

## 2<sup>nd</sup> Year Assignment Test 15 Version O

1. Sara is investigating the variation in the daily maximum gust,  $t$  kn, for Camborne in June and July 1987. She used the large data set to select a sample of size 20 from the June and July data for 1987. Sara selected the first value using a random number from 1 to 4 and then selected every third value after that.

(a) State the sampling technique used by Sara.

(b) From your knowledge of the large data set explain why this process may not generate a sample of size 20.

The data Sara collected are summarised as follows  $n = 20, \sum t = 374, \sum t^2 = 7600$

(c) Calculate the standard deviation. Give your answer to 3 s.f.

---

2. Harriet believes that the random sample  $S$ , representing total daily hours of sunshine from the large data set, can be modelled by a discrete uniform distribution, once  $S$  has been rounded to the nearest integer.

a) Write down the probability distribution of  $S$

b) Using this model, find the probability that the total daily hours of sunshine is less than 10

c) State what makes Harriet's assumption very unlikely

d) Suggest a refinement to Harriet's model.

---

3) Joshua compares the amount of rain in 2015 between Heathrow and the city X on the continent of Asia using the Large Data Set.

(a) Write down the name of the city X that Joshua compares with Heathrow.

At random, he selects 8 data points about the daily total rainfall, in mm, in May 2015 for the two cities.

These 8 data points are shown below.

Heathrow: 7.0 0.2 1.2 tr 0.8 6.8 0.2 4.2

City X: 6.0 0.0 20.7 9.0 14.3 0.5 0.0 0.4

(b) Explain what is meant by the reading 'tr'.

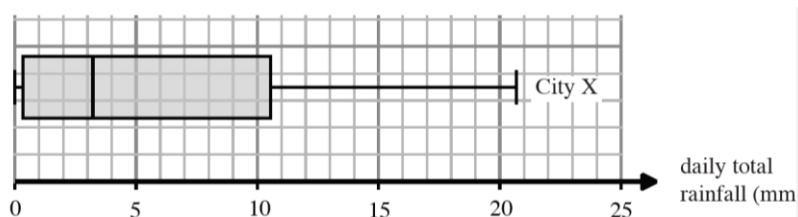
(c) State **one**

(i) advantage

(ii) disadvantage

of Joshua using 8 data points from the large data set for his comparisons.

The diagram below shows a box-plot for the data collected by Joshua on the rainfall in the city X in May 2015.



Draw another box-plot to represent the data collected by Joshua for Heathrow. In your data processing, take 'tr' to mean 0.0 mm of rainfall and ignore outliers.

(e) Compare the amount of rainfall in May 2015 between Heathrow and the city X.

## 2<sup>nd</sup> Year Assignment Test 15 Version P

1. Sara is investigating the variation in the daily maximum gust,  $t$  kn, for Camborne in June and July 1987. She used the large data set to select a sample of size 19 from the June and July data for 1987. Sara selected the first value using a random number from 1 to 4 and then selected every third value after that.

(a) State the sampling technique used by Sara.

(b) From your knowledge of the large data set explain why this process may not generate a sample of size 19.

The data Sara collected are summarised as follows  $n = 19, \sum t = 377, \sum t^2 = 7600$

(c) Calculate the standard deviation. Give your answer to 3 s.f.

---

2. Harriet believes that the random sample  $S$ , representing total daily hours of sunshine from the large data set, can be modelled by a discrete uniform distribution, once  $S$  has been rounded to the nearest integer.

a) Write down the probability distribution of  $S$

b) Using this model, find the probability that the total daily hours of sunshine is less than 9

c) State what makes Harriet's assumption very unlikely

d) Suggest a refinement to Harriet's model.

---

3) Joshua compares the amount of rain in 2015 between Heathrow and the city X on the continent of Asia using the Large Data Set.

(a) Write down the name of the city X that Joshua compares with Heathrow.

At random, he selects 8 data points about the daily total rainfall, in mm, in May 2015 for the two cities.

These 8 data points are shown below.

Heathrow: 7.0 0.1 1.3 tr 0.7 6.8 0.3 4.2

City X: 6.0 0.0 20.7 9.0 14.3 0.5 0.0 0.4

(b) Explain what is meant by the reading 'tr'.

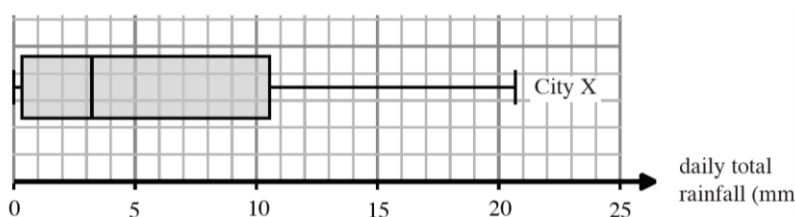
(c) State **one**

(i) advantage

(ii) disadvantage

of Joshua using 8 data points from the large data set for his comparisons.

The diagram below shows a box-plot for the data collected by Joshua on the rainfall in the city X in May 2015.



Draw another box-plot to represent the data collected by Joshua for Heathrow. In your data processing, take 'tr' to mean 0.0 mm of rainfall and ignore outliers.

(e) Compare the amount of rainfall in May 2015 between Heathrow and the city X.

## 2<sup>nd</sup> Year Assignment Test 15 Version Q

1. Sara is investigating the variation in the daily maximum gust,  $t$  kn, for Camborne in June and July 1987. She used the large data set to select a sample of size 18 from the June and July data for 1987. Sara selected the first value using a random number from 1 to 4 and then selected every third value after that.

(a) State the sampling technique used by Sara.

(b) From your knowledge of the large data set explain why this process may not generate a sample of size 18.

The data Sara collected are summarised as follows  $n = 18, \sum t = 288, \sum t^2 = 7600$

(c) Calculate the standard deviation. Give your answer to 3 s.f.

2. Harriet believes that the random sample  $S$ , representing total daily hours of sunshine from the large data set, can be modelled by a discrete uniform distribution, once  $S$  has been rounded to the nearest integer.

a) Write down the probability distribution of  $S$

b) Using this model, find the probability that the total daily hours of sunshine is less than 8

c) State what makes Harriet's assumption very unlikely

d) Suggest a refinement to Harriet's model.

3) Joshua compares the amount of rain in 2015 between Heathrow and the city X on the continent of Asia using the Large Data Set.

(a) Write down the name of the city X that Joshua compares with Heathrow.

At random, he selects 8 data points about the daily total rainfall, in mm, in May 2015 for the two cities. These 8 data points are shown below.

Heathrow:	7.0	0.15	1.1	tr	0.9	6	0.25	5
City X:	6.0	0.0	20.7	9.0	14.3	0.5	0.0	0.4

(b) Explain what is meant by the reading 'tr'.

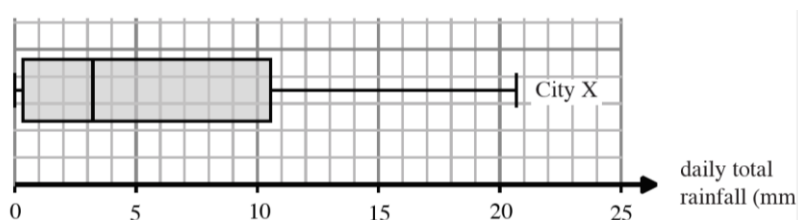
(c) State **one**

(i) advantage

(ii) disadvantage

of Joshua using 8 data points from the large data set for his comparisons.

The diagram below shows a box-plot for the data collected by Joshua on the rainfall in the city X in May 2015.



Draw another box-plot to represent the data collected by Joshua for Heathrow. In your data processing, take 'tr' to mean 0.0 mm of rainfall and ignore outliers.

(e) Compare the amount of rainfall in May 2015 between Heathrow and the city X.

## 2<sup>nd</sup> Year Assignment Test 15 Version R

1. Sara is investigating the variation in the daily maximum gust,  $t$  kn, for Camborne in June and July 1987. She used the large data set to select a sample of size 17 from the June and July data for 1987. Sara selected the first value using a random number from 1 to 4 and then selected every third value after that.

(a) State the sampling technique used by Sara.

(b) From your knowledge of the large data set explain why this process may not generate a sample of size 17.

The data Sara collected are summarised as follows  $n = 17, \sum t = 200, \sum t^2 = 7600$

(c) Calculate the standard deviation. Give your answer to 3 s.f.

2. Harriet believes that the random sample  $S$ , representing total daily hours of sunshine from the large data set, can be modelled by a discrete uniform distribution, once  $S$  has been rounded to the nearest integer.

a) Write down the probability distribution of  $S$

b) Using this model, find the probability that the total daily hours of sunshine is less than 7

c) State what makes Harriet's assumption very unlikely

d) Suggest a refinement to Harriet's model.

3) Joshua compares the amount of rain in 2015 between Heathrow and the city X on the continent of Asia using the Large Data Set.

(a) Write down the name of the city X that Joshua compares with Heathrow.

At random, he selects 8 data points about the daily total rainfall, in mm, in May 2015 for the two cities.

These 8 data points are shown below.

Heathrow: 7.0 tr 1.6 tr 0.4 5.5 0.4 5.5

City X: 6.0 0.0 20.7 9.0 14.3 0.5 0.0 0.4

(b) Explain what is meant by the reading 'tr'.

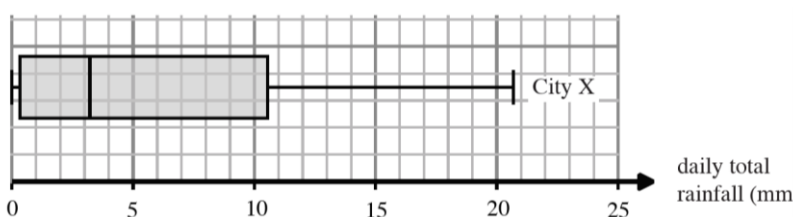
(c) State **one**

(i) advantage

(ii) disadvantage

of Joshua using 8 data points from the large data set for his comparisons.

The diagram below shows a box-plot for the data collected by Joshua on the rainfall in the city X in May 2015.



Draw another box-plot to represent the data collected by Joshua for Heathrow. In your data processing, take 'tr' to mean 0.0 mm of rainfall and ignore outliers.

(e) Compare the amount of rainfall in May 2015 between Heathrow and the city X.

## Answers Version O

1. a) Systematic

b) In the Large Data Set some days have gaps because the data was not recorded.

c) 5.51

2. a)

s	0	1	2	3	4	5	6	7	8	9	10	11	12
P(S=s)	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$

s	13	14	15	16	17	18	19	20	21	22	23	24	
P(S=s)	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	

b)  $\frac{2}{5}$

c) None of these cities are in the Arctic Circle so there will always be some hours when the sun has set – so the maximum number will never be 24

d) The number of hours of sunshine will vary from month to month and place to place

So e.g. use a non-uniform distribution.

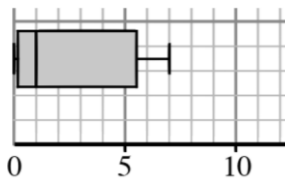
3. a) Beijing

b) tr means a rainfall of less than 0.05 mm

c) i) large data set has many more than 8 data points, so using 8 points is easier to process, it is quicker to process and requires less analysis

ii) large data set has many more than 8 data points, so using 8 points may not very representative and may lead to inaccurate and unreliable conclusions

d)



e) Heathrow had less rainfall on average than City X as the median is lower; Heathrow had less variation in the amount of rainfall it received than City X, as the Interquartile range and the range is smaller

## Answers Version P

1. a) Systematic

b) In the Large Data Set some days have gaps because the data was not recorded.

c) 2.51

2. a)

s	0	1	2	3	4	5	6	7	8	9	10	11	12
P(S=s)	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$

s	13	14	15	16	17	18	19	20	21	22	23	24	
P(S=s)	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	

b)  $\frac{9}{25}$

c) None of these cities are in the Arctic Circle so there will always be some hours when the sun has set – so the maximum number will never be 24

d) The number of hours of sunshine will vary from month to month and place to place

So e.g. use a non-uniform distribution.

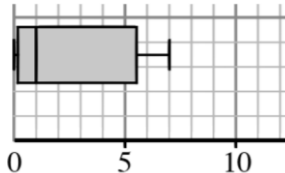
3. a) Beijing

b) tr means a rainfall of less than 0.05 mm

c) i) large data set has many more than 8 data points, so using 8 points is easier to process, it is quicker to process and requires less analysis

ii) large data set has many more than 8 data points, so using 8 points may not very representative and may lead to inaccurate and unreliable conclusions

d)



e) Heathrow had less rainfall on average than City X as the median is lower; Heathrow had less variation in the amount of rainfall it received than City X, as the Interquartile range and the range is smaller

## Answers Version Q

1. a) Systematic

b) In the Large Data Set some days have gaps because the data was not recorded.

c) 12.9

2. a)

s	0	1	2	3	4	5	6	7	8	9	10	11	12
P(S=s)	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$

s	13	14	15	16	17	18	19	20	21	22	23	24	
P(S=s)	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	

b)  $\frac{8}{25}$

c) None of these cities are in the Arctic Circle so there will always be some hours when the sun has set – so the maximum number will never be 24

d) The number of hours of sunshine will vary from month to month and place to place

So e.g. use a non-uniform distribution.

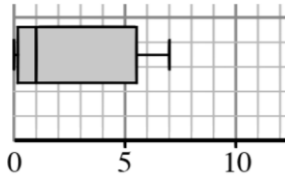
3. a) Beijing

b) tr means a rainfall of less than 0.05 mm

c) i) large data set has many more than 8 data points, so using 8 points is easier to process, it is quicker to process and requires less analysis

ii) large data set has many more than 8 data points, so using 8 points may not very representative and may lead to inaccurate and unreliable conclusions

d)



e) Heathrow had less rainfall on average than City X as the median is lower; Heathrow had less variation in the amount of rainfall it received than City X, as the Interquartile range and the range is smaller

## Answers Version R

1. a) Systematic

b) In the Large Data Set some days have gaps because the data was not recorded.

c) 17.6

2. a)

s	0	1	2	3	4	5	6	7	8	9	10	11	12
P(S=s)	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$

s	13	14	15	16	17	18	19	20	21	22	23	24	
P(S=s)	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	

b)  $\frac{7}{25}$

c) None of these cities are in the Arctic Circle so there will always be some hours when the sun has set – so the maximum number will never be 24

d) The number of hours of sunshine will vary from month to month and place to place

So e.g. use a non-uniform distribution.

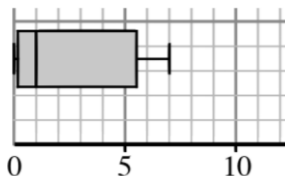
3. a) Beijing

b) tr means a rainfall of less than 0.05 mm

c) i) large data set has many more than 8 data points, so using 8 points is easier to process, it is quicker to process and requires less analysis

ii) large data set has many more than 8 data points, so using 8 points may not very representative and may lead to inaccurate and unreliable conclusions

d)



e) Heathrow had less rainfall on average than City X as the median is lower; Heathrow had less variation in the amount of rainfall it received than City X, as the Interquartile range and the range is smaller