

2008년 통계자료에디팅(SDE)
국제회의 출장결과 보고

2008. 5.

통계개발원

2008년 통계자료에디팅(SDE) 국제회의 출장결과 보고

I. 출장개요

가. 출장목적

- UNECE가 주관하여 개최하는 “통계 자료 에디팅 국제 회의 (UNECE Work Session on Statistical Data Editing)”에 참석하여 통계자료 에디팅에 대한 최근 연구동향을 파악하고 연구과제를 수행함에 있어 질적 향상을 도모하고자 함

나. 출장자

- 연구기획실 이의규 사무관, 심규호 주무관

다. 출장기간

- 2008. 4. 20. ~ 4. 25. (5박 6일간, 기내 1박)

라. 출장지

- 오스트리아 비엔나 (오스트리아 통계청)

II. 회의개요

가. 회의명

- 통계자료 에디팅에 관한 UNECE 국제회의

나. 회의목적

- 각국 통계청에서 생산하는 통계자료의 품질 향상
- 실제 자료를 활용한 에디팅 사례 및 기법 소개
- 데이터 에디팅과 관련한 협력, 경험 교환 및 새로운 기법 개발

다. 회의일시 및 장소

- 2008. 4. 21. ~ 4. 23. (3일간)
- 오스트리아 통계청

라. 참가자 : 69명

- 22개국 63명
- UNECE 등 국제기구 6명

마. 주요의제

- 전자 자료 수집을 통하여 획득된 자료의 에디팅
- 행정자료와 통합자료의 에디팅
- 데이터 에디팅을 통한 품질향상
- 새로운 에디팅 기법
- 결과에 근거한 에디팅
- 센서스

Ⅲ. 주제별 발표내용

가. 전자 자료 수집을 통하여 획득된 자료의 에디팅 (Editing of data acquired through electronic data collection)

1) 개요

- 전자 자료 수집은 수집 시점에서 자료의 에디팅을 가능하게 함
- 웹 조사(web survey)에서의 에디팅에 대한 일반적인 지침
- 전자 자료수집과 다른 수집방식 간의 에디팅 또는 에디팅 결과의 비교

2) 세부 발표주제

- 사업체 구조 통계를 위한 혼합방식 설문지의 선별적 자동 에디팅
- 인터넷 자료수집 시스템에서의 자료 에디팅
- 2011년 영국과 웨일스 센서스에서 혼합방식수집의 에디팅에 대

한 초고

- 2006년 캐나다 센서스에서 온라인 편집(edit)의 영향
- 새로운 자료수집방법을 도입할 때 변화를 보이는 전통적인 편집 시스템의 사용
- EDR(Electronic Data Reporting) 방법을 통한 에디팅 업무를 개선하려는 노력

3) 주요발표 내용

- 사업체구조통계를 위한 혼합방식 설문지 선별적 자동 에디팅
 - 네덜란드 통계청은 응답부담을 줄이기 위해 2006년부터 사업체구조통계(SBS)에 대한 전자 설문을 재설계하여 사용 (기업의 일부는 여전히 종이설문으로 받음)
 - 이러한 변경은 선별적 자동에디팅 절차의 개선을 야기함
 - 결측치에 대한 처리를 변경하였으며 자동 에디팅 프로그램 SLICE의 최신 버전을 이행
 - 논문에서는 이러한 변경의 의미를 제시하고 SLICE의 성능을 검토하며 혼합방식의 효과를 조사함

나. 행정자료와 통합소스의 에디팅 (Editing administrative data and combined sources)

1) 개요

- 통계작성기관은 통계 품질을 향상시키고 비용과 응답자 부담을 줄이기 위해 점점 행정자료에 의존
- 그러나 행정자료는 통계자료로 사용을 위해 설계된 것이 아니어서 광범위한 처리와 에디팅을 요구
- 여러 개의 데이터 소스를 합치는 것은 통계 데이터 편집의 새로운 도전

2) 세부 발표주제

- 행정자료에 근거한 노동비용에 관한 이탈리아 단기조사에서의

편집과정

- 사업체 통계 생산을 위해 사용된 행정자료의 E&I(Editing and Imputation)
- 사업체 구조통계에 대한 소스의 통합에 있어서 E&I의 역할
- ISEE(Integrated System for editing and Estimation)에서의 예측과 응답대체: 통합된 자료소스의 효율적 사용을 위한 틀
- 행정자료의 품질 - 사업체 등록 통계의 관리를 위한 도전
- 프랑스 사업체 구조통계의 향후 시스템: 추정치의 역할
- 이탈리아의 EU-SILC(European Survey on Income and Living Conditions) 경험에서 조사와 행정자료의 통합
- VAT(Value Added Tax) 자료에 대한 에디팅 전략

3) 주요발표 내용

- 행정자료에 근거한 노동비용에 관한 이탈리아 단기 조사에서의 편집과정
 - 임금과 노동비용에 관한 이탈리아의 분기조사는 단기 통계를 생산하기 위해 조사자료와 결합되는 행정소스의 사용사례
 - 통계변수로 쓰여 지기 전에 행정 개념과 정의에서의 빈번한 변경을 따라잡기 위해 행정자료에 대한 예비 검토 과정 필요
 - 행정자료가 요구된 통계변수로 옮겨진 후 가능한 예외적인 값을 찾아내기 위해 하나 이상의 전통적인 검토 절차가 필요
 - 단위 무응답에 대한 응답대체 과정 필요
 - 시계열 분석 및 다른 통계 소스와의 수준비교를 통해 매크로 자료 검증 과정을 논함

다. 데이터 에디팅을 통한 품질개선 (Improvement of quality through data editing)

1) 개요

- 데이터 에디팅의 질은 자료에서 오류를 효율적으로 찾아내는 능력뿐만 아니라 데이터의 품질과 조사과정에 대한 정보를 제공하

는 것과도 관련됨

- 조사 데이터의 품질과 조사과정의 향상을 위한 자료 에디팅의 사용과 관련한 모든 문제를 다룸

2) 세부 발표주제

- 자료품질의 향상을 위한 수정의 유용성
- E&I 수용과 뉴질랜드 통계청에서의 최상의 업무처리
- 횡단면 사업체조사에서의 E&I
- 스위스 연방 통계청 내의 EDIMBUS-RPM(Editing and Imputation in Business Survey-Recommendation Practices Manual)의 이행
- 핀란드 통계청의 에디팅 프로젝트
- 규모가 큰 사업체에 대한 자동 에디팅 절차의 향상된 접근법

3) 주요발표 내용

- 자료품질의 향상을 위한 수정의 유용성
 - 독일의 도매업조사는 12,000 기업에 자동 데이터 전송, 인터넷 설문, 종이 설문으로 수집됨
 - 최적의 설문은 참여를 용이하게 하고 동시에 자료의 품질을 올려야 함
 - 수정은 쉽게 기록될 수 있고 이는 설문지나 에디팅 과정을 향상시키는 데 유용함
 - 수정은 측정 오류를 탐지하는 데 유용할 수 있음. 양(+)^의 수정은 과소 포함을, 음(-)^의 수정은 과대 포함을 나타냄
 - 수정은 정확성의 개선을 나타내고 조사의 비용 배분의 기초로 사용될 수 있음

라. 새로운 방법 (New and emerging methods)

1) 개요

- 데이터 에디팅과 임퓨테이션에 대한 새로운 방법과 기술을 다룸
- 데이터 에디팅 이론과 최신 방법론, 소프트웨어 툴의 소개

2) 세부 발표주제

- 연속형 다변량자료에서의 이상치와 영향력이 큰 오류의 검출을 위한 모형
- 선형 편집 제약 하에서 수치자료의 보정된 응답대체
- 사업체조사에서 편집규칙 변화의 영향을 평가하는 일반적인 틀
- 선형제약 하에서 다항 응답대체에 대한 다양한 접근방법의 평가
- 다항 임퓨테이션에 의한 예비 실업률의 추정
- 통계자료의 자동 에디팅: DIA 소프트웨어의 새로운 버전
- 통계자료의 에디팅과 임퓨테이션

3) 주요발표 내용

- 연속형 다변량자료에서의 이상치와 영향력이 큰 오류의 탐지를 위한 모형
 - 모든 에러를 식별하기 보다는 영향력 있는 에러에 집중해야 함
 - 에디팅으로 인한 비용절감 요구는 영향력이 있는 관측값과 이상치의 식별과 처리의 문제로 확대되어야 함
 - 대부분 이상치의 탐색은 일변량이나 이변량 방법을 통해 수행되어지곤 하나 경제자료의 이상치는 일반적으로 다변량 속성을 가짐
 - 다변량 이상치 탐색법은 일반적으로 자료의 중심과 관측값의 거리의 개념을 사용함
 - 에러 확률과 크기의 추정을 허락하는 연속형 변수의 다변량 에러 모형을 제안함

마. 결과에 근거한 에디팅 (Editing based on Results)

1) 개요

- 제공후 자료의 에디팅에 관련된 문제를 다룸
- 결과에 근거한 에디팅, 선별적 에디팅, 그래픽 자료 분석 등과 관련된 주제

- 매크로 에디팅의 개발, 적용, 평가 등

2) 세부 발표주제

- 하나의 매크로 에디팅 접근방법의 개발
- 선별적 에디팅에서의 국소, 전체 스코어 함수
- 스웨덴 통계청에서의 사례연구
- 통계 데이터 에디팅의 새로운 틀
- 마이크로데이터의 포스트 에디팅(post-editing)
- 2009년 농업 총조사 편집(edit)을 위해 행정자료의 사용

3) 주요발표 내용

- 하나의 매크로 에디팅 접근방법의 개발
 - 미국 에너지 정보국(EIA)은 석유시장 자료를 수집하고 제공
 - 논문은 조사 자료의 매크로 에디팅의 새로운 접근방법을 개발하고 있는 EIA의 현 업무를 소개
 - 경제시계열 또는 ARMA 모델을 사용하여 다음 달 석유제품 수요의 예측값 또는 구간 예측값을 제공함
 - 모형으로부터 유도된 구간 예측은 예비조사 추정값과 비교하여 잠재적인 이상치를 탐색함

바. 센서스 (Censuses)

1) 개요

- 대부분의 나라에서 인구 센서스는 중요한 정치적 경제적인 의사 결정을 위한 기초를 제공하기 위해 주기적으로 수행
- 센서스 자료에 적용되는 데이터 정제과정은 최종 결과물의 충분한 품질을 보증하기 위해 필수적임
- 최근 전통적인 센서스의 수행에 대한 저항이 있으며 몇몇 나라는 순전히 등록기반의 센서스로 전환할 요구에 직면
- 최근 센서스뿐만 아니라 가까운 장래에 계획된 센서스에 대한 에디팅과 임퓨테이션 주제를 다룸

2) 세부 발표주제

- 오스트리아의 등록기반 센서스의 응답대체와 추정
- 캐나다 센서스 E&I: 2011년 계획과 2006년으로부터 학습된 교훈
- 대체 가구에서 성별과 연령의 응답대체의 평가
- 2011년 영국 센서스 에디팅 전략
- 2007년 미국 농업 총조사의 E&I 시스템의 개선

3) 주요발표 내용

- 오스트리아의 등록기반 센서스의 응답대체와 추정
 - 2001년 전통적인 센서스로부터 2010년 등록 기반 센서스로의 변환으로 오스트리아 통계청은 자료 수집, 자료 에디팅 및 임퓨테이션에 관련한 새로운 도전에 직면
 - 주요한 임무 중의 하나가 불완전한 자료를 추정하는 문제임
 - 직업과 교육과 같은 변수의 추정을 예를 들어 설명
 - 직업은 센서스의 핵심문제이나 불행하게 어느 행정 등록에도 포함되지 않음
 - 노동력조사 자료에 근거하여 핫덱(hot-deck) 추정방법을 사용
 - 직업을 포함하는 2001년 전통적인 센서스와의 비교에서 등록 기반 센서스의 테스트는 만족할 만한 결과를 보임

IV. 회의결과

가. 전자 자료 수집을 통하여 획득된 자료의 에디팅

- 1) 이 주제에서 다루어진 논의는 센서스와 사업체조사를 포함
 - 전자 자료 수집은 전통적 자료수집방식의 몇 가지 문제에 대한 솔루션을 제공하는 반면 이는 몇몇 새로운 문제를 보임
 - 항목무응답의 가시적인 감소를 보이며 일반적으로 전자 자료 수

- 집은 종이 설문지보다 더 좋은 품질을 나타낸다고 여겨짐
- 발표에서 에디팅을 실시하는 시점을 정의하는 문제를 다룸
 - 평범한 외형과 느낌이 반복적인 사업체조사를 위한 전자식 설문지에 응답자의 친근성을 높이고 따라서 품질이 향상될 것이라 제안함
 - 전자 자료 수집을 이행하기 전에 유용성 검증을 수행할 것을 제안함
- 2) 참가자는 응답 과정에서 유발되는 모드효과를 줄이기 위해 전자 수집과 종이설문 수집에서의 에디팅과 에디팅과정의 차이를 어떻게 조정하여야 할지를 토의함
- 전자 설문지는 모드 효과에 대해 더 파악하고 그것들을 제한하기 위한 방법을 찾기 위해 더 많은 실험이 요구됨
 - 유의한 모드 효과가 있는 경우, 데이터는 각 데이터 수집 모드를 통해 획득되는 데이터 세트들에 근거하여 전자식과 종이 설문지에 대해 별도로 응답대체가 되어야 함
- 3) 자료품질 향상과 전자식 수집에서의 응답률 향상목적이 갈등을 보이는 문제점을 토의함
- 응답자의 이전 자료를 이용하는 것과 에디팅 부담을 낮게 유지하는 자동 에디팅은 전자식 자료 보고를 촉진하는 한 방법으로서 추천됨
 - 몇몇 참가자는 전자식 자료수집에 너무 많은 에디팅을 만드는 것을 경고하였으며 그렇지 않으면 응답자는 좌절감을 느낄지도 모르고 전자식 자료 보고에 더 저항할 것이라 함
 - 경험에 의하면, 현재 인터넷 데이터 수집은 늘 유의한 비용절감을 보여주지는 않으며 높은 데이터 품질을 얻는 하나의 모드로 보여짐
- 4) 참가자들은 전자 수집 시스템에서 이행되고 있는 일반화된 편집

시스템에서 보통/표준 에디트 유형들이 얼마나 가용한지와 과거 자료의 사용에 대해 토의함

- 전자식 설문지에서 이행되는 자료 편집규칙들은 포스트 에디팅에서 사용되는 것보다 더 단순해야한다고 제안
- 이전에 보고된 자료를 사용함에 있어서 응답 대체된 자료를 응답자에게 공급하지 않음을 확인하는 것이 중요
- 메시지, 자동스킵, 하나의 응답을 허락하는 응답 버튼은 응답자에게 원치 않는 행동을 이끌 수 있다는 경고가 있었음

5) 참가자는 또한 응답부담, 무응답 위험과 함께 데이터 품질의 균형을 맞추기 위해 수집 내에 수행될 에디팅을 얼마나 할지를 결정하기 위해 전자 수집을 위해 관리되어야 하는 것을 논의

- 에디팅 프로세스를 평가하기 위해 하나의 에디트가 실패할 때의 모든 경우가 기록되어야 함
- 모든 응답의 완전한 기록이 보존될 것을 제안

6) 전통적인 데이터 수집모드는 하나의 코더와 다른 코더 사이의 일관성과 관련된 코딩의 문제에 직면하고 있음이 강조, 이러한 문제들은 코딩이 직접적으로 응답자에 의해 이루어지는 전자 수집에서 더욱 커질 수 있음

7) 향후 논의될 필요가 있는 문제

- 종이설문지에서 사용되는 것뿐만 아니라 전자 설문지 설계와 전자설문지에서의 에디팅을 향상시키기 위해 응답자로부터의 피드백에 대한 메타데이터를 정의하고 사용하는 방법들
- 전자설문에서 어떤 관련된 타당성 점검규칙으로 엔트리에 채울 때 응답자가 덜 정확하게 보고하는가에 대한 문제

나. 행정자료와 통합소스의 에디팅

- 1) 통계의 품질을 향상시키고 비용과 응답부담을 줄이기 위해 행정 자료에 의지하나 행정자료는 통계데이터로서 사용을 위해 원래 설계된 것이 아니어서 광범위한 프로세싱과 에디팅이 요구되며 데이터의 다양한 소스의 통합은 새로운 도전을 줌

- 2) 행정 소스로부터 얻어진 데이터의 품질은 바람직한 수준이 아닐 수 있으며 행정소스의 통계 사용을 위해 자료를 조정하는 방법들이 고려됨
 - 계절효과를 극복하기 위해 시계열모형의 사용이 추정치를 만드는 데 제안되었으며 통합된 행정자료에 대한 매크로 에디팅을 위한 시계열 방법이 장점을 가질 수 있음
 - 효과적인 에디팅과 임퓨테이션 전략의 사용은 통계 데이터베이스의 품질 구축을 위해 제안되었으며 이보다 이것을 책임질 수 있는 하나의 팀이 실존하는 것을 제안함
 - 추천된 대부분의 방법들은 대량 임퓨테이션, 마이크로 에디팅, 스코어에 근거한 선택적 에디팅임

- 3) 행정자료와 관련된 이슈
 - 행정소스의 적시성은 통계사업과정 사이클과 반드시 일치하지 않음
 - 많은 행정소스는 일반적으로 연간 자료를 생성하는 반면에 월간 자료에 대한 수요가 있음
 - 사업체는 종종 기업체 수준과 같이 상위 총계 수준에 대한 자료만을 제공하나 설립체 수준에 대한 수요가 있음
 - 행정소스는 통계인에게 관심있는 모든 단위를 커버하지 않을 수 있고(단지 법적인 단위), 모든 활동을 커버하지 않을 수도 있음

- 4) 행정등록과 기록은 다중 소스의 통합을 유도
 - 통합된 데이터 소스에 대한 품질 고려는 하나의 소스 자료에 대한 품질고려에 더하여 링킹 데이터에서의 에러와 불일치에 강조

되어야 함

- 조사과정의 별개의 단계에서 행정자료의 병합은 여러가지 장단점을 초래
- 다중 행정등록이 하나의 통계등록을 구축하기 위해 사용되어질 수 있음. 예를 들면 사업체 등록은 법인 사업체 등록과 Tax 등록, 사회보장등록과 다양한 산업체등록으로부터 구축될 수 있음
- 일반적으로 종합적인 정보를 얻기 위한 최선의 방법이나 이것은 불일치하는 단위의 문제와 다양한 행정소스사이의 시의성 문제를 초래함

5) 행정자료의 에디팅을 토의할 때 다음과 같은 점이 도출

- 적시성이 중요함. 오랜 기간 후에 기업체를 콜백(call back)하는 것은 불가능함
- 어떤 통계청은 예측치를 예비치로 사용하는 것으로 적시성을 극복하기 위해 노력함. 이 방법은 안정성과 예측을 허용하는 과거 자료의 충분한 양을 갖고 있는 것을 기대하나 그러한 안정성이 대기업에서 드물다는 것을 강조
- 행정소스로부터 얻어지는 정보의 품질은 또한 관련된 행정기관과의 관계에 의존함

6) 다음과 같은 점들이 품질 평가 관련 토론에서 이루어짐

- 표본조사는 행정소스의 품질을 추정하고 개별적인 행정등록자료에 대한 앞으로의 에디팅과 임퓨테이션 접근법을 결정하기 위해 도움이 됨

7) 일반 토론의 마지막은 매스 임퓨테이션(mass imputation)을 다루었으며 두 개의 반론이 있었음

- 몇몇 참가자는 다량 응답대체된 자료가 비현실적인 일관성 수준으로 정확성/품질의 잘못된 인식을 준다는 것을 지적함
- 다른 참가자는 매스 임퓨테이션이 행정과 결합된 자료소스를 통

해 얻어진 자료의 일관성을 유지하는데 도울 수 있다고 변호하였으며 완전한 자료와 함께 하나의 데이터 셋의 가용성으로 연구자에 의해 통계분석을 위한 기회를 부여한다고 하였음

다. 데이터 에디팅을 통한 품질개선

- 1) 이 주제에 대한 발표는 다음과 같은 테마를 포함
 - 통계사업과정을 모니터링하고 개선하는 것
 - 품질구성부문을 균형잡는 것
 - 지원도구를 개발하는 것
 - 단위에러의 자동에디팅

- 2) 참가자는 품질개선을 위한 수정의 유용성에 대해 토론
 - 상대적 수정은 설문지의 개선에 대하여 유용한 정보를 제공하며 이 정보는 정보수집과정 동안 과소 또는 과대 커버리지를 나타냄. 상대적인 수정은 원 자료와 그럴듯한 값 사이의 변화량으로 정의됨
 - 상대적 수정은 쉽게 계산될 수 있고 설명될 수 있고 그것들은 측정문제와 관련한 힌트를 줄 수 있으며 데이터 에디팅 과정의 최적을 위해 사용될 수 있음
 - 상대적 수정은 에디팅 프로세스의 비용을 분배하는데 사용될 수 있다고 제안

- 3) 통계청 내에서 데이터 에디팅 프레임을 만드는 것은 에디팅과 임퓨테이션 과정을 통한 자료품질 개선에 필요한 사전조건임
 - 데이터 에디팅과 임퓨테이션의 목적은 통계사업 과정을 개선하는 것 뿐 만아니라 사용자에게 그럴듯한 자료와 자료품질에 대한 정보를 제공하는 것임
 - 에디팅과 임퓨테이션 전략은 사업계획, 표준과 지침과 훈련을 포함해야 함

- 에디팅과 임퓨테이션 프로젝트의 관리는 시니어 매니저와 담당 통계작성자가 뿐 아니라 방법론자들을 포함하여야 함
- 커뮤니케이션 전략을 통해 에디팅과 임퓨테이션 과정이 자료품질을 개선한다는 가치를 확신시켜야 함
- 담당 통계작성자를 확신시키는 것에 추가하여 그들에게 수행되어야 할 임무와 강력한 IT 도구들을 보여주는 적절한 지침서를 제공하는 것이 필요하며 그렇지 않으면 에디팅과 임퓨테이션은 이론적인 토론 수준에 머무를 위험이 있음

4) 횡단면적인 사업체조사에서의 에디팅과 임퓨테이션(EDIMBUS)에 대한 유럽 공동 프로젝트의 이행에 대한 보고를 토의

- 이 프로젝트에서 개발된 권고 실행 매뉴얼은 통계데이터 처리틀을 개발하는데 유용한 기초임
- EDIMBUS 프로젝트의 향후 연구는 다음에 초점을 둘 것임
 - 유럽통계 시스템 내 그리고 국가적 수준에서 권고된 실행매뉴얼의 전파
 - 파일롯 세베이 내에 권고된 실행매뉴얼의 이행의 촉진
 - 에디팅과 임퓨테이션 방법론과 절차 그리고 매뉴얼에서 상세하게 설명한 원리에 대한 교육
 - 개별국가통계청에 의해 사업체조사에서의 에디팅과 임퓨테이션에 대한 높은 수준의 전략의 개발
 - 적절한 수준에서 전문가그룹에 의한 권고 실행매뉴얼의 주기적인 갱신
- 권고 실행 매뉴얼의 전자 버전은 인터넷 <http://edimbus.istat.it>에서 얻을 수 있음

라. 새로운 방법

- 1) 새로운 기술과 소프트웨어 도구뿐 아니라 조직적인 이행 문제와 영향을 검토하는 데 있어서 데이터 에디팅 이론과 새로운 방법론에서의 두드러진 기여 등을 토의

2) 자동 에디팅에 관한 토론

- 확률오류(random error)는 편의를 초래할 가능성이 적어 계통오류(systematic error)보다 덜 중요하나 분산추정에 대한 랜덤에러의 영향은 더 고려되어야 함
- 얼마나 편집해야 하는 지에 대한 경계를 정하는 것은 확률적 오류를 수정할 때 필요함
- 계통오류의 존재는 데이터 셋에서 많은 유사한 오류가 있음을 암시하며 이러한 경우 오류의 지속적인 자동 수정보다는 설문지 설계나 조사과정의 다른 부분에 대한 수정에서 해답이 찾아질 수 있음

3) 임퓨테이션과 관련하여 논의

- 예측을 위한 시계열모형의 사용이 제안되었으나 가용한 시계열의 제한된 기간으로 인해 시계열방법을 사용하고자 했던 통계청의 경험은 그리 긍정적이지 않음

4) 회의에서는 또한 오픈 소스, 상업적인 기성품과 같은 소프트웨어를 개발하는 다양한 방법을 검토

- 상업적인 판매자는 통계청의 모든 요구를 커버하는 소프트웨어를 생산할 것 같지 않으며 특별하게 에디팅과 임퓨테이션 시장은 상업적으로 매력적일 만큼 크지 않음
- 이상적으로 소프트웨어는 오픈소스 소프트웨어(OSS) 모드에서 공유되어야 하며 OSS는 상이한 통계청과 상이한 IT 플랫폼사이의 소프트웨어와 구성요소의 이식가능하게 함
- 몇몇 나라에서 통계청이 오픈소스로서 조직 내에서 개발한 가용한 소프트웨어를 만드는 것을 막는 법적인 문제가 있으나 대조적으로 다른 나라들은 정부기관이 오픈소스로 모든 소프트웨어를 생산하도록 강요함
- 통계청간에 통계소프트웨어의 공유를 위한 범위와 모델을 파악

할 목적으로 통계정보시스템 관리(MSIS)에 대한 UNECE/
Eurostat/OECD 운영그룹의 감독 하에 태스크 포스가 만들어졌음

마. 결과에 근거한 에디팅(포스트-에디팅)

- 1) 이 주제 하에서 발표된 논문은 다음과 같은 논의점을 다룸
 - 매크로 에디팅
 - 선별적 에디팅
 - 포스트 에디팅

- 2) 참가자는 선별적 에디팅에서 스코어 함수의 사용을 토의
 - 공분산효과는 다변량 스코어 함수를 사용할 때 중요할 것임
 - 도매거래의 경우 기업들은 시장수요에 매우 다이내믹하게 반응하기 때문에 이전년도의 값이 가능한 값을 결정하는데 적합하지 않을 것임
 - 과거자료의 결손은 스코어 함수의 정의를 어렵게 함
 - 몇몇 참가자는 통계적인 모형을 근거로 예측을 통하여 행정자료 소스에 대한 스코어 함수의 정의를 제안

- 3) 토론은 매크로 에디팅과 관련한 다음과 같은 이슈들을 꺼냄
 - 매크로 에디팅은 예를 들면 그래픽 방법과 같이 다른 방법들에 의해 수반되어야 함

- 4) 포스트-에디팅
 - 이미 발표된 결과들은 포스트-에디팅 후에 바뀔 수 있음
 - 공표된 자료의 변경은 일찍 파악할 수 없었던 의도에서 사용될 때 일어날 수 있음

바. 센서스

- 1) 최근 전통센서스로부터 등록기반 센서스 쪽으로 움직임을 보여 왔음
 - 센서스의 목적은 정치적·경제적 의사결정을 위한 굳건한 기초를 제공하는 것이며 참조자료의 종합적인 집단을 만드는 것임
 - 그러므로 데이터 에디팅은 충분한 품질을 보증하기 위해 필수적

- 2) 다음과 같은 논의들이 일반토의에서 이루어짐
 - 이전 센서스로부터의 정보를 사용하는 학습과정은 에디팅과 임퓨테이션 전략을 개발하는 데 일반적인 방침으로서 제공됨
 - 통계전문가는 데이터 정제 과정을 위한 충분한 정보가 제공되어야 하며 이 정보는 또한 센서스 사용자에게 유용할 것임
 - 몇몇 경험에서 일부 지역에서는 등록기반 접근이 잘 되었으나 다른 지역에서는 품질 감소가 있었음
 - 센서스 커버리지 서베이와 같은 사후 연구는 과정을 향상하며 결과를 평가하는 유용한 도구일 수 있다고 제안됨. 몇몇 참가자는 에디팅과 임퓨테이션이 사후조사의 데이터에 적용되어야 한다고 생각
 - 센서스 레코드를 현재와 과거의 레코드 간에 연결하는 것은 무응답군을 처리하는 데 도움이 되나 몇몇 통계청은 비밀보호에 해가 되지 않도록 하기 위해 센서스 레코드를 링크하는 데 조심스러움

사. 향후 의제에 대한 추천 및 기타사항

- 1) 참가자는 약 18개월의 주기로 회의가 개최되어 질 것을 추천

- 2) 향후 워크세션의 의제를 추천
 - 자동 에디팅과 임퓨테이션 그리고 소프트웨어 적용들
 - 행정자료 에디팅
 - E/I 행정자료와 센서스자료

- 기업/인구 통계에 대한 업무처리
- E/I 방법들을 이행하는 성공적인 전략
- 새로운 방법들
- 데이터 에디팅과 임퓨테이션의 품질 영향을 측정하는 지표
- 선별적 에디팅과 매크로에디팅

3) 통계자료편집에 대한 지식베이스(K-Base)는 UNECE 웹사이트 www.unece.org/stats/k-base에서 가용

4) 스위스 대표단은 통계데이터 편집에서 다음 회의를 주관할 것을 제안하였으며 참가자는 2009년 후반기에 회의 개최를 추천