

- 4 Two groups of students have carried out a fieldwork study about the relationship between traffic congestion and land use intensity in urban areas. Tables 4a and b show the data collected by them at different sites in an urban district on a weekday.

**Table 4a Data collected by Team A**

Survey site	Recorded by	Plot ratio	Average waiting time for vehicles to pass an intersection (sec.)
1	Peter at 8 am	10.1	93
2	Mary at 8 am	8.6	83
3	David at 8 am	6.5	68
4	Victor at 8 am	8.3	75
5	Ivan at 8 am	9.3	78
6	Irene at 8 am	9.6	80
7	Joyce at 8 am	7.8	80
8	Dolly at 8 am	7.9	75

**Table 4b Data collected by Team B**

Survey site	Recorded by	Plot ratio	Average waiting time for vehicles to pass an intersection (sec.)
9	Fanny at 8 am	6.3	57
10	Carol at 8 am	4.2	30
11	Jim at 8 am	2.0	18
12	Iris at 8 am	10.5	90
13	Brian at 8 am	4.2	35
14	Jennifer at 8 am	9.3	82
15	Keith at 8 am	8.0	76
16	Janet at 8 am	3.6	34

- a i Set a hypothesis for this field study. (2 marks)
- ii Refer to the hypothesis set by you. Which is the independent variable? Which is the dependent variable? (1 mark)
- b Refer to Tables 4a and b on p. 139.
- i Which set of data is more useful for testing the hypothesis? (1 mark)
- ii Explain your choice in Question bi with data support. (2 marks)
- c i With reference to the data set chosen in Question bi, calculate the Spearman's rank correlation coefficient between the two variables. Show your calculations. (3 marks)

$$r = 1 - \frac{6 \sum d^2}{n^3 - n}$$

- ii Is the hypothesis accepted or rejected? Explain your answer. (3 marks)
- d To what extent does the choice of time and day to carry out the fieldwork study affect the reliability of the data? (6 marks)

## Answers

- 4 a i • traffic congestion is not more serious with increasing land use intensity/  
waiting time for vehicles does not increase with plot ratio ( $H_0$ )
- OR
- traffic congestion is more serious with increasing land use intensity/waiting time for vehicles increases with plot ratio ( $H_1$ ) 2 (2)
- (One mark is given to correct variables: 'traffic congestion/waiting time for vehicles' and 'land use intensity/plot ratio'; another mark is given to the relationship: 'is not more serious' or 'does not increase'/'is more serious' or 'increases')
- ii • independent variable: land use intensity/  
plot ratio 0.5
- dependent variable: traffic congestion/  
waiting time for vehicles 0.5 (1)
- b i Table 4b/Data collected by Team B 1 (1)
- ii • plot ratio is an indicator of land use intensity 1
- the range of plot ratio in Table 4b (2–10.5) is larger than that in Table 4a (6.5–10.1) 1
- that means there are more variations in land use intensity in Table 4b 1
- the students can thus find out the effects of different land use intensity on traffic congestion/waiting time for vehicles 1 (2)

c i

Plot ratio	Rank	Average waiting time	Rank	$d$	$d^2$
2.0	1	18	1	0	0
3.6	2	34	3	-1	1
4.2	3.5	30	2	1.5	2.25
4.2	3.5	35	4	-0.5	0.25
6.3	5	57	5	0	0
8.0	6	76	6	0	0
9.3	7	82	7	0	0
10.5	8	90	8	0	0

$$\begin{aligned}
 r &= 1 - [(6 \times 3.5) \div (8^3 - 8)] \\
 &= 1 - (21 \div 504) \\
 &= 0.96 \qquad \qquad \qquad 3 \ (3)
 \end{aligned}$$

- ii
- the null hypothesis ( $H_0$ ) is rejected/the alternative hypothesis ( $H_1$ ) is accepted 1 (1)
  - the high value of the coefficient (0.96) shows that the sample observation is not a result purely of chance 1
  - there is a very strong relationship between the variables 1
  - the positive value indicates that the relationship is positive 1
  - this means that traffic congestion is more serious with increasing land use intensity/waiting time for vehicles increases with plot ratio 1 (2)

- d
- in this study, data about the waiting time for vehicles are collected to reflect the traffic conditions
  - as the waiting time is both time- and day-dependent, the choice of time and day will affect the reliability of the data collected

Arguments on how the choice of time affects the data collected:

- first, data should be collected at the same time
- to keep the time factor constant
- the data is not comparable if it is collected at different times
- second, data should be collected during rush hours
- such as 8 am or 6 pm
- traffic conditions vary at different times of a day
- waiting time is usually very short at night or in the early morning
- there is no traffic congestion at all even if land use intensity is very high in these non-peak hours

Arguments on how the choice of day affects the data collected:

- data should be collected on weekdays
- business trips/commuting are frequent on weekdays
- transport flow is high and traffic congestion is common
- waiting time is therefore higher on weekdays than at weekends/on holidays
- traffic congestion is less serious at weekends/on holidays even if land use intensity is very high

Other factors affecting the data collected:

- types of land use
- type/width/capacity of the roads
- time for traffic lights to change signal
- pedestrian flow
- traffic accidents/construction work on roads

(or any reasonable answers) (Max. 6)

(Arguments should be well-elaborated, with appropriate geographical concepts and terms.)