

자료에서 빠진 값은 0과 어떻게 다를까?

똑똑 · 에디터 모아 · 발행 2026.09.30

0은 해당 변수의 수량이나 측정값이 영으로 관찰된 값이고, 결측은 값을 알 수 없거나 기록하지 않은 표시입니다.

0은 해당 변수의 수량이나 측정값이 영으로 관찰된 값이고, 결측은 값을 알 수 없거나 기록하지 않은 표시입니다. 결측을 제외한 평균 11도도 “3일 평균”은 아닙니다.

쉬운 설명

0은 해당 변수의 수량이나 측정값이 영으로 관찰된 값이고, 결측은 값을 알 수 없거나 기록하지 않은 표시입니다. 응답 거부·측정기 고장·해당 없음은 서로 다른 결측 사유입니다.

핵심

- 0은 해당 변수의 수량이나 측정값이 영으로 관찰된 값이고, 결측은 값을 알 수 없거나 기록하지 않은 표시입니다.
- 응답 거부·측정기 고장·해당 없음은 서로 다른 결측 사유입니다.
- 결측을 제외한 평균 11도도 “3일 평균”은 아닙니다.

핵심 질문 · 대상과 범위부터 읽기

0은 해당 변수의 수량이나 측정값이 영으로 관찰된 값이고, 결측은 값을 알 수 없거나 기록하지 않은 표시입니다. 둘을 바꾸면 평균과 추세가 달라집니다. 기온 0°C는 실제로 측정할 수 있는 값입니다. “기온 자료 없음”은 측정기가 꺼졌거나 기록을 잃었다는 뜻일 수 있습니다. 둘을 같은 0으로 입력하면 없는 관측을 새로운 관측으로 만들어 버립니다.

개념 구분 · 같아 보이는 말을 나누기

응답 거부·측정기 고장·해당 없음은 서로 다른 결측 사유입니다. 결측을 모두 0으로 넣는 것도, 조용히 삭제하는 것도 근거가 필요합니다. 10도와 12도가 기록된 이들의 평균은 11도입니다. 화요일 기록을 알 수 없으므로 월~수요일 3일 평균은 계산할 수 없습니다. 화요일을 0도로 넣어 7.3도라 보고하면 실제보다 낮게 보이는 숫자를 발명한 셈입니다.

적용 사례 · 가상의 장면에서 판단하기

가상 기온 기록이 월요일 10도, 화요일 결측, 수요일 12도라면 화요일을 0도로 넣어 평균 7.3도를 계산하는 것은 근거 없는 대체입니다. 결측의 이유를 남기면 대체가 달라집니다. 조사 문항이 해당하지 않는 사람에게 빈칸이 생긴 경우와 응답을 거부한 경우는 같은 현상이 아닙니다. 후자에서는 누가 빠졌는지에 따라 비율이 치우칠 수 있습니다.

비교표 · 이 글의 세 가지 판단 기준

확인 항목	본문에서 확인한 것	읽을 때 주의할 점
핵심 범위	0은 실제 관측·집계한 값이 영(예: 대출 0건 또는 기온 0°C)이라는 뜻이고, 결측은 관측값을 얻지 못했거나 공개하지 않았다는 표시입니다.	응답 거부·측정기 고장·해당 없음은 서로 다른 결측 사유입니다.
가상 사례	가상 기온 기록이 월요일 10도, 화요일 결측, 수요일 12도라면 화요일을 0도로 넣어 평균 7.3도를 계산하는 것은 근거 없는 대체입니다.	결측을 제외한 평균 11도도 “3일 평균”은 아닙니다.
후속 확인	자료 코드북에서 결측 기호를 찾고 0·해당 없음·미응답을 각각 다른 표시로 남깁니다.	결측을 추정해 채웠다면 추정 방법을 공개합니다.

본문 설명을 비교하기 위해 직접 만든 표입니다. 기관 원문의 도표나 실제 조사 결과는 아닙니다.

한 단계 더 · 조건을 바꾸면 답도 달라집니다

그래프에서는 자료가 없는 날을 선으로 자동 연결하면 연속적인 변화가 있었던 것처럼 보일 수 있습니다. 빈 구간이나 별도 기호로 표시하고 분석에 포함한 관측 수를 적습니다. 결측을 추정해 채웠다면 추정 방법을 공개합니다. 결측을 제외한 평균 11도도 “3일 평균”은 아닙니다. 관측된 이들의 평균이라는 범위를 밝혀야 합니다.

깊이 읽기 · 다른 조건으로 시험하기

자료가 10, 빈칸, 12라면 결측을 제외한 평균 11은 관측된 값의 평균입니다. 빈칸을 0으로 넣은 7.33은 다른 가정을 추가한 값입니다. 빈칸이 실제 11이었다면 3일 평균은 11이지만 이를 알 근거는 없으므로 추정과 관측을 섞지 않습니다. 시간대별 전력 사용 자료에서 밤 2시 값 0은 기기가 꺼졌다는 관측일 수도 있고 통신 장애로 기록이 빠진 것일 수도 있습니다. 자료 제공자의 결측 기호와 설명을 확인하고, 의심스러운 0이 어떻게 생성되었는지 물어야 합니다.

직접 확인 · 원문과 사례에 적용하기

자료 코드북에서 결측 기호를 찾고 0·해당 없음·미응답을 각각 다른 표시로 남깁니다. 관측된 이들의 평균은 11도입니다. 3일 전체의 참평균은 화요일 값이 없어서 알 수 없습니다.

직접 생각해 보기

10, 결측, 12를 0으로 대체하지 않을 때 계산 가능한 평균과 이름은?

확인 설명

관측된 이틀의 평균은 11도입니다. 3일 전체의 참평균은 화요일 값이 없어서 알 수 없습니다.

자료에서 빠진 값은 0과 어떻게 다를까?

0은 해당 변수의 수량이나 측정값이 영으로 관찰된 값이고, 결측은 값을 알 수 없거나 기록하지 않은 표시입니다. 둘을 바꾸면 평균과 추세가 달라집니다. 응답 거부·측정기 고장·해당 없음은 서로 다른 결측 사유입니다. 결측을 모두 0으로 넣는 것도, 조용히 삭제하는 것도 근거가 필요합니다.

10, 결측, 12를 0으로 대체하지 않을 때 계산 가능한 평균과 이름은?

관측된 이틀의 평균은 11도입니다. 3일 전체의 참평균은 화요일 값이 없어서 알 수 없습니다.

빈칸을 0으로 넣어 평균을 계산해도 되나요?

결측을 제외한 평균 11도도 “3일 평균”은 아닙니다. 관측된 이틀의 평균이라는 범위를 밝혀야 합니다. 자료 코드북에서 결측 기호를 찾고 0·해당 없음·미응답을 각각 다른 표시로 남깁니다.

원문과 출처

• 미국 인구조사국 · 결측 자료 연구 — <https://www.census.gov/topics/research/stat-research/expertise/missing-data.html>: 결측은 관측한 0과 다르며 별도 처리 대상; 본문의 가상 수치나 학교 사례를 제공한 자료는 아닙니다.

• 미국 인구조사국 · 분포 읽기 학습 자료 — <https://www.census.gov/programs-surveys/sis/activities/math/census-in-counties.html>: 중앙값·평균과 분포를 함께 비교하는 방법; 본문의 가상 수치나 학교 사례를 제공한 자료는 아닙니다.

제작 방식과 정정 안내: <https://www.dokdok.co/about#editorial>

온라인 원문 · 수정된 내용은 아래 페이지에서 확인하세요.

<https://www.dokdok.co/reading/missing-versus-zero>