



Examiners' Report Principal Examiner Feedback

Summer 2025

Pearson Edexcel GCSE
Statistics Foundation Paper 1
(1ST0_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 students at:

www.pearson.com/uk

Summer 2025

Publications Code 1ST0_1F_2506_ER

All the material in this publication is copyright

© Pearson Education Ltd 2025

GCSE (9 – 1) Statistics – 1ST0 Principal Examiner Feedback – Foundation Paper 1

Introduction

General comments

Most responses indicated that the paper was attempted in full within the time allowed. The first six questions were generally answered successfully. For the remaining questions, the quality and completeness of responses varied. In some cases, handwriting made it difficult to determine whether marks could be awarded. Question 7 was more challenging, with many responses lacking clear explanations for the use of a calculation or diagram. Later questions were often partially completed, allowing for some marks to be gained.

Question 1

This question assessed understanding of five key probability terms, the ability to mark probabilities on a scale, state the probability of an event, and calculate the expected number of outcomes for landing on a 3 using a spinner. The context was familiar, and most responses showed engagement with all parts of the question.

In part (a), 'Impossible' was frequently selected as the correct answer. Parts (b) and (c) were generally answered correctly, although some responses showed difficulty interpreting the scale, which differed from the commonly used 0 to 1 scale.

In part (d), many responses correctly stated the probability of landing on a 3. Answers were often given as fractions, though some used decimals or percentages. Where these were not equivalent or clearly stated, marks could not be awarded. A common incorrect response was $\frac{3}{8}$, which, when used consistently, allowed for follow-through in part (e), where the expected number of outcomes was calculated.

In part (e), typical misconceptions included confusing the values 3 and 8, or expressing the answer as a fraction of 80 or a multiple of 80.

Question 2

The context of football was familiar, and most responses showed engagement with the question. In part (a), some responses identified only one diagram or did not select both correct options, pictogram and pie chart, as suitable ways to display the data. Parts (b) and (c) were generally answered correctly.

In part (d), many responses identified a valid limitation, such as the data being limited to one club, one year, or a small sample of players. Some responses referred to players occupying multiple positions, but these did not meet the requirements of the question, as the stem stated that each player had a single identified playing position.

Question 3

The familiar context of leisure activities helped engagement with this question and correct interpretation of the data shown in the pictogram. In part (a), most responses correctly identified “Boat” as the modal tour type. Common incorrect answers included other tour types or the number of people on the boat tour.

In part (b), slightly more responses successfully calculated the difference between the number of people choosing the boat and bus tours. Part (c) was more challenging, with many responses providing a reason or posing a question rather than stating a clear hypothesis. Some responses were too general, such as “the type of tour people use will change.”

In part (d), several responses did not fully address the question and instead described next steps in the statistical enquiry cycle, referencing diagrams or calculations. More successful responses identified factors to keep consistent, such as location, time, or day, or gave general statements like “keep the method of data collection the same.” Some responses included two valid factors in a single statement, for example, “collect the data on a Saturday morning.”

Question 4

This question focused on identifying data types and evaluating questionnaire design.

In part (a), ‘qualitative’ was correctly selected as the data type in just under half of the responses. In part (b), many responses considered the question suitable or highlighted issues such as limited detail for local residents or the grouping of “4 or more”. More successful responses identified specific issues such as the absence of a time frame, overlapping response boxes, or the lack of a zero option. To support accessibility, responses that recognised the question as unsuitable and attempted to give a reason were awarded one mark.

In part (c), fewer responses identified that the question was biased or leading. Unsuccessful responses often focused on the lack of opportunity for respondents to

expand on their answers, rather than addressing the structure or wording of the question.

Question 5

This probability question involved completing a sample space diagram for the totals from rolling two dice. Most responses successfully filled in the missing values. Where marks were not awarded, common issues included calculating differences instead of totals, errors in addition, or leaving the diagram incomplete.

In parts (b) and (c), responses that had completed the diagram were able to follow through using a denominator of 36. In part (b), a frequent misconception was the use of $\frac{5}{36}$ as the probability. In part (c), some responses showed difficulty interpreting “8 or more”, with errors such as counting only the number of 8s or giving the probability of scoring exactly 8.

Question 6

This question involved calculating statistical values and interpreting comparisons. In part (a), most responses correctly calculated the mean, and in part (b), the range. Common errors included using other average measures for the mean, omitting the total before dividing, or incorrectly identifying the highest and lowest values when calculating the range.

In part (c), responses often lacked specificity. Some did not state which year had the highest or lowest mean, or failed to clarify that the comparison was based on the mean. A number of responses compared the ranges instead, or made incorrect conclusions when comparing the means.

In part (d), the focus was on comparing the ranges. Some responses provided a general conclusion or comparison, which was accepted. Where the mean had not been compared in part (c), marks were awarded for a correct comparison in this part.

Question 7

This extended response question focused on evaluating a plan to investigate employee views on company activities. Each part required identification of an element from the plan and a comment on its suitability, with one mark awarded for the reason and one for stating whether it was suitable.

Responses were most successful in the section on data collection. These often began by stating whether the method was suitable or not, followed by comments on each bullet point. In the remaining parts, responses were less consistent. Some did not clearly state the suitability or failed to explain their reasoning. Others repeated the wording from the plan without evaluating its appropriateness.

Some responses included vague phrases such as “might be suitable” or “mostly suitable” without linking these to specific points, which limited the clarity of the evaluation. Structured responses that grouped points and clearly stated suitability were more effective and scored higher.

If a valid reason was given, a mark was awarded, and if this was linked to suitability, a second mark was given. The most successful responses followed this approach, clearly organising their answers and identifying multiple points per section.

Question 8

This question focused on interpreting statistical measures and comparing distributions. In part (a), many responses described how to calculate the median rather than explaining its advantage over the mean. Some responses noted that the median uses only part of the data. To gain the mark, responses needed to identify the presence of outliers or explain that the median is not affected by extreme values. This was achieved in a small proportion of responses.

In part (b), fewer responses recognised the need to order the data and identify the lower and upper quartiles to calculate the interquartile range (IQR). Some responses incorrectly calculated the range instead.

Part (c) required a comparison of distributions. Many responses focused on suitcase weights and drew on personal experience, suggesting that suitcases might be lighter due to using toiletries or heavier due to buying souvenirs. Correct responses identified the command word “compare” and noted that the Sunday flight had a smaller IQR. Some responses included additional explanations, which were accepted but not required.

Question 9

This question focused on sampling within the Statistical Enquiry Cycle, specifically stratified sampling. Many responses did not clearly explain why stratified sampling was appropriate, often referring to general strengths of sampling rather than addressing

stratification directly. Some responses recognised that the method would include individuals from each department, but marks were not awarded where this was followed by incorrect statements such as selecting equal numbers or lacking further explanation.

In part (b), some responses identified the need to use the figures 50 and 120, but did not calculate the total number of employees to determine the correct proportion from manufacturing. A few responses showed an understanding of proportionality but were affected by incorrect approximations, such as using a 10:10:20:10 ratio. Another common error involved rounding the total to 250 and then calculating $250 \div 50 = 5$ followed by $120 \div 5 = 24$.

Question 10

This question focused on interpreting revenue data across years and quarters. Most responses correctly identified the year and quarter with the greatest revenue in part (a). Where marks were not awarded, responses often named the quarter but omitted the year.

In part (b), fewer responses successfully linked the use of a 4-point moving average to seasonal patterns or trends over time. Some responses included relevant ideas but placed them in the wrong part of the question or added unnecessary details that contradicted the correct reasoning.

Parts (c) and (d) commonly featured point-by-point comparisons rather than identifying overall trends. In part (c), some responses noted that the data peaked in Q2 but did not generalise across the dataset. In part (d), many responses focused on Q2 in 2022 as having the highest revenue, rather than making a broader comparison across years. Some responses included speculative explanations, which were not required.

Question 11

This question assessed interpretation and construction of statistical diagrams, as well as understanding of data distribution and sampling methods. In parts (a) and (b), around half of the responses successfully interpreted the histogram by adding a suitable scale and/or correctly completing values in the table. Many responses also plotted one or both bars accurately.

Parts (c) and (d) were more challenging. Only a small proportion of responses identified the type of skew, and fewer were able to interpret it. Some responses referred to the spread of times but did not link this to the position of the median. In part (d), many responses focused on the tag itself, suggesting it might fall off or malfunction. However,

the question stated that the time was recorded when the competitor passed the finish line, so such responses did not gain marks. More successful responses referred to the accuracy of the digital tag or its role in reducing human error.

Part (e) was more accessible, with over a third of responses gaining at least one mark. The mark scheme allowed credit for identifying that a systematic sample could save time, reduce cost, or be quicker or easier. Each valid point was awarded once.

Question 12

This question focused on interpreting percentage change and analysing data presented in a choropleth map. In part (a), a small proportion of responses correctly calculated the percentage change and identified that it fell within the 10–12% category. Many responses showed difficulty with the method, either not starting the calculation or not arriving at comparable values.

Part (b) was more accessible, with most responses correctly identifying Yorkshire and Humber as the region with the lowest percentage change. In part (c), a good number of responses challenged the conclusion by noting that the percentage change fell within a range or by referencing specific values that could apply to either region.

Part (d) was more challenging. Few responses recognised that percentage change and total number of vehicles are different measures. One mark was often awarded where responses identified that Scotland had the darkest shading or was the only region with a percentage change of 14% or more. Some responses incorrectly focused on total vehicle numbers rather than the percentage change.

Question 13

Question 13 was the final of the two common questions shared with the higher-tier paper. Responses from the foundation tier showed varied success across the different parts.

Parts (a) and (b) were more accessible. In part (a), around half of the responses correctly explained secondary data, often by stating it is data collected by someone else. Some responses attempted to give examples or referred to the internet, or discussed advantages and disadvantages of primary or secondary data, which did not meet the requirements of the question. In part (b), most responses successfully completed the stem-and-leaf diagram. Common errors included missing values, unordered leaves, or incorrect or missing keys.

Parts (c), (d), and (e) were more challenging. In part (c), few responses correctly identified the median age of 35 for Best Actress from 2000 to 2022. Some responses used the wrong diagram or struggled to count inwards to find the middle value. It was rare to see the ages written out in a list or to see recognition that the 12th value was needed.

In part (d), fewer than 7% of responses explained why the interquartile range (IQR) was more appropriate than the range. Many responses incorrectly stated that the range was too large or did not attempt the question. Successful responses identified the presence of outliers, often referencing the age of 83 from the Best Actor diagram. Very few responses correctly identified the lower and upper quartile values needed to calculate the IQR. Writing out the ages could have supported candidates who struggled to interpret the stem-and-leaf diagram.

Part (f) was an extended response question. Most responses attempted the task, with varying levels of success. The table provided helped candidates transfer values from parts (c) and (e) to support their comparisons. Some responses focused on changes over time for either Best Actor or Best Actress, but did not make direct comparisons or clearly state which data set or statistical measure was being discussed. The most successful responses were well structured, comparing medians and IQRs before 2000 and again from 2000 to 2022. Interpretations in context were often included after comparisons, which was an effective approach.

Question 14

In the final question, most responses showed an understanding of the task but struggled to interpret the full set of Spearman's rank correlation coefficients provided in the table. Where responses identified Experts 2 and 5, they often included a valid reason, typically noting that these experts had the highest rank values.

Summary

Based on performance in this paper, responses would benefit from the following areas of focus:

- Read each question carefully before attempting an answer to ensure full understanding of what is being asked.
- Show working to support final answers, particularly in calculation-based questions.

- Review how to make comparisons and practise how to communicate these clearly and accurately in responses.
- Practise interpreting statistical calculations and diagrams in context, ensuring that conclusions are based on the data provided.
- Develop understanding of when different calculations, diagrams, and approaches are appropriate, and be able to explain this clearly.
- Give a decision when required, and support it using information from the question.
- Communicate decisions with clarity and confidence. Avoid vague phrases such as “mostly appropriate”, “sometimes appropriate”, or “appropriate and not appropriate”, as these do not meet the criteria for a decision linked to a reason. Where both positive and negative points are included, clearly state which is which (e.g. “Appropriate because it reduces time; not appropriate because it excludes those who do not finish the race”).
- Practise extended response questions, and learn how to break them down using the steps outlined in the question where applicable.
- Review key terms and processes within the Statistical Enquiry Cycle, especially those covered in Question 7, such as data collection, calculations, and diagrams, and understand when each may or may not be suitable.
- Include calculator working on the question paper to support answers and allow for method marks to be awarded.