)	19	최소경로
(6,3)=1	(6,2)	(5,2)	(5,3)	30	최소경로

위와 같은 최소경로 탐색결과를 바탕으로 보조감추기셀을 찾은 결과, $a_{41}, a_{43}, a_{51}, a_{52}, a_{62}, a_{72}, a_{82}$ 가 보조감추기셀(Complementary Suppression Cell : CSC)로 결정되었다. 그러나 이 경우 각각의 PDC에 대한 최소경로의 탐색에서 나아가 감추어지는 셀 전체의 관점에서 보면, 5번행에서 과대감추기(Over Suppression)가 발생되었음을 알 수 있다. 따라서 과대감추기셀인 a_{52} 는 보조감추기셀에서 제외한다.

〈표4-6〉 보조셀감추기(CCS) 결과(세세분류)

산업분류 종사자규모	D2721	D27211	D27212	D27213	D27219
계	217	22	128	32	35
5 ~ 9	72	9	40	9	14
10 ~ 19	79	5	53	9	12
20 ~ 49	47	CSC(5)	CSC(25)	CSC(10)	CSC(7)
50 ~ 99	CSC(11)	0	CSC(7)	PDC(2)	PDC(2)
100 ~ 199	4	PDC(2)	PDC(2)	0	0
200 ~ 299	0	0	0	0	0
300 ~ 499	PDC(2)	0	PDC(1)	PDC(1)	0
500명 이상	PDC(2)	PDC(1)	0	PDC(1)	0

〈표4-6〉 보조셀 감추기 결과(세세분류)(계속)

산업분류 종사자규모	D2722	D27221	D27222	D27229
계	467	182	240	45
5 ~ 9	147	51	75	21
10 ~ 19	137	53	75	9
20 ~ 49	101	42	49	10
50 ~ 99	45	CSC(23)	20	PDC(2)
100 ~ 199	CSC(21)	CSC(7)	12	PDC(2)
200 ~ 299	CSC(7)	PDC(2)	CSC(4)	PDC(1)
300 ~ 499	CSC(5)	PDC(2)	CSC(3)	0
500명 이상	CSC(4)	PDC(2)	PDC(2)	0

〈표4-6〉 보조셀 감추기 결과(세세분류)(계속)

산업분류 종사자규모	D2729	D27290
계	64	64
5 ~ 9	22	22
10 ~ 19	16	16
20 ~ 49	21	21
50 ~ 99	CSC(3)	CSC(3)
100 ~ 199	PDC(1)	PDC(1)
200 ~ 299	PDC(1)	PDC(1)
300 ~ 499	0	0
500명 이상	0	0

4. 타당성 검증

보조셀감추기(CCS)방법을 적용해본 결과, 기존의 SODC¹¹⁾방법과 본 절에서 적용한 방법에 어떠한 차이가 있는지에 대해 비교해 그 차이에 대해 살펴보고자 한다. 또한 보조셀감추기 방법에 의해 적용된 위의 결과가 선형관련식에 의해 재계산 되지 않는지를 확인하는 검증을 실시하고자 한다.

먼저, 소분류 D272의 종사자규모별 자료에 대해 보조셀감추기(CCS)방법을 적용한 결과와 기존 적용하고 있는 SODC방법의 결과에 대해 「감추어진 셀 수」와 「감추어진 정보량」 측면에서 각각 비교해 보았다. 본 절에서 적용한 보조셀감추기 방법이 기존 방법보다 감추어진 셀 수와 「감추어진 정보량」은 각각 증가하는 것을 <표4-7>을 통해 알 수 있다. 그러나 보조셀감추기 방법으로 노출 위험이 있는 민감한 셀(PDC)을 비밀보호하기 위해 추가적으로 제한된 정보량은 약 6.8%로, SODC방법이 민감한 셀(PDC)을 선형관련식으로 재계산해낼 수 있다는 한계점을 극복한다는 점에서 높은 수준은 아니다.

<표4-7> CCS와 SODC방법 적용결과 비교 (감추어진 셀의 수)

산업분류		D272	D2721	D2722	D2729
원자료	셀수(A)	24	8	8	8
CCS방법	감추어진 셀 수(B)	10	4	3	3
	$C=(B/A) \times 100$	41.7	50.0	37.5	37.5
SODC방법	감추어진 셀수(D)	4	2	0	2
	$E=(B/A) \times 100$	16.7	25.0	0.0	25.0
Difference	C - E	25.0	25.0	37.5	12.5

11) SODC : Suppressing Only the Disclosure Cells

〈표4-8〉 CCS와 SODC방법 적용결과 비교 (감추어진 정보량)

산업분류		D272	D2721	D2722	D2729
원자료	정보량(H)	748	217	467	64
CCS방법	감추어진 정보량(I)	40	19	16	5
	$J=(I/H) \times 100$	5.4	8.8	3.4	7.8
SODC방법	감추어진 정보량(K)	6	4	0	2
	$L=(K/H) \times 100$	0.8	1.8	0.0	3.1
Difference	J - L	4.6	7.0	3.4	4.7

본 절에서 제시한 실증분석 결과, 보조셀 감추기(CCS) 방법은 집계표내 선형관련식에 의해 감추어진 셀 값이 재계산될 수 있는 SODC방법의 한계점을 보완하였다. 그러나 추가적으로 보조셀을 찾아 감추기 해야 하는 등 제한되는 셀 수는 평균 2배 수준, 정보량은 구간에 따라 차이를 보이고 있으나 SODC방법의 결과 보다는 증가한다. 이와 같은 감추어진 정보량의 증가는 보조셀감추기(CCS)방법의 문제점이 아니라, SODC방법에서 충분한 비밀보호를 위해 정보를 제한하지 않은 차이로 기인된다.

다음으로는 보조셀감추기 결과가 집계표내 행과 열의 선형관련식에 의해 재계산되는지 <표4-5>를 이용하여 디버깅(debugging) 과정을 살펴본다.

〈표4-5〉 보조셀감추기(CCS) 결과 (세분류)

산업분류 종사자규모	D272	D2721	D2722	D2729
계	748	217	467	64
5 ~ 9	241	72	147	22
10 ~ 19	232	79	137	16
20 ~ 49	169	47	101	21
50 ~ 99	59	CSC(11)	45	CSC(3)
100 ~ 199	26	CSC(4)	21	PDC(1)
200 ~ 299	8	0	CSC(7)	PDC(1)
300 ~ 499	7	PDC(2)	CSC(5)	0
500명 이상	6	PDC(2)	CSC(4)	0

〈표4-5〉에서 감추어진 셀은 10개의 셀 a_{41} , a_{51} , a_{71} , a_{81} , a_{62} , a_{72} , a_{82} , a_{43} , a_{53} , a_{63} 이다. 선형관련식으로 표현하면 다음과 같은데, 이들 식을 통해 감추어진 셀 값에 대한 유일한 해(Unique solution)를 찾을 수 없으며, 보조감추기셀(CSC) 값의 합도 최소이기 때문에 보조셀감추기(CCS)가 최적(Optimal)으로 이루어진 것으로 평가한다.

$$a_{41} + a_{43} = 14$$

$$a_{51} + a_{53} = 5$$

$$a_{62} + a_{63} = 8$$

$$a_{71} + a_{72} = 7$$

$$a_{81} + a_{82} = 6$$

$$a_{41} + a_{51} + a_{71} + a_{81} = 19$$

$$a_{62} + a_{72} + a_{82} = 16$$

$$a_{43} + a_{53} + a_{63} = 5$$

$$a_{ij} \geq 1$$

<표4-6>에 대한 결과도 위와 같은 동일한 방법을 적용하여 선형관련 식으로 나타낼 수 있으며, 이 경우도 유일한 해(Unique solution)를 찾을 수 없다.

5. 실증분석 결과 요약

실증분석을 통해 얻은 결론을 요약하면 다음과 같다. 집계표 구성형태를 어떻게 할 것인가의 측면에서 살펴보면, 첫째로 기존 집계표 형태를 유지하는 방안과 둘째로 집계표 구성형태를 재조정하는 방안을 생각해 볼 수 있다.

먼저 기존 집계표 형태를 유지하기 위해서는 노출위험이 있는 셀만 감추는 기존방법(Suppression Only the Disclosure Cells: SODC)을 보완한 보조셀 감추기(CCS)방법을 적용하여 추가적인 셀 감추기를 실시해야 한다. 그러나 이 방법을 적용할 때, 모든 세부분류가 산업별로 동일한 기준으로 적용되고 있기 때문에, 산업별 특성에 따라 특정 셀에서 0 값 또는 기준(base) 이하의 노출위험이 있는 셀이 편중되어 나타나고 있다. 이와 같은 특징은 보조셀 감추기(CCS) 방법 적용 시 실행 가능한 감추기경로(feasible circuit)를 찾지 못하는 경우가 나타날 수 있기 때문에 특정 집계표 전체를 감추기 해야 하는 문제점이 발생한다. 따라서 산업별 또는 세부 분류별 분포를 미리 파악하여 「분류구조변경(Cell Collapsing)방법」과 「보조셀 감추기(CCS)방법」을 병행하여 노출위험 셀을 감추기 하는 방법을 적용해야 한다.

두 번째로 집계표 구성형태를 재조정 하는 방법을 살펴보자. 기존 집계표의 대부분은 사업체수(빈도데이터)와 기타 항목(크기데이터)이 혼합되어 있다. 그러나 노출위험 셀을 결정하고 비밀보호를 적용하는

기준변수(key variable)는 빈도데이터인 사업체수에 의존하고 있기 때문에 자료의 종류에 따라 다양한 비밀보호 기법을 적용하는데 한계점을 갖게 된다. 따라서 집계표 구성형태를 재검토하여 빈도데이터와 크기데이터를 구분하여 집계표를 작성하는 것이 바람직하다고 판단된다.

끝으로 노출위험이 있는 셀을 결정하는 기준(Base)에 대해 살펴보면, 통계청에서는 그 기준을 2로 정의하고 있다. 그러나 이 경우 특정 업체의 정보를 알고 있을 경우 다른 업체의 정보가 추정가능하기 때문에 비밀보호가 효과적으로 이루어졌다고 보기 어렵다. 따라서 최소한 기준(Base)을 3이상으로 상향하는 것을 검토할 필요가 있다.

제 5절 결론

본 연구는 집계표 데이터의 비밀보호방법에 관한 실증연구로서, 미국 국립보건통계센터(National Center for Health Statistics)에서 동 분야에 대한 선행연구를 수행한 Dr. Lawrence Henry COX의 연구를 바탕으로 하였다. 이 연구의 초점은 집계표 데이터의 비밀보호 기법에 관한 고찰을 통해 통계작성자가 이 문헌연구와 예시 등을 바탕으로, 추후 집계표 데이터의 비밀보호기법을 적용할 때 용이하게 접근할 수 있도록 방법론을 소개하고, 통계청 산업구조통계를 중심으로 집계표 데이터의 비밀보호 현황분석과 한계점을 도출하는데 있다. 또한, 한계점을 극복하기 위한 새로운 기법을 선정하여 실증분석을 통해 우리나라의 실정에 적합한 비밀보호 기법을 도출하는데 있다.

이에 따라 제2절에서는 전통적으로 적용하고 있는 비밀보호기법 선행연구 결과를 참조하여 방법론을 소개하였다. 외국의 적용사례를 통한 우리나라에 적합한 방법론 선정을 위해, 먼저 우리나라와 미국의

제조업통계를 중심으로 집계표의 자료형태를 비교하여 차이점을 분석하였다. 이는 비밀보호기법이 집계표를 구성하고 있는 자료형태 즉, 빈도데이터(Count Data) 및 크기데이터(Magnitude Data) 여부에 따라 각각 다르게 적용하고 있기 때문이다. 자료형태에 따른 비밀보호 대상을 찾기 위해 노출위험이 있는 민감한 셀을 탐색하는 방법과 이 셀들을 보호할 수 있는 방법론도 본 절에서 살펴보았다.

제3절에서는 통계청에서 작성중인 산업구조통계의 비밀보호 현황을 광업 및 제조업통계조사를 중심으로 살펴보았다. 광업 및 제조업통계조사는 노출위험 셀을 결정하는 기준(Base)을 2로 정하고 노출위험이 있는 셀만 감추기 하는 SODC(Suppression Only the Disclosure Cells) 방법을 적용하고 있다. 그러나 이 방법은 집계표내의 선형관련식을 이용하여 대부분의 감추기한 셀 값을 재계산 할 수 있기 때문에 민감한 셀을 비밀보호 하는데 한계점이 있음을 발견하였다.

제4절에서는 제3절에서 도출된 한계점을 보완하기 위해, 광업 및 제조업 통계조사 23개 중분류중 매출액, 유형자산 비중과 소분류내 세분류 및 세세분류별 분포를 고려하여 실증분석대상 중분류를 선정하여 보완된 비밀보호 기법을 적용하고 그 결과를 분석하였다. 실증분석 결과, 기존 집계표 형태를 유지하기 위해서는 노출위험이 있는 셀만 감추는 기존의 SODC방법을 보완한 보조셀감추기(CCS)방법을 적용하여 추가적인 셀 감추기를 실시해야 한다는 결론을 도출하였다. 그러나 모든 세부분류가 산업별로 동일한 기준으로 적용되고 있기 때문에, 산업별 특성에 따라 특정 셀에서 0 값 또는 기준(base) 이하의 노출위험이 있는 셀이 편중되어 나타나고 있다. 이와 같은 특징은 보조셀감추기(CCS) 방법 적용 시 실행가능한 감추기경로(feasible circuit)을 찾지 못하는 경우가 나타날 수 있기 때문에 특정 집계표 전체를 감추기 해야 하는 문제점이 발생한다. 따라서 산업별 또는 세부 분류별 분포를 미

리 파악하여 「분류구조변경(Cell Collapsing)방법」과 「보조셀 감추기(CCS)방법」을 병행하여 노출위험 셀을 감추기 하는 방법을 적용해야 한다.

그러나 현재 제공되고 있는 산업구조통계의 집계표 구성형태가 빈도데이터(count data)인 사업체수와 크기데이터(magnitude data)인 기타 항목이 혼합되어 있기 때문에, 노출위험 셀을 결정하는 기준변수(key variable)는 빈도데이터인 사업체수에 의존할 수밖에 없다. 이에 따라 현재의 집계표형태를 가지고는 적용할 수 있는 비밀보호 기법이 매우 한정적이다. 따라서 집계표 구성형태에 따라 다양하고 효과적인 비밀보호기법을 적용하기 위해서는 빈도데이터와 크기데이터를 구분하여 구성하는 형태로 집계표를 재조정하고, 이에 해당하는 다양한 비밀보호기법을 적용하는 것에 관련한 추가적인 후속연구가 필요할 것이다.

정보에 대한 이용자와 응답자는 창과 방패의 관계이다. 정보공개를 확대하는 것에 앞서 중요하게 생각해야 하는 것은 정보의 원천인 응답자의 민감한 정보를 보호하는 것이다. 고갈되어 가는 원천에서 좋은 정보의 생산을 기대하기는 어렵기 때문이다. 본 연구를 계기로 현재 적용중인 비밀보호기법이 재검토되어 응답자의 개별정보 노출에 대한 불안감을 제거함은 물론 정보이용자와 정보제공자(응답자)간의 중재기능으로 본 연구결과가 활용되기를 기대한다.

부 록

A. 종사자규모별 주요지표의 보조셀감추기방법 적용 결과

산업별		종사자규모별	2006			
			사업체수	종사자수	급여액	출하액
제 1차 금속산업	D27	계	3,173	117,484	3,858	88,721
		5 ~ 9	1,028	6,954	128	1,709
		10 ~ 19	967	13,113	278	3,794
		20 ~ 49	769	23,538	567	10,955
		50 ~ 99	248	16,982	475	9,330
		100 ~ 199	98	13,624	437	9,348
		200 ~ 299	26	6,350	203	3,011
		300 ~ 499	17	6,303	220	7,157
		500명 이상	20	30,620	1,550	43,417
		제 1차 철강산업	D271	계	1,805	74,240
5 ~ 9	625			4,235	78	1,060
10 ~ 19	540			7,255	154	2,066
20 ~ 49	418			13,016	323	6,448
50 ~ 99	130			8,729	248	5,427
100 ~ 199	57			7,739	264	5,302
200 ~ 299	11			2,682	92	1,543
300 ~ 499	10			3,985	155	5,584
500명 이상	14			26,599	1,393	33,165
제철, 제강 및 합금철 제조업	D2711			계	47	26,522
		5 ~ 9	12	79	1	23
		10 ~ 19	8	98	3	46
		20 ~ 49	7	216	5	132
		50 ~ 99	2	X	X	X
		100 ~ 199	4	532	23	338
		200 ~ 299	1	X	X	X
		300 ~ 499	2	X	X	X
		500명 이상	11	24,537	1,312	31,384

산업별	종사자규모별	2006				
		사업체수	종사자수	급여액	출하액	
냉간 압연 및 압출제품 제조업	D27122	계	77	3,565	122	3,252
		5 ~ 9	19	129	3	69
		10 ~ 19	21	284	6	103
		20 ~ 49	21	703	19	656
		50 ~ 99	9	713	23	276
		100 ~ 199	4	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	2	X	X	X
		500명 이상	0	0	0	0
철강선 제조업	D27123	계	135	5,137	161	3,668
		5 ~ 9	55	375	8	147
		10 ~ 19	35	474	10	145
		20 ~ 49	17	537	15	245
		50 ~ 99	12	794	24	422
		100 ~ 199	12	1,709	61	1,046
		200 ~ 299	2	X	X	X
		300 ~ 499	2	X	X	X
		500명 이상	0	0	0	0
철강관 제조업	D2713	계	695	18,481	466	6,791
		5 ~ 9	221	1,511	27	290
		10 ~ 19	217	2,917	60	716
		20 ~ 49	178	5,543	125	1,681
		50 ~ 99	55	3,578	94	1,414
		100 ~ 199	17	2,260	70	1,373
		200 ~ 299	4	1,125	34	336
		300 ~ 499	1	X	X	X
		500명 이상	2	X	X	X
주철관 제조업	D27131	계	33	1,503	38	280
		5 ~ 9	6	44	1	7
		10 ~ 19	9	115	2	28
		20 ~ 49	10	314	7	48
		50 ~ 99	4	347	8	54
		100 ~ 199	3	405	9	44
		200 ~ 299	1
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		종사자규모별	2006			
			사업체수	종사자수	급여액	출하액
강관 제조업	D27132	계	662	16,978	428	6,511
		5 ~ 9	215	1,467	26	283
		10 ~ 19	208	2,802	57	688
		20 ~ 49	168	5,229	117	1,633
		50 ~ 99	51	3,231	86	1,361
		100 ~ 199	14	1,855	60	1,328
		200 ~ 299	3	847	23	236
		300 ~ 499	1
		500명 이상	2
		기타 철강산업	D2719	계	704	16,248
5 ~ 9	275			1,848	35	470
10 ~ 19	225			3,035	67	920
20 ~ 49	148			4,483	121	3,175
50 ~ 99	35			X	X	X
100 ~ 199	15			1,937	68	1,385
200 ~ 299	2			X	X	X
300 ~ 499	3			X	X	X
500명 이상	1			X	X	X
절단가공 및 표면처리강재 제조업	D27191	계	698	16,181	453	11,685
		5 ~ 9	273	1,832	34	469
		10 ~ 19	221	2,984	66	918
		20 ~ 49	148	4,483	121	3,175
		50 ~ 99	35	2,324	64	2,376
		100 ~ 199	15	1,937	68	1,385
		200 ~ 299	2
		300 ~ 499	3	1,269	44	2,129
		500명 이상	1
그외 기타 철강산업	D27199	계	6	67	1	4
		5 ~ 9	2
		10 ~ 19	4	51	1	2
		20 ~ 49	0	0	0	0
		50 ~ 99	0	0	0	0
		100 ~ 199	0	0	0	0
		200 ~ 299	0	0	0	0
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		종사자규모별	2006			
			사업체수	종사자수	급여액	출하액
제 1차 비철금속산업	D272	계	748	26,147	743	24,156
		5 ~ 9	241	1,641	31	504
		10 ~ 19	232	3,184	72	1,318
		20 ~ 49	169	5,044	120	3,529
		50 ~ 99	59	4,252	120	2,888
		100 ~ 199	26	3,741	118	3,008
		200 ~ 299	8	1,946	60	1,084
		300 ~ 499	7	2,318	65	1,573
		500명 이상	6	4,021	158	10,252
		비철금속 제련, 정련 및 합금	D2721	계	217	6,315
5 ~ 9	72			478	9	209
10 ~ 19	79			1,116	27	683
20 ~ 49	47			1,402	35	2,155
50 ~ 99	11			X	X	X
100 ~ 199	4			553	15	554
200 ~ 299	0			0	0	0
300 ~ 499	2			X	X	X
500명 이상	2			X	X	X
동 제련, 정련 및 합금 제조업	D27211	계	22	1,180	48	5,098
		5 ~ 9	9	61	1	22
		10 ~ 19	5	73	2	37
		20 ~ 49	5	X	X	X
		50 ~ 99	0	0	0	0
		100 ~ 199	2	X	X	X
		200 ~ 299	0	0	0	0
		300 ~ 499	0	0	0	0
		500명 이상	1	X	X	X
알루미늄 제련, 정련 및 합금	D27212	계	128	2,853	75	2,196
		5 ~ 9	40	272	5	101
		10 ~ 19	53	754	18	489
		20 ~ 49	25	X	X	X
		50 ~ 99	7	X	X	X
		100 ~ 199	2	X	X	X
		200 ~ 299	0	0	0	0
		300 ~ 499	1	X	X	X
		500명 이상	0	0	0	0

산업별		종사자규모별	2006			
			사업체수	종사자수	급여액	출하액
연 및 아연제련, 정련 및 합금	D27213	계	32	1,694	54	3,030
		5 ~ 9	9	56	1	19
		10 ~ 19	9	133	3	56
		20 ~ 49	10	321	6	98
		50 ~ 99	2	X	X	X
		100 ~ 199	0	0	0	0
		200 ~ 299	0	0	0	0
		300 ~ 499	1	X	X	X
		500명 이상	1	X	X	X
기타 비철금속제련,정련 및 합금	D27219	계	35	588	15	841
		5 ~ 9	14	89	2	67
		10 ~ 19	12	156	3	101
		20 ~ 49	7	X	X	X
		50 ~ 99	2	X	X	X
		100 ~ 199	0	0	0	0
		200 ~ 299	0	0	0	0
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
비철금속압연,압출 및 연신제품	D27222	계	467	18,357	516	12,429
		5 ~ 9	147	1,019	19	256
		10 ~ 19	137	1,854	41	577
		20 ~ 49	101	3,050	72	1,129
		50 ~ 99	45	3,252	90	1,948
		100 ~ 199	21	X	X	X
		200 ~ 299	7	X	X	X
		300 ~ 499	5	X	X	X
		500명 이상	4	X	X	X
동 압연, 압출 및 연신제품 제조업	D27221	계	182	7,891	216	7,316
		5 ~ 9	51	356	7	109
		10 ~ 19	53	750	18	281
		20 ~ 49	42	1,257	27	512
		50 ~ 99	23	X	X	X
		100 ~ 199	7	X	X	X
		200 ~ 299	2	X	X	X
		300 ~ 499	2	X	X	X
		500명 이상	2	X	X	X

산업별	종사자규모별	2006				
		사업체수	종사자수	급여액	출하액	
알루미늄 압연, 압출 및 연신제품	D27222	계	240	9,214	262	4,433
		5 ~ 9	75	521	9	123
		10 ~ 19	75	988	21	267
		20 ~ 49	49	1,446	35	464
		50 ~ 99	20	1,400	38	503
		100 ~ 199	12	1,799	54	627
		200 ~ 299	4	X	X	X
		300 ~ 499	3	X	X	X
		500명 이상	2	X	X	X
기타 비철금속 압연, 압출 및 연신제품	D27229	계	45	1,252	38	679
		5 ~ 9	21	142	3	24
		10 ~ 19	9	116	2	29
		20 ~ 49	10	347	10	153
		50 ~ 99	2	X	X	X
		100 ~ 199	2	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
기타 제 1차 비철금속산업	D2729	계	64	1,475	37	563
		5 ~ 9	22	144	3	39
		10 ~ 19	16	214	5	58
		20 ~ 49	21	592	13	245
		50 ~ 99	3	X	X	X
		100 ~ 199	1	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
기타 제 1차 비철금속산업	D27290	계	64	1,475	37	563
		5 ~ 9	22	144	3	39
		10 ~ 19	16	214	5	58
		20 ~ 49	21	592	13	245
		50 ~ 99	3	X	X	X
		100 ~ 199	1	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0

산업별		종사자규모별	2006			
			사업체수	종사자수	급여액	출하액
금속 주조업	D273	계	620	17,097	407	3,970
		5 ~ 9	162	1,078	19	145
		10 ~ 19	195	2,674	51	411
		20 ~ 49	182	5,478	124	978
		50 ~ 99	59	4,001	107	1,015
		100 ~ 199	15	2,144	55	1,037
		200 ~ 299	7	1,722	51	384
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
		철강 주조업	D2731	계	392	12,320
5 ~ 9	75			519	9	60
10 ~ 19	115			1,595	30	231
20 ~ 49	141			4,255	97	691
50 ~ 99	43			2,858	80	643
100 ~ 199	12			X	X	X
200 ~ 299	6			X	X	X
300 ~ 499	0			0	0	0
500명 이상	0			0	0	0
선철주물 주조업	D27311			계	217	7,566
		5 ~ 9	36	265	5	28
		10 ~ 19	62	869	16	116
		20 ~ 49	83	2,546	56	401
		50 ~ 99	24	1,648	46	337
		100 ~ 199	7	X	X	X
		200 ~ 299	5	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
		강주물 주조업	D27312	계	175	4,754
5 ~ 9	39			254	4	33
10 ~ 19	53			726	14	115
20 ~ 49	58			1,709	41	291
50 ~ 99	19			1,210	34	307
100 ~ 199	5			X	X	X
200 ~ 299	1			X	X	X
300 ~ 499	0			0	0	0
500명 이상	0			0	0	0

산업별		종사자규모별	2006			
			사업체수	종사자수	급여액	출하액
비철금속 주조업	D2732	계	228	4,777	104	1,652
		5 ~ 9	87	559	9	85
		10 ~ 19	80	1,079	21	179
		20 ~ 49	41	1,223	27	287
		50 ~ 99	16	1,143	28	372
		100 ~ 199	3	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
		알루미늄주물 주조업	D27321	계	138	2,985
5 ~ 9	51			325	5	47
10 ~ 19	48			653	12	115
20 ~ 49	25			781	18	211
50 ~ 99	12			862	20	246
100 ~ 199	2		
200 ~ 299	0			0	0	0
300 ~ 499	0			0	0	0
500명 이상	0			0	0	0
동주물 주조업	D27322			계	39	849
		5 ~ 9	17	110	2	18
		10 ~ 19	10	124	3	23
		20 ~ 49	10	283	7	58
		50 ~ 99	1
		100 ~ 199	0	0	0	0
		200 ~ 299	1
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
		기타 비철금속 주조업	D27329	계	51	943
5 ~ 9	19			124	2	20
10 ~ 19	22			302	6	41
20 ~ 49	6			159	3	17
50 ~ 99	3			189	5	37
100 ~ 199	1		
200 ~ 299	0			0	0	0
300 ~ 499	0			0	0	0
500명 이상	0			0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		종사자규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
제 1차 금속산업	D27	계	3,173	64,609	25,282	30,990
		5 ~ 9	1,028	1,190	523	533
		10 ~ 19	967	2,693	1,127	1,091
		20 ~ 49	769	8,095	2,820	2,304
		50 ~ 99	248	6,935	2,524	2,331
		100 ~ 199	98	6,838	2,566	2,371
		200 ~ 299	26	1,982	1,046	888
		300 ~ 499	17	6,006	1,329	3,732
		500명 이상	20	30,869	13,348	17,740
		제 1차 철강산업	D271	계	1,805	42,038
5 ~ 9	625			722	338	369
10 ~ 19	540			1,449	633	671
20 ~ 49	418			4,993	1,488	1,526
50 ~ 99	130			4,196	1,295	1,406
100 ~ 199	57			3,586	1,715	1,481
200 ~ 299	11			851	672	455
300 ~ 499	10			4,683	1,007	3,337
500명 이상	14			21,557	12,008	15,888
제철, 제강 및 합금철 제조업	D2711			계	47	21,230
		5 ~ 9	12	18	5	6
		10 ~ 19	8	32	19	9
		20 ~ 49	7	120	17	39
		50 ~ 99	2	X	X	X
		100 ~ 199	4	265	72	302
		200 ~ 299	1	X	X	X
		300 ~ 499	2	X	X	X
		500명 이상	11	20,103	11,686	15,124

산업별		종사자규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
제철 및 제강업	D27111	계	29	20,909	12,000	15,501
		5 ~ 9	7	5	3	2
		10 ~ 19	3	17	4	3
		20 ~ 49	2
		50 ~ 99	1
		100 ~ 199	2
		200 ~ 299	1
		300 ~ 499	2
		500명 이상	11	20,103	11,686	15,124
		합금철 제조업	D27112	계	18	321
5 ~ 9	5			12	3	4
10 ~ 19	5			15	15	6
20 ~ 49	5			117	12	30
50 ~ 99	1		
100 ~ 199	2		
200 ~ 299	0			0	0	0
300 ~ 499	0			0	0	0
500명 이상	0			0	0	0
철강압연·압출 및 연신제품	D2712			계	359	7,200
		5 ~ 9	117	188	86	79
		10 ~ 19	90	292	98	139
		20 ~ 49	85	1,128	333	324
		50 ~ 99	38	1,152	388	502
		100 ~ 199	21	1,676	500	650
		200 ~ 299	4	281	342	128
		300 ~ 499	4	2,483	295	1,968
		500명 이상	0	0	0	0
열간 압연 및 압출제품 제조업	D27121	계	147	1,748	532	621
		5 ~ 9	43	41	20	26
		10 ~ 19	34	100	39	39
		20 ~ 49	47	406	152	153
		50 ~ 99	17	667	174	251
		100 ~ 199	5	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		종사자규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
냉간 압연 및 압출제품 제조업	D27122	계	77	2,482	806	1,300
		5 ~ 9	19	32	34	14
		10 ~ 19	21	81	24	54
		20 ~ 49	21	539	119	105
		50 ~ 99	9	214	65	109
		100 ~ 199	4	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	2	X	X	X
		500명 이상	0	0	0	0
철강선 제조업	D27123	계	135	2,970	704	1,868
		5 ~ 9	55	115	32	39
		10 ~ 19	35	111	35	45
		20 ~ 49	17	183	62	66
		50 ~ 99	12	271	149	142
		100 ~ 199	12	784	249	352
		200 ~ 299	2	X	X	X
		300 ~ 499	2	X	X	X
		500명 이상	0	0	0	0
철강관 제조업	D2713	계	695	4,657	2,229	1,976
		5 ~ 9	221	189	103	85
		10 ~ 19	217	485	238	226
		20 ~ 49	178	1,193	499	548
		50 ~ 99	55	1,053	398	380
		100 ~ 199	17	792	594	265
		200 ~ 299	4	199	154	130
		300 ~ 499	1	X	X	X
		500명 이상	2	X	X	X
주철관 제조업	D27131	계	33	152	126	147
		5 ~ 9	6	5	2	4
		10 ~ 19	9	17	11	11
		20 ~ 49	10	29	19	14
		50 ~ 99	4	31	26	40
		100 ~ 199	3	28	16	17
		200 ~ 299	1
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		종사자규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
강관 제조업	D27132	계	662	4,505	2,104	1,828
		5 ~ 9	215	185	100	81
		10 ~ 19	208	469	226	215
		20 ~ 49	168	1,163	480	534
		50 ~ 99	51	1,022	372	340
		100 ~ 199	14	764	578	248
		200 ~ 299	3	156	102	67
		300 ~ 499	1
		500명 이상	2
		기타 철강산업	D2719	계	704	8,951
5 ~ 9	275			327	144	200
10 ~ 19	225			640	278	297
20 ~ 49	148			2,551	639	615
50 ~ 99	35			X	X	X
100 ~ 199	15			853	550	263
200 ~ 299	2			X	X	X
300 ~ 499	3			X	X	X
500명 이상	1			X	X	X
절단가공 및 표면처리강재 제조업	D27191			계	698	8,950
		5 ~ 9	273	326	143	199
		10 ~ 19	221	639	277	294
		20 ~ 49	148	2,551	639	615
		50 ~ 99	35	1,922	471	498
		100 ~ 199	15	853	550	263
		200 ~ 299	2
		300 ~ 499	3	1,666	480	1,208
		500명 이상	1
		그외 기타 철강산업	D27199	계	6	1
5 ~ 9	2		
10 ~ 19	4			1	1	3
20 ~ 49	0			0	0	0
50 ~ 99	0			0	0	0
100 ~ 199	0			0	0	0
200 ~ 299	0			0	0	0
300 ~ 499	0			0	0	0
500명 이상	0			0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		종사자규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
제 1차 비철금속산업	D272	계	748	19,971	4,751	4,650
		5 ~ 9	241	376	132	105
		10 ~ 19	232	985	339	268
		20 ~ 49	169	2,501	954	487
		50 ~ 99	59	2,129	820	599
		100 ~ 199	26	2,458	608	712
		200 ~ 299	8	886	238	232
		300 ~ 499	7	1,323	322	395
		500명 이상	6	9,312	1,339	1,852
		비철금속 제련, 정련 및 합금	D2721	계	217	9,487
5 ~ 9	72			165	44	32
10 ~ 19	79			522	163	111
20 ~ 49	47			1,445	612	160
50 ~ 99	11			X	X	X
100 ~ 199	4			395	162	100
200 ~ 299	0			0	0	0
300 ~ 499	2			X	X	X
500명 이상	2			X	X	X
동 제련, 정련 및 합금 제조업	D27211	계	22	4,466	842	477
		5 ~ 9	9	14	7	4
		10 ~ 19	5	30	7	12
		20 ~ 49	5	X	X	X
		50 ~ 99	0	0	0	0
		100 ~ 199	2	X	X	X
		200 ~ 299	0	0	0	0
		300 ~ 499	0	0	0	0
		500명 이상	1	X	X	X
알루미늄 제련, 정련 및 합금	D27212	계	128	1,804	404	319
		5 ~ 9	40	75	26	21
		10 ~ 19	53	384	107	73
		20 ~ 49	25	X	X	X
		50 ~ 99	7	X	X	X
		100 ~ 199	2	X	X	X
		200 ~ 299	0	0	0	0
		300 ~ 499	1	X	X	X
		500명 이상	0	0	0	0

산업별	종사자규모별	2006				
		사업체수	주요생산비	부가가치	유형자산	
연 및 아연제련, 정련 및 합금	D27213	계	32	2,492	661	626
		5 ~ 9	9	14	5	2
		10 ~ 19	9	40	16	11
		20 ~ 49	10	74	26	13
		50 ~ 99	2	X	X	X
		100 ~ 199	0	0	0	0
		200 ~ 299	0	0	0	0
		300 ~ 499	1	X	X	X
		500명 이상	1	X	X	X
기타 비철금속제련,정련 및 합금	D27219	계	35	724	72	56
		5 ~ 9	14	61	6	5
		10 ~ 19	12	68	33	16
		20 ~ 49	7	X	X	X
		50 ~ 99	2	X	X	X
		100 ~ 199	0	0	0	0
		200 ~ 299	0	0	0	0
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
비철금속압연,압출 및 연신제품	D2722	계	467	10,103	2,581	3,013
		5 ~ 9	147	187	73	62
		10 ~ 19	137	426	155	143
		20 ~ 49	101	884	265	276
		50 ~ 99	45	1,297	692	486
		100 ~ 199	21	X	X	X
		200 ~ 299	7	X	X	X
		300 ~ 499	5	X	X	X
		500명 이상	4	X	X	X
동 압연, 압출 및 연신제품 제조업	D27221	계	182	5,963	1,524	1,500
		5 ~ 9	51	86	24	21
		10 ~ 19	53	211	74	58
		20 ~ 49	42	417	104	112
		50 ~ 99	23	X	X	X
		100 ~ 199	7	X	X	X
		200 ~ 299	2	X	X	X
		300 ~ 499	2	X	X	X
		500명 이상	2	X	X	X

산업별	종사자규모별	2006				
		사업체수	주요생산비	부가가치	유형자산	
알루미늄 압연, 압출 및 연신제품	D27222	계	240	3,566	943	1,427
		5 ~ 9	75	86	40	34
		10 ~ 19	75	195	72	81
		20 ~ 49	49	345	129	139
		50 ~ 99	20	407	106	187
		100 ~ 199	12	495	144	185
		200 ~ 299	4	X	X	X
		300 ~ 499	3	X	X	X
		500명 이상	2	X	X	X
기타 비철금속 압연, 압출 및 연신제품	D27229	계	45	574	114	86
		5 ~ 9	21	15	9	7
		10 ~ 19	9	20	9	4
		20 ~ 49	10	122	33	26
		50 ~ 99	2	X	X	X
		100 ~ 199	2	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
기타 제 1차 비철금속산업	D2729	계	64	381	191	158
		5 ~ 9	22	24	15	11
		10 ~ 19	16	37	20	14
		20 ~ 49	21	172	77	50
		50 ~ 99	3	X	X	X
		100 ~ 199	1	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
기타 제 1차 비철금속산업	D27290	계	64	381	191	158
		5 ~ 9	22	24	15	11
		10 ~ 19	16	37	20	14
		20 ~ 49	21	172	77	50
		50 ~ 99	3	X	X	X
		100 ~ 199	1	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0

산업별	종사자규모별	2006				
		사업체수	주요생산비	부가가치	유형자산	
금속 주조업	D273	계	620	2,601	1,374	1,208
		5 ~ 9	162	92	53	59
		10 ~ 19	195	259	155	152
		20 ~ 49	182	600	378	292
		50 ~ 99	59	609	409	326
		100 ~ 199	15	794	243	178
		200 ~ 299	7	245	136	201
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
철강 주조업	D2731	계	392	1,355	962	832
		5 ~ 9	75	36	24	31
		10 ~ 19	115	139	92	73
		20 ~ 49	141	406	282	222
		50 ~ 99	43	364	279	198
		100 ~ 199	12	X	X	X
		200 ~ 299	6	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
선철주물 주조업	D27311	계	217	873	595	556
		5 ~ 9	36	16	11	18
		10 ~ 19	62	66	49	44
		20 ~ 49	83	229	169	131
		50 ~ 99	24	194	142	106
		100 ~ 199	7	X	X	X
		200 ~ 299	5	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
강주물 주조업	D27312	계	175	482	367	275
		5 ~ 9	39	20	13	13
		10 ~ 19	53	73	43	29
		20 ~ 49	58	177	113	91
		50 ~ 99	19	170	137	93
		100 ~ 199	5	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0

산업별		종사자규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
비철금속 주조업	D2732	계	228	1,245	412	376
		5 ~ 9	87	56	29	28
		10 ~ 19	80	120	63	80
		20 ~ 49	41	194	96	70
		50 ~ 99	16	245	130	128
		100 ~ 199	3	X	X	X
		200 ~ 299	1	X	X	X
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
알루미늄주물 주조업	D27321	계	138	484	220	231
		5 ~ 9	51	31	17	18
		10 ~ 19	48	79	40	53
		20 ~ 49	25	146	67	50
		50 ~ 99	12	179	70	80
		100 ~ 199	2
		200 ~ 299	0	0	0	0
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
동주물 주조업	D27322	계	39	121	87	61
		5 ~ 9	17	12	6	4
		10 ~ 19	10	16	7	4
		20 ~ 49	10	39	20	11
		50 ~ 99	1
		100 ~ 199	0	0	0	0
		200 ~ 299	1
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0
기타 비철금속 주조업	D27329	계	51	640	105	84
		5 ~ 9	19	14	6	6
		10 ~ 19	22	26	16	22
		20 ~ 49	6	10	8	9
		50 ~ 99	3	22	16	26
		100 ~ 199	1
		200 ~ 299	0	0	0	0
		300 ~ 499	0	0	0	0
		500명 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

B. 출하액규모별 주요지표의 보조셀감추기방법 적용 결과

산업별		출하액규모별	2006			
			사업체수	종사자수	급여액	출하액
제 1차 금속산업	D27	계	3,173	117,484	3,858,282	88,720,949
		5억 미만	282	2,564	35,375	88,797
		5 ~ 10	435	4,002	67,594	322,549
		10 ~ 50	1,342	19,551	405,746	3,225,060
		50 ~ 100	408	11,118	277,122	2,876,762
		100 ~ 300	373	16,305	430,108	6,416,958
		300 ~ 400	74	5,206	148,359	2,558,808
		400 ~ 1000	162	14,746	444,863	9,749,943
		1000 ~ 5000	68	11,316	421,053	12,742,127
		5000억 이상	29	32,676	1,628,062	50,739,945
제 1차 철강산업	D271	계	1,805	74,240	2,708,160	60,594,683
		5억 미만	160	1,504	22,022	51,234
		5 ~ 10	264	2,432	40,528	196,217
		10 ~ 50	756	10,428	221,599	1,779,890
		50 ~ 100	209	5,427	141,308	1,482,073
		100 ~ 300	200	8,250	217,612	3,408,533
		300 ~ 400	49	2,636	74,466	1,713,896
		400 ~ 1000	102	8,061	248,113	6,023,474
		1000 ~ 5000	47	7,731	303,428	8,442,434
		5000억 이상	18	27,771	1,439,084	37,496,932
제철, 제강 및 합금철 제조업	D2711	계	47	26,522	1,383,708	32,920,911
		5억 미만	3	50	665	565
		5 ~ 10	4	34	378	3,418
		10 ~ 50	12	143	3,266	26,951
		50 ~ 100	4	185	6,998	29,240
		100 ~ 300	3	64	1,899	38,985
		300 ~ 400	0	0	0	0
		400 ~ 1000	5	337	11,059	282,120
		1000 ~ 5000	6	X	X	X
		5000억 이상	10	X	X	X

산업별		출하액규모별	2006			
			사업체수	종사자수	급여액	출하액
강관 제조업	D27132	계	662	16,978	427,569	6,510,907
		5억 미만	72	666	8,381	23,991
		5 ~ 10	85	789	12,904	63,634
		10 ~ 50	297	4,391	90,794	678,477
		50 ~ 100	82	2,181	50,464	578,223
		100 ~ 300	83	3,483	85,009	1,333,339
		300 ~ 400	8	588	15,933	280,552
		400 ~ 1000	26	2,503	72,508	1,436,726
		1000 ~ 5000	7	1,690	62,870	962,441
		5000억 이상	2
기타 철강산업	D2719	계	704	16,248	454,182	11,688,606
		5억 미만	68	629	11,097	21,554
		5 ~ 10	129	1,204	20,431	95,382
		10 ~ 50	292	3,792	82,947	687,907
		50 ~ 100	70	1,731	49,321	501,143
		100 ~ 300	64	2,390	67,964	1,147,861
		300 ~ 400	21	998	29,101	739,197
		400 ~ 1000	39	1,792	49,978	2,374,303
		1000 ~ 5000	18	1,868	76,411	3,261,836
		5000억 이상	3	1,844	66,932	2,859,423
절단가공 및 표면처리강재 생산업	D27191	계	698	16,181	453,217	11,684,950
		5억 미만	65	591	10,635	20,384
		5 ~ 10	127	X	X	X
		10 ~ 50	291	X	X	X
		50 ~ 100	70	1,731	49,321	501,143
		100 ~ 300	64	2,390	67,964	1,147,861
		300 ~ 400	21	998	29,101	739,197
		400 ~ 1000	39	1,792	49,978	2,374,303
		1000 ~ 5000	18	1,868	76,411	3,261,836
		5000억 이상	3	1,844	66,932	2,859,423
그외 기타 철강산업	D27199	계	6	67	965	3,656
		5억 미만	3	38	462	1,170
		5 ~ 10	2	X	X	X
		10 ~ 50	1	X	X	X
		50 ~ 100	0	0	0	0
		100 ~ 300	0	0	0	0
		300 ~ 400	0	0	0	0
		400 ~ 1000	0	0	0	0
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		출하액규모별	2006			
			사업체수	종사자수	급여액	출하액
제철 및 제강업	D27111	계	29	25,914	1,364,880	32,507,442
		5억 미만	1	X	X	X
		5 ~ 10	3	X	X	X
		10 ~ 50	7	94	2,122	17,184
		50 ~ 100	1	X	X	X
		100 ~ 300	1	X	X	X
		300 ~ 400	
		400 ~ 1000	1	X	X	X
		1000 ~ 5000	5	X	X	X
		5000억 이상	10	23,970	1,285,630	30,924,030
		합금철 제조업	D27112	계	18	608
5억 미만	2			X	X	X
5 ~ 10	1			X	X	X
10 ~ 50	5			49	1,144	9,767
50 ~ 100	3			X	X	X
100 ~ 300	2			X	X	X
300 ~ 400	0					
400 ~ 1000	4			X	X	X
1000 ~ 5000	1			X	X	X
5000억 이상	0			0	0	0
철강압연·압출 및 연신제품	D2712			계	359	12,989
		5억 미만	17	159	1,879	5,124
		5 ~ 10	40	356	5,973	29,578
		10 ~ 50	141	1,769	37,150	349,310
		50 ~ 100	46	1,059	28,392	320,689
		100 ~ 300	45	1,741	49,127	802,092
		300 ~ 400	20	1,050	29,432	694,147
		400 ~ 1000	31	3,151	104,276	1,830,643
		1000 ~ 5000	16	2,434	90,334	2,602,555
		5000억 이상	3	1,270	57,816	2,559,955
		열간 압연 및 압출제품 제조업	D27121	계	147	4,287
5억 미만	8			94	820	2,296
5 ~ 10	24			248	4,264	17,490
10 ~ 50	51			748	15,773	117,041
50 ~ 100	18			419	10,039	114,081
100 ~ 300	22			827	23,486	413,405
300 ~ 400	9			522	12,752	305,792
400 ~ 1000	9			618	20,395	510,564
1000 ~ 5000	6			811	34,728	793,563
5000억 이상	0			0	0	0

산업별		출하액규모별	2006			
			사업체수	종사자수	급여액	출하액
냉간 압연 및 압출제품 제조업	D27122	계	77	3,565	121,547	3,251,709
		5억 미만	3	X	X	X
		5 ~ 10	5	36	618	3,304
		10 ~ 50	30	428	9,632	85,588
		50 ~ 100	9	294	9,998	68,252
		100 ~ 300	10	350	9,214	168,402
		300 ~ 400	6	298	10,229	217,982
		400 ~ 1000	6	500	13,901	338,027
		1000 ~ 5000	6	767	24,825	1,192,779
		5000억 이상	2	X	X	X
철강선 제조업	D27123	계	135	5,137	160,575	3,668,152
		5억 미만	6	X	X	X
		5 ~ 10	11	72	1,091	8,784
		10 ~ 50	60	593	11,745	146,681
		50 ~ 100	19	346	8,355	138,356
		100 ~ 300	13	564	16,427	220,285
		300 ~ 400	5	230	6,451	170,373
		400 ~ 1000	16	2,033	69,980	982,052
		1000 ~ 5000	4	856	30,781	616,213
		5000억 이상	1	X	X	X
철강관 제조업	D2713	계	695	18,481	465,891	6,791,073
		5억 미만	72	666	8,381	23,991
		5 ~ 10	91	838	13,746	67,839
		10 ~ 50	311	4,724	98,236	715,722
		50 ~ 100	89	2,452	56,597	631,001
		100 ~ 300	88	4,055	98,622	1,419,595
		300 ~ 400	8	588	15,933	280,552
		400 ~ 1000	27	2,781	82,800	1,536,408
		1000 ~ 5000	7	X	X	X
		5000억 이상	2	X	X	X
주철관 제조업	D27131	계	33	1,503	38,322	280,166
		5억 미만	0	0	0	0
		5 ~ 10	6	49	842	4,205
		10 ~ 50	14	333	7,442	37,245
		50 ~ 100	7	271	6,133	52,778
		100 ~ 300	5	572	13,613	86,256
		300 ~ 400	0	0	0	0
		400 ~ 1000	1
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		출하액규모별	2006			
			사업체수	종사자수	급여액	출하액
제 1차 비철금속산업	D272	계	748	26,147	743,172	24,156,073
		5억 미만	52	515	6,193	16,043
		5 ~ 10	77	745	12,399	56,870
		10 ~ 50	295	3,967	77,995	755,995
		50 ~ 100	108	2,148	52,219	750,232
		100 ~ 300	118	3,881	104,319	2,091,985
		300 ~ 400	20	X	X	X
		400 ~ 1000	47	4,695	135,116	2,893,741
		1000 ~ 5000	21	3,585	117,625	4,299,693
		5000억 이상	10	X	X	X
비철금속 제련, 정련 및 합금	D2721	계	217	6,315	190,652	11,164,291
		5억 미만	11	X	X	X
		5 ~ 10	18	246	3,506	12,636
		10 ~ 50	79	902	18,345	214,678
		50 ~ 100	39	587	13,467	264,743
		100 ~ 300	38	736	19,742	669,194
		300 ~ 400	8	X	X	X
		400 ~ 1000	13	678	21,248	882,112
		1000 ~ 5000	6	584	18,794	975,065
		5000억 이상	5	X	X	X
동 제련, 정련 및 합금 제조업	D27211	계	22	1,180	47,553	5,097,829
		5억 미만	2	X	X	X
		5 ~ 10	1	X	X	X
		10 ~ 50	7	54	1,283	17,280
		50 ~ 100	4	52	1,214	24,969
		100 ~ 300	1	X	X	X
		300 ~ 400	1	X	X	X
		400 ~ 1000	2	X	X	X
		1000 ~ 5000	2	X	X	X
		5000억 이상	2	X	X	X
알루미늄 제련, 정련 및 합금	D27212	계	128	2,853	74,784	2,195,555
		5억 미만	3	X	X	X
		5 ~ 10	8	72	1,023	5,982
		10 ~ 50	43	438	8,512	117,646
		50 ~ 100	27	418	9,813	185,484
		100 ~ 300	31	570	14,971	545,814
		300 ~ 400	5	527	15,659	162,845
		400 ~ 1000	7	388	11,855	540,248
		1000 ~ 5000	4	X	X	X
		5000억 이상	0	0	0	0

산업별		출하액규모별	2006			
			사업체수	종사자수	급여액	출하액
연 및 아연제련, 정련 및 합금	D27213	계	32	1,694	53,589	3,029,562
		5억 미만	2	X	X	X
		5 ~ 10	5	X	X	X
		10 ~ 50	13	183	3,138	32,369
		50 ~ 100	4	63	1,405	26,613
		100 ~ 300	4	X	X	X
		300 ~ 400	0	0	0	0
		400 ~ 1000	2	X	X	X
		1000 ~ 5000	0	0	0	0
		5000억 이상	2	X	X	X
기타 비철금속제련,정련 및 합금	D27219	계	35	588	14,726	841,345
		5억 미만	4	43	736	1,557
		5 ~ 10	4	45	775	2,580
		10 ~ 50	16	227	5,412	47,383
		50 ~ 100	4	54	1,035	27,677
		100 ~ 300	2	X	X	X
		300 ~ 400	2	X	X	X
		400 ~ 1000	2	X	X	X
		1000 ~ 5000	0	0	0	0
		5000억 이상	1	X	X	X
비철금속압연,압출 및 연신제품	D27222	계	467	18,357	515,907	12,428,609
		5억 미만	38	296	3,959	12,281
		5 ~ 10	45	368	6,646	33,657
		10 ~ 50	190	2,636	50,896	475,729
		50 ~ 100	64	1,482	36,487	455,909
		100 ~ 300	69	2,787	75,342	1,234,673
		300 ~ 400	11	X	X	X
		400 ~ 1000	30	3,594	100,713	1,775,749
		1000 ~ 5000	15	3,001	98,831	3,324,628
		5000억 이상	5	X	X	X
동 압연, 압출 및 연신제품 제조업	D27221	계	182	7,891	216,454	7,316,333
		5억 미만	16	149	2,109	4,691
		5 ~ 10	10	82	1,515	7,193
		10 ~ 50	67	935	18,065	179,701
		50 ~ 100	25	461	11,270	178,631
		100 ~ 300	33	1,117	27,465	560,330
		300 ~ 400	3	X	X	X
		400 ~ 1000	17	1,769	48,697	965,443
		1000 ~ 5000	8	X	X	X
5000억 이상	3	X	X	X		

산업별		출하액규모별	2006			
			사업체수	종사자수	급여액	출하액
알루미늄 압연, 압출 및 연신제품	D27222	계	240	9,214	261,854	4,433,370
		5억 미만	16	110	1,338	5,043
		5 ~ 10	29	242	4,302	21,850
		10 ~ 50	105	1,507	28,791	259,375
		50 ~ 100	35	919	22,597	246,057
		100 ~ 300	28	1,306	36,893	508,431
		300 ~ 400	7	X	X	X
		400 ~ 1000	13	1,825	52,016	810,306
		1000 ~ 5000	5	1,406	44,893	815,492
		5000억 이상	2	X	X	X
기타 비철금속 압연, 압출 및 연신제품	D27229	계	45	1,252	37,599	678,906
		5억 미만	6	37	512	2,547
		5 ~ 10	6	44	829	4,614
		10 ~ 50	18	194	4,040	36,653
		50 ~ 100	4	102	2,620	31,221
		100 ~ 300	8	364	10,984	165,912
		300 ~ 400	1	X	X	X
		400 ~ 1000	0	0	0	0
		1000 ~ 5000	2	X	X	X
		5000억 이상	0	0	0	0
기타 제 1차 비철금속산업	D2729	계	64	1,475	36,613	563,173
		5억 미만	3	X	X	X
		5 ~ 10	14	131	2,247	10,577
		10 ~ 50	26	429	8,754	65,588
		50 ~ 100	5	79	2,265	29,580
		100 ~ 300	11	358	9,235	188,118
		300 ~ 400	1	X	X	X
		400 ~ 1000	4	423	13,155	235,880
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0
기타 제 1차 비철금속산업	D27290	계	64	1,475	36,613	563,173
		5억 미만	3	X	X	X
		5 ~ 10	14	131	2,247	10,577
		10 ~ 50	26	429	8,754	65,588
		50 ~ 100	5	79	2,265	29,580
		100 ~ 300	11	358	9,235	188,118
		300 ~ 400	1	X	X	X
		400 ~ 1000	4	423	13,155	235,880
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0

산업별		출하액규모별	2006			
			사업체수	종사자수	급여액	출하액
금속 주조업	D273	계	620	17,097	406,950	3,970,193
		5억 미만	70	545	7,160	21,520
		5 ~ 10	94	825	14,667	69,462
		10 ~ 50	291	5,156	106,152	689,175
		50 ~ 100	91	3,543	83,595	644,457
		100 ~ 300	55	4,174	108,177	916,440
		300 ~ 400	5	X	X	X
		400 ~ 1000	13	1,990	61,634	832,728
		1000 ~ 5000	0	0	0	0
		5000억 이상	1
철강 주조업	D2731	계	392	12,320	302,712	2,317,748
		5억 미만	37	307	4,302	11,858
		5 ~ 10	50	480	9,054	37,373
		10 ~ 50	187	3,654	75,989	455,094
		50 ~ 100	66	2,711	64,658	469,661
		100 ~ 300	40	3,108	82,715	693,521
		300 ~ 400	4	526	16,752	135,948
		400 ~ 1000	8	1,534	49,242	514,293
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0
선철주물 주조업	D27311	계	217	7,566	189,230	1,470,348
		5억 미만	21	189	3,159	6,877
		5 ~ 10	25	262	4,821	19,131
		10 ~ 50	99	1,985	40,326	241,400
		50 ~ 100	39	1,524	35,506	262,599
		100 ~ 300	23	1,698	45,374	386,418
		300 ~ 400	3	452	14,177	101,305
		400 ~ 1000	7	1,456	45,867	452,618
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0
강주물 주조업	D27312	계	175	4,754	113,482	847,400
		5억 미만	16	118	1,143	4,981
		5 ~ 10	25	218	4,233	18,242
		10 ~ 50	88	1,669	35,663	213,694
		50 ~ 100	27	1,187	29,152	207,062
		100 ~ 300	17	1,410	37,341	307,103
		300 ~ 400	1
		400 ~ 1000	1
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		출하액규모별	2006			
			사업체수	종사자수	급여액	출하액
비철금속 주조업	D2732	계	228	4,777	104,238	1,652,445
		5억 미만	33	238	2,858	9,662
		5 ~ 10	44	345	5,613	32,089
		10 ~ 50	104	1,502	30,163	234,081
		50 ~ 100	25	832	18,937	174,796
		100 ~ 300	15	1,066	25,462	222,919
		300 ~ 400	1
		400 ~ 1000	5	456	12,392	318,435
		1000 ~ 5000	0	0	0	0
		5000억 이상	1
		알루미늄 주물 주조업	D27321	계	138	2,985
5억 미만	15			91	1,269	4,741
5 ~ 10	32			254	3,862	22,914
10 ~ 50	59			897	17,567	134,095
50 ~ 100	16			566	12,539	114,814
100 ~ 300	11			644	15,695	159,553
300 ~ 400	1		
400 ~ 1000	4			364	9,311	229,838
1000 ~ 5000	0			0	0	0
5000억 이상	0			0	0	0
동주물 주조업	D27322			계	39	849
		5억 미만	6	41	409	1,712
		5 ~ 10	5	32	508	3,711
		10 ~ 50	19	229	5,402	42,758
		50 ~ 100	6	172	4,329	40,691
		100 ~ 300	2
		300 ~ 400	0	0	0	0
		400 ~ 1000	1
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0
		기타 비철금속 주조업	D27329	계	51	943
5억 미만	12			106	1,180	3,209
5 ~ 10	7			59	1,243	5,464
10 ~ 50	26			376	7,194	57,228
50 ~ 100	3			94	2,069	19,291
100 ~ 300	2		
300 ~ 400	0			0	0	0
400 ~ 1000	0			0	0	0
1000 ~ 5000	0			0	0	0
5000억 이상	1		

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		출하액규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
제 1차 금속산업	D27	계	3,173	64,609,305	25,282,343	30,990,026
		5억 미만	282	41,453	52,022	116,485
		5 ~ 10	435	165,844	158,271	192,179
		10 ~ 50	1,342	2,051,180	1,196,434	1,296,553
		50 ~ 100	408	1,984,482	911,558	884,715
		100 ~ 300	373	4,699,446	1,767,107	1,680,906
		300 ~ 400	74	1,935,759	663,142	739,204
		400 ~ 1000	162	7,425,694	2,401,935	2,468,939
		1000 ~ 5000	68	9,535,974	3,297,505	2,889,028
		5000억 이상	29	36,769,473	14,834,369	20,722,017
제 1차 철강산업	D271	계	1,805	42,037,894	19,157,092	25,132,337
		5억 미만	160	24,975	29,804	70,232
		5 ~ 10	264	103,168	94,507	133,457
		10 ~ 50	756	1,117,804	674,054	778,416
		50 ~ 100	209	1,023,066	473,636	498,338
		100 ~ 300	200	2,534,414	900,606	962,410
		300 ~ 400	49	1,310,743	421,336	459,146
		400 ~ 1000	102	4,546,556	1,486,613	1,509,710
		1000 ~ 5000	47	6,208,858	2,244,494	2,136,169
		5000억 이상	18	25,168,310	12,832,042	18,584,459
제철, 제강 및 합금철 제조업	D2711	계	47	21,229,727	12,102,779	15,648,650
		5억 미만	3	581	2,084	17,662
		5 ~ 10	4	1,722	2,023	449
		10 ~ 50	12	14,836	12,127	17,376
		50 ~ 100	4	13,408	21,207	5,297
		100 ~ 300	3	34,465	3,181	5,481
		300 ~ 400	0	0	0	0
		400 ~ 1000	5	223,002	64,700	67,128
		1000 ~ 5000	6	X	X	X
		5000억 이상	10	X	X	X

산업별		출하액규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
제철 및 제강업	D27111	계	29	20,908,681	11,999,689	15,500,760
		5억 미만	1	X	X	X
		5 ~ 10	3	X	X	X
		10 ~ 50	7	9,288	7,894	12,334
		50 ~ 100	1	X	X	X
		100 ~ 300	1	X	X	X
		300 ~ 400	
		400 ~ 1000	1	X	X	X
		1000 ~ 5000	5	X	X	X
		5000억 이상	10	19,773,825	11,556,410	14,898,793
합금철 제조업	D27112	계	18	321,046	103,090	147,890
		5억 미만	2	X	X	X
		5 ~ 10	1	X	X	X
		10 ~ 50	5	5,548	4,233	5,042
		50 ~ 100	3	X	X	X
		100 ~ 300	2	X	X	X
		300 ~ 400	0			
		400 ~ 1000	4	X	X	X
		1000 ~ 5000	1	X	X	X
		5000억 이상	0	0	0	0
철강압연·압출 및 연신제품	D2712	계	359	7,200,275	2,042,077	3,788,933
		5억 미만	17	1,680	3,430	4,820
		5 ~ 10	40	17,064	12,478	19,693
		10 ~ 50	141	241,951	109,788	127,063
		50 ~ 100	46	228,293	95,713	114,923
		100 ~ 300	45	577,334	229,396	243,978
		300 ~ 400	20	517,821	186,153	161,796
		400 ~ 1000	31	1,338,073	475,849	695,702
		1000 ~ 5000	16	1,867,387	701,253	581,950
		5000억 이상	3	2,410,672	228,017	1,839,008
열간 압연 및 압출제품 제조업	D27121	계	147	1,748,217	531,855	620,688
		5억 미만	8	760	1,536	1,114
		5 ~ 10	24	9,287	8,121	10,878
		10 ~ 50	51	73,811	44,274	41,727
		50 ~ 100	18	77,728	38,040	48,960
		100 ~ 300	22	305,595	112,071	109,148
		300 ~ 400	9	219,445	93,057	74,514
		400 ~ 1000	9	418,826	90,659	161,373
		1000 ~ 5000	6	642,765	144,097	172,974
		5000억 이상	0	0	0	0

산업별		출하액규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
냉간 압연 및 압출제품 제조업	D27122	계	77	2,481,761	806,258	1,300,475
		5억 미만	3	X	X	X
		5 ~ 10	5	2,237	1,070	1,596
		10 ~ 50	30	58,993	27,724	43,234
		50 ~ 100	9	44,133	25,810	26,751
		100 ~ 300	10	130,273	39,218	52,071
		300 ~ 400	6	173,563	46,866	50,446
		400 ~ 1000	6	234,432	101,474	160,430
		1000 ~ 5000	6	768,304	394,150	176,457
		5000억 이상	2	X	X	X
철강선 제조업	D27123	계	135	2,970,297	703,964	1,867,770
		5억 미만	6	X	X	X
		5 ~ 10	11	5,540	3,287	7,219
		10 ~ 50	60	109,147	37,790	42,102
		50 ~ 100	19	106,432	31,863	39,212
		100 ~ 300	13	141,466	78,107	82,759
		300 ~ 400	5	124,813	46,230	36,836
		400 ~ 1000	16	684,815	283,716	373,899
		1000 ~ 5000	4	456,318	163,006	232,519
		5000억 이상	1	X	X	X
철강관 제조업	D2713	계	695	4,656,569	2,229,242	1,975,669
		5억 미만	72	11,108	14,305	30,140
		5 ~ 10	91	35,003	33,697	28,038
		10 ~ 50	311	441,815	282,811	287,076
		50 ~ 100	89	427,081	211,489	195,913
		100 ~ 300	88	1,018,440	423,283	406,293
		300 ~ 400	8	184,043	102,746	106,600
		400 ~ 1000	27	1,117,955	425,138	425,088
		1000 ~ 5000	7	X	X	X
		5000억 이상	2	X	X	X
주철관 제조업	D27131	계	33	151,655	125,626	147,407
		5억 미만	0	0	0	0
		5 ~ 10	6	2,124	2,134	2,254
		10 ~ 50	14	19,403	18,128	14,164
		50 ~ 100	7	33,828	19,853	14,916
		100 ~ 300	5	54,135	34,163	53,895
		300 ~ 400	0	0	0	0
		400 ~ 1000	1
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		출하액규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
강관 제조업	D27132	계	662	4,504,914	2,103,616	1,828,262
		5억 미만	72	11,108	14,305	30,140
		5 ~ 10	85	32,879	31,563	25,784
		10 ~ 50	297	422,412	264,683	272,912
		50 ~ 100	82	393,253	191,636	180,997
		100 ~ 300	83	964,305	389,120	352,398
		300 ~ 400	8	184,043	102,746	106,600
		400 ~ 1000	26	1,075,790	373,790	362,910
		1000 ~ 5000	7	735,408	263,712	300,844
		5000억 이상	2
		기타 철강산업	D2719	계	704	8,951,323
5억 미만	68			11,606	9,985	17,610
5 ~ 10	129			49,379	46,309	85,277
10 ~ 50	292			419,202	269,328	346,901
50 ~ 100	70			354,284	145,227	182,205
100 ~ 300	64			904,175	244,746	306,658
300 ~ 400	21			608,879	132,437	190,750
400 ~ 1000	39			1,867,526	520,926	321,792
1000 ~ 5000	18			2,438,175	838,482	616,911
5000억 이상	3			2,298,097	575,554	1,650,981
절단가공 및 표면처리강재 생산업	D27191			계	698	8,949,882
		5억 미만	65	11,334	9,087	16,375
		5 ~ 10	127	X	X	X
		10 ~ 50	291	X	X	X
		50 ~ 100	70	354,284	145,227	182,205
		100 ~ 300	64	904,175	244,746	306,658
		300 ~ 400	21	608,879	132,437	190,750
		400 ~ 1000	39	1,867,526	520,926	321,792
		1000 ~ 5000	18	2,438,175	838,482	616,911
		5000억 이상	3	2,298,097	575,554	1,650,981
		그외 기타 철강산업	D27199	계	6	1,441
5억 미만	3			272	898	1,235
5 ~ 10	2			X	X	X
10 ~ 50	1			X	X	X
50 ~ 100	0			0	0	0
100 ~ 300	0			0	0	0
300 ~ 400	0			0	0	0
400 ~ 1000	0			0	0	0
1000 ~ 5000	0			0	0	0
5000억 이상	0			0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		출하액규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
제 1차 비철금속산업	D272	계	748	19,970,901	4,751,212	4,649,814
		5억 미만	52	7,762	9,003	22,554
		5 ~ 10	77	26,353	29,942	30,565
		10 ~ 50	295	524,495	238,197	257,731
		50 ~ 100	108	562,312	190,378	182,298
		100 ~ 300	118	1,625,800	494,329	387,915
		300 ~ 400	20	X	X	X
		400 ~ 1000	47	2,343,599	619,631	682,509
		1000 ~ 5000	21	3,327,116	1,053,011	752,859
		5000억 이상	10	X	X	X
비철금속 제련, 정련 및 합금	D2721	계	217	9,486,869	1,979,394	1,478,966
		5억 미만	11	X	X	X
		5 ~ 10	18	4,691	7,269	6,053
		10 ~ 50	79	145,332	69,927	56,863
		50 ~ 100	39	200,461	66,327	69,672
		100 ~ 300	38	527,333	143,906	81,468
		300 ~ 400	8	X	X	X
		400 ~ 1000	13	742,680	155,259	104,775
		1000 ~ 5000	6	800,341	187,003	98,627
		5000억 이상	5	X	X	X
동 제련, 정련 및 합금 제조업	D27211	계	22	4,466,477	842,382	477,180
		5억 미만	2	X	X	X
		5 ~ 10	1	X	X	X
		10 ~ 50	7	12,161	5,037	3,001
		50 ~ 100	4	17,461	7,467	12,077
		100 ~ 300	1	X	X	X
		300 ~ 400	1	X	X	X
		400 ~ 1000	2	X	X	X
		1000 ~ 5000	2	X	X	X
		5000억 이상	2	X	X	X
알루미늄 제련, 정련 및 합금	D27212	계	128	1,804,046	403,992	319,313
		5억 미만	3	X	X	X
		5 ~ 10	8	2,826	3,152	3,909
		10 ~ 50	43	80,075	38,038	36,764
		50 ~ 100	27	142,134	45,431	47,591
		100 ~ 300	31	435,786	112,445	68,021
		300 ~ 400	5	108,964	54,476	27,319
		400 ~ 1000	7	481,753	65,554	59,877
		1000 ~ 5000	4	X	X	X
		5000억 이상	0	0	0	0

산업별		출하액규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
연 및 아연제련, 정련 및 합금	D27213	계	32	2,492,086	660,536	626,071
		5억 미만	2	X	X	X
		5 ~ 10	5	X	X	X
		10 ~ 50	13	23,434	9,053	5,065
		50 ~ 100	4	20,112	6,342	5,899
		100 ~ 300	4	X	X	X
		300 ~ 400	0	0	0	0
		400 ~ 1000	2	X	X	X
		1000 ~ 5000	0	0	0	0
		5000억 이상	2	X	X	X
기타 비철금속제련,정련 및 합금	D27219	계	35	724,260	72,484	56,402
		5억 미만	4	551	1,186	2,835
		5 ~ 10	4	1,052	923	1,220
		10 ~ 50	16	29,662	17,799	12,033
		50 ~ 100	4	20,754	7,087	4,105
		100 ~ 300	2	X	X	X
		300 ~ 400	2	X	X	X
		400 ~ 1000	2	X	X	X
		1000 ~ 5000	0	0	0	0
		5000억 이상	1	X	X	X
비철금속압연,압출 및 연신제품	D27222	계	467	10,103,143	2,580,842	3,012,573
		5억 미만	38	5,931	6,585	7,655
		5 ~ 10	45	16,868	16,907	18,396
		10 ~ 50	190	341,613	139,766	164,202
		50 ~ 100	64	340,311	116,292	106,530
		100 ~ 300	69	962,573	295,359	276,992
		300 ~ 400	11	X	X	X
		400 ~ 1000	30	1,447,593	377,442	505,068
		1000 ~ 5000	15	2,526,775	866,008	654,232
		5000억 이상	5	X	X	X
동 압연, 압출 및 연신제품 제조업	D27221	계	182	5,963,144	1,524,327	1,499,879
		5억 미만	16	2,136	2,583	2,352
		5 ~ 10	10	3,271	4,003	4,621
		10 ~ 50	67	122,343	58,885	68,271
		50 ~ 100	25	138,943	39,623	27,113
		100 ~ 300	33	450,473	117,294	90,210
		300 ~ 400	3	X	X	X
		400 ~ 1000	17	789,299	214,533	305,580
		1000 ~ 5000	8	X	X	X
		5000억 이상	3	X	X	X

산업별		출하액규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
알루미늄 압연, 압출 및 연신제품	D27222	계	240	3,565,817	942,702	1,426,851
		5억 미만	16	2,681	2,622	4,217
		5 ~ 10	29	11,279	10,618	11,323
		10 ~ 50	105	195,916	67,491	85,371
		50 ~ 100	35	178,573	68,197	68,184
		100 ~ 300	28	378,785	143,123	169,747
		300 ~ 400	7	X	X	X
		400 ~ 1000	13	658,294	162,909	199,488
		1000 ~ 5000	5	639,167	183,662	253,150
		5000억 이상	2	X	X	X
기타 비철금속 압연, 압출 및 연신제품	D27229	계	45	574,182	113,813	85,843
		5억 미만	6	1,114	1,380	1,086
		5 ~ 10	6	2,318	2,286	2,452
		10 ~ 50	18	23,354	13,390	10,560
		50 ~ 100	4	22,795	8,472	11,233
		100 ~ 300	8	133,315	34,942	17,035
		300 ~ 400	1	X	X	X
		400 ~ 1000	0	0	0	0
		1000 ~ 5000	2	X	X	X
		5000억 이상	0	0	0	0
기타 제 1차 비철금속산업	D2729	계	64	380,889	190,976	158,275
		5억 미만	3	X	X	X
		5 ~ 10	14	4,794	5,766	6,116
		10 ~ 50	26	37,550	28,504	36,666
		50 ~ 100	5	21,540	7,759	6,096
		100 ~ 300	11	135,894	55,064	29,455
		300 ~ 400	1	X	X	X
		400 ~ 1000	4	153,326	86,930	72,666
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0
기타 제 1차 비철금속산업	D27290	계	64	380,889	190,976	158,275
		5억 미만	3	X	X	X
		5 ~ 10	14	4,794	5,766	6,116
		10 ~ 50	26	37,550	28,504	36,666
		50 ~ 100	5	21,540	7,759	6,096
		100 ~ 300	11	135,894	55,064	29,455
		300 ~ 400	1	X	X	X
		400 ~ 1000	4	153,326	86,930	72,666
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0

산업별		출하액규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
금속 주조업	D273	계	620	2,600,510	1,374,039	1,207,875
		5억 미만	70	8,716	13,215	23,699
		5 ~ 10	94	36,323	33,822	28,157
		10 ~ 50	291	408,881	284,183	260,406
		50 ~ 100	91	399,104	247,544	204,079
		100 ~ 300	55	539,232	372,172	330,581
		300 ~ 400	5	X	X	X
		400 ~ 1000	13	535,539	295,691	276,720
		1000 ~ 5000	0	0	0	0
		5000억 이상	1
철강 주조업	D2731	계	392	1,355,021	961,821	831,548
		5억 미만	37	4,663	7,299	15,592
		5 ~ 10	50	18,377	18,920	9,609
		10 ~ 50	187	255,633	200,825	173,274
		50 ~ 100	66	281,601	187,401	140,294
		100 ~ 300	40	399,852	288,883	248,598
		300 ~ 400	4	79,899	58,964	49,128
		400 ~ 1000	8	314,996	199,529	195,053
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0
선철주물 주조업	D27311	계	217	873,378	595,152	556,097
		5억 미만	21	1,917	4,984	11,652
		5 ~ 10	25	9,301	9,735	4,667
		10 ~ 50	99	134,199	108,321	97,212
		50 ~ 100	39	152,698	108,451	80,974
		100 ~ 300	23	225,565	157,622	138,437
		300 ~ 400	3	62,241	41,381	41,194
		400 ~ 1000	7	287,457	164,658	181,961
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0
강주물 주조업	D27312	계	175	481,643	366,669	275,451
		5억 미만	16	2,746	2,315	3,940
		5 ~ 10	25	9,076	9,185	4,942
		10 ~ 50	88	121,434	92,504	76,062
		50 ~ 100	27	128,903	78,950	59,320
		100 ~ 300	17	174,287	131,261	110,161
		300 ~ 400	1
		400 ~ 1000	1
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

산업별		출하액규모별	2006			
			사업체수	주요생산비	부가가치	유형자산
비철금속 주조업	D2732	계	228	1,245,489	412,218	376,327
		5억 미만	33	4,053	5,916	8,107
		5 ~ 10	44	17,946	14,902	18,548
		10 ~ 50	104	153,248	83,358	87,132
		50 ~ 100	25	117,503	60,143	63,785
		100 ~ 300	15	139,380	83,289	81,983
		300 ~ 400	1
		400 ~ 1000	5	220,543	96,162	81,667
		1000 ~ 5000	0	0	0	0
		5000억 이상	1
		알루미늄 주물 주조업	D27321	계	138	483,570
5억 미만	15			2,090	3,019	5,952
5 ~ 10	32			12,875	10,812	15,444
10 ~ 50	59			87,976	48,712	51,588
50 ~ 100	16			77,851	38,802	45,563
100 ~ 300	11			103,671	56,817	38,825
300 ~ 400	1		
400 ~ 1000	4			175,838	52,375	59,981
1000 ~ 5000	0			0	0	0
5000억 이상	0			0	0	0
동주물 주조업	D27322			계	39	121,492
		5억 미만	6	821	859	935
		5 ~ 10	5	2,199	1,536	747
		10 ~ 50	19	29,078	13,160	9,393
		50 ~ 100	6	27,030	13,913	6,427
		100 ~ 300	2
		300 ~ 400	0	0	0	0
		400 ~ 1000	1
		1000 ~ 5000	0	0	0	0
		5000억 이상	0	0	0	0
		기타 비철금속 주조업	D27329	계	51	640,427
5억 미만	12			1,142	2,038	1,220
5 ~ 10	7			2,872	2,554	2,357
10 ~ 50	26			36,194	21,486	26,151
50 ~ 100	3			12,622	7,428	11,795
100 ~ 300	2		
300 ~ 400	0			0	0	0
400 ~ 1000	0			0	0	0
1000 ~ 5000	0			0	0	0
5000억 이상	1		

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!

보조셀 감추기
경로 없음!
산업분류
통합 필요!