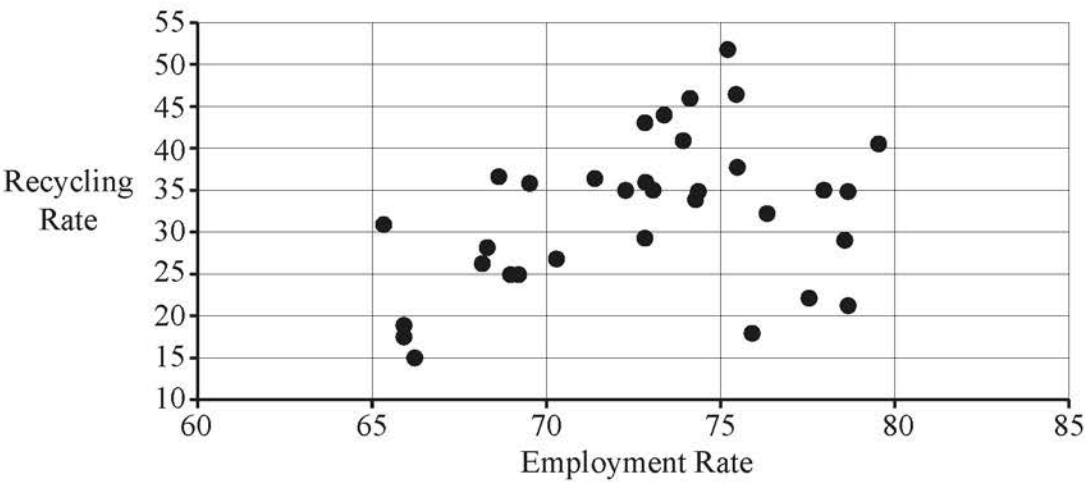


**13** The pre-release material contains information concerning median house prices, recycling rates and employment rates. Fig. 13.1 shows a scatter diagram of recycling rate against employment rate for a random sample of 33 regions.



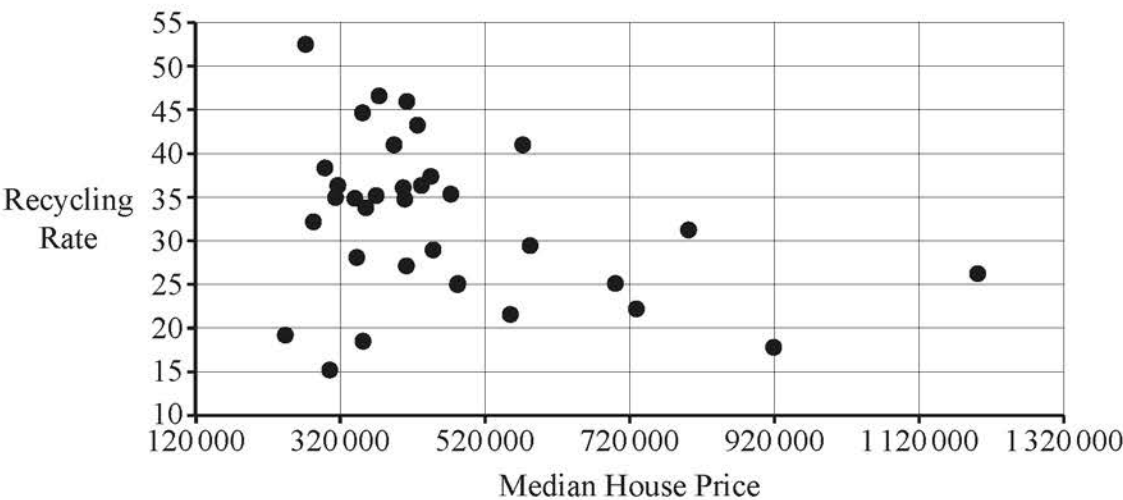
**Fig. 13.1**

The product moment correlation coefficient for this sample is 0.37154 and the associated  $p$ -value is 0.033.

Lee conducts a hypothesis test at the 5% level to test whether there is any evidence to suggest there is positive correlation between recycling rate and employment rate. He concludes that there is no evidence to suggest positive correlation because  $0.033 \approx 0$  and  $0.37154 > 0.05$ .

**(a)** Explain whether Lee’s reasoning is correct. **[2]**

Fig. 13.2 shows a scatter diagram of recycling rate against median house price for a random sample of 33 regions.



**Fig. 13.2**

The product moment correlation coefficient for this sample is  $-0.33278$  and the associated  $p$ -value is  $0.058$ .

Fig. 13.3 shows summary statistics for the median house prices for the data in this sample.

Statistics	
n	33
Mean	465467.9697
$\sigma$	201236.1345
s	204356.2606
$\Sigma x$	15360443
$\Sigma x^2$	8486161617387
Min	243500
Q1	342500
Median	410000
Q3	521000
Max	1200000

**Fig. 13.3**

- (b) Use the information in Fig. 13.3 and Fig. 13.2 to show that there are at least two outliers. [2]
- (c) Describe the effect of removing the outliers on
- the product moment correlation coefficient between recycling rate and median house price,
  - the  $p$ -value associated with this correlation coefficient,
- in each case explaining your answer. [2]

All 33 items in the sample are areas in London. A student suggests that it is very unlikely that only areas in London would be selected in a random sample.

- (d) Use your knowledge of the pre-release material to explain whether you think the student's suggestion is reasonable. [1]