

**Mean, Standard Deviation and Other Data Measures (From OCR 4732)**

**Q1, (Jun 2006, Q7)**

In a UK government survey in 2000, smokers were asked to estimate the time between their waking and their having the first cigarette of the day. For heavy smokers, the results were as follows.

Time between waking and first cigarette	1 to 4 minutes	5 to 14 minutes	15 to 29 minutes	30 to 59 minutes	At least 60 minutes
Percentage of smokers	31	27	19	14	9

Times are given correct to the nearest minute.

- (i) Assuming that 'At least 60 minutes' means 'At least 60 minutes but less than 240 minutes', calculate estimates for the mean and standard deviation of the time between waking and first cigarette for these smokers. [6]
- (ii) Find an estimate for the interquartile range of the time between waking and first cigarette for these smokers. Give your answer correct to the nearest minute. [4]
- (iii) The meaning of 'At least 60 minutes' is now changed to 'At least 60 minutes but less than 480 minutes'. Without further calculation, state whether this would cause an increase, a decrease or no change in the estimated value of
  - (a) the mean, [1]
  - (b) the standard deviation, [1]
  - (c) the interquartile range. [1]

**Q2, (Jan 2008, Q8)**

The stem-and-leaf diagram shows the age in completed years of the members of a sports club.

<b>Male</b>		<b>Female</b>
8 8 7 6	1	6 6 6 7 7 8 8 9
7 6 5 5 3 3 2 1	2	1 3 3 4 5 7 8 8 9 9
9 8 4 4 3	3	2 3 3 4 7
5 2 1	4	0 1 8
9 0	5	0

Key: 1 | 4 | 0 represents a male aged 41 and a female aged 40.

- (i) Find the median and interquartile range for the males. [3]
- (ii) The median and interquartile range for the females are 27 and 15 respectively. Make two comparisons between the ages of the males and the ages of the females. [2]
- (iii) The mean age of the males is 30.7 and the mean age of the females is 27.5, each correct to 1 decimal place. Give one advantage of using the median rather than the mean to compare the ages of the males with the ages of the females. [1]

**Q3, (Jan 2007, Q8)**

In the 2001 census, the household size (the number of people living in each household) was recorded. The percentages of households of different sizes were then calculated. The table shows the percentages for two wards, Withington and Old Moat, in Manchester:

	Household size						
	1	2	3	4	5	6	7 or more
Withington	34.1	26.1	12.7	12.8	8.2	4.0	2.1
Old Moat	35.1	27.1	14.7	11.4	7.6	2.8	1.3

- (i) Calculate the median and interquartile range of the household size for Withington. [3]
- (ii) Making an appropriate assumption for the last class, which should be stated, calculate the mean and standard deviation of the household size for Withington. Give your answers to an appropriate degree of accuracy. [6]

The corresponding results for Old Moat are as follows.

Median	Interquartile range	Mean	Standard deviation
2	2	2.4	1.5

- (iii) State one advantage of using the median rather than the mean as a measure of the average household size. [1]
- (iv) By comparing the values for Withington with those for Old Moat, explain briefly why the interquartile range may be less suitable than the standard deviation as a measure of the variation in household size. [1]
- (v) For one of the above wards, the value of Spearman's rank correlation coefficient between household size and percentage is  $-1$ . Without any calculation, state which ward this is. Explain your answer. [2]

**Q4, (Jun 2014, Q3)**

The table shows information about the numbers of people per household in 280 900 households in the north-west of England in 2001.

Number of people	1	2	3	4	5 or more
Number of households	86 900	92 500	45 000	37 100	19 400

- (i) Taking '5 or more' to mean '5 or 6', calculate estimates of the mean and standard deviation of the number of people per household. [5]
- (ii) State the values of the median and upper quartile of the number of people per household. [2]

**Q5, (Jun 2008, Q6)**

Last year Eleanor played 11 rounds of golf. Her scores were as follows:

79, 71, 80, 67, 67, 74, 66, 65, 71, 66, 64.

(i) Calculate the mean of these scores and show that the standard deviation is 5.31, correct to 3 significant figures. [4]

(ii) Find the median and interquartile range of the scores. [4]

This year, Eleanor also played 11 rounds of golf. The standard deviation of her scores was 4.23, correct to 3 significant figures, and the interquartile range was the same as last year.

(iii) Give a possible reason why the standard deviation of her scores was lower than last year although her interquartile range was unchanged. [1]

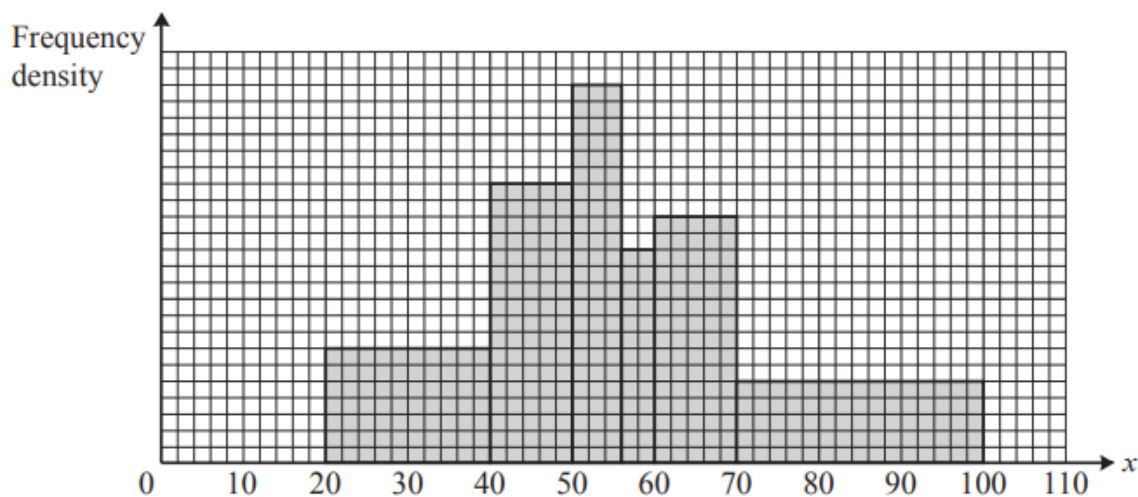
In golf, smaller scores mean a better standard of play than larger scores. Ken suggests that since the standard deviation was smaller this year, Eleanor’s overall standard has improved.

(iv) Explain why Ken is wrong. [1]

(v) State what the smaller standard deviation does show about Eleanor’s play. [1]

**Q6, (Jan 2013, Q6)**

The masses,  $x$  grams, of 800 apples are summarised in the histogram.



(i) On the frequency density axis, 1 cm represents  $a$  units. Find the value of  $a$ . [3]

(ii) Find an estimate of the median mass of the apples. [4]