1 | Name | Height(m) | No. of Siblings | Eye colour | Highest Degree

Jeanne | 1.75... | 2 | brown | PhD

⋮ | ⋮ | ⋮ | ⋮ | ⋮

"Identifier" | · numerical · continuous IR · with unit. · ordered · distance · bounded | · numerical · discrete/quantities (IN, ZI) | "categorical" | "ordinal"

2 get data: No of Siblings: $\underline{x} = (0, 1, 1, 2, 1, 0, 0, \ldots)$ item x_i

$x_2 = 1$

0	1	2
III	III	I

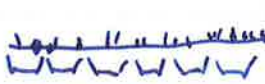
→ sample size: $n = \sum_j n_j$

absolute frequency of $x_i = 0$: $= n_{(x_i = 0)} = 3$

$x_i = 1$ $= 3$

relative frequency $f_{x=0} = \frac{n_{(x=0)}}{n}$ etc.

↳ between 0 and 1. often in %

3  → Temperature → "Digitize" (Audio, ...)/Quantize

cont → discrete

$n_{(24 \leq T < 25)}$

4 (arithmetic) mean / "avg"

· if all were the same & the sum was the same

$\langle x \rangle = \bar{x} = \frac{\sum_i x_i}{n} = \frac{\sum_i x_i}{\sum_i 1} = \frac{293}{7} = 41.9$

max: 50, 55, 45

min: 50, 35, 45 | 15

mode: ↑ → the most frequent value

· (2, 4)

weighted average. coffees/day 0, 1, 2, 3, 4 4 2

median: the "middle value"
at least 50% are $\geq \tilde{x}$
& at least 50% are $\leq \tilde{x}$

→ 50 n = odd.

→ 6 points → 45 .. 50 → $\frac{45+50}{2}$ "midpoint"
50%

~~mode: most frequent value.~~

5 Quantiles: e.g. $\tilde{x}_{25\%}$ 25th percentile
first quartile

↳ at least 25% below or equal to $x_{25\%}$
of the points

at least 75% above or equal to $x_{25\%}$

find it: $n * 25\% = q$ → not an integer → take the $\tilde{x}_{25\%} = x_2$
next above (ceiling)

e.g. (1, 2, 3, 4, 5) $1.5 \rightarrow \boxed{2}$
20%

→ is an integer → average x_q and x_{q+1}

6

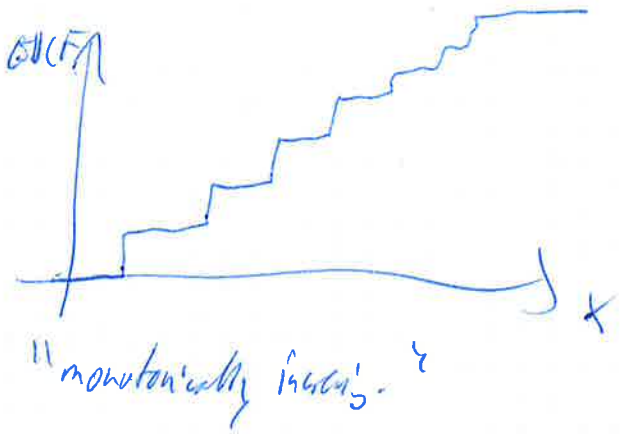
$$\begin{cases} \text{if } n \cdot \alpha = \text{integer} & \tilde{x}_\alpha = x_{\lceil n \cdot \alpha \rceil} \\ \text{else} & \tilde{x}_\alpha = \frac{x_{n\cdot\alpha} + x_{n\cdot\alpha}}{2} \end{cases}$$

(7) ECDF(x) : fraction of data points $\leq x$

$ECDF(47) = \frac{3}{7}$

$ECDF(7) = 0$

$ECDF(60) = 1$



"monotonically increasing"

1 2 3 4 5 6 7

(8) → max, min, median, $\tilde{x}_{25\%}$, $\tilde{x}_{75\%}$

$n \cdot \alpha = 1.75 \rightarrow 2$ → 21

$n \cdot \alpha = 5.25 \rightarrow 6$ → 5 ✓

(9) → outliers any point $> \tilde{x}_{75} + 1.5\,(\tilde{x}_{75\%} - \tilde{x}_{25\%})$ (max limit)

or $< \tilde{x}_{25} - 1.5\,(IQR)$ (min limit)

INTER QUARTILE RANGE

↳ T ⊥ at max/min limit + extra "outliers"