Correlation

Correlation looks at the link or relationship between two variables.

- Zero correlation —▸there is no relationship between the points (points are scattered everywhere)
- Positive correlation —▸both variables increase in the same direction
- Negative correlation—▸the variables change in opposite directions

Correlation Coefficient - r

The correlation coefficient ranges in value between [-1,1].

- $r = -1$: very strong (or perfect) negative correlation
- $r = 1$: very strong (or perfect) positive correlation
- $r = 0$: zero correlation

Please copy Pg. 255 from you yellow book into your notes.

Estimating the Correlation Coefficient - (r)

Apart from estimating the correlation coefficient by eye, a graphical method can also be used.

Method: ("line of best fit")

1) Draw a dotted line through the center of a set of points on a scatter plot.

2) Draw a rectangle, the smallest one possible, that contains all the points, except outliers.

(Outlier - A point that is far away from the other points in a scatter plot)

3) Measure the sides accurately (yes - you need a ruler!) and calculate the correlation coefficient using the following formula:

$$r \approx \pm\left(1 - \frac{\text{length of the shorter side}}{\text{length of the longer side}}\right)$$

* Use + when the correlation is positive

* Use - when the correlation is negative