S1 Edexcel

Sheet #4

Chapter 1
Organising Data

1- Solomon A question 5

The number of patients attending a hospital trauma clinic each day was recorded over several months, giving the data in the table below.

Number of patients	10 - 19	20 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 69
Frequency	2	18	24	30	27	14	5

These data are represented by a histogram.

Given that the bar representing the 20 - 29 group is 2 cm wide and 7.2 cm high,

(a) calculate the dimensions of the bars representing the groups

    (i) 30 - 34

    (ii) 50 - 69   **(6 marks)**

(b) Use linear interpolation to estimate the median and quartiles of these data.   **(6 marks)**

The lowest and highest numbers of patients recorded were 14 and 67 respectively.

(c) Represent these data with a boxplot drawn on graph paper and describe the skewness of the distribution.

                                                                        (5 marks)

2- Solomon B question 5

For a project, a student asked 40 people to draw two straight lines with what they thought was an angle of 75° between them, using just a ruler and a pencil. She then measured the size of the angles that had been drawn and her data are summarised in this stem and leaf diagram.

Angle	(6 | 4 means 64°)	Totals
4	1	(1)
4		(0)
5	0 2 4	(3)
5	5 8 9	(3)
6	1 1 3 3 4	(5)
6	5 5 7 8 9	(5)
7	0 1 1 2 3 3 4 4 4	(9)
7	5 6 6 7 7 9 9	(7)
8	0 1 1 3 4	(5)
8	5 6	(2)

(a) Find the median and quartiles of these data. **(4 marks)**

Given that any values outside of the limits $Q_1 - 1.5(Q_3 - Q_1)$ and $Q_3 + 1.5(Q_3 - Q_1)$ are to be regarded as outliers,

(b) determine if there are any outliers in these data, **(3 marks)**

(c) draw a box plot representing these data on graph paper, **(3 marks)**

(d) describe the skewness of the distribution and suggest a reason for it. **(2 marks)**

3- Solomon D question 7

Jane and Tahira play together in a basketball team. The list below shows the number of points that Jane scored in each of 30 games.

39	19	28	30	18	21	23	15	34	24
29	17	43	12	24	25	41	19	26	40
45	23	21	32	37	24	18	15	24	36

(a) Construct a stem and leaf diagram for these data.   **(3 marks)**

(b) Find the median and quartiles for these data.   **(4 marks)**

(c) Represent these data with a boxplot.   **(3 marks)**

Tahira played in the same 30 games and her lowest and highest points total in a game were 19 and 41 respectively. The quartiles for Tahira were 27, 31 and 35 respectively.

(d) Using the same scale draw a boxplot for Tahira's points totals.   **(2 marks)**

(e) Compare and contrast the number of points scored per game by Jane and Tahira.

(3 marks)

4- Solomon F question 5

Each child in class 3A was given a packet of seeds to plant. The stem and leaf diagram below shows how many seedlings were visible in each child's tray one week after planting.

Number of seedlings	(2 | 1 means 21)	Totals
0	0 2	(2)
0		(0)
1	1	(1)
1	5 7	(2)
2	0 1 3 3 4	(5)
2	5 7 7 7 8 9 9	(7)
3	0 0 0 1 2 2 4	(7)
3	5 6 8 8	(4)
4	1 3 4	(3)

(a)   Find the median and interquartile range for these data.   **(5 marks)**

(b)   Use the quartiles to describe the skewness of the data. Show your method clearly.

(3 marks)

The mean and standard deviation for these data were 27.2 and 10.3 respectively.

(c)   Explaining your answer, state whether you would recommend using these values or your answers to part *(a)* to summarise these data.

(2 marks)

Outliers are defined to be values outside of the limits $Q_1 - 2s$ and $Q_3 + 2s$ where s is the standard deviation given above.

(d)   Represent these data with a boxplot identifying clearly any outliers.   **(6 marks)**

5- Solomon I question 4

A College offers evening classes in GCSE Mathematics and English.

In order to assess which age groups were reluctant to use the classes, the College collected data on the age in completed years of those currently attending each course. The results are shown in this back-to-back stem and leaf diagram.

Totals	Mathematics	Age	English	Totals
(6)	9 9 9 8 8 7	1	9 9	(2)
(8)	8 5 3 1 1 1 0 0	2	0 1 3 5 5 8	(6)
(7)	7 6 6 4 2 2 1	3	2 3 7 9	(4)
(4)	9 7 5 4	4	0 2 6 8 9	(5)
(3)	8 6 0	5	0 3 7 7	(4)
(2)	5 2	6	2 4 4 8	(4)
(0)		7	1	(1)

Key: 1 | 3 | 2 means age 31 doing Mathematics and age 32 doing English

(a) Find the median and quartiles of the age in completed years of those attending the Mathematics classes.

(4 marks)

(b) On graph paper, draw a box plot representing the data for the Mathematics class.

(3 marks)

The median and quartiles of the age in completed years of those attending the English classes are 25, 41 and 57 years respectively.

(c) Draw a box plot representing the data for the English class using the same scale as for the data from the Mathematics class.

(3 marks)

(d) Using your box plots, compare and contrast the ages of those taking each class.

(4 marks)

6- Solomon K question 3

A soccer fan collected data on the number of minutes of league football, m, played by each team in the four main divisions before first scoring a goal at the start of a new season. Her results are shown in the table below.

m (minutes)	Number of teams
$0 \leq m < 40$	36
$40 \leq m < 80$	28
$80 \leq m < 120$	10
$120 \leq m < 160$	4
$160 \leq m < 200$	5
$200 \leq m < 300$	4
$300 \leq m < 400$	2
$400 \leq m < 600$	3

(a) Calculate estimates of the mean and standard deviation of these data. **(8 marks)**

(b) Explain why the mean and standard deviation might not be the best summary statistics to use with these data.

(2 marks)

(c) Suggest alternative summary statistics that would better represent these data.

(1 mark)

A cyber-cafe recorded how long each user stayed during one day giving the following results.

Length of stay (minutes)	0 –	30 –	60 –	90 –	120 –	240 –	360 –
Number of users	15	31	32	23	17	2	0

(a)   Use linear interpolation to estimate the median and quartiles of these data.   **(6 marks)**

The results of a previous study had led to the suggestion that the length of time each user stays can be modelled by a normal distribution with a mean of 72 minutes and a standard deviation of 48 minutes.

(b)   Find the median and quartiles that this model would predict.   **(7 marks)**

(c)   Comment on the suitability of the suggested model in the light of the new results.

(2 marks)