

$$Y = \beta_0 + \beta_1 X$$

Slope and intercept estimates

Correlation Coefficient

$$R = \frac{S_{xy}}{S_{xx}}$$

$r$ , the sample correlation coefficient, is an estimate for ('rho') the population correlation coefficient.

By definition

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

For a group of 15 students the following table shows the average numbers of hours per week spent on study and their final results in the corresponding examination.

No. of hours studied ( $x$ )	16	17.5	11.5	13.5	15	12.5	20.5
Exam mark ( $y$ )	77	85	48	59	75	41	95
No. of hours studied ( $x$ )	16.5	13.5	22	18.5	17	19.5	19.5
Exam mark ( $y$ )	80	70	99	85	83	97	89

## Some useful values:

Summary statistics for these data are:

- ▶ Sum of **x** data = 247.5
- ▶ Sum of squares of **x** data = 4218.75
- ▶ Sum of **y** data = 1155
- ▶ Sum of squares of **y** data = 92999
- ▶ Sum of products of **x** and **y** data = 19750.5

- ▶ Calculate the least squares regression line of  $y$  on  $x$ .
- ▶ Use the calculated line to predict examination marks for students who studied for (i) 16 hours. (ii) 20 hours.
- ▶ State, without calculation, which of these predictions is likely to be the more accurate. Explain your choice.