Page 7

1. Mean = 466.2 (1 dp)
Median = 367
Lower quartile = 276
Upper quartile = 649
IQ range = 373
Range = 840

2. Mean = 53.6 (1 dp)
Median = 52
Lower quartile = 32
Upper quartile = 70
IQ range = 38
Range = 69

3. Mean = 6.43 (2 dp)
Median = 6.365
Lower quartile = 4.27
Upper quartile = 8.82
IQ range = 4.55
Range = 11.05

4. Mean = 24.7 (1 dp)
Median = 21
Lower quartile = 11
Upper quartile = 37
IQ range = 26
Range = 97

5. a) Mean = 5.26 (2 dp)
Median = 5.1
Lower quartile = 3.2
Upper quartile = 7.0
Range = 7.3

b) Mean = 526 (3 sf)
Median = 510
Lower quartile = 320
Upper quartile = 700
Range = 730

All 100 times as each piece of data is multiplied by 100.

6. a) Mean = 54.5 (1 dp)
Median = 53
Lower quartile = 25
Upper quartile = 77
Range = 86

b) Mean = 354.5 (1 dp)
Median = 353
Lower quartile = 325
Upper quartile = 377
Range = 86

Measures of the middle have 300 added but measues of spread are unchanged.

Page 11

7. Std. Dev. = 6.66 (2 dp)
Variance = 44.41 (2 dp)

8. Std. Dev. = 5.46 (2 dp)
Variance = 29.84 (2 dp)

9. Mean = 58.75 (2 dp)
Std. Dev. = 16.80 (2 dp)
proportion within one
Std. Dev. is 14 out of 20 = 0.7

Page 11 cont...

10. Mean = 28.45 (2 dp)
Std. Dev. = 6.81 (2 dp)
Proportion within one
Std. Dev. is 12 out of 20 = 0.6

11. a) Mean = 164.60 (2 dp)
Std. Dev. = 8.80 (2 dp)

b) Proportion within one
Std. Dev. is 22 out of 30 = 0.73

12. a) Each measure is doubled so
68, 68, 90, 90, 90, 92, 112, 112, 112, 114, 114, 114, 134, 134, 146.

b) Mean only plus 35 so
69, 69, 80, 80, 80, 81, 91, 91, 91, 92, 92, 92, 102, 102, 108.

c) Mean only subtract 53 to 0 so
$^{-}$19, $^{-}$19, $^{-}$8, $^{-}$8, $^{-}$8, $^{-}$7, 3, 3, 3, 4, 4, 4, 14, 14, 20.

Page 15

13. I wonder if left-handed students in Year 11 at my school gain less achievement credits than right-handed students in art? I will compare the median total art achievement NCEA 1 credits of this year's Year 12 students at my school to the handedness of students.

14. I wonder if Year 11 students who drink caffeine complete more homework than students who do not? I will compare the median number of hours homework of Year 11 students at our school for students who drink less than five caffeine drinks a week compared to students who drink more than five caffeine drinks a week.

15. I wonder if there is less participation in sport per week from students who live in large cities compared to students who live in small towns in New Zealand? I will compare the median total hours participation in sport of Year 11 students on the CensusAtSchool data base for those living in large cities (over 300 000 population) to students in small cities (under 100 000 population).

16. I wonder if the number of texts sent per week is greater for girls than boys? I will compare the median number of text message sent per day of Year 11 girls and boys at our school.

Page 15 cont...

17. I wonder if the median hour of going to bed is greater for teenagers than adults? I will compare the median bedtime on week nights of teenagers at our school to their parents.

18. I wonder if students who usually eat a high starch breakfast gain more credits in NCEA compared to students who have a small breakfast? I will compare the median total achievement NCEA 1 credits of this year's Year 12 students at our school for students who regularly eat muesli, porridge or weet-bix to students who eat something else or nothing.

19. I wonder if students at single sex schools get more achievement standards in Year 11 than students at coeducational schools? I will compare the median total achievement NCEA 1 credits of this year's Year 12 students at the boys and girls schools in my town to students at the two coeducational schools in my town.

20. I wonder if soccer players are lighter than rugby players? I will compare the median weight of boys at my school that play soccer to boys that play rugby.

Page 19

Other correct answers are possible.

21. This may be biased as
- from 3pm to 5pm very few people with full-time jobs will be near Countdown.
- you may approach some people who do not live in Napier.
- you may only approach friendly looking people.

22. This may be biased as
- a significant section of the population either does not have a phone or has an unlisted number.
- Wellington people are unlikely to be representative of New Zealand as a whole.
- you have no control on who answers the questions. They may not be of voting age. A lot of teenagers will answer the phone.

Page 19 cont...

Other correct answers are possible.

23. This may be biased as
- this form class may get more (or less) homework than other form classes.
- the students are likely to exaggerate the amount of homework they do to impress the Principal.
- the students may lie to impress each other.

24. This may be biased as
- different students may be more or less serious about filling in the questionnaire.
- some racial groups (e.g. Chinese) will have less chance of being selected, yet may have different homework hours.
- you have little control on the supervision and the instructions given to the students completing the questionnaire.

Page 20

25. Population. All students in the school.
Method of sampling.
Systematic sampling. Select every nth member of the school roll by going through the form registers of the entire school. Each person selected is asked to complete a questionnaire no matter whether they regularly use the school canteen or not. The school roll divided by 30 is how n is calculated.

26. Population. All Year 11 students who sat NCEA 1 mathematics last year at your school.
Method of sampling.
Systematic sampling. Select every 5th student in a Year 12 form class. If a person selected did not sit NCEA 1 mathematics last year, then they are asked to give the questionnaire to the next person on the form register.
Note: This method may be biased as any student who has left school has no chance of being selected.

27. Population. All teenagers living in the Wellington region.

Page 20 Q27 cont...

27. Method of sampling. Simple Random.
Use random numbers to first select the page and then position on the page of about 100 phone numbers listed in the white pages. Ring the number and ask to speak to the teenager who is closest to the phone.

28. Population. All leaves on the sunny and shady sides of a pohutukawa tree.
Method of sampling. Cluster.
Select one branch on the sunny side and measure every leaf. Then select a similar branch with the same diameter on the shady side of the tree and measure every leaf.
Note: This method may be biased because you are less likely to select branches that are difficult to reach.

Page 23

29. Need 30 pairs of random integers. The first integer in each pair is from 1 to 13 which selects the class. The second integer 1 to 24 selects the student in the class.

30. Need one or two samples. If one sample then the same 30 days are selected for each of the two cities. If two samples then different days are selected. For each random number 1 to 365 the corresponding month and day must be worked out.

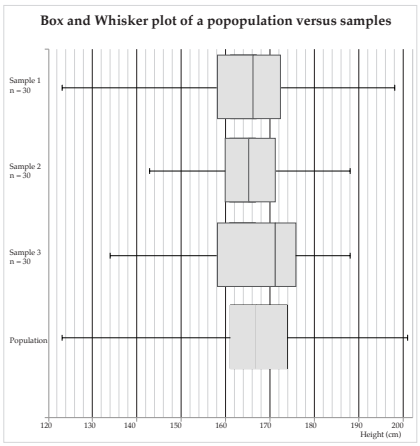
Page 24

31. These are the author's results. Two of the three medians are close to the population median but the other statistics vary. When we take samples of size 30 we get an estimate of the population median that varies with the random process of taking a sample.

	Pop. data	S. 1	S. 2	S. 3
Min.	123	123	143	134
LQ	161	158	160	158
Med.	167	166	165	171
UQ	174	172	171	176
Max.	201	198	188	188

Page 24 Q31 cont...

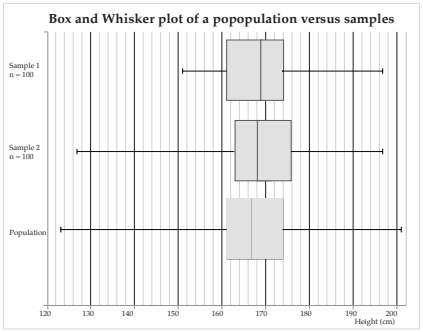
31.



32. These are the author's results with n = 100. A sample size of 100 has the medians and quartiles close to the population parameters.

One median of one sample is 3 cm above the population median and the minimum of the same sample is 151 cm.

	Pop.	S. 1	S. 2
Min.	123	127	151
LQ	161	163	161
Med.	167	168	170
UQ	174	176	174
Max.	201	197	197



Page 30 These are the author's results.

33. Sample size 15.
The medians vary by up to 5 cm from the population median with eight results 3 cm or more away.

The upper quartiles are consistent but the lower quartile varies by 9 cm. These sample statistics were a poor predictor of the population median.

Table next page.

Page 30 Q33 cont...

	S. 1	S. 2	S. 3	S. 4	S. 5	S. 6	S. 7	S. 8	S. 9	S.10
n	15	15	15	15	15	15	15	15	15	15
Med.	170	171	170	170	170	171	172	170	169	169
LQ	166	168	166	165	165	163	164	163	159	163
UQ	172	174	175	175	175	177	178	177	176	176
Min.	155	156	156	156	156	151	160	154	127	127
Max.	193	193	193	201	201	191	191	191	191	179

34. Sample size 30.
The medians vary by only 3 cm from the population median with six results within 1 cm. The upper quartile is not as consistent and varies by 6 cm. With samples of 30, the sample variation was reduced for the median but quartiles were not consistent.

	S. 1	S. 2	S. 3	S. 4	S. 5	S. 6	S. 7	S. 8	S. 9	S.10
n	30	30	30	30	30	30	30	30	30	30
Med.	168	165	164	166	166	164	164	166	168	168
LQ	159	159	159	160	160	158	158	160	161	162
UQ	172	170	169	171	172	174	176	174	174	174
Min.	140	140	140	140	140	134	149	149	149	148
Max.	187	185	195	195	195	190	190	190	197	197

Page 31 These are the author's results.

35. Sample size 50.
The medians vary by up to 3 cm from the population median with seven results 1 cm away.
The lower quartile is consistent with all measures within 1 cm but the upper quartile varies by 6 cm.
With samples of 50, the sample variation was reduced for the median but the samples were still a poor predictor of population extremes.

	S. 1	S. 2	S. 3	S. 4	S. 5	S. 6	S. 7	S. 8	S. 9	S.10
n	50	50	50	50	50	50	50	50	50	50
Med.	166	166	169	170	169	166	166	166	166	166
LQ	161	160	162	162	162	162	162	161	162	161
UQ	175	172	176	176	175	172	170	171	172	172
Min.	149	149	127	127	127	127	127	123	123	123
Max.	193	190	193	193	197	197	197	197	197	197

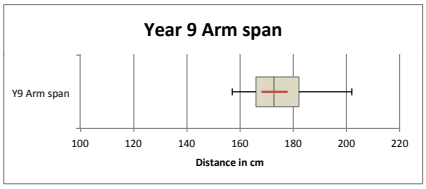
Page 31

36. Sample size 100.
The medians vary by only 1 cm from the population median. The lower quartile is very consistent with results ±1 cm. The upper quartile is not as consistent with three results 3 cm away from the population UQ.
With samples of 100 the sampling variation is reduced.

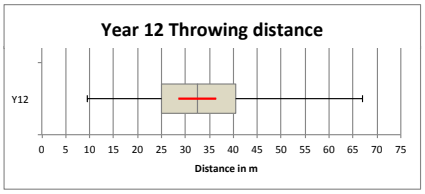
	S. 1	S. 2	S. 3	S. 4	S. 5	S. 6	S. 7	S. 8	S. 9	S.10
n	100	100	100	100	100	100	100	100	100	100
Med.	167	168	168	167	168	167	167	167	167	167
LQ	162	161	161	161	162	161	162	161	161	161
UQ	176	176	175	175	175	171	171	171	172	172
Min.	134	134	127	127	127	134	134	134	134	134
Max.	198	195	195	195	195	189	189	189	198	198

Page 46

37. We would expect to find the population median in the region
173 + 4.9 = 178 cm to
173 – 4.9 = 168 cm (0 dp).
This large 9.8 cm interval is caused by the very small sample size and the wide spread of results as shown by the 16 cm Inter quartile range (IQR).

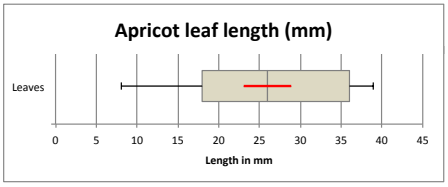


38. We would expect to find the population median in the region
32.5 + 3.9 = 36.4 m (1 dp) to
32.5 – 3.9 = 28.6 m (1 dp). The interval we expect to find the median in is still almost 8 metres as the sample size is small. The wide spread of results is shown by the Inter quartile range (IQR) = 15.3 m.

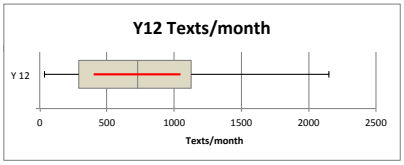


Page 47

39. We would expect to find the population median of the apricot leaf length in the region
26 + 2.9 = 28.9 mm to
26 – 2.9 = 23.1 mm (1 dp).
This interval of values is relatively small despite the wide spread in the data (IQR = 18 mm) because the sample size is 85.



40. We would expect to find the population median Y12 texts per month in the region
725 + 323 = 1048 texts to
725 – 323 = 402 texts (0 dp).
This huge interval of texts reflects the small sample size (n = 15) and wide spread of the data (IQR = 835).
As shown by the bar on the graph the prediction for the population median would be from almost half the median to 1.5 times the median.

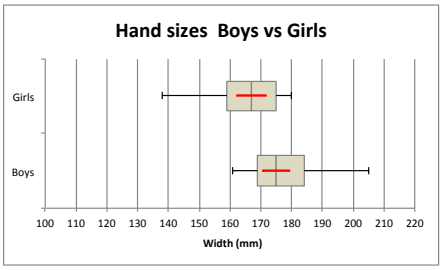


Page 50

41. We would expect to find the population median of boys' hand sizes in the region
175 + 4.6 = 180 mm to
175 – 4.6 = 170 mm (0 dp).
We would expect to find the population median of girls' hand sizes in the region
167 + 5 = 172 mm to
167 – 5 = 162 mm (0 dp).
The largest hand size median we would expect for girls is 172 mm while the smallest hand size median for boys is 170 mm. Therefore it is possible that the median hand size for boys in the population could be smaller than the median hand size for girls.

Page 50 Q41 cont...

Therefore even though the boys' sample median was larger we **cannot make a call** that this difference in sample medians is reflected in the population medians. The informal confidence interval bars on the graph overlap. See graph below.



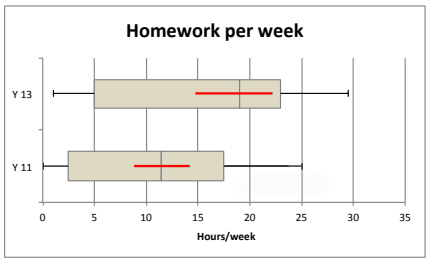
42. We would expect to find the population median of females' wages in the interval
695 + 59 = $754 to
695 – 59 = $636.
The confidence interval of the males' wages is larger caused by the high spread (IQR = $455) and smaller sample, so we would expect to find the population median in the region
850 + 142 = $992 to
850 – 142 = $708.

The highest median wage we would expect for females is $754 while the lowest median wage for males is $708. Therefore it is possible by chance that the median wage for males could be smaller than the median wage for females. Therefore even though the males' sample median was larger we **cannot make the call** that this difference is reflected in the population. The informal confidence interval bars on the graph overlap. See graph below.

Page 51

43. We would expect to find the Y11 homework hours population median in the interval
11.5 + 2.7 = 14.2 hours to
11.5 – 2.7 = 8.8 hours.
We would expect to find the Y13 homework hours population median in the interval
18.5 + 3.7 = 22.2 hours to
18.5 – 3.7 = 14.8 hours.

The highest Y11 median hours for homework is 14.2 hours while the lowest median hours for Y13 is 14.8 hours, which is more. Therefore it is reasonable to conclude on this data that Y13 students tend to do more homework than Y11 students. The informal confidence interval bars on the graph do not overlap. See graph below.

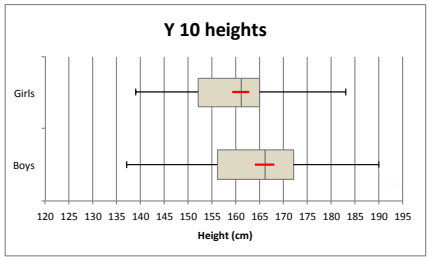


44. We would expect to find the Y10 boys' heights population median in the region
166 + 2.1 = 168 cm to
166 – 2.1 = 164 cm (0 dp).
We would expect to find the Y10 girls' heights population median in the region
161 + 1.7 = 163 cm to
161 – 1.7 = 159 cm (0 dp).

The tallest girls' median height is 163 cm while the shortest boys' median height is 164 cm, which is more.

Therefore it is reasonable to conclude on this data that the Y10 boys' median is larger than the Y10 girls' median. The informal confidence interval bars on the graph do not overlap. See the graph following.

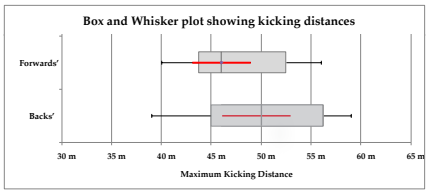
Page 51 Q44 cont...



Page 52

45. We would expect to find the population median for the backs' kicks to be in the region
50 + 3.9 = 53.9 m to
50 – 3.9 = 46.1 m (1 dp).
We would expect to find the population median for the forwards' kicks to be in the region
46 + 3.2 = 49.2 m to
46 – 3.2 = 42.8 m (1 dp).

The backs' kicks appeared to travel further, but the median for the backs' kicks could be as low as 46.1 m and the forwards' kicks could have a median as high as 49.2 m, so even though at first glance the backs' kicks appeared to travel further it could be that the population median is lower. Therefore we are unable on this data to make a call that the difference exists in the population. If we study the graph we can see the overlap between the informal confidence interval bars for each median.



Pages 53 - 56 Practice Internal Assessment Task – Statistical Inference 2.9

The student must use the statistical enquiry cycle which means: posing an appropriate comparison question from a data set; selecting random samples; selecting and using appropriate displays and measures; discussing sample distributions; discussing sampling variability, including the variability of estimates; making an inference; and communicating findings in a conclusion.

Concepts	Achieved	Merit	Excellence
	To Achieve students must show evidence of using each component of the statistical enquiry cycle to make an inference.	To Achieve with Merit, in addition to the requirements for Achievement, the student must relate their findings to the context and population and make supporting statements.	To Achieve with Excellence, in addition to Merit, students must show evidence of integrating statistical and contextual knowledge throughout the process. They must reflect on the process and may consider other explanations.
Investigative question	Specified the purpose of the investigation or has a clear investigative question. E.g. *Do students in Year 9 in New Zealand have heavier bags than students in Year 12?*	The purpose or question is **linked** to the situation being investigated. E.g. *I want to know if the year level at school affects the weight of the student's school bag in all of New Zealand. I intend to use the data at the CensusAtSchool website to represent all the students in New Zealand. Specifically I will compare the median bag weights of Year 9 students to the median bag weights of Year 12 students to see if Year 9 students tend to have heavier bags.*	The purpose or question is **relevant** to the situation being investigated. E.g. *I want to know if the year level at school tends to affect the weight of the student's school bag for students at New Zealand schools. I cannot sample students from all over New Zealand so I intend to sample from the large data set on the CensusAtSchool website as I believe it will adequately reflect the New Zealand student population. I think younger students carry material for all subjects but I believe that older students are more efficient in what they choose to bring to school. My question is 'Are the median weight of Year 9 students' bags heavier than the median weight of Year 12 students' bags?'*
Random sample	Selected a random sample with evidence of how this done. The description of selection is sufficient and relevant to the investigative question. E.g. *Generates a simple random sample of 50 students from the cleaned data from the Year 9 data set and the another from the Year 12 data set.*	Selected a random sample with evidence how this was done. Reference to why they used the sample method **or** the sample size is made. E.g. *In addition to the sample description, the student gives contextual reasons for deciding on the use of a simple random sample or a sample size of 50.*	Selected a random sample with evidence how this was done. Reference to why they used the sample method **and** the sample size is made. E.g. *In addition to the sample description, the student gives contextual reasons for deciding on the use of a simple random sample and a sample size of 50.*
Statistics and display	Selected and used appropriate displays and measures. Included is relevant information for making an inference. E.g. *A box and whisker plot is drawn correctly for each sample.*	Included in the displays is relevant information for making an inference. E.g. *Has calculated and plotted informal confidence intervals for the median and displayed this either in the box and whisker plot or separately.*	Included relevant information for making an inference. E.g. *Has calculated and plotted informal confidence intervals for the median and displayed this either in the box and whisker plot or separately. Has calculated the expected minimum and maximum differences between population medians.*

Concepts	Achieved	Merit	Excellence
Discussed the sample distribution	Discussed the sample distributions. E.g. *Has identified at least two comparative features of the sample distributions (shape, overlap, shift, spread, middle 50% or unusual features).*	Discussed the sample distributions. E.g. *Has identified at least **two** comparative features of the sample distributions (shape, overlap, shift, spread, middle 50% or unusual features) and has linked comments to the investigative question and the population.*	Discussed the sample distributions, integrating statistical and contextual knowledge. E.g. *Has identified at least **three** comparative features of the sample distributions (shape, overlap, shift, spread, middle 50%, unusual or interesting features) and has used contextual knowledge to link comments to the investigative question and the population.*
Discussed the sample variability	Discussed the sample variability. E.g. *Has identified that different samples will give different estimates of population parameters.*	Discussed the sample variability. E.g. *Has identified that different samples will give different estimates of population parameters. Has considered the effect of at least one aspect, for example sample size. Has related comments, in context, to the interval, for example that it would be acknowledged that such an interval would contain the population median in most cases.*	Discussed the sample variability. E.g. *Has identified that different samples will give different estimates of population parameters. Has considered the effect of at least one aspect, for example sample size **giving reasons as to why the sample size was selected**. Has related comments, in context, to the interval, for example that it would be acknowledged that such an interval would contain the population median in most cases.*
Made a correct inference	Made a correct inference. E.g. *Has stated, in context for at least one case, a belief that the population median will lie within a correctly calculated interval.*	Made a supported correct inference. E.g. *Has stated, in context, for both sets of data, a belief that the population median will lie within a correctly calculated interval.*	Made a supported correct inference. E.g. *Has stated, in context, for both sets of data, a belief that the population median will lie within a correctly calculated interval.*
Communicated findings clearly	Communicated findings clearly. E.g. *The report covers points above and it also includes an answer to the investigative question which is supported by comments in context.*	Communicated findings clearly. Conclusions are supported by comments in context which are consistent with the investigation. E.g. *There is an overlap between the intervals that we expect to contain the population medians. This means the population median for Year 9 bags may be smaller or greater than the Year 12 bags therefore we cannot conclude that there is a difference in the median weights of the bags for the Year 9 and 12 students represented on the Census At School data set.*	Communicated findings clearly in context. They have justified their inference in context, or they have reflected about the process, or they have considered other explanations. E.g. *There is an overlap between the intervals that we expect to contain the population medians. Therefore we cannot conclude that median weights of Year 9 bags is greater or less than the median weight of Year 12 students' bags represented on the Census At School data set. They have reflected on the process by discussing sample size, the reliability of estimates on the bag weights and whether the data set reflects New Zealand students.*

Final grades will be decided using professional judgement based on a holistic examination of the evidence provided against the criteria in the Achievement Standard.