
2022년 지역별고용조사 표본개편 보고서

2022. 7.



조사관리국
표본과

1 조사개요

- **(조사목적)** 지역 고용정책 수립에 필요한 시군구 단위의 고용통계를 작성하고, 시도별 고용구조 및 산업 직업에 대한 세분화된 자료를 생산하여 제공
- **(주기·기간)** 반기(4월, 10월) 주기, 조사월의 15일이 포함된 1주를 대상으로 조사
- **(조사대상)** 조사대상 주간에 국내 상주하는 만 15세 이상인 자
- **(조사항목)** 취업자, 실업자, 비경제활동인구, 맞벌이가구 등 32개 항목
- **(공표)**

시군구 주요 고용지표	: 상·하반기 조사 후 4개월 이내
취업자의 산업·직업별 특성	: 상·하반기 조사 후 6개월 이내
기혼여성의 고용현황	: 상반기 조사 후 7개월 이내
맞벌이 및 1인 가구 고용현황	: 하반기 조사 후 8개월 이내

2 표본설계 개요 (※ 2021년 하반기 기준)

- **(표본체계)** 5년 고정표본
- **(설계층)** 총 229개 시군구
 - 특광역시는 구군 단위(세종시 포함, 75개), 그 외 지역은 시군 단위(154개)
- **(표본규모)** 총 11,695개 조사구(추출률 3.2%)
 - 경제활동인구조사 조사구 1,737개와 지역별고용조사 조사구 9,958개로 구성
 - 조사구 당 20개 가구를 조사하여 총 234,000여 가구 조사
 - 229개 설계층별 모집단 규모, 목표 RSE 등을 고려하여 표본배분
 - 목표 상대표준오차(이하 RSE)는 실업자수 기준 전국 2%, 시도 18% 이하
- **(추출방법)** 층화2단집락추출
 - 1단계(PSU=조사구) : 일반조사구에 대한 확률비례계통추출(가구수 크기 기준)
 - 2단계(SSU=가구) : 추출된 조사구 내에서 시작가구를 단순임의추출

3 표본설계 연혁

- ('12 개편) 특광역시 설계층을 공표 수준(시도)에 맞춰 구군 → 시도로 축소
- ('13 확대) 소지역 고용통계에 대한 외부 수요에 따라 전체 표본규모를 확대하고 특히 특광역시 표본배분 비율을 상향 조정(전체 중 약 '12년 20.2% → '13년 26.3%)
- ('19 추가) 특광역시 자치구군 통계 작성을 위해, 특광역시 설계층을 시도 → 구군 수준으로 복원하고 표본규모를 대폭 확대('17년 2,506개 → '19년 4,015개)

< 주요 표본설계 연혁 >

구분	'08 설계	'09 설계	'10 설계	'11 설계	'12 개편	'13 확대	'15 조정	'17 개편	'19 추가*	
추출틀	'05 인총	"	"	"	'10 인총	"	"	'15 인총	'16 등록	
설계층	전체	232개	232개	232개	230개	163개	163개	162개	161개	229개
	특광	구군	"	"	"	시도	"	"	"	구군
	도	시군	"	"	"	"	"	"	"	"
표본규모	전체	8,786	8,800	8,800	8,839	8,721	9,949	9,958	10,063	11,695
	특광	2,010	1,916	1,987	1,987	1,758	2,614	2,454	2,506	2,454
	도	6,773	6,884	6,793	6,852	6,963	7,335	7,504	7,557	7,495

* 특광역시 자치구군 표본설계('19년 2월) + 추가 표본설계('19년 8월)

4 표본개편 배경 및 방향

① 모집단 변화를 반영하기 위한 최신 표본추출틀 활용

- (배경) 현행 표본은 '15년('17 개편) 혹은 '16년('19 추가) 인총 기반 표본추출틀에서 추출된 것으로, 표본의 대표성 저하 등 노후화로 인한 문제가 발생 중
- 인구주택총조사 결과를 보면, 지난 5년간 고령(65세 이상) 인구, 1인가구 비율, 아파트거주가구 비율 등 고용 활동과 관련성이 높은 특성들이 크게 변화함

< 「'20년 인구주택총조사 등록센서스 방식 집계 결과」 中 >

주요 지표	단위	'15년(A)	'16년	'17년	'18년	'19년	'20년(B)	증감(B-A)
성비	여자 100명당	99.7	99.6	99.6	99.5	99.5	99.3	-0.4
수도권인구	비율(%)	49.5	49.5	49.6	49.8	50	50.2	0.7
생산연령(15~64)인구	비율(%)	72.9	72.8	72.5	72.2	71.9	71.3	-1.6
고령(65+)인구	비율(%)	13.2	13.6	14.2	14.8	15.5	16.4	3.2
평균가구원수	명	2.53	2.51	2.47	2.44	2.39	2.34	-0.19
1인가구	비율(%)	27.2	27.9	28.6	29.3	30.2	31.7	4.5
아파트거주가구	비율(%)	48.1	48.6	49.2	50.1	51.1	51.5	3.4

- (방향) '20년 11월 기준 등록센서스 및 인총 표본조사를 기초로 구축된 최신 표본추출틀을 활용, 그간 모집단 변화를 반영하여 표본의 대표성을 확보

② 고용 특화 모집단을 활용한 표본추출

- (배경) 종합모집단 구축 계획('14)에 따라 경제활동통계등록부(이하 경활등록부)*가 구축('20) 되었으며, 이를 이용해 고용과 직접 관련있는 변수 생성이 가능해짐

* 사회보험, 국세자료 등 25종 행정자료를 활용, 행정기관에 신고된 모든 노동관계(경제단위와 노동자)를 기록한 통계등록부로서, 노동의 투입(기업체-종사자)과 공급(가구-취업자) 양측에서 관측이 가능

- 경제활동인구조사에서는 7차 연동('21년 도입)부터 계통추출 시 내재적 층화효과를 얻기 위한 분류지표로서 경활등록부의 등록고용률을 활용

< 「가구부문 경상조사 7차 연동 표본설계 보고서」 中 >

▶ 설계층별로 3종 산업(농림어업, 광제조업, 서비스업)별 고용률과 분류지표 간 정준상관분석 실시 결과, 경활등록부 등록고용률의 정준상관계수가 가장 높게 나온 설계층이 27개 중 15개로 최다

분류변수		1순위	2순위	3순위	4순위
인구	경활등록부 등록고용률	15*	1	1	
	30~40대 인구비율	4	7	1	3
	15~29세 인구비율		5		1
	65세이상 인구비율		1	5	
가구	농림어가비율	7	3		
	가구주연령		1	2	1

- (방향) 경활등록부와 표본추출틀을 연계하여 조사구 단위의 등록고용률 변수를 생성하고 이를 설계층별 분류지표로 활용하여 표본추출

③ 공표수준별 정도를 고려한 표본규모 산정

- (배경) 실업률 상대표준오차가 25% 미만 특광역시 구군은 75개 중 5개(6.7%)인 반면, 도지역 시군은 154개 중 102개(66.2%)로, 정도 측면에서는 도지역이 더 미흡

- 한편, '13년 및 '19년 표본설계 결과, 특광역시를 중심으로 표본조사구 수가 큰 폭으로 증가('08년 대비, 현재 +99.8%)한 상황

< 표본조사구 규모 연혁 >

구분	'08년(A)	'09년	'10년	'11년	'12년	'13년	'15년	'17년	'19년(B)	'08 대비 증감(B-A), (률)
전국	8,786	8,800	8,800	8,839	8,721	9,949	9,958	10,063	11,695	2,909 (33.1)
특광	2,010	1,916	1,987	1,987	1,758	2,614	2,454	2,506	4,015	2,005 (99.8)
도	6,776	6,884	6,813	6,852	6,963	7,335	7,504	7,557	7,680	904 (13.3)

- (방향) 시군구별 통계의 정도 제고를 목표로, 자원 효율성 및 지방청 업무량을 고려하여 설계층별 표본규모를 적정하게 조정·배분

1 모집단

① 목표모집단

- 목표모집단은 조사기간(4·10월의 15일이 포함된 1주)에 국내 상주하는 만 15세 이상인 자로서, 작성 기준에 따라 아래에 해당되는 자는 제외

< 조사대상에서 제외되는 경우 >

1) 의무복무병(군인, 의경)	2) 영내거주 직업군인
3) 해외상주 취업(학)자	4) 형이 확정된 교도소 수감자
5) 요양소, 기도원 수용자	

② 조사모집단

- **(범위)** 실제 표본추출을 위한 조사모집단은 2020년 11월 1일 기준 등록센서스 결과로 파악된, 국내 상주하는 만 15세 이상인 모든 내·외국인
 - 목표모집단의 범위를 고려하여, 기숙시설조사구와 특수시설조사구(군대병영, 병원, 교도소 등)는 제외하고, 일반조사구(1, 2, A) 내 15세 이상 가구원을 대상으로 함
- **(규모·포함률)** 전체 규모는 2015년 대비 1,308천명(3.1%) 증가한 43,828천명이며, 등록센서스상 15세 이상 인구 대비 조사모집단의 포함률은 96.7%
 - 조사구특성별로 보면, 보통조사구의 15세 이상 인구는 1,065천명(-5.2%) 감소한 반면, 아파트조사구는 2,387천명(9.8%) 증가하여 전체 중 55.7%를 차지
 - 섬조사구의 규모는 큰 폭으로 감소(-16.9%)

< 조사구특성별 조사모집단 분포 >

구분	2015년		2020년			
	조사구* 개	15세 이상 인구 천명	조사구* 개	15세 이상 인구 천명	증감	증감률
전체	691,955	43,273	384,669	45,316	2,043	4.7
조사모집단	682,414	42,520	385	43,828	1,308	3.1
아파트(A)	320,307	22,015	195,840	24,402	2,387	10.8
보통(1)	360,488	20,422	187,883	19,356	-1,065	-5.2
섬(2)	1,619	83	912	69	-14	-16.9
그 외	9,541	753	17	744	-9	-1.2
기숙사회(9)	9,541	753	17	744	-9	-1.2

* 2015년 조사구는 30±10가구 기준으로 작성되었으며, 2020년 조사구는 60±20가구 기준임

- (지역별 증감) 2015년 대비 인구증가율은 세종(76.8%), 제주(12.5%), 경기(10.2%), 충남(5.6%), 충북(4.7%) 순이며, 이들 지역은 인구 유입이 많은 지역임
- 반면에 부산(-24.0%), 서울(-1.4%), 울산(-1.0%) 등 특광역시 중 상당수가 감소세로 전환되었고, 도지역에서는 경북(-0.4%)과 전북(-0.2%)에서 인구 감소가 시작됨

< 지역별 조사모집단 분포 >

	2010년	2015년		2020년			
	천명	천명	증감	증감률(%)	천명	증감	증감률(%)
전국	39,900	42,520	2,619	6.6	43,828	1,308	3.1
서울	8,298	8,513	215	2.6	8,392	-122	-1.4
부산	2,887	2,955	68	2.4	2,883	-72	-2.4
대구	2,013	2,075	62	3.1	2,057	-17	-0.8
인천	2,179	2,411	232	10.6	2,498	86	3.6
광주	1,182	1,227	46	3.9	1,232	5	0.4
대전	1,217	1,262	45	3.7	1,253	-9	-0.7
울산	868	954	86	9.9	945	-9	-1.0
세종*		155	155		274	119	76.8
경기	9,148	10,240	1,092	11.9	11,286	1,046	10.2
강원	1,200	1,253	52	4.4	1,294	41	3.3
충북*	1,233	1,300	67	5.5	1,361	61	4.7
충남*	1,660	1,704	44	2.6	1,799	95	5.6
전북	1,450	1,516	66	4.6	1,513	-2	-0.2
전남	1,430	1,492	62	4.4	1,499	6	0.4
경북	2,157	2,231	74	3.4	2,221	-10	-0.4
경남	2,557	2,741	184	7.2	2,770	29	1.1
제주	421	490	69	16.4	551	61	12.5

* 2010년 기준 조사모집단에서의 세종시는 충남 등 이전 행정구역에 포함됨

② 표본추출틀

- (섬조사구) 조사비용·난이도를 고려하여 섬 조사구를 표본추출틀에서 제외하되, 조사모집단 중 섬조사구가 차지하는 비율이 높은 지역은 예외적으로 포함
- 섬조사구에 거주하는 만 15세 이상 인구의 구성비는 용진군(64.9%), 신안군(49.6%), 완도군(38.4%), 진도군(10.6%), 통영시(5.6%) 순으로 높았으며,
- 그 외 지역은 3% 미만으로, 조사모집단에서 섬조사구가 차지하는 비중이 낮음
- '17년 개편 당시 기준을 준용, 구성비 10%를 초과하는 4개 시군(용진, 신안, 완도, 진도)은 섬조사구를 포함하고, 나머지 시군구의 섬조사구는 표본추출틀에서 제외

< 섬조사구 비중 상위 10개 시군구 >

시도	시군	조사구			15세 이상 인구		
		전체 명	섬 개	구성비 %	전체 명	섬 명	구성비 %
인천광역시	옹진군	182	109	59.9	15,927	10,331	64.9
전라남도	신안군	357	195	54.6	31,265	15,514	49.6
전라남도	완도군	451	186	41.2	41,717	16,030	38.4
전라남도	진도군	298	55	18.5	25,645	2,730	10.6
경상남도	통영시	1,082	92	8.5	106,875	5,942	5.6
전라남도	여수시	2,103	82	3.9	228,769	6,051	2.6
전라북도	부안군	463	14	3.0	43,390	983	2.3
충청남도	보령시	831	20	2.4	85,150	1,761	2.1
인천광역시	강화군	528	10	1.9	57,068	675	1.2
전라남도	고흥군	638	13	2.0	53,824	636	1.2

○ (포함률) 최종 표본추출틀의 조사모집단 대비 포함률은 전국 기준 99.9%

< 지역별 조사모집단 대비 표본추출틀 포함률 >

지역	조사모집단 천명	표본추출틀 천명	포함률 %	지역	조사모집단 천명	표본추출틀 천명	포함률 %
전국	43,828	43,803	99.9	경기	11,286	11,286	100.0
서울	8,392	8,392	100.0	강원	1,294	1,294	100.0
부산	2,883	2,883	100.0	충북	1,361	1,361	100.0
대구	2,057	2,057	100.0	충남	1,799	1,796	99.9
인천	2,498	2,497	100.0	전북	1,513	1,511	99.8
광주	1,232	1,232	100.0	전남	1,499	1,490	99.5
대전	1,253	1,253	100.0	경북	2,221	2,221	100.0
울산	945	945	100.0	경남	2,770	2,763	99.7
세종	274	274	100.0	제주	551	548	99.4

○ (등록 고용률 생성) 경찰등록부에서 '20년 하반기 조사대상 기간에 활동했던 등록취업자를 조사구 단위로 집계한 후 표본추출틀에 연계

- 경찰등록부는 연간 존재했던 모든 노동관계를 포함하므로, 조사대상 기간 내 활동 여부를 판단한 후 집계하고, 일자리 특성은 주된 일자리 기준으로 작성

< 조사구 단위 등록 고용률 생성 예시 >

등록센서스 조사구번호	연계키 개인(가구원)	경찰등록부					('20년 10월 기준)		
		기업체	시작일	종료일	종사상지위	산업분류	일자리	취업자	
11012****A	김○○	A	20180508	20201231	임금일용외	서비스업	주	1	
	이○○	B	20190101	20200110	임금일용외	도소매업		0	
		C	20200501	20200520	임금일용	도소매업			
	박○○	D	D	20170801	20200817	임금일용외	광제조업	부	1
		D	D	20200820	20201030	임금일용	광제조업		
	E	E	20190901	20201231	자영자	서비스업	주		

1 총화 · 분류지표 개요

- (총화변수) 75개 특광역시 구군(세종시 포함)과 154개 도지역 시군으로 구성된, 전체 229개 시군구를 설계층으로 사용
- (분류지표) 설계층 내 추출단위들을 관심변수와 연관성이 높은 변수에 따라 정렬한 후 계통추출함으로써 내재적 총화효과를 기대
 - 조사구의 특성(연령구성, 주택유형 등)이 가구원의 경제활동상태(고용, 실업, 비경)에 미치는 영향을 설계층별로 분석하고, 연관성이 높은 특성을 분류지표로 선정

2 분류지표 분석 · 적용

① 변수 전처리

- (종속변수) 가구원별로 경제활동여부, 경제활동상태 이상 2가지 분류변수 생성

< 종속변수에 대한 분류변수 생성 >

경제활동여부 1. 경제활동(취업 or 실업) 2. 비경제활동	경제활동상태 1. 취업 2. 실업 3. 비경제활동
--	---

- (명목형 독립변수) 독립변수 중 주택유형(5개), 산업분류(21개), 직업분류(9개)는 명목척도로서, 분석 가용성을 위해 구성비, 유사성 등을 고려하여 병합
 - 주택유형은 시군구 간 구성 차이가 큰 점을 고려하여 병합을 시군구별로 하며, 구성비가 10%(대형A는 5%) 이상인 유형은 유지하고 그 미만은 유사 유형에 병합

< 명목척도인 독립변수 병합 >

주택유형(시군구 내 병합) <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="padding: 2px;">유형</th> <th style="padding: 2px;">병합 기준</th> </tr> <tr> <td style="padding: 2px;">단독</td> <td style="padding: 2px;">10% 미만 → 기타</td> </tr> <tr> <td style="padding: 2px;">기타</td> <td style="padding: 2px;">10% 미만 → 단독</td> </tr> <tr> <td style="padding: 2px;">소형A</td> <td style="padding: 2px;">10% 미만 → 중형</td> </tr> <tr> <td style="padding: 2px;">중형A</td> <td style="padding: 2px;">10% 미만 → 소형</td> </tr> <tr> <td style="padding: 2px;">대형A</td> <td style="padding: 2px;">5% 미만 → 중형</td> </tr> </table>	유형	병합 기준	단독	10% 미만 → 기타	기타	10% 미만 → 단독	소형A	10% 미만 → 중형	중형A	10% 미만 → 소형	대형A	5% 미만 → 중형	산업분류 1. 농림어 2. 광+제조+전기+수도+건설 3. 도소매+운수+숙박음식 4. 정보+금융 5. 부동산+전문+사업 6. 공공+교육+보건 7. 여가+개인+자가	직업분류 1. 관리+전문 2. 사무 3. 서비스+판매 4. 농림어업 5. 기계+기능 6. 단순노무
유형	병합 기준													
단독	10% 미만 → 기타													
기타	10% 미만 → 단독													
소형A	10% 미만 → 중형													
중형A	10% 미만 → 소형													
대형A	5% 미만 → 중형													

② 1·2차 분류지표 분석 및 선정

- (분석방법) 행정구역 및 주택유형에 따라 가구원의 경제활동여부·산업분류·직업분류에 유의미한 차이가 있는지를 교차분석(카이제곱)을 통해 검정
- (분석결과) 경제활동여부와의 분석 결과, 분석 대상 시군구 중 약 1/3에서만 통계적 차이가 존재하였으나, 산업분류·직업분류는 대부분 시군구에서 통계적 차이가 존재
- (선정결과) 행정구역 및 주택유형 모두 1·2차 분류지표로 선정

③ 기타 분류지표 분석 및 선정

- (분석방법) 가구원 경제활동상태(취업/실업/비경)는 속성값이 3개인 명목척도이므로, 조사구 특성 항목(13종, 비척도)을 독립변수로 하는 다항로지스틱회귀모형 적용
 - 설계층별 독립변수 유의성·설명력의 차이를 반영하기 위해 시군구별로 분석

< 기타 분류지표에 대한 분석 모형 >

회귀모형	$\log\left(\frac{p(Y=i)}{p(Y=\text{비경})}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_{13} x_{13} + \epsilon, \quad i = \text{취업, 실업}$	
	< 인구사회 >	< 가구 >
독립변수 ($x_1 \dots x_{13}$)	1529세비율, 3040대비율, 5064세비율, 65세이상비율 여성비율, 유배우비율 대출비율, 3040대대출비율, 등록고용률	농림어가비율 자가비율 1인가구비율, 평균가구원수

- (분석결과) 모형 적합도에 유의미한 영향을 준다고 판단된 시군구(총 229개)가 가장 많은 독립변수는 등록고용률(139개)이며, 그 다음은 농림어가비율(102개)
 - 특광역시와 도지역-시에서는 등록고용률이, 도지역-군에서는 농림어가비율이 경제활동상태에 대한 주요 설명변수인 것으로 확인
- (선정방법) 시군구별 모형 적합도에 유의미한 영향을 준다고 판단된 독립변수를 대상으로 전진선택법(forward selection)*을 활용, 1~3단계 독립변수를 순서대로 1~3순위 기타 분류지표로 선정
 - * 모형에 포함되지 않은 독립변수 중에서 스코어 카이제곱검정(score chi-square test) 통계량이 가장 큰 것을 단계적으로 추가해 가는 방법
- (선정결과) 가장 많은 시군구에서 1순위 기타 분류지표로 선정된 독립변수는 등록고용률로, 총 229개 시군구 중 90개(39.3%)에서 1순위로 선택됨
 - 특광역시에서는 전체 75개 구군 중 48개(64.0%)에서 등록고용률이 1순위로 선택

- 특광역시 → 시 → 군으로 갈수록 농림어업 취업자 비중이 높아지는데, 그에 따라 등록고용률 대신 농림어가비율이 1순위로 선택되는 빈도가 높아지는 경향

< 행정구역 × 독립변수별 1순위 선정 횟수 >

특광역시 : 총 75개 구군			도지역-시 : 총 77개 시			도지역-군 : 총 77개 군		
순위	독립변수명	횟수	순위	독립변수명	횟수	순위	독립변수명	횟수
1	등록고용률	48	1	등록고용률	35	1	농림어가비율	21
2	가구원수	4	2	농림어가비율	12	2	자가비율	8
3	자가비율	4	3	유배우비율	7	3	1인가구비율	8
4	농림어가비율	3	4	여성비율	6	4	등록고용률	7
5	3040세대대졸비율	3	5	15-29세비율	4	5	유배우비율	6

< 독립변수별 선정 횟수 >

	전국 : 총 229개 시군구			
	합계	1순위	2순위	3순위
등록고용률	131	90	28	13
농림어가비율	110	36	39	35
유배우비율	49	15	15	19
자가비율	40	15	12	13
1인가구비율	41	12	11	18
평균가구원수	43	11	17	15
15-29세인구비율	34	11	10	13
여성비율	47	9	17	21
65세인구비율	39	8	17	14
대졸이상비율	50	6	21	23
3040세대대졸이상비율	36	6	15	15
3040대비율	27	6	12	9
50-64세비율	39	4	15	20

④ 분류지표 최종 적용

- (범주·구분점 생성) 연속형 변수인 기타 분류지표(총 13종)에 대해 시군구별로 군집분석(k-means)을 수행하여 2개, 3개로 범주화하고 각 구분점(1, A·B)을 생성
- (적용) 기타 분류지표의 우선순위, 시군구별 최대 범주수 등을 고려하여 최종 기타 분류지표 개수 및 구분점 개수를 차등 적용

1 표본규모 산정 요인 검토

① 모집단(15세 이상 인구) 규모 변화

- 2015년 표본추출틀 대비 모집단(15세 이상 인구) 규모 증감률은 하남(79.5%), 세종(76.8%), 화성(47.9%) 순이며, 대구 서구(-15.1%) 등 3곳은 10% 이상 감소

< 모집단규모 증감률이 ±10%를 초과한 주요 시군구 >

시도	시군구	가구			15세 이상 인구		
		2015년 개	2020년 개	증감률 %	2015년 명	2020년 명	증감률 %
10% 이상 증가							
<i>상위 10개</i>							
경기	하남시	56,418	113,020	100.3	129,451	232,363	79.5
세종	세종시	76,244	140,697	84.5	154,962	273,942	76.8
경기	화성시	211,110	334,700	58.5	467,448	691,411	47.9
부산	강서구	34,032	52,544	54.4	74,793	105,994	41.7
대구	달성군	67,699	101,076	49.3	152,743	209,229	37.0
경기	김포시	123,324	177,855	44.2	282,848	380,159	34.4
경기	시흥시	158,894	213,108	34.1	349,915	443,363	26.7
경기	광주시	111,397	146,834	31.8	253,727	319,720	26.0
충북	진천군	27,245	37,039	35.9	57,919	72,379	25.0
인천	연수구	109,792	146,574	33.5	255,603	318,589	24.6
10% 이상 감소							
부산	영도구	52,805	51,114	-3.2	112,360	100,292	-10.7
인천	동구	27,293	25,575	-6.3	60,893	53,892	-11.5
대구	서구	81,828	74,917	-8.4	179,717	152,576	-15.1

② 이전 조사에서의 극단가중값

- 2021년 하반기 조사의 사후가중값의 분포를 검토한 결과, 불균등 가중값 효과(UWE* > 1.3) 혹은 최대/최소 비율(MMR > 10)이 높게 나타난 시군구는 총 12개

* 불균등 가중 효과(Unequal Weighting Effect), 불균등한 가중값으로 인해 추정의 표본오차가 증가한 정도를 의미하는 것으로, UWE가 1.44이면 표본오차가 1.2배(=√1.44) 증가하였음을 의미

③ 이전 조사에서의 RSE

- (전국 단위) 2021년 하반기 조사 기준, 전국 단위 고용률 RSE는 0.2% 실업률 RSE는 1.6%로, 목표한 정도(실업자수 기준 2.0%)를 충족
- (시도 단위) 시도 단위의 고용률 RSE는 최소 0.5% ~ 최대 1.9%이고, 실업률 RSE는 최소 3.4% ~ 최대 19.0%으로, 세종(19.0%)를 제외한 모든 시도에서 목표 정도(18.0%)를 충족

- (시군구 단위) 고용률 RSE는 전체적인 수준이 낮고, 시도 간·시도 내 편차도 안정적인 수준이나, 실업률 RSE는 전체적인 수준이 높고, 편차가 매우 큼
 - 실업률 RSE가 「표본설계 및 관리지침」 상 신뢰도 낮음($RSE \geq 50\%$)에 해당하는 시군구는 총 229개 중 29개(12.7%)로, 모두 도지역임
 - 신안, 태안, 울릉 등 구직활동이 희소한 지역에서 특히 높은 편

② 표본규모 산정

① 산정 방법

- (기본 방향) 설계층별 목표 정도(실업률 RSE)를 기준으로 표본규모를 산정하되, 지역과 인구규모에 따라 목표를 차등 설정하고, 기존 규모 대비 $\pm 10\%$ 이내에서 목표 정도에 점진적으로 접근하도록 설계
 - 통계 작성 목적과 공표수준(시군구)을 고려하면, 설계층별 목표 정도를 고정하여 표본규모를 산정하고 그 합으로 전체 표본규모를 구하는 것이 이상적
 - 하지만 현재까지의 표본규모 산정·배분 방법(전체 표본규모를 먼저 결정한 후 배분)과 달라, 해당 방법 적용 시 표본규모 변동이 과대할 것으로 예상
 - 한편, 구직활동이 구조적으로 희소한 지역에서 실업률 기준 목표 정도를 강행 적용하는 것은 달성 가능성 측면에서 비현실적일 것으로 판단
 - 따라서, 설계층별 목표 정도를 기초로 표본규모를 산정하되,
 1. 지역과 인구규모에 따라 목표 정도에 차등을 적용
 2. 목표 정도를 엄격하게 적용하지 않고, 상단과 하단을 중심으로 관리
 3. 설계층별 표본규모 증감률이 $\pm 10\%$ 를 넘지 않는 수준에서 조정
- (표본규모 산정 방법) 본 통계는 반기마다 조사되므로, 계속조사에서의 표본규모 산정 방법을 활용해 설계층별로 산정

< 계속조사에서의 표본규모 결정 >

$$n_{h,1} = n_{h,0} \times \left(\frac{RSE_{h,0}}{RSE_{h,1}} \right)^2, \quad n_1 = \sum_{h=1}^{229} n_{h,1}$$

$RSE_{h,0}$: 이전 조사에서의 h 설계층 RSE
 $n_{h,0}$: 이전 조사에서의 h 설계층 표본규모
 $RSE_{h,1}$: 다음 조사에서의 h 설계층 목표 RSE
 $n_{h,1}$: 다음 조사에서의 h 설계층 목표 표본규모
 n_1 : 다음 조사에서의 전체 표본규모

- (이전 표본규모) '21년 상반기 기준 별도표본 + 경활표본 (전국 기준 11,695개)
- (이전 RSE) '21년 상하반기 실업률 RSE 중 큰 값으로 설정
 - 고용률 RSE는 절대적인 수준이 낮고 편차가 적어, 실업률을 관심변수로 설정
 - 구직활동에 계절성이 있음을 고려하여 상하반기 실업률 RSE 중 최댓값 활용

② 그룹별 목표 RSE 설정

- (지역×인구 그룹) 15세 이상 인구 및 실업률 RSE에 대해 군집분석 수행한 후, 특광역시 및 도지역을 각각 4개¹⁾와 5개²⁾로 범주화
 - 1) 15만 미만 / 15~30만 미만 / 30~40만 미만 / 40만 이상
 - 2) 5만 미만 / 5~15만 미만 / 15~30만 미만 / 30~60만 미만 / 60만 이상
- 분산분석 결과, 그룹별 실업률 RSE의 차이가 유의한 것으로 확인
- (그룹별 목표 RSE 상하한 설정) 「표본설계 및 관리지침」 과 지역×인구규모별 이전 실업률 RSE를 종합적으로 고려하여, RSE 상하한을 그룹별로 차등 설정

< 그룹별 목표 RSE 상하한 기준 >

지역	그룹		이전 실업률 RSE			목표 실업률 RSE	
	15세 이상 인구	시군구수	최소	중위	최대	하한	상한
특광역시	-15만	15	15.4	23.8	36.2	18.75	25.00
	-30만	29	13.2	19.4	28.9	18.75	25.00
	-40만	19	14.5	17.8	23.9	16.00	20.00
	40만+	11	12.1	16.1	18.7	15.00	18.75
도지역	-05만	58	19.4	48.0	100.3	20.00	60.00
	-15만	46	19.2	31.7	100.5	20.00	50.00
	-30만	23	15.7	22.2	34.8	18.75	25.00
	-60만	14	15.0	17.3	22.2	16.00	20.00
	60만+	9	14.0	15.1	16.8	14.00	17.50

③ 그룹 내 설계층별 목표 RSE 조정

- (1단계) 이전 실업률 RSE이 상한을 넘는 시군구는 RSE를 절삭하되, 그 결과로 인한 설계층의 표본규모 증가율이 10%를 넘지 않는 수준에서 재조정
- (2단계) 1단계 조정 결과로 증가한 그룹 내 표본규모 총합이 이전 수준으로 복원될 때까지, 중위수 기준, 그 이하의 RSE를 점근적으로 상향 조정
- (3단계) 그룹 내 RSE 최솟값이 하한을 넘을 때까지 RSE를 점근적으로 상향 조정 후, 설계층의 표본규모 감소율이 10%를 넘지 않는 수준에서 재조정

< 목표 RSE 조정 주요 결과 >

그룹		목표 RSE가 이전 RSE 대비 3% 이상 하향 조정된 시군구	목표 RSE가 이전 RSE 대비 3% 이상 상향 조정된 시군구
지역	규모		
특광	-15만	부산 : 중, 서, 동 대구 : 중 인천 : 강화	서울 : 종로 부산 : 영도, 강서, 기장 인천 : 동 울산 : 동
	-30만	대구 : 서, 달성 대전 : 대덕	서울 : 용산, 성동, 도봉, 금천 부산 : 동래, 사하, 연제, 사상 인천 : 계양 울산 : 중, 남, 북, 울주
	-40만	대구 : 북 서울 : 서초	서울 : 중랑, 성북, 양천, 구로, 동작 부산 : 해운대 인천 : 연수
	40만+	-	서울 : 노원, 관악 인천 : 남동, 부평, 서
도	-05만	강원 : 화천 충청 : 괴산, 서천, 청양 전라 : 무주, 임실 경상 : 의성, 영덕, 청도, 예천, 의령, 합천	경기 : 과천 충청 : 증평, 계룡 경상 : 울진
	-15만	경기 : 가평 충청 : 보령, 부여, 태안 전라 : 김제, 고흥, 해남	충청 : 진천 경상 : 통영
	-30만	강원 : 강릉 충청 : 충주, 아산 전라 : 군산, 익산, 목포, 여수, 순천	경기 : 광명, 오산, 군포, 안성 강원 : 춘천, 원주 경상 : 경산, 진주, 양산
	-60만	경기 : 파주, 김포 경상 : 포항	경기 : 의정부, 남양주 경상 : 구미, 김해
	60만+	-	경기 : 수원, 성남, 부천, 안산, 고양, 용인, 화성 충청 : 청주 경상 : 창원

④ 최종 표본규모 산정

- (표본규모 산정) 결정된 설계층별 목표 RSE를 이용해 표본규모 산정
 - 설문조사구는 이전 RSE를 목표 RSE로 그대로 활용하여 표본규모 산정
- (산정 요인을 고려한 조정) 설계층별 모집단 규모 변화, 극단가중값 발생 여부를 고려하여 산정된 표본규모를 일부 조정
 - 1단계 : 설계층별 이전 표본규모 대비 증감률 상하한을 ±10%으로 적용
 - 2단계 : 모집단 규모가 10% 이상 증가한 설계층은 상하한을 -5% ~ +10%로 설정하고, 10% 이상 감소한 설계층은 -10% ~ 5%로 설정
 - 3단계 : 극단가중값이 나타난 설계층은 표본규모가 유지 혹은 증가될수 있도록 상하한을 0% ~ 10%로 설정

- (최종 표본규모 및 예상 RSE) 전체 표본규모는 139개 감소한 11,556개 조사구
- '21년 하반기 조사와 분산이 동일하다는 가정하에, '22년 하반기 고용조사의 전국 단위 고용률 RSE 및 실업률 RSE는 각각 0.2%, 1.6%로 예상

< 최종 표본규모 및 예상 RSE >

	2021년 하반기				2022년 하반기				
	표본규모(조사구)		고용률RSE	실업률RSE	표본규모(조사구)			고용률RSE (예상)	실업률RSE (예상)
	전체	별도	%	%	전체	별도	증감	%	%
전국	11,695	9,958	0.2	1.6	11,556	9,819	-139	0.2	1.6
서울	1,243	1,036	0.6	3.4	1,191	984	-52	0.6	3.5
부산	740	632	0.8	5.1	717	609	-23	0.8	5.2
대구	457	367	1.0	8.8	475	385	18	1.0	8.6
인천	547	448	0.8	5.3	518	419	-29	0.8	5.5
광주	325	244	1.0	8.5	324	243	-1	1.0	8.5
대전	323	242	1.0	9.1	328	247	5	1.0	9.0
울산	280	208	1.1	7.3	254	182	-26	1.1	7.7
세종	100	46	1.9	19.0	96	42	-4	1.9	19.4
경기	1,667	1,460	0.5	3.5	1,610	1,403	-57	0.5	3.6
강원	832	724	0.8	8.8	822	714	-10	0.8	8.9
충북	582	492	0.9	11.4	574	484	-8	0.9	11.5
충남	734	644	0.8	8.4	756	666	22	0.8	8.3
전북	733	634	0.8	12.1	762	663	29	0.8	11.9
전남	972	882	0.6	9.7	991	901	19	0.6	9.6
경북	1,087	979	0.7	7.2	1,094	986	7	0.7	7.2
경남	943	844	0.7	6.6	911	812	-32	0.7	6.8
제주	130	76	1.3	16.0	133	79	3	1.3	15.8

① 조사구 · 가구 추출

- (조사구) 설계층별 분류지표 적용, 조사구의 가구수를 크기 척도로 표본조사구를 확률비례계통추출
 - 설계층별로 1~100의 시드값을 입력하여 100번 표본추출 후 표본조사구의 특성 항목이 표본추출틀과 가장 유사한 시드값을 선택하여 최종 표본추출
 - 표본추출틀과 ^{1~100번} 표본조사구 집합에서 조사구 특성 항목 각각의 산술평균을 구한 후, 두 값의 차이의 절댓값을 계산
 - 기타 분류지표와의 절대오차합을 이용해 5개의 시드값을 1차 선정한 후, 모든 (13종) 특성 항목과의 절대오차합을 이용해 최종 시드값 선택
- (가구) 표본조사구별로 단순임의추출방법으로 시작가구를 선정, 시작가구로부터 동일 방향 동선에 따라 인접한 20가구를 조사대상가구로서 조사

② 경찰표본 대체

- (개념) 추출된 11,556개 표본조사구 중 1,737개를 추출하여 경찰표본으로 대체
 - 경찰조사구는 시도 단위로 설계된 표본이므로, 단순 통합할 경우 시군구 수준에서 대표성 문제가 발생 가능
 - 따라서, 추출된 표본조사구를 시군구 수준에서 재추출하여 경찰표본으로 대체
- (추출·대체 방법) 경찰표본 1,737개 중 하나를 무작위 추출한 후, 동일 시군구 내에서 가장 유사한 특징을 가진 표본조사구를 선택하여 대체하는 과정 반복

③ 표본추출틀과의 조사구 특성 비교

- 표본추출틀, 최초 표본조사구, 최종 표본조사구(경찰대체) 각각의 조사구 특성 항목의 산술평균을 비교하여, 표본 대표성에 문제가 없는지 점검

① 설계가중값

- 별도표본 중에서 경활표본 크기만큼 조사구를 추출하여 대체하였으므로, 설계가중값은 별도표본과 경활표본을 통합하여 작성

$$w_{hij}^d = \frac{M_h}{n_h M_{hi}} \times \frac{M_{hi}}{m_{hi}} = \frac{M_h}{n_h \times 20} = w_{hijk}^d$$

$\begin{array}{ccc} \uparrow & & \uparrow \\ \text{조사구} & \text{추출률} & \text{가구} & \text{추출률} \\ \text{역수} & & \text{역수} & \end{array}$

- w_{hij}^d : h 층 i 조사구 j 가구의 설계가중값
- h : 설계층(229개 시군, $h = 1, \dots, 229$)
- i : 표본 조사구($i = 1, \dots, n_h$)
- j : 표본 가구($j = 1, \dots, m_{hi}$)
- k : 표본 가구 내 조사대상 가구원($k = 1, \dots, l_{hij}$)
- M_h : h 층의 전체 가구 수
- M_{hi} : h 층 i 표본 조사구의 전체 가구 수
- n_h : h 층의 표본 조사구 수
- m_{hi} : h 층 i 표본 조사구 내 표본 가구 수(20가구)

② 사후가중값

- 설계층·성·연령대별로 사후층화하여 현재 인구에 맞게 조정

$$w_{hijk} = w_{hijk}^d \times \frac{N_c}{\sum_c w_{hijk}^d}$$

$\begin{array}{cc} \uparrow & \uparrow \\ \text{설계가중값} & \text{사후보정계수} \end{array}$

- w_{hijk} : h 층의 i 조사구 j 가구 k 가구원의 사후가중값
- w_{hijk}^d : h 층의 i 조사구 j 가구 k 가구원의 설계가중값
- h : 설계층(229개 시도 및 시군, $h = 1, \dots, 229$)
- i : 표본 조사구($i = 1, \dots, n_h$)
- j : 표본 가구($j = 1, \dots, m_{hi}$)
- k : 표본 가구 내 조사대상 가구원($k = 1, \dots, l_{hij}$)
- N_c : 사후층 c 의 현재인구
- $\sum_c w_{hijk}^d$: 사후층 c 의 현재인구 추정값(설계가중값으로 추정)
- c : 사후층

3 추정

○ 총합 추정

$$\hat{Y} = \sum_h \hat{Y}_h$$

$$\hat{Y}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} \sum_{k=1}^{l_{hij}} w_{hijk} \cdot y_{hijk}$$

· w_{hijk} : h 층의 i 번째 표본조사구 내 j 번째 가구의 k 번째 가구원의 사후가중값

○ 표본오차 추정

$$\hat{V}(\hat{Y}) = \sum_{h=1}^H \frac{n_h}{n_h - 1} (1 - f_h) \sum_{i=1}^{n_h} (y_{hi.} - \bar{y}_{h..})^2$$

- h : 설계층
- N_h : h 층의 모집단 조사구수
- n_h : h 층의 표본 조사구수
- f_h : 조사구 추출률(= n_h/N_h)

$$\cdot y_{hi.} = \sum_{j=1}^{m_{hi}} w_{hij} y_{hij}$$

$$\cdot \bar{y}_{h..} = \left(\sum_{i=1}^{n_h} y_{hi.} \right) / n_h$$

$$SE(\hat{Y}_h) = \sqrt{\widehat{Var}(\hat{Y}_h)}$$

$$RSE = \frac{SE(\hat{Y}_h)}{\hat{Y}_h} \times 100 (\%)$$