

1. ‘표본추출에서 60대 이상 응답자가 다른 연령대에 비해 상당히 많았다’는 해석에 대해

- 이번 에스티아이 서울시장 보궐선거 여론조사의 60대 이상 응답 샘플은, 여론조사심의위원회 기준인 가중값 배율 0.7~1.5 사이에서 적절하게 추출되었습니다.

가중값 배율이란?

여론조사시 인구비례와 정확히 맞게 샘플을 추출하기 힘들기 때문에, 그 범위를 정해놓은 것을 말함.

$$\text{조사완료 샘플 수} \times \text{가중값 배율} = \text{목표 샘플 수(인구비례)}$$

$$\text{목표 샘플 수(인구비례)} \div \text{조사완료 샘플 수} = \text{가중값 배율}$$

가중값 배율을 0.7 ~ 1.5로 한다는 의미는,

예를 들어 인구비율대로 계산했을 때 ‘60대 이상’이 전체 1,000명 중 250명이라고 한다면,

$$250 \div 0.7 = 357.1$$

$$250 \div 1.5 = 166.7$$

즉 167명 ~ 357명 사이에서 60대 이상 샘플을 추출해야 한다는 의미임.

- 2021년 2월 행정안전부 주민등록인구통계 비율대로 계산했을 때, 60대 이상 인구비율은 약 27.2%이고, 본 조사에서 60대 이상 샘플 수는 336명으로 가중값 배율은 0.81입니다.

목표 샘플 수	조사완료 샘플 수	가중값 배율
272	336	0.81

- 실제 0.8대 수준의 가중값 배율 적용은 타 조사에서도 여럿 나타나고 있습니다. 아래 참고 자료 1), 2)에서 보듯 60대 이상의 가중값 배율이 본 조사보다 더 낮은 조사도 있으며, 참고자료 3)의 경우에도 다른 연령대의 가중값 배율이 0.8대로 나타나고 있습니다. 참고자료 4)의 경우에도 60대 이상 샘플의 비율이 본 조사와 크게 차이하지 않습니다.

(참고 자료)

1) 3.8 - 3.12 리얼미터-YTN 정기조사 (3월 2주차 주간집계)

연령대	목표 샘플 수	조사완료 샘플 수	가중값 배율
60대	385	525	0.73
70대 이상	325	357	0.91
60대 이상 합계	710	882	0.80

위 조사는 총 2,500명 대상 조사로 60대 이상 샘플 비율이 35.3%임.

2) 3.9 - 3.10 리얼미터-오마이뉴스 전국 대통령선거 정당지지도

연령대	목표 샘플 수	조사완료 샘플 수	가중값 배율
60대	154	198	0.78
70대 이상	130	156	0.83
60대 이상 합계	284	354	0.80

위 조사는 총 1,000명 대상 조사로 60대 이상 샘플 비율이 35.4%임.

3) 3.6 - 3.7 원지코리아컨설팅-아시아경제 서울시장 보궐선거 조사

연령대	목표 샘플 수	조사완료 샘플 수	가중값 배율
40대	180	215	0.84

4) 3.13 - 3.14 리얼미터-문화일보 서울시장 보궐선거 조사

연령대	목표 샘플 수	조사완료 샘플 수	가중값 배율
60대 이상	272	325	0.84

위 조사는 총 1,000명 대상 조사로 60대 이상 샘플 비율이 32.5%임

- 원래 인구 비율 자체가 60대 이상이 27.2%로 가장 많기 때문에, 단순히 샘플수가 얼마나 많이 잡혔는지 보다는 가중값 배율을 살펴보아야 합니다.
- ‘표본추출에서 60대 이상이 다른 연령대에 비해 많았다’는 해석은, 여론조사 전체 결과에 ‘60대 이상 연령층의 응답결과’가 많이 반영되었다는 식으로 읽힐 여지가 있습니다.
그러나 본 조사를 비롯하여 현재 보도되는 모든 여론조사는 성, 연령, 지역별 가중치를 적용하여 통계를 산출하도록 되어 있습니다. 즉, 어떤 연령대의 응답이던지, 그 연령대의 인구 비율 만큼만 반영된다는 의미입니다.
본 조사에서 60대 이상이 336명이 응답했지만, 가중치를 적용하여 총 272명의 인구비율 만큼만 전체 결과 통계에 반영되었습니다. 그러므로 ‘60대 이상의 의견이 많이 반영된 결과다’는 해석은 맞지 않습니다.

2. 질문의 표현이 bias를 만들 수 있다는 분석에 대해

- 위 분석의 취지는, 본 조사 1번과 2번 문항 질문에서 ‘오세훈 후보가 범야권 단일후보로 출마한다면’, ‘안철수 후보가 범야권 단일후보로 출마한다면’과 같이 ‘오세훈’, ‘안철수’를 언급해 준 것이 영향을 미쳤다는 것으로 보입니다.
- 그러나 본 조사 설문에서는 아래와 같이 1번 문항 앞에 범여권 단일 후보로 박영선 후보를 가정한다는 설명을 먼저 해 주고 있으며, 이 설명에는 ‘오세훈’, ‘안철수’ 이름은 언급되지 않고 있습니다.

※ 이번 서울시장 보궐선거에서 더불어민주당과 열린민주당 등 범여권이 더불어민주당 **박영선 후보**로 단일화하고, 국민의힘, 국민의당 등이 범야권 후보 단일화를 한다는 가정 하에 질문드리겠습니다.

Q1. 만약 국민의힘 **오세훈 후보**가 범야권 단일후보로 출마한다면 귀하께서는 다음 중 누구를 지지하시겠습니까? (로테이션)

- ① 더불어민주당 박영선 후보
- ② 국민의힘 오세훈 후보
- ③ 기타 다른 후보
- ④ 지지 후보 없음
- ⑤ 잘 모르겠다

...

- 또한 본 조사와 같은 질문 방식은 타 조사에서도 나타나고 있습니다. 오히려 아래 참고자료의 타 기관조사에서는 ‘안철수’, ‘오세훈’의 이름이 더 많이 언급되고 있습니다.

(참고자료)

3.8 - 3.9 KBS-한국리서치 서울시장 보궐선거 조사

문8. 만약 오세훈 후보와 안철수 후보가 오세훈 후보로 단일화할 경우, 다음 후보 중 누구에게 투표하시겠습니까?

- 1. 더불어민주당 박영선
- 2. 국민의힘 오세훈
- 97. 기타 후보
- 98. 적합한 후보가 없다
- 99. 모름/무응답

문9. 만약 오세훈 후보와 안철수 후보가 안철수 후보로 단일화할 경우, 다음 후보 중 누구에게 투표하시겠습니까?

- 1. 더불어민주당 박영선
- 2. 국민의당 안철수
- 97. 기타 후보

98. 적합한 후보가 없다

99. 모름/무응답

- 본 조사는 서울특별시 여론조사심의위원회에 사전 문항 신고를 거친 후 진행된 조사로, 여론조사심의위원회에서도 본 문항을 검토 후, 문제가 없다고 판단하였습니다.
- 위의 내용들을 종합적으로 고려해볼 때 본 조사의 질문이 bias를 유발했다는 주장은 근거가 부족해 보입니다.