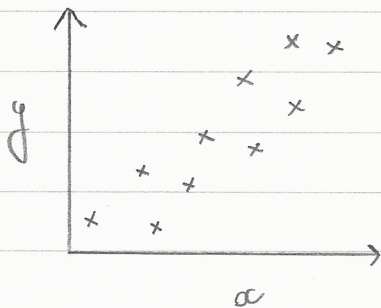# Bivariate Data

Product
- You can have strong, moderate, weak, or perfect correlation.
- The Product Moments Correlation Coefficient (PMCC) is a measure of this:

$$-1 \leq r \leq 1$$

$r = 1$ means all points lie on line w/ positive gradient

$r = 0$ means no correlation

$r = -1$ means all points lie on line w/ negative gradient

- Around 0.5 is moderate, 0.02 is weak.

Essentially, you have a mean point $\left(\frac{\Sigma x}{n}, \frac{\Sigma y}{n} \right)$ and for each point, you measure the deviation from the mean point.

⇩

(one point can have a large effect).

Line of Regression (aka line of best fit)
- If the PMCC can be found, so can the equation of the line of regression. Both can be found using FX 991 EX

$$y = a + bx$$

→ rate of change (defined in context)

↘ initial value (defined in context)

Interpolation & Extrapolation
- Interpolation is using data within the range.
- Extrapolation is predicting using data ∴ dangerous ∵ there is no data.

Histograms
+ Shows distribution away.
+ Can find frequency easily.
+ Estimate mean
+ Find modal class easily.

− Doesn't indicate raw data.
≈ Estimates for mean

Stem & leaf
+ Show groups and clusters easily.
+ See outliers easily.

− Cannot work out things like mean easily.

---

24 students in a gym class balance on one leg for as long as they can, and their times are recorded to the **nearest second.**

a i Find the missing frequency. **9**
   ii Give the width and height of the missing bar on the histogram. $4 = $ width
        $12 \div 4 = 3$ ~ height

b Estimate the number of students who could stand on one leg for

   i Less than 5.5 seconds, $\frac{9}{4}$ 0.75
       $\frac{18}{8} + \frac{9}{4} \approx 10$
   ii More than 10 seconds.
       $6 + \frac{1.5}{4} \times 6 = 6 + \frac{9}{4}$
       $= 6 + 8.25$
       $= 14.25$
       ≈ 14

area $\propto$ f

Look at scale factor
− perhaps to used to fit nicely

anything would round to 5

Time (seconds)	Continuity correction	Frequency	Area
0–4	$0 \leq t < 4.5$	9	18
5–7	$4.5 \leq t < 7.5$	3	6
8–11	$7.5 \leq t < 11.5$	6	12
12–16	$11.5 \leq t < 16.5$	6	area =
≥17	$16.5 \leq t$	0	

− 24

$4.5 \times 4 = 18$

6

12

$5 \times 2.4 = 12$

18 implies scale factor 2

$6 \times \frac{5}{2} = 12$