

GIS 기반 통계 공표구역 획정에 대한 해외사례

통계지리정보과

박진순



목 차

1. 배경 및 목적
2. 영국의 집계구 획정
3. 해외사례 비교
4. 해외사례 시사점

1. 집계구 획정의 배경 및 목적

1-1 배경

- 현재 최소 공표단위로 행정구역 사용
 - 동읍면에 적게는 수천에서 많게는 4만명의 인구 거주
 - 동질적 지역개념으로 보기 어려움
- 보다 세분화된 지역의 통계 수요 증가
 - 지역계획 및 도시계획 등 수립
 - 근린지역 정보, 기업 및 자영업자 등 상권 분석, 판매전략 수립
- 통계구는 통계 집계단위로 고려 미흡
 - 개인정보 보호
 - 통계적 유의성 확보

1-1 배경

통계구 (2002년)

- ◇ 읍면동 내에서 기초단위구를 묶어 통계구 설정
 - 유지, 관리를 고려 전국 10만개 내외(91,064개)로 구획
 - 지형지물 및 지역특성을 고려
 - * 실제 기초단위구별 특성을 충분히 배려하지 않음

- ◇ 가구수 기준으로 구획하여 통계 집계구로서의 고려 미흡
 - 특광역시 200-400 가구
 - 기타시지역 150-300 가구
 - 읍지역 100-200 가구
 - 면지역 10-150 가구

1-2 목 적

- 기초단위구 기반으로 통계를 공표할 수 있는 최소 집계단위 마련
 - 개인정보 보호 및 통계적 유의성을 확보한 소지역통계 생산
 - GIS 분야에서 다양한 통계자료의 제공 및 분석 가능
 - 통계의 품질 제고 및 향후 통계 발전에 기여

현재 진행 상황

2006년 국가 GIS 일반연구분야 지원사업으로 선정

- “집계구”에 대한 연구용역 수행 : 서울시정개발연구원
- 집계구 획정 기준 설정, 획정을 위한 알고리즘 개발
- 대내외적 객관성 확보
- 2007년 3월 완료 예정

2. 영국의 통계 공표구역

1. 통계구역 구분체계

- 잉글랜드와 웨일즈는 동일 체계, 스코틀랜드와 북아일랜드는 독립적 체계
- 법적, 행정경계(ward, parish/community)를 유지하면서 그 밑에 Enumeration District(ED) 설정
- 2001년 센서스부터 공표구역으로 Output Area(OA) 별도 설정
- OA는 행정경계를 따르므로 손쉽게 더 큰 지역단위로 연계 가능
- ED를 통해 수집된 센서스 자료는 개별 주소단위로 입력되며 OA를 기본단위로 집계 공표됨
- 2001년 센서스의 경우 최소 인구 규모 설정 : 100명, 40가구
 - * 표준테이블 자료(교차분석 자료)는 1,000명, 400가구

2 영국의 통계 공표구역

2. 조사구역(ED)

- 1961년-1991년 센서스까지 조사 및 공표구역으로 사용
 - * 2001년부터 조사단위로만 사용
- 잉글랜드/웨일즈 지역에 116,919개 설정, 4,840개 특수 ED 설정
- 조사원의 업무량 균등 및 조사의 효율성 고려
- 주요 간선도로, 강, 철도 등을 경계로 설정
- 1991년부터 GIS기술 발달로 디지털화, 2001년 센서스부터 디지털 구획
- 공표구역 목적이 아니었으므로 문제점 내포
 - * 형상과 인구 규모가 다양하게 나타남
 - * 사회적 동질성이 고려되지 않음
 - * 경계 및 업무량 변화에 따른 ED 경계 변화로 시계열 유지 곤란

2 영국의 통계 공표구역

3. 공표구역(Output Area)

□ 도입목적 및 특성

- 최저 인구 기준 미달 문제의 해결
- 우편번호 지역과 센서스 지역의 연동이 가능하도록 하기 위함
- 공표 구역의 인구, 사회적 동질성, 형상 표준화를 위한 목적
- 소지역 수준에서 중앙 및 지방정부가 수행할 경제·사회적 데이터 제공

* 정책효과에 대한 소지역 수준에서의 모니터링 수행

- 2001년 센서스 결과 OA는 175,434개, 평균인구 297명, 평균가구수는 124호
- 번호체계는 ward내에서 고유번호 4자리를 순서대로 부여하고 county, district를 포함하여 8자리를 부여

* 00 AA BB 0001
 COUNTY DISTRICT WARD OA일련번호

2 영국의 통계 공표구역

□ OA 경계 설정

➤ OA 기본단위 및 설정기준

○ 기본단위로 우편번호 구역 사용

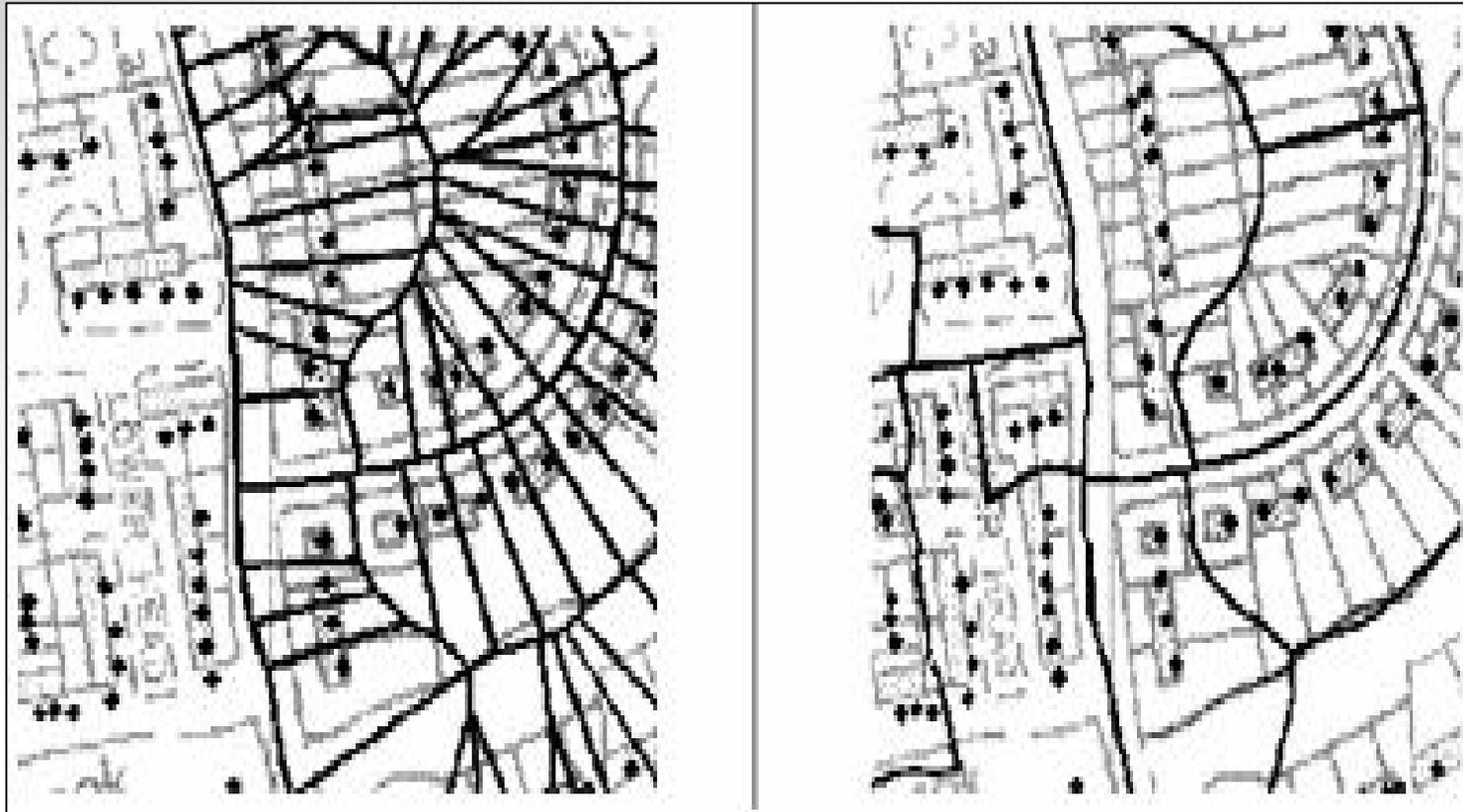
- 우편번호 구역이 이전부터 지역 구분, 통계, 서비스 제공 등에 널리 활용됨
- 센서스 자료, 다른 사회통계 자료와 통합 가능

○ 우편번호구역 설정 방법

- ED를 통해 수집된 통계 자료를 주소에 따라 입력
- 각 주소위치를 이용하여 티센 폴리곤을 형성하고 인접한 티센 폴리곤 중 같은 번호끼리 통합하여 우편번호 구역 경계를 설정
- 형성된 티센 폴리곤을 간선도로, 수로 등과 같은 지형 데이터 및 ward/parish 경계 레이어와 중첩
- ward/parish 경계에 걸쳐 있는 구역은 분할하여 ward/parish 경계를 준수

2 영국의 통계 공표구역

우편번호 구역 설정



2 영국의 통계 공표구역

➤ Output Area 구획 기준

○ 인구 규모

- 최소 100명 : 공표과정에서 개인정보 보호
- 인구규모의 표준화 : 최적 125가구(약 312명)

○ 형상조건

- 가급적 단순하고 균일한 형태가 되는 것을 우선으로 지역 구분과 활용의 편의성 추구
- 인접성

○ 사회적 동질성 ; 내부의 동질성 최적화

- 주택소유형태 : 자가소유, 공공임대, 사적임대
- 주택형태 : 단독주택, 공동주택 등

○ ward/parish의 행정구역 경계 준수

2. 영국의 통계 공표구역

➤ Output Area 경계 설정 방법

○ Automated Zoning Procedure(AZP)

- 우편번호구역을 컴퓨터를 통해 반복 결합하는 방법, 최적화된 조합 찾음
- ward 단위로 설정작업 수행, 3단계로 수행

초기화 단계

임의 zone 선택

1	2	3	4
6	7	6	5
9	10	11	12
16	15	14	13

인구규모 만족 grouping

1	2	3	4
6	7	6	5
9	10	11	12
16	15	14	13

임의 zone 선택

1	2	3	4
6	7	6	5
9	10	11	12
16	15	14	13

인구규모 만족 grouping

1	2	3	4
6	7	6	5
9	10	11	12
16	15	14	13

흰색, 회색그룹 규모미만

1	2	3	4
6	7	6	5
9	10	11	12
16	15	14	13

인접한 빨간그룹에 포함 모든 그룹이 인구규모 만족

1	2	3	4
6	7	6	5
9	10	11	12
16	15	14	13

2 영국의 통계 공표구역

최적화 단계

청색그룹에 인접한 zone ‘?’ 표시

1	2	3	4 ?
6	7	6 ?	5
9	10	11 ?	12
16	15	14 ?	13

Zone 4 청색그룹으로 이동

1	2	3	4 ?
6	7	6 ?	5
9	10	11 ?	12
16	15	14 ?	13

노란그룹 인구규모 미달로 거부, 원위치

1	2	3	4
6	7	6 ?	5
9	10	11 ?	12
16	15	14 ?	13

Zone 6,11,14 이동 반복할 때
6은 동질성(임대, 자가) 나빠짐,
11은 녹색그룹의 인접성 파괴,
14는 너무 큰 빨간그룹 작아지고
파란그룹은 커짐

녹색그룹에 인접한 zone에
‘?’ 표시, 인접성 or 인구규모
의 개선과 거부된 모든 작업
과정 저장

더 이상 개선의 여지가
없을 때까지 수많은 시
킷을 반복하여 최적화

1	2	3	4
6	7	6	5
9	10	11	12
16	15	14	13

1	2	3 ?	4
6	7 ?	6	5 ?
9 ?	10	11	12 ?
16	15 ?	14 ?	13

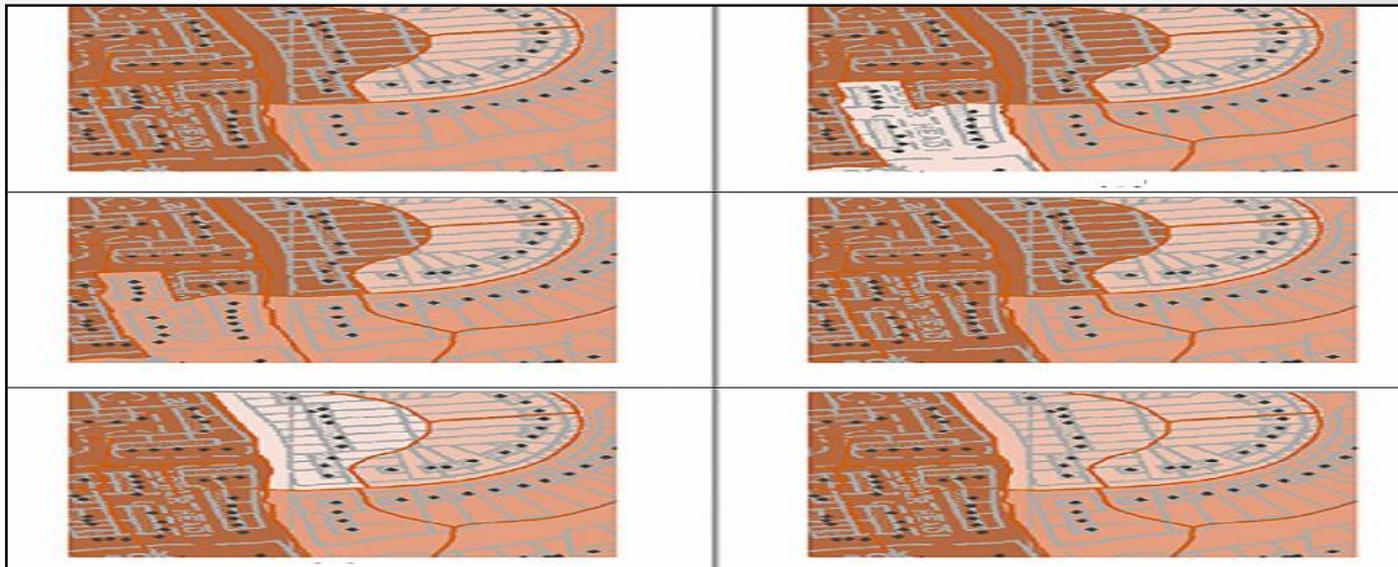
1	2	3	4
6	7	6	5
9	10	11	12
16	15	14	13

2. 영국의 통계 공표구역

반복 단계

- 초기화 과정의 첫번째 섹션에서 시작점으로 임의의 다른 우편번호구역으로 설정하여 앞에서 소개한 모든 과정을 동일하게 수행한다.
- 이 과정을 총 100회 실행하여 그 중 최적성을 판단하여 OA의 경계로 확정한다.
 - * 최적성 판단 : AZP프로그램의 계산
 - * 기준 : 인구규모, 내부의 동질성, 형상의 단순성

예) OA 경계 획정



2. 영국의 통계 공표구역

○ Automated Zoning Procedure(AZP) 점수 계산

- 최적의 조합은 OA 설정기준인 인구 규모, 형상, 사회적 동질성 세가지 조건을 점수화하여 이 지수가 가장 낮은 조합을 최적의 OA로 판단한다.
- 각 기준의 지수는 현재 조합의 각 부문별 점수를 이전까지의 각 기준별 최고 점수로 나눈값에 로그를 취해 계산한다.
- 세가지 지수의 가중평균을 구한 조합지수를 통해 최적의 조합을 판단하며, 0보다 적으면 이전 조합들에 비해 더 최적화된 조합이 된다

	인구점수 (작을수록 좋음)	형상점수 (작을수록 좋음)	동질성 점수 (클수록 좋음)
이전 반복 최고점수	8,000	20백만	0.3
현재 반복 점수	2,000	25백만	0.4
비율	0.25	1.25	0.75 (빈대로 계산)

2. 영국의 통계 공표구역

□ Super Out Areas(SOA)

▶ 도입 배경

○ 공간 규모면에서 OA와 지방정부의 중간 정도에 해당하는 크기

- SOA 도입 이전까지는 근린통계의 공간단위로 선거구 사용
- 선거구의 경우 잦은 경계 변동으로 시계열 분석이 어려움
- 인구가 적은 지역의 통계자료 미공표

○ 일관성 있고 영구적인 공간 단위의 도입 요구됨

○ 근린통계를 위한 안정적이고 통일된 단위로서, OA를 기초로 만든 공간단위

○ Lower layer, Middle layer, Upper layer 3개 층으로 이루어 짐

- Upper layer는 2004-2006년 테스트 기간을 거쳐 통해 타당성, 경계 조정 등 검토 후 최종 확정할 예정

2. 영국의 통계 공표구역

➤ Lower layer

- 최소인구 1,000명, 평균인구 1,500명으로 4-6개 정도의 OA를 묶어 설정
- 2004년 2월 설정, 34,373개 구축
- ward 경계를 준수하며 컴퓨터를 통해 자동 설정

➤ Middle layer

- 최소인구 5,000명, 평균인구 7,200명으로 몇 개 Lower layer 를 묶어 설정
- 2004년 8월 설정, 7,193개 구축
- 2003년 기준 지방정부 경계를 준수하며 지방정부 및 기관의 검토를 거쳐 확정

➤ Upper layer

- 최소인구 25,000명 선으로 계획되고 있으며 세부사항 검토 중

2. 영국의 통계 공표구역

□ LowerLayer OA의 유지보수

- 우편번호에 의한 우편포인트 분석/추정, 2002년은 광범위한 우편번호 개정
 - 가구수 0인 OA가 적고 2006년 400가구미만은 0.25%, 1200가구이상은 0.69%
 - 이 추세가 계속되면 2011년까지 대략 2.5% 미만이 유지보수 대상임, 보증은 없음
 - 해마다 1-2% 우편번호 변경
- 안정성이 필요하므로 유지보수 비율을 전체의 5%미만으로 유지
 - 각 층에서 합리적으로 정해진 규모 및 체계를 유지

년	0명 거주	400가구미만	600가구이상	1,200가구이상
2001	0	116(0.0343%)	3,556(10.344%)	49(0.143%)
2002	22	197(0.573%)	3,639(10.585%)	53(0.154%)
2003	0	103(0.300%)	3,965(11.534%)	74(0.215%)
2004	0	94(0.273%)	4,329(12.592%)	112(0.326%)
2005	0	96(0.279%)	4,672(13.590%)	170(0.495%)
2006	0	86(0.250%)	5,135(14.937%)	239(0.695%)

3. 해외사례 비교

3. 집계구의 구분과 특징

	미국	영국	캐나다	호주	일본
센서스 기간	10년	10년	5년	5년	5년
최소 공표구역 (도입년도)	Block (1940)	Output Area (2001)	Dissemination Area (2001)	Census Collection District (2011 부터 mesh block)	Grid Square (1973)
공표구역 설정기준	-최소면적 0.69acre -최소폭 70feet -한면 이상 도로에 인접	-인구:최소 100명, 최적 312명 -형상: 중심간 거리제곱을 최소화 - 내부동질성 -타 공간단위 경계	-인구: 최적 500명 -형상: 정방형 우선 - 도로근접성 -타 공간단위 경계 - 가시적 경계	-조사원 1인의 10일간 업무 량 -SLA 및 타지 구분 경계 -가시적 지형 지물	-국세조사: 행정구역 체계 -Grid Square : 위도40분, 경도1도에 해당부분을 80*80 균등 분할
공표구역 인구규모	없음	평균 297명	400-700명	도시지역평균 220명	없음

3. 집계구의 구분과 특징

	미국	영국	캐나다	호주	일본
상위 계층 개인보호 규정	표본조사는 블록 그룹 (600-3000명) 단위 일부 교차분석 자료는 센서스트랙 (1500-8000명) 단위	교차분석 자료는 1000명 이상	최소 인구규모 기준 400명 이상	최소 인구규모 기준 100명 이상	표본조사는 시구정촌 단위로 공표

참고) 조사구역

조사구역 인구규모	제한없음 0명 포함	지역별 다양 50명 이하	농촌 125명 도시 650명	최소 100명 원주민 80명	평균 50명
최소 조사구역	Block	Enumeration District	Enumeration Area	Census Collection District	Enumeration District

4. 해외사례 시사점

- 조사구와 공표구의 분리 강화
 - GIS 기술의 발달을 바탕으로 행정구역외 새로운 지역 구분 체계 도입
 - 조사구는 통계조사의 효율성 극대화 목적으로 이용
 - 공표구는 개인정보 보호, 통계적 유의성 확보, 수요자 중심의 소지역 통계 생산 목적으로 이용
- 설정 기준으로 인구 규모가 가장 중요한 조건으로 작용하고 있으며
형상 조건, 내부의 동질성, 도로의 접근성이 추가적으로 반영
- 행정구역 체계의 경계를 반드시 따름
 - 공표 데이터가 기존 행정구역 체계와 쉽게 연동 가능
 - 데이터의 다양한 활용 가능
- 자동으로 집계구를 획정할 수 있는 알고리즘 개발
 - 인구 규모, 형상, 동질성 등 최적안 선택

4. 해외사례 시사점

➤ 개인정보 보호를 위한 제도적 장치 마련

- 통계자료 유형에 따라 공표구역 단위를 다르게 적용

- 최소 인구 기준을 설정하여 일정 규모 이상 구역만 공표

* 인구에 대해 외부에서는 작은 규모를 원하지만 개인의 프라이버시 문제로 내부적으로는 규모가 커질 수밖에 없음

➤ 통계 자료의 폭넓은 이용을 위해 데이터 유형에 따라 공표구역을 다르게 적용

- 개인정보 보호와 통계적 유의성이 확보되는 데이터는 최소단위까지 제공

- 공표구역과 행정구역 사이에 최소 집계구보다 큰 상위 집계구 획정

➤ 통계이외의 통계 목적외 선거구 구획, 도시지역 구분 등에 이용

4. 해외사례 시사점

➤ MESH, GRID 개념 도입

- 행정경계 변동, 지형지물 변화, 도시화지역 확대 등 변화에 융통성 발휘
- 개인정보, 항목, 재구획 등 많은 장애물에 의해 당분간 기초단위구가 체계가 유지 되겠지만 시계열 유지를 위해 필요

➤ 이용자가 원하는 다양한 통계 제공을 위한 통계생산부서의 적극성 필요

- 대외적 객관성을 확보한 집계구 기준 마련
- 통계의 정확성과 시의성 동시 만족 어려움
- 가능한 많은 항목을 공개하여 이용자의 만족도 제고