

Solutions to analysis questions

Question 1

Part 1

$$\bar{x} \pm 2s = 5.93 \pm 2(.21)$$

$$5.51 \rightarrow 6.35$$

mean = 5.47%
standard deviation = 0.19%

The mean Basic rate (5.47%) is lower than the mean Standard rate (5.93%) and therefore better for customers.

$$z = \frac{x - \bar{x}}{s}$$

$$= \frac{5.56 - 5.47}{0.19}$$

$$= 0.47$$

Part 2

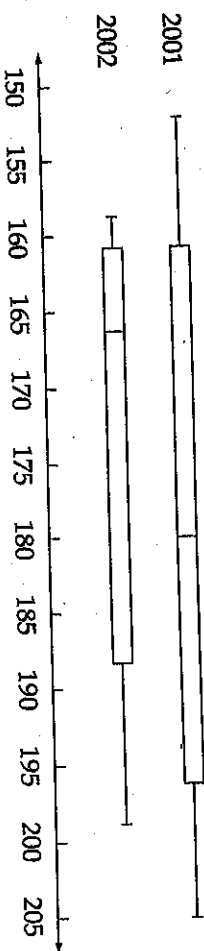
The z score for Commonwealth Bank's Basic Rate is lower. Since low interest rates are desirable when lending, Commonwealth Bank is more competitive in the Basic Rate market.

Part 2

- Independent variable is Standard rate.
- Basic rate = $0.79(6.05) + 0.72$
 $= 5.4995 \approx 5.50\%$
- $4.8 = 0.79(\text{Standard}) + 0.72$
 $4.08 = 0.79(\text{Standard})$
Standard rate $\approx 5.15\%$

Part 3

Prices in thousands of dollars



b. 2001: Symmetrical

2002: Positively skewed

c. 2001: Median = \$180 000

2002: Median = \$166 000

Median price for 2002 is \$14 000 lower than median price for 2001.

d. 2001: IQR = 195 000 - 161 000 = \$34 000

2002: IQR = 188 000 - 161 000 = \$27 000

The 2001 prices have a greater spread than the 2002 prices.

Question 2

Part 1

a. i. 32

ii. 980

b. $\frac{22}{912} = .0241 = 2.41\%$

Part 2

a. Teachers $\frac{22}{68} \approx 32\%$

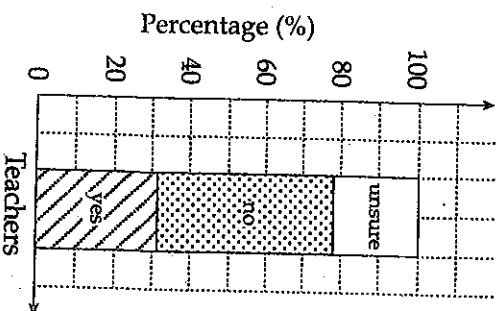
$\frac{32}{68} \approx 47\%$

$\frac{14}{68} \approx 21\%$

Yes

No

Unsure



b. As more students were surveyed, you cannot compare the number of students saying 'No' with the number of teachers.

47% of teachers said 'No', 4% of students said 'No', therefore a larger proportion of teachers said 'No'.

Part 3

- a. Seasonal, no trend
- b. 7-day cycle with increased hours on Friday, Saturday and Sunday.
- c. i. 4/02: $\frac{2.9 + 0.5 + 0.6}{3} \approx 1.3$
 ii. 5/02: Median (2.9, 0.5, 0.6, 0.8, 1.2) : 0.8
 14/02: Median (0.6, 0.7, 1.3, 2.1, 3.3) : 1.3
- d. 7 points are needed to smooth the data as there is a 7-day (weekly) cycle.

Question 3

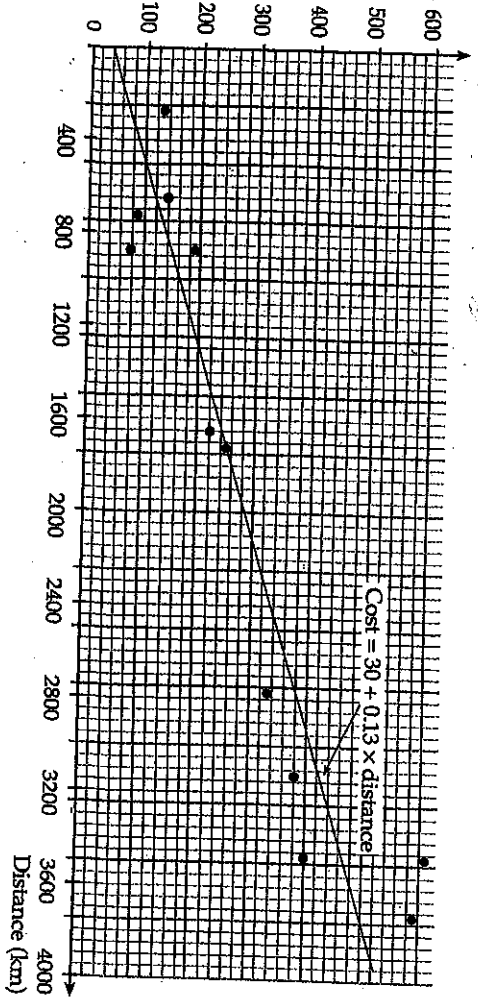
Part 1

- a. Range : 3773 - 254 = 3519 km
- b. $\frac{\sum x}{12} = \frac{3427}{12} = 285.58 \approx \286.00
- c. $\frac{\sum x + a}{13} < 270, \therefore 3427 + a < 3510$
 $\therefore a < 83$

The cost of the flight would need to be less than \$83.00

Part 2

- a. Cost (\$)



- b. \$0.13 = 13 cents is the increase in cost per kilometre of distance.

- c. \$30 is the initial fixed cost per flight.
- e. Total cost is \$30 plus 13 cents per kilometre
- d. $C = 30 + 0.13(2270) = 325.10$
 Cost = \$325

- e. i. Strong linear relationship between cost and distance
- ii. 86.7%
- f. i. The residual for Sydney is negative which means the actual cost of flying to Sydney is less than that predicted by the regression line.

ii. Darwin: Actual : (3773, 589)
 Predicted: $C = 30 + 0.13(3773)$
 $= 520.49$
 Residual = 589 - 520 = 69

iii. Residual plot appears to be a random collection of points, roughly evenly spread around the horizontal axis.

Question 4

Part 1

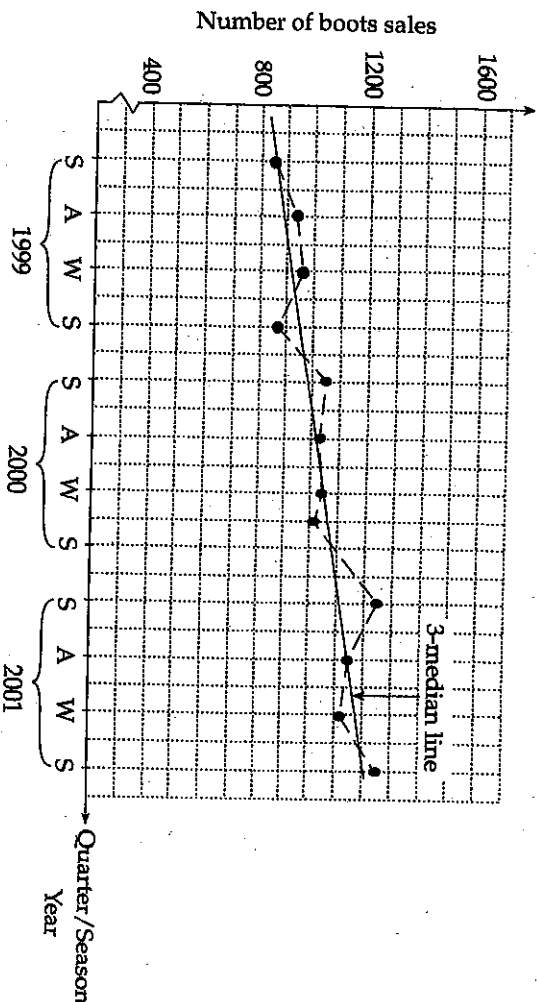
- a. $\frac{659 + 1324 + 1450 + 1299}{4} = 1183$
- b. $4 - (0.52 + 1.15 + 1.02) = 1.31$

- c. In the summer, sales are about half (52%) of the average for the year.
- d. Upward Trend
 Seasonal

Part 2

- a. $\frac{446}{0.52} \approx 858$ $\frac{1180}{1.15} \approx 1026$
 Summer 1999 Autumn 2000

b.



Median points are (2.5, 918) (6.5, 1033) (10.5, 1206)

Part 3

- a. $N = 839 + 32t$
 Winter 2002: $t = 15$
 $N = 839 + 32(15) = 1319$
- b. $1319 \times 1.31 = 1728$

Part 4

- a. 4 as there are 4 seasons or 4 quarters
 b. The values are less influenced by outliers.

Question 5

Part 1

- a. Class A — negatively skewed
 Class B — bimodal
- b. $\frac{22+1}{2} = 11.5$; so between 11th and 12th number: $\frac{37+38}{2} = 37.5$
 $\frac{19+1}{2} = 10$; 10th number = 24

Part 2

a.

0*	2								
0.	7	9							
1*	0	1	2	2	3	4			
1.	5	5	5	6	6	6	7	7	8
									9

OR

0*	2								
0.	7	9							
0f	2								
0s	7								
0.	9								
1*	0	1							
1f	2	2	3						
1s	4	5	5	5					
	6	6	6	7	7				
	8	9							

- b. $Q_1 = 11, Q_3 = 16$

IQR = 5

- c. $Q_1 - 1.5(5) = 3.5 \therefore 2$ is an outlier

$Q_3 + 1.5(5) = 23.5 \therefore$ no other outliers

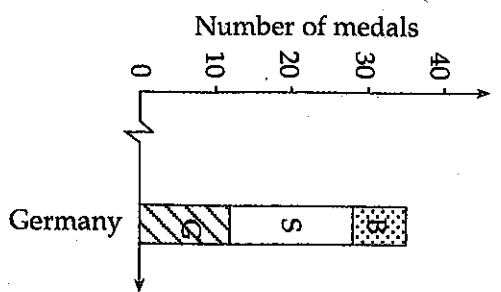
Part 3

Number the students in each class 1, 2, 3, ...
 Class A has 22 students, Classes B and C each have 19. So use a random number from 1 to 22 for Class A, and a random number from 1 to 19 for each of Classes B and C.

Question 6

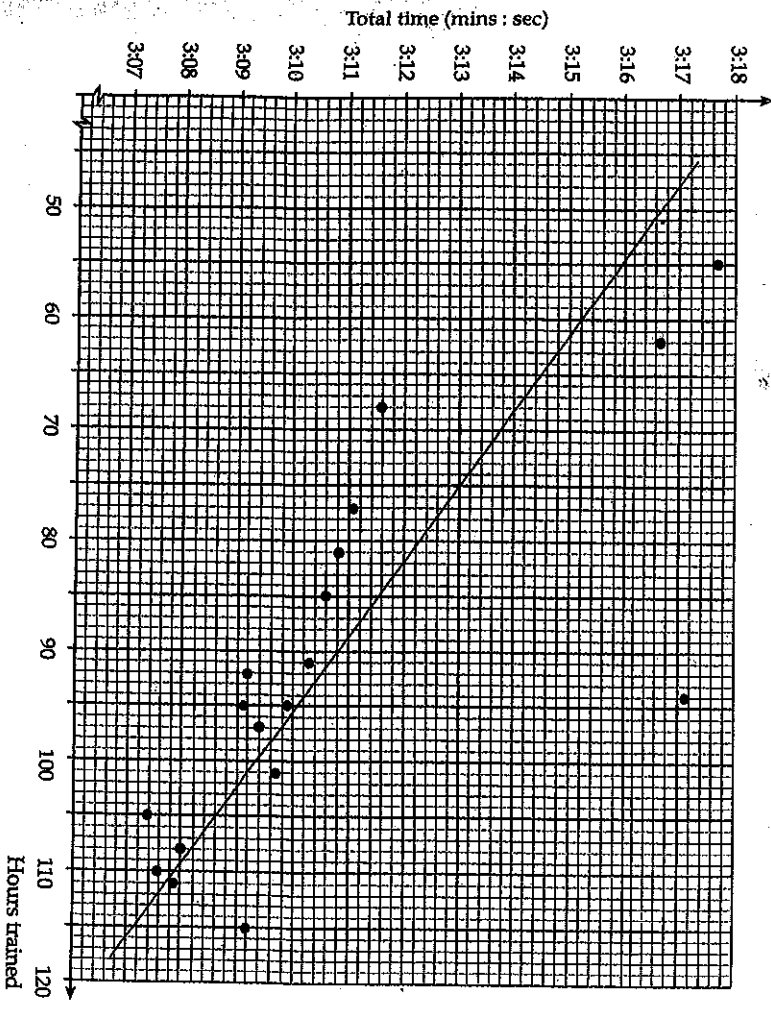
- a. USA
 b. Germany $12 + 16 + 7 = 35$

Part 2



Part 3

- a. Monaco
- b. $a = 24.41$ $b = -0.15$
- c.

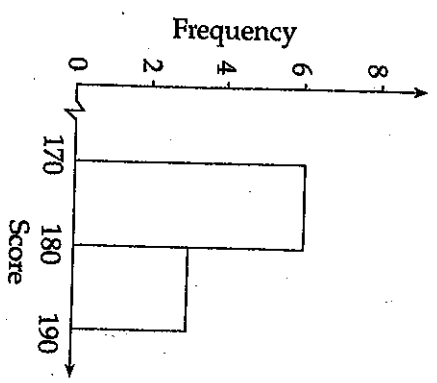


- d. Linear model does not appear to be appropriate, because of uneven distribution of points about regression line, and because of existence of an outlier.
- e. While training for longer hours appears to improve times, it is no guarantee of winning the gold medal (evidence in Monaco team). Also as the hours of training increase, the times tend to level off.

Part 4

Class interval	Frequency
170-	6
180-	3
190-	2
Total	30

- a. one correct
- b. others correct
- i.



- ii. Closely resembles a symmetrical distribution