

# The average salary is a lie.

D8A.ACADEMY/RESOURCES/MEAN-MEDIAN-MODE

## THE THREE

|        |                   |
|--------|-------------------|
| mean   | sum / count       |
| median | the middle value  |
| mode   | the most frequent |

## WHEN EACH LIES

|                 |                     |
|-----------------|---------------------|
| mean + outlier  | ← one CEO skews it. |
| median + tiny n | jumpy under 30 rows |
| mode + decimals | every value unique  |

## IN SQL

|                     |                   |
|---------------------|-------------------|
| AVG(salary)         | the mean          |
| PERCENTILE_CONT(.5) | the median        |
| GROUP BY + LIMIT 1  | the mode, by hand |

## IN PANDAS

|             |                     |
|-------------|---------------------|
| s.mean()    | mean                |
| s.median()  | median              |
| s.mode()[0] | mode, first if tied |

## PICK ONE

|                   |              |
|-------------------|--------------|
| salaries, prices  | median       |
| categories, sizes | mode         |
| symmetric data    | mean is fine |

## GOTCHAS

|                    |                     |
|--------------------|---------------------|
| mean > median      | ← right skew. look. |
| mean of ratios     | weight it first     |
| "the average user" | often nobody        |

## WORTH MEMORIZING

### 01 Same team, two truths

Nine analysts earn 45–52k. One director earns 260k. Watch the mean work for the director.

```

SELECT
  AVG(salary)           AS mean_salary,   -- 72,400
  PERCENTILE_CONT(0.5)
    WITHIN GROUP (ORDER BY salary)
                    AS median_salary  -- 48,000
FROM team;
```

→ Both numbers are correct. Only one describes a person who exists. Report medians for money.

### 02 Trap: the wrong average

Each average has one failure mode, and dashboards hit all three.

| YOU USED | IT BREAKS WHEN                 | REACH FOR             |
|----------|--------------------------------|-----------------------|
| mean     | outliers exist (money, delays) | median                |
| median   | you need totals back           | mean, plus a note     |
| mode     | values are continuous          | binned mode or median |

→ Cheap tell in one second: mean far above median means right skew, a few big values dragging it up.