

2021 경제총조사 표본설계

2020. 9.

강정희 (통계청 표본과)

 **목 차****I. 경제총조사 개요 및 변동사항** 2**II. 표본설계** 7

- ① 조사개요 7
- ② 표본설계 기본방향 8
- ③ 모집단 정의 및 현황 9
- ④ 표본추출틀 주요항목 분포 12
- ⑤ 매출액 분석 및 대체 22
- ⑥ 층화 및 표본추출 25

III. 향후 고려사항 31

I 경제총조사 개요 및 변동사항

1 경제총조사 개요

- 조사목적
 - ✓ 국가 전체 산업의 고용, 생산, 비용 등 정보 제공 및 각종 통계의 모집단 자료로 활용
- 조사연혁(조사주기 5년)
 - ✓ 2011년 제1회 경제총조사, 2016년 제2회 경제총조사
- 조사기준
 - ✓ 2020년 12월 31일
- 조사기간
 - ✓ 2021년 6월~7월

○ 조사대상

- ✓ 약 620만 사업체(추정)

○ 조사방법

- ✓ 면접조사 및 인터넷(모바일) 조사 병행

○ 결과공표

- ✓ 잠정('21년 말), 확정('22년 6월)

○ 조사항목

- ✓ 조사표 : 업종별 조사표 9종, 본사일괄조사 1종
- ✓ 조사항목 : 36개(공통항목 13개, 산업별 특성항목 23개)
- ✓ 신규항목 : 디지털 플랫폼 이용여부(공통), 무인결제기기 도입여부(음식주점, 숙박, 소매), 배달서비스 활용 여부(음식, 소매)

2 경제총조사 변동사항

○ 조사 대상 사업체 개념의 범위 확대

- ✓ 등록센서스기반으로 전환하면서 물리적 장소 여부 관계없이 조사대상에 포함
* 기존에는 물리적 장소가 있는 사업체에 한정

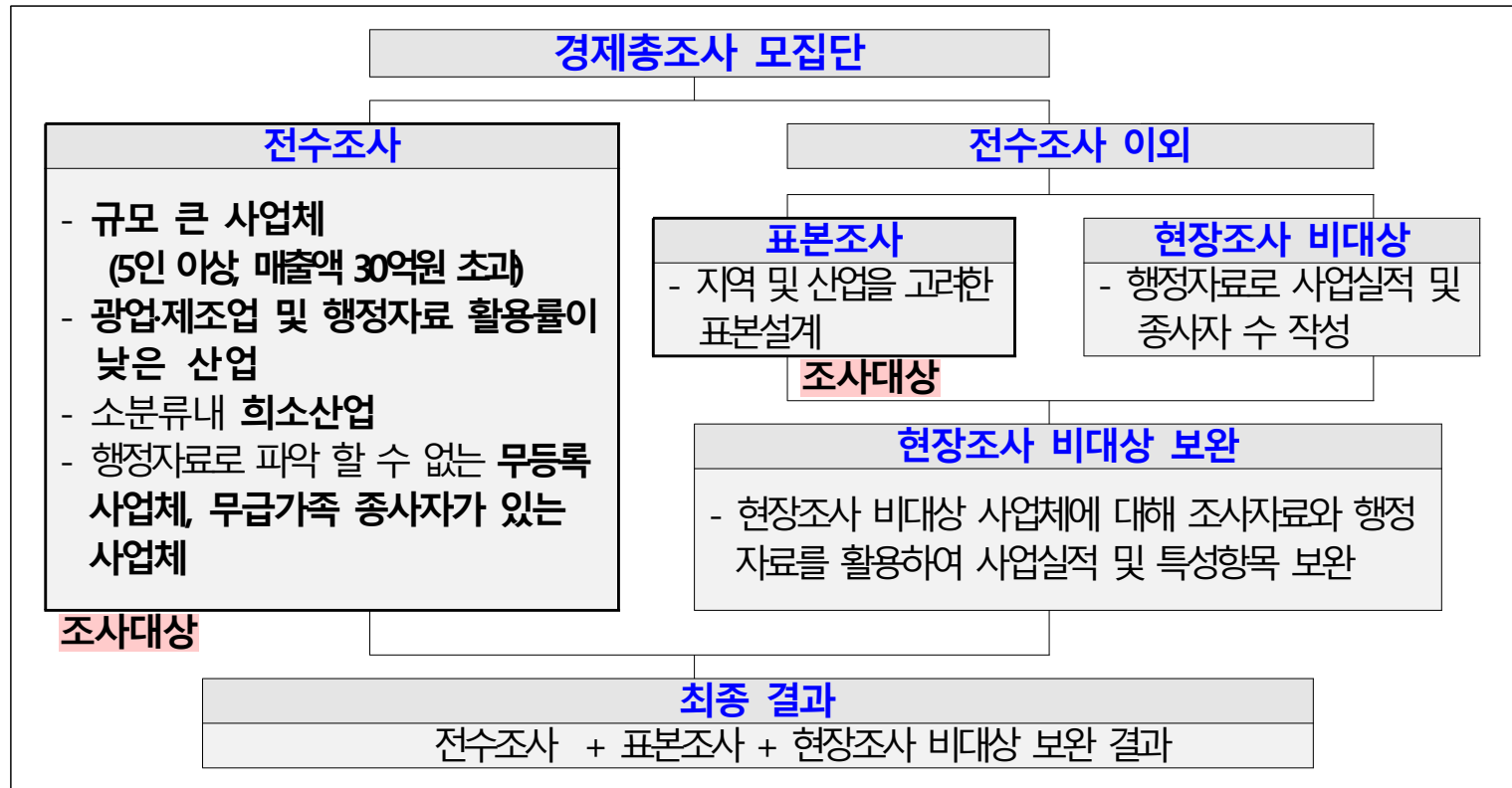
○ 표본추출틀 전환

- ✓ 전국사업체조사 기반에서 기업통계등록부*(SBR)로 전환
* 조사자료와 행정자료를 연계하여 우리나라 모든 기업체(사업체, 사업자)를 포괄하는 경제통계 부문 통합 모집단

○ 행정자료 활용 및 현장조사 대상 규모 축소

- ✓ 현장조사 미 수행 사업체 등 행정자료(과세자료) 적극 활용
- ✓ 전수조사 및 표본조사 병행으로 현장조사 대상 축소

< 경제총조사 조사방법 요약 >



< 주기별 경제총조사 비교 >

	2011년 경제총조사	2016년 경제총조사	2021년 경제총조사(예정)
모집단 (포괄범위)	<ul style="list-style-type: none"> - 물리적 장소 있는 사업체만 포함 - 전국사업체조사 336만개 사업체 	<ul style="list-style-type: none"> - 물리적 장소 있는 사업체만 포함 - 전국사업체조사 387만개 사업체 	<ul style="list-style-type: none"> - 물리적 장소 여부 상관 없이 포함 - 기업통계등록부(SBR) 약 620만개 사업체
현장조사 범위	<ul style="list-style-type: none"> - 3개 대분류 표본조사 * G도소매업 숙박음식점 S개인서비스업 - 나머지 산업 전수조사 	<ul style="list-style-type: none"> - 전(全)산업 전수조사 	<ul style="list-style-type: none"> - 대규모 사업체, 주요 업종, 무등록사업체 등은 전수조사 - 그 외 표본조사
결과공표	<ul style="list-style-type: none"> - 잠정 : 2011년 12월 - 확정 : 2012년 4월 	<ul style="list-style-type: none"> - 잠정 : 2016년 12월 - 확정 : 2017년 6월 	<ul style="list-style-type: none"> - 잠정 : 2021년 12월 - 확정 : 2022년 6월

II 표본설계('20년 시범예행조사)

1 조사개요

- 조사대상
 - ✓ 2019년 12월말 현재 조사지역에서 산업활동을 하는 모든 사업체

- 조사지역
 - ✓ 서울시 구로구, 부산시 수영구

- 조사기간
 - ✓ 조사기준일 : 2019.12.31.
 - ✓ 조사대상기간 : 2019.1.1.~12.31.
 - ✓ 조사실시기간 : 2020.7.6.~8.14.

2 표본설계 기본방향

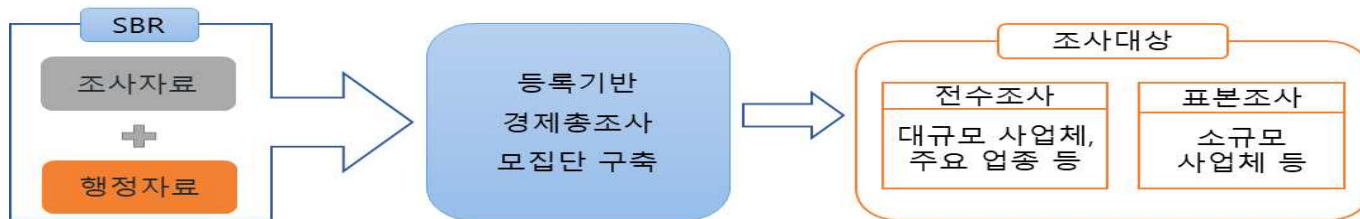
○ 표본추출틀

- ✓ 최신 기업통계등록부를 활용하며 물리적 장소 여부 관계없이 국내에서 영업활동하는 모든 사업체로 설정

○ 현장조사 대상

- ✓ 전수층*과 표본층으로 분류하여 표본설계

* 산업별 특성, 자료의 활용도와 중요도, 조사자료와 행정자료의 일치율 등을 고려



- ✓ 공표수준* 고려 표본설계

* 마이크로데이터(MD) 제공 수준 : 읍면동 × 산업중분류

3 모집단 정의 및 현황

○ 조사모집단

- ✓ 2019년 12월말 현재 서울시 구로구, 부산시 수영구에서 산업활동을 하는 모든 사업체*, 비활동 사업체 및 일부 산업 제외**

* 사업체수(구로구 47,644, 수영구 17,399, 전체 65,043)

** 비활동 · 폐업, 개인 부동산임대업, 개인 농림어업, 국제기구 등

○ 표본추출틀

- ✓ '18년 기준 기업통계등록부에서 조사지역의 전수조사* 대상 등을 제외한 사업체(34,352)

* 종사자수 5인 이상

* 제조업(상세 통계 필요)

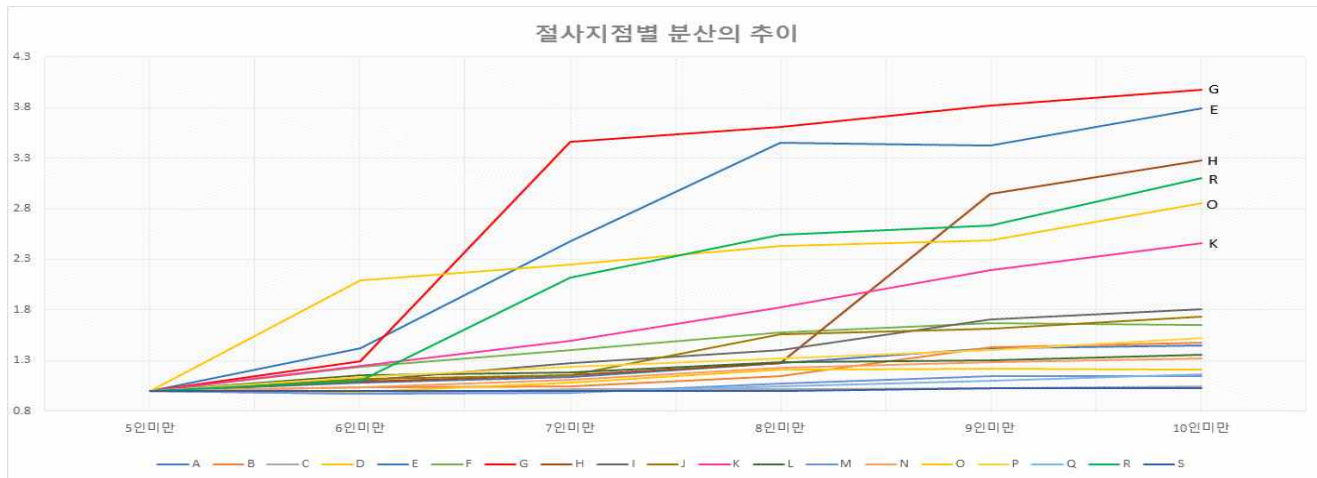
* 조사자료와 행정자료의 사업실적 일치율이 낮은 산업(A, B, D, E, J, K, P, Q)

* 추적조사 대상(무등록, 무급가족종사자) 사업체

* 추적조사 비대상(희소 산업·사업체)

○ 전수조사 대상 기준 (경제부문 등록센서스기법 심층연구, 2019.12.)

- ✓ 이전 주기와 차이 최소화
- ✓ 주요변수의 비대칭적 분포(왜도)에 대한 통계 품질 제고
 - (종사자수 5인) 산업대분류별 동일 전수층, 매출액 추정량 효율성 증가, 이전 주기 통계와 종사자 규모와 동일 기준 활용 가능



- (희소산업 및 사업체) 전국 사업체수 3,500개 미만 산업소분류 38개,
전국 사업체수 100개 미만인 산업세세분류 90개,
매출액 30억 이상 사업체

< 모집단 사업체 현황 >

	전체		구로구		수영구	
	사업체수	비중(%)	사업체수	비중(%)	사업체수	비중(%)
전체	65,043	100.0	47,644	100.0	17,399	100.0
- 전수대상(①~⑦)	30,691	47.2	22,416	47.0	8,275	47.6
① 5인이상	11,166	17.2	8,823	18.5	2,343	13.5
② 광업제조업	4,355	6.7	3,733	7.8	622	3.6
③ 사업실적 일치율 낮음	5,927	9.1	4,680	9.8	1,247	7.2
④ 무등록사업체	2,075	3.2	811	1.7	1,264	7.3
⑤ 무급가족종사자	5,549	8.5	3,145	6.6	2,404	13.8
⑥ 희소산업 등	556	0.9	424	0.9	132	0.8
⑦ 기타	1,063	1.6	800	1.7	263	1.5
- 표본추출률	34,352	52.8	25,228	53.0	9,124	52.4

4 표본추출틀 주요항목 분포

○ 조직구분별 주요항목 분포

	사업체수		임금종사자수		매출액	
	합계	비중(%)	합계	비중(%)	합계	비중(%)
구로구	25,228	100.0	21,625	100.0	6,558,407	100.0
개인	19,436	77.0	10,868	50.3	4,038,772	61.6
법인	5,792	23.0	10,757	49.7	2,519,635	38.4
수영구	9,124	100.0	6,117	100.0	1,185,835	100.0
개인	8,050	88.2	4,153	67.9	899,132	75.8
법인	1,074	11.8	1,964	32.1	286,703	24.2

○ 산업대분류별 주요항목 분포

	사업체수(개)		임금종사자수(명)		매출액(백만원)	
	합계	비중(%)	합계	비중(%)	합계	비중(%)
구로구(서울)	25,228	100.0	21,625	100.0	6,558,407	100.0
F 건설업	1,516	6.0	1,177	5.4	343,939	5.2
G 도매소매업	12,893	51.1	11,300	52.3	5,026,668	76.7
H 운수창고업	380	1.5	289	1.3	81,146	1.2
I 숙박음식점업	2,853	11.3	2,898	13.4	317,491	4.9
L 부동산업	1,852	7.3	1,409	6.5	238,187	3.6
M 전문과학기술서비스업	1,752	7.0	1,894	8.7	253,969	3.8
N 사업시설관리업	961	3.8	878	4.1	126,491	1.9
O 공공국방행정	6	0.0	6	0.0	498	0.0
R 예술여가서비스업	781	3.1	484	2.3	62,645	1.0
S 수리및기타개인서비스업	2,234	8.9	1,290	6.0	107,373	1.6

○ 산업대분류별 주요항목 분포(계속)

	사업체수(개)		임금종사자수(명)		매출액(백만원)	
	합계	비중(%)	합계	비중(%)	합계	비중(%)
수영구(부산)	9,124	100.0	6,117	100.0	1,185,835	100.0
F 건설업	823	9.1	537	8.8	110,882	9.4
G 도매소매업	3,333	36.4	1,661	27.2	682,136	57.5
H 운수창고업	219	2.4	169	2.7	20,652	1.8
I 숙박음식점업	1,582	17.3	1,648	26.9	122,078	10.3
L 부동산업	834	9.1	658	10.8	98,440	8.3
M 전문과학기술서비스업	406	4.4	357	5.8	51,450	4.3
N 사업시설관리업	348	3.8	254	4.2	33,943	2.9
O 공공국방행정	0	0.0	0	0.0	0	0.0
R 예술여가서비스업	294	3.3	180	3.0	15,598	1.3
S 수리및기타개인서비스업	1,285	14.1	653	10.8	50,656	4.3

○ 읍면동별 주요항목 분포

	사업체수(개)		임금종사자수(명)		매출액(백만원)	
	합계	비중(%)	합계	비중(%)	합계	비중(%)
구로구(서울)	25,228	100.0	21,625	100.0	6,558,407	100.0
신도림동	1,581	6.3	1,705	7.9	373,766	5.7
구로1동	465	1.8	299	1.4	86,759	1.3
구로2동	6,107	24.2	6,306	29.2	2,966,645	45.2
구로3동	4,348	17.2	4,950	22.9	978,980	14.9
구로4동	1,386	5.5	781	3.6	188,051	2.9
구로5동	2,804	11.1	2,628	12.2	620,984	9.5
고척1동	1,512	6.0	1,098	5.1	353,900	5.4
고척2동	1,127	4.5	453	2.1	154,441	2.4
개봉1동	1,263	5.0	704	3.3	181,534	2.8

○ 읍면동별 주요항목 분포(구로구 계속)

	사업체수(개)		임금종사자수(명)		매출액(백만원)	
	합계	비중(%)	합계	비중(%)	합계	비중(%)
개봉2동	709	2.8	439	2.0	98,306	1.5
개봉3동	650	2.6	325	1.5	85,429	1.3
오류1동	798	3.2	531	2.5	130,579	2.0
오류2동	1,084	4.3	675	3.1	170,458	2.6
수궁동	737	2.9	447	2.1	105,444	1.6
가리봉동	657	2.6	284	1.3	63,131	1.0

○ 읍면동별 주요항목 분포(계속)

	사업체수(개)		임금종사자수(명)		매출액(백만원)	
	합계	비중(%)	합계	비중(%)	합계	비중(%)
수영구(부산)	9,124	100.0	6,117	100.0	1,185,835	100.0
남천1동	1,359	14.9	1,174	19.2	184,358	15.5
남천2동	501	5.5	293	4.8	42,702	3.6
수영동	1,020	11.2	587	9.6	120,910	10.2
망미1동	840	9.2	395	6.5	91,853	7.7
망미2동	462	5.1	286	4.7	81,232	6.9
광안1동	972	10.7	614	10.0	127,289	10.7
광안2동	957	10.5	783	12.8	122,496	10.3
광안3동	1,083	11.9	726	11.9	125,679	10.6
광안4동	469	5.1	274	4.5	54,465	4.6
민락동	1,461	16.0	985	16.1	234,851	19.8

○ 사업체규모별 층 분포

사업체수(개)	읍면동×중분류		읍면동×소분류	
	층수	비중(%)	층수	비중(%)
합계	535	100.0	1,513	100.0
≤ 10	194	36.3	854	56.4
≤ 20	88	16.5	261	17.3
≤ 30	49	9.2	121	8.0
< 50	58	10.8	128	8.5
≥ 50	146	27.3	149	9.9

○ 사업체규모별 읍면동 분포

	읍면동수		비중(%)	
	구로구	수영구	구로구	수영구
전체	15	10	100.0	100.0
규모3(<750)	5	3	33.3	30.0
규모4(<1400)	5	6	33.3	60.0
규모5(<2500)	2	1	13.3	10.0
규모6(≥2500)	3	-	20.0	-

* 읍면동별 사업체수 규모

규모	1	2	3	4	5	6
사업체수	<100	<250	<750	<1400	<2500	≥2500
백분위수	10%	25%	50%	75%	90%	

○ 사업체규모별 주요항목 분포

	사업체수		임금종사자수		매출액	
	계	비중(%)	계	비중(%)	계	비중(%)
구로구	25,228	100.0	21,625	100.0	6,558,407	100.0
규모3(<750)	3,218	12.8	1,794	8.3	439,069	6.7
규모4(<1400)	5,658	22.4	3,144	14.5	825,063	12.6
규모5(<2500)	3,093	12.3	2,803	13.0	727,666	11.1
규모6(≥2500)	13,259	52.6	13,884	64.2	4,566,609	69.6
수영구	9,124	100.0	6,117	100.0	1,185,835	100.0
규모3(<750)	1,432	15.7	853	13.9	178,399	15.0
규모4(<1400)	6,231	68.3	4,279	70.0	772,585	65.2
규모5(<2500)	1,461	16.0	985	16.1	234,851	19.8

○ 임금종사자규모별 주요항목 분포

	사업체수		임금종사자수		매출액	
	합계	비중(%)	합계	비중(%)	합계	비중(%)
구로구	25,228	100.0	21,625	100.0	6,558,407	100.0
0인	13,431	53.2	-	-	1,744,450	26.6
1인	5,490	21.8	5,490	25.4	1,429,328	21.8
2~4인	6,307	25.0	16,135	74.6	3,384,629	51.6
수영구	9,124	100.0	6,117	100.0	1,185,835	100.0
0인	5,611	61.5	-	-	439,101	37.0
1인	1,761	19.3	1,761	28.8	240,082	20.2
2~4인	1,752	19.2	4,356	71.2	506,652	42.7

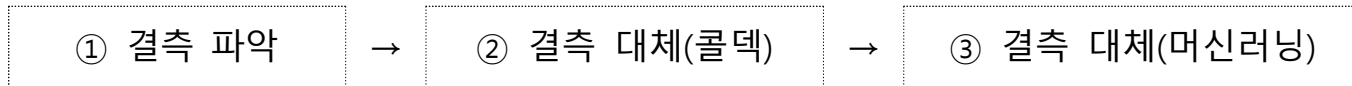
5 매출액 분석 및 대체(Imputation)

○ 주요 보조변수 탐색 분석

- ✓ (상관분석) 주요항목으로 활용할 종사자 기준 선정을 위한 산업중분류별 매출액과 임금종사자(총종사자) 상관분석(임금종사자 > 총종사자)
- ✓ (분산분석) 매출액과 주요 보조변수의 평균의 분산 분석
임금종사자규모, 조직구분이 평균의 유의미한 차이를 보임

종속변수	독립변수	수준	F Value	Pr > F
매출액	임금종사자규모	3	3309.04	<.0001
	조직구분	2	2420.72	<.0001

○ 매출액 결측 대체



- ① (결측파악) 매출액 결측 사업체(34,352개 중 1,898개 : 구로구 1,390개, 수영구 508개)
- ② (콜덱대체) 전년기준 매출액으로 대체(1,898개 중 43개(0.1%))
- ③ (머신러닝 대체) 제2차 경총 시험조사*(‘19년) 결과에 따라 머신러닝방법으로 대체
 - * 평균대체, 머신러닝대체, 중위수대체, 회귀대체, 계층적하트덱대체 중 머신러닝대체방법 (랜덤포레스트)의 대체결과값이 참값에 근사하게 나타남

⇒ 주요변수를 활용하여 랜덤포레스트 1000번 시행한 후 평균값으로 대체

사업체 수	결측수 (비중)	매출액					
		대체 전			대체 후		
		총계	CV	왜도	총계(증감률)	CV	왜도
34,352	1,855 (5.4)	7,748,375	174.3	3.1	8,301,616 (7.1)	167.9	3.2

< 머신러닝(랜덤포레스트) >

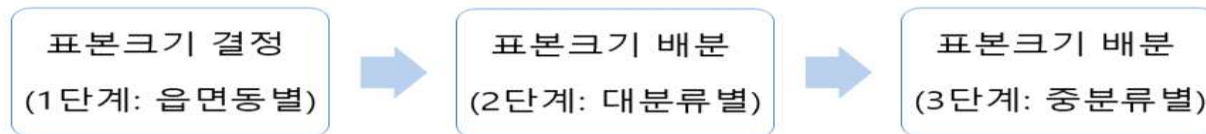
- ✓ 머신러닝 기법 중 하나로 Bootstrap과 Aggregation 방법의 혼합
즉 여러개의 의사결정트리를 랜덤으로 만들고 이들의 결과값을 조합하는 형태로
결측값을 대체함
- ✓ test data set이 얼마나 좋느냐에 따라 효율이 결정되나, 과정 설명이 어려운 단점이 있음
- ✓ 연속변수, 범주변수 모두 사용 가능
- ✓ 많은 결측에는 효율적이지 못함

6 증화 및 표본추출

○ 증화

- ✓ 읍면동(25) × 중분류(23) 통계생산 및 MD 제공 고려

○ 표본 배분



① 읍면동별 표본 배분(총표본수 요청 40,000개)

- ✓ 읍면동별 사업체수 제공근비례배분
- ✓ (전수) 모집단 사업체수 < 배분 사업체수
- ✓ (최소표본) 읍면동별 사업체수 규모에 따라 최소 표본크기 설정

규모1 (<100)	규모2 (<250)	규모3 (<750)	규모4 (<1400)	규모5 (<2500)	규모6 (≥2500)
30개	60개	80개	100개	125개,	150개

② 읍면동 내 대분류별 표본 배분

- ✓ 읍면동 내 대분류별 사업체수 제공근비례배분
- ✓ (전수) 읍면동별 모집단크기 < 125개 & 대분류 모집단 크기 < 20개,
읍면동별 모집단크기 ≥ 125개 & 대분류 모집단 크기 < 5개
- ✓ (최소표본) 표본크기가 100개 미만이면 8개, 100개 이상이면 12개

③ 읍면동×대분류 내 중분류별 표본 배분

- ✓ 읍면동×대분류 내 중분류별 사업체수 제공근비례배분
- ✓ (전수) 중분류별 모집단 사업체수가 10개 미만

○ 분류지표

✓ 읍면동×중분류(534개) 층별 일반선형모형(GLM) 분석

* (분석변수) 임금종사자규모(0인/1인/2-4인), 조직구분(개인/법인), 세세분류

< 분석결과 유의미한 층수 >

	전체	임금종사자규모	조직구분	세세분류
구로구	328	184	178	125
F 건설업	30	19	17	7
G 도매소매업	45	39	35	16
H 운수창고업	30	14	12	11
I 숙박음식점업	28	22	19	17
L 부동산업	15	4	3	13
M 전문과학기술서비스업	57	33	31	17
N 사업시설관리업	43	11	21	11
O 공공국방행정	5	5	5	5
R 예술여가서비스업	30	9	16	12
S 수리및기타개인서비스업	45	28	19	16

< 분석결과 유의미한 증수 >

	전체	임금종사자규모	조직구분	세세분류
수영구	206	111	111	78
F 건설업	20	10	10	8
G 도매소매업	29	23	16	12
H 운수창고업	20	9	9	6
I 숙박음식점업	18	17	13	12
L 부동산업	10	6	6	9
M 전문과학기술서비스업	32	14	15	11
N 사업시설관리업	27	6	11	11
R 예술여가서비스업	20	6	12	2
S 수리및기타개인서비스업	30	20	19	7

< 층별 R²값 평균 >

	임금종사자 규모	조직구분	세세분류
F 건설업	0.24	0.16	0.45
G 도매소매업	0.33	0.12	0.41
H 운수창고업	0.44	0.24	0.40
I 숙박음식점업	0.26	0.04	0.23
L 부동산업	0.08	0.05	0.36
M 전문과학기술서비스업	0.33	0.14	0.39
N 사업시설관리업	0.30	0.19	0.41
O 공공국방행정	0.00	0.00	0.00
R 예술여가서비스업	0.24	0.07	0.46
S 수리및기타개인서비스업	0.32	0.07	0.30

○ 표본추출

- ✓ 층별(읍면동×중분류) 임금종사자규모, 조직구분, 세세분류, 매출액(내림)순으로 정렬하여 계통추출

	조사 모집단 (a+b)	전수 (a)	표본추출률 (b)	표본수 (c)				조사 대상 (a+c)
					전수층	표본층	추출률(%)	
전체	65,043	30,691	34,352	9,939	919	9,020	28.9	40,630
전수	14,369	14,369	-	-	-	-	-	14,369
표본	50,674	16,322	34,352	9,939	919	9,020	28.9	26,261
구로구	47,419	22,191	25,228	6,507	533	5,974	25.8	28,698
수영구	17,104	7,980	9,124	3,432	386	3,046	37.6	11,412

Ⅲ 향후 고려사항

- 표본추출틀
 - ✓ SBR의 품질 제고 및 관리 방안

- 조사 및 항목 대체 방법
 - ✓ 물리적장소가 없는 사업체, 본사일괄조사 사업체 등
 - ✓ 현장조사 미대상 사업체의 항목 작성('20년 연구용역 진행중)

- 표본설계
 - ✓ 주요항목(매출액) 결측 대체시 유효표본수 확대 방안
 - ✓ 시범예행조사 결과 공표수준(읍면동×산업중분류)별 주요항목의 상대표준오차(RSE) 분석
 - ✓ 시범예행조사 응답률 검토