Pearson Edexcel

Examiners' Report
Principal Examiner Feedback

November 2021

Pearson Edexcel GCSE
In Statistics (1ST0) Foundation Tier
Paper 1F

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your candidates at: www.pearson.com/uk

November 2021
Publications Code 1ST0_1F_2111_ER

GCSE (9 - 1) Statistics 1ST0

Principal Examiner feedback – Foundation Paper 1

General comments

This was once again a limited entry due to the extraordinary circumstances the year 2021 has presented. The questions in which most candidates excelled were once again those which required routine calculations to be performed, graphs or charts completed, or information extracted from them. It was however noted that questions asking candidates to explain, assess or interpret have improved slightly from previous entries.

It is worth mentioning here that centres must impress on their candidates the importance of legible handwriting as a few responses were very difficult indeed to read; we cannot award marks when the handwriting is completely illegible.

Question 1

Parts (a) and (b).

It was very rare to see an error in this question asking candidates to extract information from a pictogram and to complete one row.

Part (c)

We were quite lenient in this part of the question as it was not easy to articulate precisely why the pictogram key is unsuitable to show 22 rackets, so we accepted any explanation that demonstrated understanding that that the key was hard to use. Most candidates managed to explain this adequately.

Question 2

Part (a) Virtually every candidate achieved the mark for writing down the correct value extracted from the time series graph.

Part (b) Just over half the candidates achieved both marks in this part. It is worth mentioning that the allocation of 2 marks for this part indicates that some working is

required, and it was interesting to note that those candidates just writing the answer down (despite having obtained the correct value in part (a)) scored no marks because their answer was incorrect with no working.

Parts (c) and (d). Candidates were asked to describe a trend and we accepted 'downwards' and 'upwards' and any similar words for part (c) and (d) respectively. Candidates who were unsuccessful in writing down a correct answer usually wrote down a value. Centres are advised that misunderstanding the demand in the question is a frequent source of error. In this case the demand word 'describe' requires the candidates to write down a description of the trend.

Question 3

Part (a) The majority of candidates were able to identify that the data is raw.

Part (b) This part of the question required candidates to check the tally chart with the raw data given in the stem to part (a). Not every candidate understood this and therefore did not check to see that the tally for 0 should have been 2 lines and lost this mark. We expect a knowledge of correct tallying, and this was the other error in the tally chart (for three days). Many candidates went into autopilot and wrote down; there are no labels, no totals and other spurious reasons.

Part (c) The smallest minority of candidates were able to articulate a reason **in context** for the given median of 3 friends. Most candidates described how to find a median, wrote the data out in full to show the median was indeed 3, or did not answer the question at all. We expect candidates to be able to show understanding of averages in context.

Part (d) Candidates were expected to deduce that because 7 days is not in the data then it cannot possibly be the mode. We required a comment 'not correct' **with** the correct supporting reason. It was not enough for example to state – nobody played a sport seven days in a week – because that does not tell the examiner if the candidate agrees or disagrees with Julie.

Part (e) Few candidates were able to give the **two** reasons required **with** the comment 'not appropriate'. When a question asks for an assessment if something is appropriate then a comment of 'not appropriate' or 'appropriate' is necessarily required as well as in this case **two** supporting reasons. We were expecting candidates to make comments on the loss of accuracy that inevitably results from grouping data, and that there were no labels on the axes but about half of candidates were unable to assess the appropriateness or a correct reason or both.

Question 4

Parts (a) and (b) Most candidates were able to write down the required answers of $\frac{1}{5}$ or equivalent and $\frac{7}{30}$ (we also accepted answers which round to [awrt] 23% or 0.23 respectively), but some candidates still write down a description such as 'unlikely' for an answer. When the demand is 'write down' a numerical answer is required, whereas if we require a description the demand word is 'describe'.

Part (c) This part was in the form of a question – Whose estimate is more reliable, Bryce's or Mary's? For this response there must be an answer to the question, which was not always seen, with a reason for the decision. Approximately half of all candidates were able to score both marks here.

Question 5

Part (a) The most common response in this part of question 5 was – 'When you pick employees at random'. We were looking for an explanation of the word random as applied to a sample, so the word random cannot be used in the explanation. In fact, the explanation was very simple – everyone has an equal chance of being picked – but very few candidates were able to score this mark. It is important to not 'overthink' a question and to write a description in straightforward language.

Parts (b) and (c) . Both parts were very well answered with virtually every candidate scoring both marks in part (c) and the mark is part (c) demonstrating that routine calculations are well understood and applied.

Part (d)(i) and (ii) Virtually every candidate recognised that £3000 must be an outlier and that it was a very different in value to the other 7 employees.

Part (e) This part was much less well answered although it is specifically mentioned in the specification – 'Understand the effect on the mean, mode and median of changes in the data, including the addition or withdrawal of a population or sample member'. [2b.01]

Few were able to understand that removal of an extreme data item would affect the mean and in what way. Many candidates worked out the new mean and then commented that it would be greater. The question specifically stated - 'Without carrying out any further calculations…'- so a decision based on this calculation scored no marks.

Part (e) About half of the candidates recognised at least one correct assessment of Bethany's method and were able to score at least one mark.

Question 6

Part (a) Most candidates scored both available marks on this question. Those who failed to score did not read the question correctly and wrote down 10% or failed to show any working and just wrote an erroneous answer down. Once again, a failure to look at the mark allocation and also show any working resulted in no marks for a few candidates.

Part (b) The command word here is 'compare'. The easiest way to score the two available marks was to write down the percentage for Japan, the percentage for Kenya **and** then write down which country was greater. Whilst about half of the number of candidates scored both marks here, the other half did not understand what was required as a response and consequently lost both marks.

Part (c)(i) and (ii) This question involved extracting information from the given bar charts and most candidates managed to score at least one out of the available four marks with a significant minority scoring 4/4 marks. As in part (b) candidates were required to identify a country and justify it by quoting the correct percentages extracted from the bar chart.

Part (d). The overwhelming majority of candidates failed to score the available mark here suggesting a lack of understanding regarding the nature and purpose of a population pyramid.

Question 7

This is an extended response question [of which there are at least three in total across the Foundation papers] and was worth a total of 6 marks. The median and modal score was 2/6 marks with a high score of 4/6 which suggests that centres still need to devote time to practising these types of questions. The demand in the questions gave very clear instructions on what was required, and it would help candidates to organise their thoughts in bullet points rather than the prose style of answer the vast majority gave. Virtually every candidate attempted the question but 15% of candidates failed to score any marks at all.

The first part of the demand asks candidates to make a judgement whether Naomi's questionnaire and conclusions are appropriate and it was worth one mark just to state 'Not appropriate'.

Three marks were available for discussing the three questions and each question and an example of a suitable way of presenting an answer would have been as follows:

- Naomi's questions are not appropriate.
- <u>Questions</u>
    1. The question is leading /biased
    2. The question is too open
    3. Etc.,

The final two marks were for discussing any two of the conclusions Naomi draws from her results.

- <u>Results/Conclusions</u>
    1. Conclusion 1 is likely to favour August
    2. Etc.
    3. Etc,

This method of presentation, or something similar is precise, easy to construct and is a positive aid to formulating a candidate's thoughts.

Question 8

Part (a) The overwhelming majority of candidates were able to draw an accurate box plot and score the available three marks here. The candidates who did not score all three marks gained at least one mark usually for the correct shape.

Part (b) By this point in the paper some of the less able candidates were unable to access the question and so it was really a case of either a candidate understood the demand or did not and failed to score the two marks here.

Part (c) The question was very clear on the demand – 'Describe the skew of the distribution represented by the box plot'. As the entry was small, not one single candidate was able to explain that the distribution was symmetrical [1 mark] because the median lies exactly between the lower and upper quartiles [I mark].

Many candidates did not attempt the question at all and those who did attempt it clearly did not have the necessary knowledge to assess the skewness in the distribution. [Skewness by inspection is covered in the specification in 2a.09]

Question 9

As this is a common question with the Higher paper, many Foundation candidates were beginning to find the questions difficult.

Part (a) (i) A small number of candidates were able to score the two marks for the 75^{th} percentile. Given that the number of counties is 48, finding 75% of this number is very straightforward and so locating the area corresponding the cumulative frequency of 36 is routine. It was noted that only a few candidates were using the given graph and drawing lines on it to help themselves.

(ii) Not one single candidate was able to articulate the meaning of the 75^{th} percentile. Explaining summary statistics in context is a feature of this examination and we will be expecting candidates to demonstrate understanding.

Part (b) Not a single candidate was able to even score one mark in this part of the question. The vast majority of responses were either complete blanks or a random value entered on the answer line. However, there was one mark available for just reading the cumulative frequency value [19] from the graph when the area is 2000 square kilometres which most candidates could have attempted easily.

Question 10

Parts (a) (i) and (ii) No candidates were able to identify the populations in either parts. Most mentioned candidates in the school for (i) and there were no credible answers for part (ii). The ability to define and identify a precise population is crucial in any statistical survey and we were strict in our expectations for the correct answers.

Part (b) A significant minority of candidates were able to explain that a trusted website is needed as a reliable source of information. Some responses did not refer to the internet at all.

Part (c) Whilst the majority of candidates could not describe the randomness of Quota and Opportunity sampling, a significant minority could at least explain why at least one of these sampling methods why the method is not random. Note however, that a response such as; 'Quota sampling because it is not random' is not going to score any marks.

Part (d) It was pleasing and encouraging to note that although approximately half of the entry did not score any marks in this part of the question, a good number of the other half scored both marks for identifying Method A with a correct supporting reason. This suggests that Quota sampling is possibly more familiar than Opportunity sampling.

Question 11

The Statistics GCSE papers will test application of Statistics in unfamiliar contexts, and so candidates must use the available data/summary statistics in the question to answer questions. In this question some candidates who are clearly also candidates of Economics relied on that knowledge rather than on the information given in the question.

With the exception of one candidate, no-one was able to make a credible attempt at the question. [The topic is covered in 2d.01 of the specification]

Part (a) Very few candidates were able to identify the correct year and month. This is another instant of where annotating the table would have helped to identify increases and decreases.

Parts (b) and (c) Only one candidate was able to calculate the simple index number and the GDP in 2010 Quarter 1. The vast majority of candidates made no attempt at all.

Part (d) This was answered slightly better with a significant minority of candidates able to score at least one mark by extracting information from the table.

Question 12

Part (a) The vast majority of candidates extracted the correct value from the Venn diagram.

Part (b) This part was answered very well with a few candidates able to extract the required information from the Venn diagram, but it would be the second only two marks those candidates scored in the whole question.

Parts (c) and (d) Neither part was answered correctly by any candidate indicating that conditional probability was not known for this group.

Summary

Based on their performance in this paper, candidates should:

- practice development of extended response questions, laying out answers in bullet points
- learn the names and attributes of different sampling methods
- learn the names of different types of data
- give statistical interpretations in the context of the scenarios in the questions
- revise conditional probability [Spec reference 3p.09]
- revise index numbers and RPI, CPI and GDP [Spec reference 2d.01]
- revise skew in distributions. [Spec reference 2a.09]

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom