

UNIT 20 *Statistics*

Overhead Slides

Overhead Slides

- 20.1 Box and Whisker Plots
- 20.2 Time Series Analysis
- 20.3 Correlation
- 20.4 Regression Lines
- 20.5 Spearman's Rank Correlation Coefficient
- 20.6 Standardised Scores

OS 20.1

Box and Whisker Plots

The goals scored in the first 11 football matches played by a *Premier League* team were:

1 0 4 2 2 3 1 2 5 0 1

These data can be represented using a *box and whisker plot*.

1. Reorder the data, starting with the smallest.

2. Identify:

	smallest value	<input style="width: 50px; height: 30px;" type="text"/>	largest value	<input style="width: 50px; height: 30px;" type="text"/>
	lower quartile	<input style="width: 50px; height: 30px;" type="text"/>	upper quartile	<input style="width: 50px; height: 30px;" type="text"/>
	median	<input style="width: 50px; height: 30px;" type="text"/>		

3. Construct a *box and whisker plot* above the scale below.



Construct a *box and whisker plot* below the scale for a team scoring

3 5 2 1 3 2 2 4 6 2 1

Compare the two sets of data.

OS 20.2

Time Series Analysis

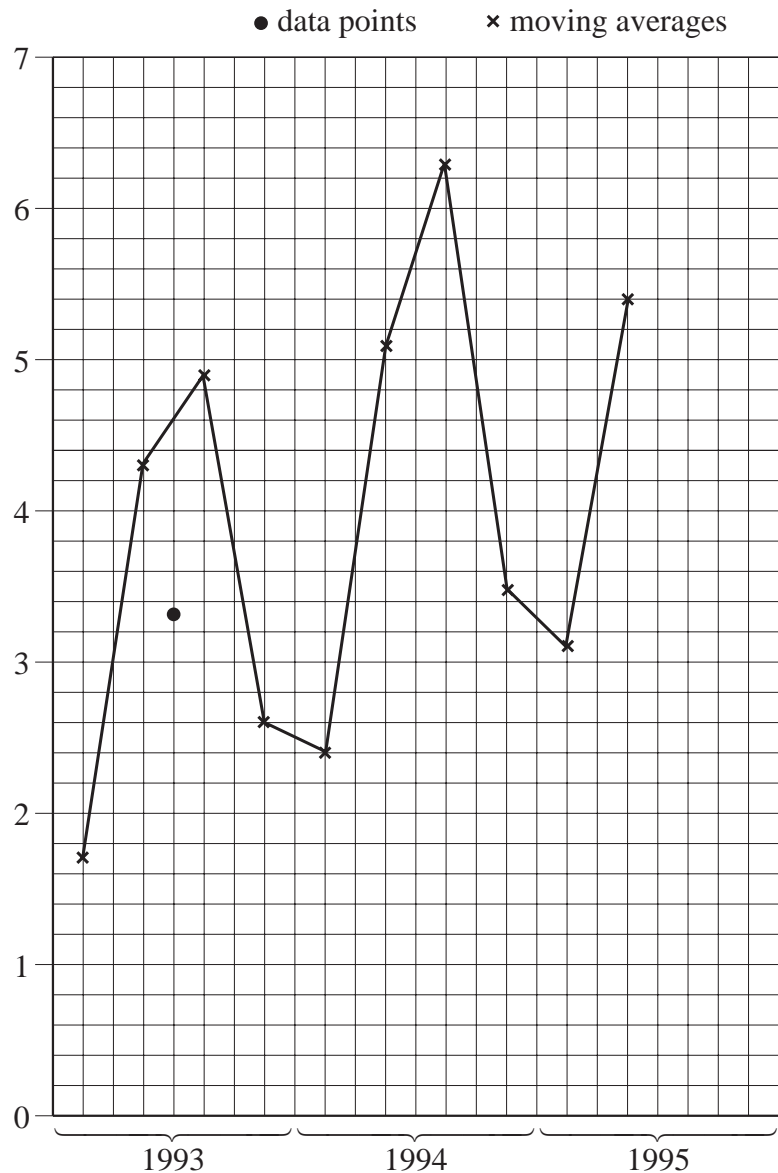
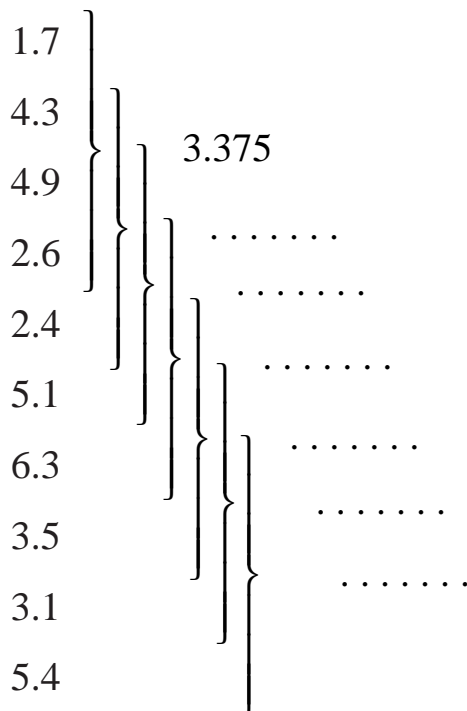
Sunshine records for a British holiday resort

(average number of hours per day)

	Jan – Mar	Apr – Jun	Jul – Sep	Oct – Dec
1993	1.7	4.3	4.9	2.6
1994	2.4	5.1	6.3	3.5
1995	3.1	5.4	?	?

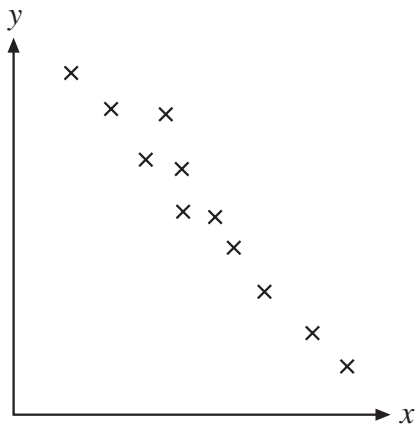
Complete the 4-point moving averages, plot them and estimate the value of the next two data points.

4-point moving averages

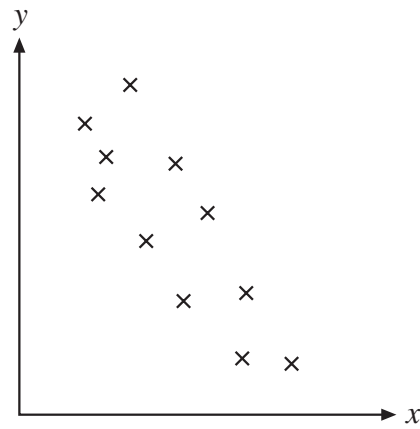


OS 20.3

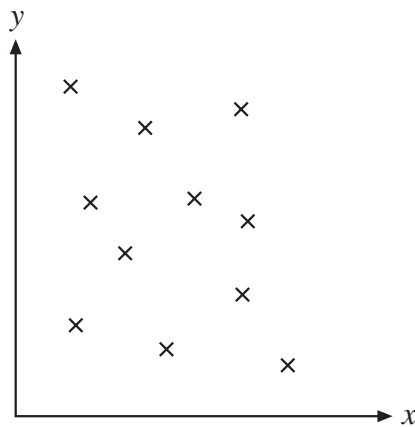
Correlation



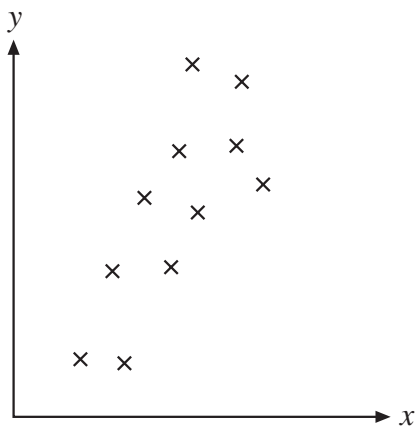
Strong *negative* correlation



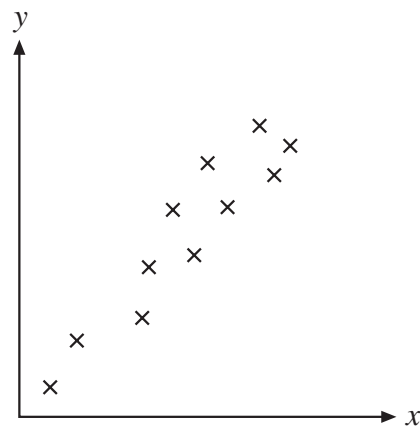
Weak *negative* correlation



No correlation



Weak *positive* correlation



Strong *positive* correlation

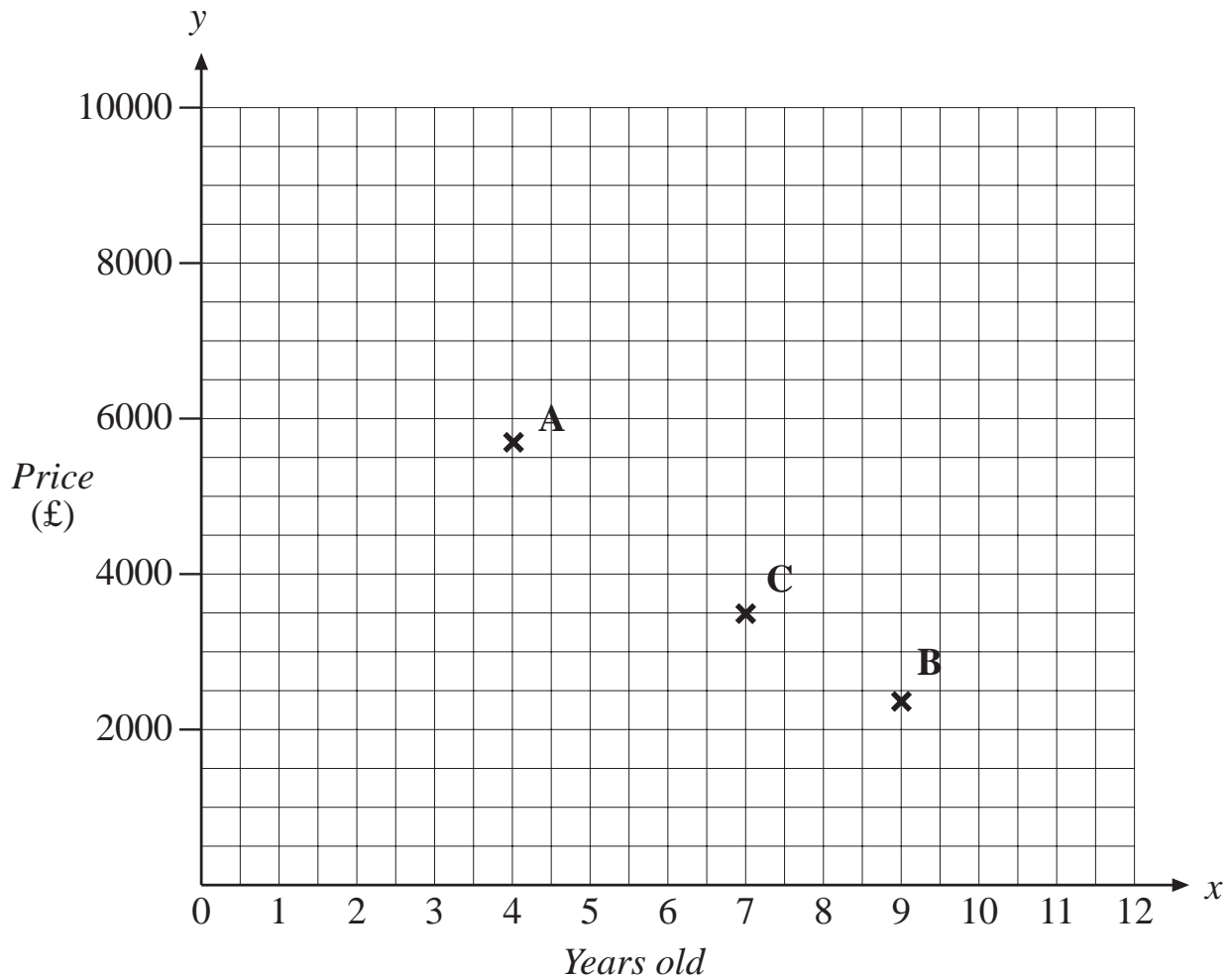
OS 20.4

Regression Lines

A garage has six *Ford Escort* cars for sale. The prices and ages are:

CAR	A	B	C	D	E	F
Age (years)	4	9	7	3	6	8
Price (£)	5700	2200	3500	6500	4180	2760

Complete the scatter diagram below.



Draw a line of best fit.

- Use it to estimate
- the price of a 1-year old *Ford Escort*
 - the price of a 12-year old *Ford Escort*.

OS 20.5

Spearman's Rank Correlation Coefficient

Attendances at some of the most popular tourist attractions in the UK are shown below for the years 1991 and 1996.

Tourist Attraction	1991	1996	Ranks		<i>d</i>	<i>d</i>²
			1991	1996		
<i>Madame Tussard's</i>	2.2	2.4				
<i>Alton Towers</i>	2.0	3.1				
<i>Tower of London</i>	1.9	1.6				
<i>National History Museum</i>	1.6	1.2				
<i>Chessington World of Adventure</i>	1.4	1.7				
<i>Science Museum</i>	1.4	1.3				
<i>London Zoo</i>	1.1	1.5				

Rank these results and use the formula

$$r = 1 - \frac{6\sum d^2}{n(n^2 - 1)} \quad (\text{with } n = 7)$$

to find *Spearman's Rank Correlation Coefficient*,

$$r = 1 - \frac{6 \times}{\times} =$$

OS 20.6*Standardised Scores*

Two components of an assessment are shown in the table:

- a particular candidate's result
- everyone's results

<i>Paper</i>	<i>Particular Candidate</i>	<i>ALL</i>	
		<i>Mean</i>	<i>Standard deviation</i>
1	63	55	12
2	75	65	20

Use the formula to convert the candidate's score on each paper to a standardised score, using a *mean* of 50 and a *standard deviation* of 15.

Formula

$$\text{Standardised score} = 50 + \left(\frac{\text{actual mark} - \text{mean mark}}{\text{standard deviation}} \times 15 \right)$$

Paper 1

Standardised score =

Paper 2

Standardised score =

On which paper did the candidate perform better?