Statistical Enquiry Cycle: Teaching Guide and Practice Questions

GCSE (9–1) Statistics

Contents

1. Introduction 3

2. Statistical enquiry cycle 4

2.1 Overview 4

2.2 Explanation of each stage of the SEC 5

3. Suggested activities 6

3.1 Using and differentiating the activities 6

3.2 Examples of activities which would take place within each stage of the SEC 7

Stage 1: Importance of initial planning 7

Stage 2: Constraints involved in sourcing appropriate data 7

Stage 3: Ways that data can be processed and presented 8

Stage 4: Results must be interpreted with reference to the context of the problem 8

Stage 5: Evaluating statistical work 8

4. Template plan for a statistical investigation 9

5. Resources for practical work 11

6. Themes for investigation 12

6.1 Theme: Money 12

Ideas for investigation 12

Suggested hypotheses 12

Data sources 12

Other teaching activity 12

6.2 Theme: Weather 13

Ideas for investigation 13

Suggested hypotheses 13

Data sources 13

Other teaching activity 13

6.3 Theme: Team sports and team competitions 14

Ideas for investigation 14

Suggested hypotheses 14

Data sources 14

Other teaching activity 14

6.4 Theme: Track and field 15

Ideas for investigation 15

Suggested hypotheses 15

Data sources 15

Other teaching activities 15

6.5 Theme: Music 16

Ideas for investigation 16

Suggested hypotheses 16

Data sources 16

Other teaching activity – with additional data source 16

7. Generic questions 17

8. Practice questions 20

Foundation tier 20

Higher tier 23

Mark scheme – Foundation 27

Mark scheme – Higher 32

Appendix: Further themes for investigation 36

Overview 36

Theme: Employment 36

Theme: Weight 38

Theme: Books 39

Theme: Holidays 40

Theme: Climate 41

Theme: Journeys 42

Theme: Puzzles and games 43

Theme: Cars 44

Theme: Earthquakes 45

Theme: Time 46

Theme: Angles and lines 47

Theme: Films 47

Theme: Food 48

Theme: Paper and pencil 49

Theme: Reaction times 50

Theme: Road transport 51

Theme: Trees 52

Theme: Houses 53

Theme: Memory 54

Theme: Estimation 54

Theme: Human body 55

1. Introduction

Within GCSE (9-1) Statistics the new Assessment Objective 3, which contributes approximately 20% of the total assessment, will be assessed through questions on the component parts of the Statistical Enquiry Cycle. Furthermore “any given question may assess one stage or more than one stage of the SEC”.

With this in mind, during the delivery of the course students will be best prepared through exposure to each of the stages of the SEC in a meaningful way. This may be achieved most effectively by students engaging with practical statistical investigations of a type similar to the previous controlled assessment tasks.

The aim of this document is to support teachers in doing just that. While some teachers will be familiar with previous controlled assessment tasks, there is now greater freedom. Teachers may use and adapt themes, create mini-projects and provide feedback to students to develop their understanding of the SEC. Those new to teaching GCSE Statistics will find these short tasks are suitable for use in the classroom in a number of ways and at different stages of the course.

Contained within this guide are a number of topic themes with suggested data sets and hypotheses. As the techniques are taught and developed during the course, it would be beneficial if they were framed within the context of the five stages of the SEC, and perhaps illustrated using a data set so that students are regularly reminded of the cycle and appreciate how each technique fits within it. This no longer needs to be an extended piece of work; parts of a statistical investigation can be explored in relation to the current topic.

Key to success will be students’ ability to transfer the AO3 skills developed through these tasks to the open/extended response exam questions. To this end, included within this support guide are a number of practice questions which exemplify the type of questions for which students need to be prepared. Example solutions illustrate how the mark schemes will be applied.

The Assessment Objectives are as follows:

|  |  |  |
| --- | --- | --- |
| Students must: | | % in GCSE |
| AO1 | Demonstrate knowledge and understanding, using appropriate terminology and notation, of standard statistical techniques used to:   * collect and represent information * calculate summary statistics and probabilities | 55 |
| AO2 | Interpret statistical information and results in context and reason statistically to draw conclusions | 25 |
| AO3 | Assess the appropriateness of statistical methodologies and the conclusions drawn through the application of the statistical enquiry cycle | 20 |
| Total | | 100% |

2. Statistical enquiry cycle

2.1 Overview

Students will be best prepared for the assessment of AO3 through genuine engagement with practical investigations. Students need to fully understand how each of the five stages of the cycle work though authentic application (including the use of technology) to a variety of tasks.

The cycle may be summarised as follows:

Note that any given question may assess one stage or more than one stage of the statistical enquiry cycle.

Students should appreciate that working through the cycle to achieve a statistical conclusion is an iterative process of testing and refinement.

Teachers should signpost where each stage of the SEC is being used while working on a task in order to develop students’ fluency with the terminology and the cycle itself.

2.2 Explanation of each stage of the SEC

Students should:

3. Suggested activities

3.1 Using and differentiating the activities

When supporting students with this work it is helpful for them to be able to discriminate between the different stages of the SEC. The activities suggested here enable teachers to introduce the SEC at an early stage in the course by signposting the content being taught to the stage of the SEC where it could apply.

Within section 4 is a selection of themes with accompanying data sets. Using these, teachers can select an appropriate activity from the list below and apply it to a chosen theme and data set. While the list is extensive, it is anticipated that teachers would select those which match their learning objectives at the time.

There is scope within these activities to match them to different levels of demand according to the group being taught:

**Low demand**

* Comparing single variables such as totals, mean, median, mode and range.
* Representations such as pictograms, bar charts, multiple bar charts, pie charts, stem and leaf diagrams, scatter diagrams and time series graphs.

**Mid demand**

* Comparing one or more variables, measures of central tendency, measures of dispersion, quartiles, moving averages and relative frequency.
* Representations such as scatter diagrams to test for correlation with or without a line of best fit, box plots, cumulative frequency diagrams and time series graphs.

**High demand**

* Consider the interrelationships between several variables using more advanced calculations including frequency density, formal identification of outliers, the normal distribution, Spearman’s Rank Correlation Coefficient, standard deviation and standardised scores.
* Representations such as comparative pie charts, time series graphs, histograms with unequal intervals and the normal distribution curve.   
  Very able students can explore and use techniques beyond the GCSE specification.

3.2 Examples of activities which would take place within each stage of the SEC

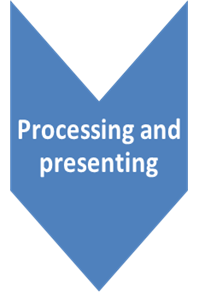
For each stage of the SEC there are activities for students to attempt. Teachers select an appropriate activity from the list below for the students to apply to their chosen theme and data set. Select the activity or activities that match your learning objectives at the time.

Stage 1: Importance of initial planning

* What are you going to investigate? Define a question or hypothesis to investigate. You can have just one question or you can consider related questions or hypotheses.
* The question(s) you plan to investigate influence the sample you collect. You need to think about what you expect the answers to your questions or hypothesis to be, and why you expect these answers. Decide what data to collect and how you will collect it:
  + Where will your data come from?
  + How reliable is your source?
  + How big a sample will you need?
  + How are you going to collect the data? Will you use all the data available? If not, what sampling method is to be used?
  + How are you going to record the data you collect?
  + What will you do about outliers or anomalies?
* What are the appropriate diagrams to draw to represent the data and what are the reasons for your choice? Develop a strategy for how to process and represent the data, giving reasons.
* What calculations will best summarise the data and address the hypothesis you are testing? Remember to consider outliers.
* Write a complete plan for an investigation. Your plan should state the hypothesis to be investigated and information on appropriate plans for data collection, diagrams and calculations.

Stage 2: Constraints involved in sourcing   
appropriate data

* Using a plan that you have developed or been given by your teacher, design the data collection sheet you need and use it to collect some primary data. Keep a record of the raw data.
* Using a plan that you have developed or been given by your teacher, research suitable sources of secondary data including the internet. Write down an acknowledgement of all the sources you have used.
* Describe the data collection process you used. Write down any problems that arose with the data and any limitations of the data. How did you get round these issues? If you were to gather similar data again, what would you do differently to avoid these issues?
* Work with a plan or hypothesis that touches on sensitive issues; for example, religion, health, drug use, income. How did working with a sensitive topic affect the data you could collect or use? How did you get round these difficulties?

Stage 3: Ways that data can be processed and presented

* Put the data you are working with into the form you need for processing. You may choose to use a spreadsheet or other software to help you organise the data.
* Produce appropriate diagrams in line with the plan or as directed by your teacher.
* Create at least one diagram in two ways: by hand and using a spreadsheet or other software. What are the advantages and disadvantages of each approach? Which is more appropriate for your purposes and why?
* Carry out calculations to compare data, in line with the plan you are using or as directed by your teacher.

Stage 4: Results must be interpreted with reference to the context of the problem

* Work with another student or your teacher to analyse a set of statistical diagrams and calculations. Discuss what the diagrams and calculations tell you about the topic you are investigating.
* Write down your conclusion(s). Your conclusions should relate to your investigation and back to the original hypothesis.   
  If appropriate, consider what predictions can be made from   
  the analysis.
* How reliable are the findings and why? Discuss the limitations of your findings.

Stage 5: Evaluating statistical work

* Evaluate your work. Did you get the results you expected? What were the weaknesses in the approach? Suggest improvements that you could make to the process, to the calculations you chose to carry out or the representations you selected and drew.
* Suggest how the process(es) could be refined to shed further light on the initial hypothesis.

4. Template plan for a statistical investigation

For each stage of the cycle there is a set of questions to which students should be able to respond.

|  |
| --- |
| **Title of the statistical investigation:**  **I will investigate the following hypothesis:** |
| Questions you need to be able to respond to regarding **stage 1 – planning**:   * Is your hypothesis clear and does it lead to a definite line of enquiry? * What questions are you hoping to answer that will help prove or disprove your hypothesis? * What data is relevant to your hypothesis? Why? * How much data will you collect/use? Why do you think this is sufficient? * Will you need to collect additional data or calculate additional values from the data given? How will you do this? * Will you be taking a sample? What sampling method is relevant? Why? * How will you avoid bias? * What problems might you have collecting the data? * How will you identify rogue values? How will you deal with these? * How will you identify outliers for your hypothesis? What will you do with these? * How will you deal with missing data? * How will you deal with duplicated data? * How will you record/manage your data/sample? * How will you process your data and present it? Decide now and justify your choice of methods. |
| Questions you need to be able to respond to regarding **stage 2 – collecting**:   * How will you record the data you collect, e.g. grouped or individual items? Why? * What problems have you considered for your data collection method? How have you addressed them? * If collecting secondary data, what were the difficulties you faced? How did you overcome these difficulties? * Have you given credit to your data sources? * Did you need to consider sensitivity in your data collection, e.g. age, personal questions, financial information ? |
| Questions you need to be able to respond to regarding **stage 3 - processing and** **presenting**:   * Will you need to put the data into tables? Will you use grouped frequency or cumulative frequency? * What **statistical calculations** are you planning to do? * Why have you chosen each technique? Why is it better than an alternative technique? * What technology will you use for your calculations? Why? * How will you know that your calculations make sense? * How will these calculations help prove your hypothesis? * What **statistical diagrams** will you draw? * How will you draw these? What technology will you use? * How will the diagrams help prove your hypothesis? |
| Questions you need to be able to respond to regarding **stage 4 – interpreting:**   * What do your calculations tell you? * What do your diagrams tell you? * What do you now know about your hypothesis as a result of your calculations and diagrams? * Can you make predictions based on what you have found out? What are they? * How certain are you of your findings? Why? |
| Questions you need to be able to respond to regarding **stage 5 – communicating and evaluating:**   * Have you communicated your investigation clearly and without waffle? * Have you used correct statistical terminology? * Have you explained your findings in a way that your audience will understand? * When evaluating your approach, have you identified any problems in your planning, calculations or diagrams? * Have you suggested improvements to your processing or representations? * How could you find out more about your initial hypothesis? |

5. Resources for practical work

|  |  |
| --- | --- |
| **Making Statistics Vital** | <http://www.making-statistics-vital.co.uk/> |
| **Stem Centre library of resources** | [www.stem.org.uk](http://www.stem.org.uk)  Free registration is required to access resources. |
| **Core Maths Support Programme** | [www.stem.org.uk/core-maths](http://www.stem.org.uk/core-maths) |
| **Nuffield Maths**  Using Data ~ Foundation  Data Handling ~ Higher  Data analysis activities ~ Level 3 | <http://www.nuffieldfoundation.org/nuffield-mathematics> |
| **Suffolk Maths**  Use the search facility to find all the activities related to statistics | <http://www.suffolkmaths.co.uk/> |
| **ICSE with Plymouth University** (formerlyRoyal Statistical Society Centre for Statistical Education)  Problem solving approach | <http://www.icse.xyz/psa/> |
| **Census at School resource library** | <http://www.censusatschool.com/resources.html> |
| **Stats for Schools** managed by the Office for National Statistics  Lesson plans and resources including data sets | <http://www.icse.xyz/stats4schools/lesson_ideas/default.html> |
| **TSM – Technology for Secondary Maths** | <http://www.tsm-resources.com/useful-files.html> |
| **Quibans**  Blog and list of activities/resources | <http://quibans.blogspot.co.uk/>  <https://docs.google.com/spreadsheets/d/1gYPMBVdrzPP-edg_RAIf_ffXD7Hj_4heuBaNY8B07Xc/edit#gid=0> |

6. Themes for investigation

6.1 Theme: Money

Ideas for investigation

The following are some possible ideas to investigate relating to money:

**•** how adult earnings have changed over time

**•** pocket money received

You can ask a variety of questions that can be investigated statistically.   
For example:

* How have adult weekly earnings changed over time?
* Do the changes vary depending on the sector in which people work?
* Do males and females earn the same amounts?
* How has the amount of pocket money received changed over time?
* Do the changes vary according to where you live?

Suggested hypotheses

1. Males earn more than females.
2. Male earnings have increased at a faster rate than females.
3. Pocket money has not increased in the last 10 years.

Data sources

The following are some data sources found at the time of publication that might prove helpful.

1. Weekly earnings for adults including a breakdown by sector:

<https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/averageweeklyearningsearn01>

1. Earnings for adults by gender, region and country:

<https://www.gov.uk/government/statistics/income-and-tax-by-gender-region-and-country-2010-to-2011>

1. Survey of pocket money over the time period 1987 to 2015 and comparison of change by region over two consecutive years:

<http://www.lloydsbankinggroup.com/globalassets/documents/media/press-releases/halifax/2015/150529---pocket-money.pdf>

Other teaching activity

Investigate pay rates for men and women, including construction of a cumulative frequency diagram in Excel and a review of the use of technology, using the following resources:

<http://www.nuffieldfoundation.org/fsmqs/level-2-data-handling#pay%20rates>

6.2 Theme: Weather

Ideas for investigation

The following are some possible ideas to investigate relating to weather:

• sunshine, rainfall or temperature

• maximum and minimum temperatures in different parts of the country

• seasonal variations in various aspects of weather

You can ask a variety of questions that can be investigated statistically.   
For example:

* + How have the different aspects of weather changed over time?
  + Do different locations have different weather?
  + How is rainfall/sunshine distributed?

Suggested hypotheses

1. Average temperatures in the United Kingdom have increased over time.
2. Rainfall is greater in the north of the UK than the south.

Data sources

The following is a data source found at the time of publication that might prove helpful.

• Pearson A level data set – weather:

<https://qualifications.pearson.com/en/qualifications/edexcel-a-levels/mathematics-2017.coursematerials.html#filterQuery=category:Pearson-UK:Category%2FSpecification-and-sample-assessments>

Use the UK weather data for 1987 and 2015 – Camborne, Heathrow, Hurn, Leeming and Leuchars.

To simplify the task, you could select just one weather station in the north and one in the south.

Other teaching activity

Music Festival – a short task asking students to use weather data to decide the best month to hold a festival:

<http://www.nuffieldfoundation.org/fsmqs/level-2-data-handling>

6.3 Theme: Team sports and team competitions

Ideas for investigation

The following are some possible ideas to investigate relating to team sports and competitions:

**•** the number of goals, points or runs scored in matches or competitions

**•** capacity of stadia and grounds

**•** ticket (and other) prices for matches

You can ask a variety of questions that can be investigated statistically.   
For example:

* + How have the points awarded in a team competition varied over time?
  + How do the capacities of stadia vary between countries and sports?
  + How have the capacities of sports stadia changed over time?
* How do costs (tickets, programmes, refreshments) vary between different football clubs in different leagues?

Suggested hypotheses

1. Points awarded in synchronised swimming competitions have increased over time.
2. Football stadia have a greater capacity than other sports stadia.
3. The average football stadium capacity in England is greater than that of other European countries.
4. The more successful your club, the greater the cost will be to watch them play.

Data sources

The following are some data sources found at the time of publication that might prove helpful.

1. Links to results for all levels of synchronised swimming competitions from 2004 onwards:

<http://www.swimming.org/britishswimming/synchro/results/>

1. Lists of football and other stadia around the world together with capacity and other details:

<http://www.stadiumguide.com/> <http://www.worldstadiums.com/europe/maps/europe.shtml>

1. Information on ticket, programme, tea and pie prices for football clubs in ten divisions of British football in 2016:

<http://www.bbc.co.uk/sport/football/37953195>

Other teaching activity

Sports injuries – decide what the most dangerous sport is:

<http://www.nuffieldfoundation.org/fsmqs/level-2-data-handling>

6.4 Theme: Track and field

Ideas for investigation

The following are some possible ideas to investigate relating to track and field:

• the times taken by boys and by girls to run races

• the statistics from past track and field events

• the ages and gender of athletes

You can ask a variety of questions that can be investigated statistically.   
For example:

* How is performance affected by age and/or gender?
* How has the performance of athletes changed over time?

Suggested hypotheses

1. 100m and 200m records are improving at the same rate.
2. Women’s 100m record times will be greater than men’s by about 1 second.

Data sources

The following is a data source found at the time of publication that might prove helpful.

• Records for all official timings for athletics events during a year:

[www.iaaf.org](http://www.iaaf.org)

Other teaching activities

* How far, fast and high? From RSSCSE project at Plymouth University:

<http://www.icse.xyz/psa/resource5.html>

* Who’s the best? Based on 2012 Olympic data:

<https://nrich.maths.org/7367>

6.5 Theme: Music

Ideas for investigation

The following are some possible ideas to investigate relating to music:

• the lengths and numbers of beats per minute of pieces of music

• the length of time a single remained at number 1 or in the top 40 in the UK singles chart

You can ask a variety of questions that can be investigated statistically.   
For example:

* + Have the playing times of singles in the UK singles chart changed over time?
  + Have the lengths of time spent at number 1 in the UK singles chart by singles changed over time?
  + Is there a relationship between the length of a piece of music and the number of beats per minute? Does this differ between genres of music? Does this differ between music with lyrics and music without lyrics?
  + How do the playing times or numbers of beats per minute in different genres of music, or music from different producers, compare?

Suggested hypotheses

1. The playing time of UK singles is shorter now than it was 20 years ago.
2. The typical length of time a song remains at number one in the singles chart has decreased in the last 40 years.
3. Longer pieces of music have fewer beats per minute.

Data sources

The following are some data sources found at the time of publication that might prove helpful.

1. Current and archive information about UK charts:

<http://www.officialcharts.com/archive-chart/_/1/1960-03-12/>

1. Data on the year of release, length and beats per minute of a large number of songs which can be sorted by artist, song title, genre and other categories:

<http://www.cs.ubc.ca/~davet/music/index.html>

Other teaching activity – with additional data source

My Music – this is a substantial task from the Bowland Maths Project. Within the Zip File is an excel spreadsheet containing the beats per minute for the longest running number one track from each year in the period 1960–2007:

<http://www.bowlandmaths.org.uk/projects/my_music.html>

7. Generic questions

**Questions**

1. What is the aim of your statistical enquiry? What do you hope to find out?
2. State your main hypothesis, with one or two additional sub-hypotheses related to your main enquiry. Outline your strategy for investigating this line of enquiry.
3. Identify the data source(s) you will use, including why they are suitable.
4. Explain how you will collect your data, including details of your sampling method and the strategies you employed to avoid bias.
5. Outline the calculations you will perform. Justify how each calculation will enable you to determine the extent to which your initial hypothesis is true.
6. Select one or more suitable data representation methods for your data. Explain why you have chosen each one and how each diagram or graph will enable you to make a decision regarding the validity of your hypothesis.
7. Describe how you have used technology in your enquiry. In what ways did this help you to achieve your aims?
8. Consider the results from your calculations and diagrams. What conclusions have you reached relating to your original hypothesis/hypotheses? To what extent are your results reliable? Explain how you know.
9. Have you used appropriate and correct statistical terminology throughout your enquiry? How has this helped to communicate your findings to your audience?
10. Evaluate your approach to this enquiry, identifying any problems which arose. In what ways could your enquiry be improved?
11. If you were going to develop this enquiry, what would your next steps be?

**Notes on answers**

1. A clear statement of what the student wishes to find out, moving from the general to the specific. E.g. “I wish to explore reaction times, specifically the differences between genders and how any differences might vary according to the type of reaction task. An example could be comparing reaction times using technology activities compared to physical activities such as the ‘dropping a ruler’ task.” Context might be given for the question posed.
2. Hypothesis is stated clearly and succinctly. Further sub-hypotheses are stated which relate to the main hypothesis. The student gives an outline of their approach which could include: identifying relevant data, where they will obtain the data from, the processes they will use to answer the question or hypothesis/hypotheses) posed and how they will represent their findings.
3. Sources of data are given. The student should explain why they have selected the data sources, perhaps referring to their relevance to the enquiry, whether they are up to date, and the size and scope of the database. Appreciation of the requirement to acknowledge any sources used. In this stage of the SEC they might include plans for any surveys or questionnaires, in which case students should explain why this is appropriate and give detail of their approach to survey design and execution. Areas of sensitivity such as age, weight or salary would be expected to be referred to, since they will impact on potential data availability.
4. A recognised method of sampling should be used. The steps taken are described in detail. The student should justify the sampling method chosen, explaining why it is appropriate in this case. The measures which have been taken to avoid bias in the sample should be described in detail. Consideration of sample size and outliers would also be expected at this stage.
5. Appropriate calculations should be selected and carried out clearly in order to determine the extent of support for the hypothesis. The student justifies their choice of calculation, including how the outcome will be used in order to determine whether their hypothesis is true. If technology is used they should explain how, e.g. use of calculator functions or Excel. They could indicate how they know their answer is correct or free of errors.
6. The student uses a technique to present the data using graphs or diagrams. The chosen method is justified with reference to how it will enable conclusions to be drawn regarding the hypothesis. The technique is appropriate for the task with no redundancy. Sufficient representations are constructed to enable comparisons to be made and thus correct conclusions can be drawn. The diagrams are used to make a decision regarding the extent to which the hypothesis is true.
7. The student describes and discusses their use of technology. This should include the benefits, such as faster processing, automation of processing, increased accuracy etc, as well as any disadvantages such as possible errors.
8. The calculations and diagrams should be interpreted coherently in the context of the task. Conclusions are drawn regarding the main and sub-hypotheses with reference to calculations/diagrams. There is evidence of a critical awareness of the extent to which any inferences can be made. Reference is made to limitations with detailed explanations, such as “while I can be confident that my results are reliable for the years 2000 and 2008, I would need to investigate further data sets to be certain that...”.
9. At this stage the student should confirm that they have communicated statistical concepts using correct terminology during each stage of the SEC, perhaps giving examples. They should review how effectively they have communicated their findings to their audience, for example how setting the results in the context of the original question posed has informed their conclusions and perhaps suggested further lines of enquiry.
10. Evaluation of the strengths and weaknesses of the approach taken. This should include suggestions for improvement, such as:

* changes which could be made to the original plan, with reasons for them stated;
* any problems which occurred with the data collection, how they were resolved and whether they could have been anticipated. Proposed alternative strategy to be used;
* issues which arose when carrying out the calculations or presenting results in diagrammatic form. Discussion of the chosen calculations and diagrams and whether they gave sufficient information to make a decision regarding the hypothesis/hypotheses). Reasons are identified for calculations or diagrams which proved to be less effective, together with suggested improvements.
* Whether the cycle has given sufficient information to determine the strength of the initial hypothesis. Suggestions given for further investigation or adaptations identified in order to gain clarity.

1. Discussion of results and where they lead in terms of further investigation. This might include a statement of whether this line of enquiry is concluded and avenues for further investigation which may have been identified. If the latter then potential next steps are clearly stated.

8. Practice questions

Questions marked with an asterisk (\*) are suitable for either tier.

Foundation tier

1. David wants to find out how much money people spend on their gardens.

David is going to send questionnaires to a sample of people included in the mailing list of a local garden centre.

1. Discuss whether or not this is an appropriate method of data collection. (2)

Here is one of the questions that David wants to put on his questionnaire.

A garden is a really important feature of a house, isn’t it?

Strongly agree Agree Don’t know Strongly disagree

1. This is not a suitable question.

Give **two** reasons why. (2)

Here is another of the questions that David wants to put on the questionnaire.

How much money do you spend on your garden?

£10–20 £20–30 £30–50

1. Discuss whether or not this is a suitable question for the questionnaire. (2)

**(Total for Question 1 is 6 marks)**

**2** Ricardo collected data about the favourite type of music of 51 people in his youth club.

The pictogram gives information about the favourite types of music.

|  |  |
| --- | --- |
| Pop |  |
| Rap |  |
| Rock |  |
| Classical |  |

Key represents 4 people

1. Which type of music was chosen by the greatest number of people? (1)
2. Which type of music was chosen by twice as many people as Classical music? (1)

Ricardo concludes that Rap is the favourite type of music of people in his city.

1. Give **two** limitations of Ricardo’s conclusion. (2)

**(Total for Question 2 is 4 marks)**

**3** Tina is investigating the performance of athletes at the Paralympics.

She has collected data for the distances thrown, in metres, by male and female athletes competing in F34 Javelin.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Male** | 14.24 | 15.00 | 15.29 | 16.96 | 17.88 | 18.16 | 18.42 | 21.86 |
| **Female** | 13.84 | 27.70 | 27.83 | 30.36 | 31.31 | 33.42 | 34.15 | 36.65 |

Tina wants to compare the average distance thrown by male athletes with the average distance thrown by female athletes.

Tina thinks that she should use the mean to compare the average distance thrown.

1. Give one reason why the mean would not be an appropriate average to use to represent the average distance thrown by female athletes. (1)

Tina plans to use a scatter diagram to compare the distances thrown by the male athletes with the distances thrown by the female athletes.

1. Discuss whether or not a scatter diagram would be a suitable diagram to use. (2)

**(Total for Question 3 is 3 marks)**

**4\*** A researcher is investigating how much money people in Boxhill spend on shopping.

One hypothesis she investigates is

“Women spend more on shopping than men.”

The researcher could find it difficult to collect information to test her hypothesis.

1. Give one difficulty the researcher could have when trying to find out how much money people spend on shopping. (1)

The researcher plans to ask 60 people about their shopping habits.

She plans to ask the first 30 males and the first 30 females who walk into the local shopping centre on Monday morning to answer her questionnaire.

(b) Give one advantage and one disadvantage of this sampling method. (2)

The manager of a large shop in the shopping centre also wants to find out about shopping habits.

He has a spreadsheet with information gathered from the store loyalty card. It contains records for over 1000 shoppers.

He plans to draw diagrams and perform calculations using the data in the spreadsheet.

(c) How will having the data in a spreadsheet help the manager?

Give **two** examples. (2)

**(Total for Question 4 is 5 marks)**

**5\*** Irina and Marta work for a local newspaper.

They have collected information on the amount of time spent listening to music each week and the IQ test score for 10 of their readers.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Reader number** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **Time spent listening to music (minutes)** | 120 | 150 | 90 | 40 | 110 | 30 | 75 | 90 | 250 | 210 |
| **IQ test score** | 110 | 115 | 94 | 103 | 111 | 102 | 123 | 116 | 125 | 131 |

Irina and Marta want to see whether there is a relationship between the amount of time spent listening to music each week and the IQ test score.

Irina wants to represent the data using box and whisker plots.

Marta wants to represent the data using a scatter diagram.

1. Which one of box and whisker plots or scatter diagram do you think they should use for their investigation?

Give a reason for your answer. (2)

The report is for a local newspaper.

Irina thinks that they should include the Spearman’s rank correlation coefficient with their report.

1. Discuss whether or not Spearman’s rank correlation coefficient would be appropriate to include. (2)

**(Total for Question 5 is 4 marks)**

**6\*** Robert is investigating weather.

He wants to investigate how the rainfall in his town has changed over the years.

He plans to collect data on yearly rainfall from the last 100 years from the internet.

(a) Explain why this is an appropriate approach to data collection. (1)

(b) Identify one possible problem which Robert may have when collecting this data. (1)

(c) Suggest a suitable diagram that Robert could use to show how the rainfall in his town has changed over the years.

Give a reason for your answer.

(2)

**(Total for Question 6 is 4 marks)**

**7\*** Louise wants to investigate how the length of pop songs (in minutes) compares to the length of rock songs (in minutes).

Here is her plan:

Use “Are rock songs longer than pop songs?” as the hypothesis.

Use random sampling to collect data from a music database on the lengths in minutes of 10 rock songs and 10 pop songs.

Draw comparative box plots to compare the length of the pop songs and the rock songs.

Comment on the suitability of each aspect of Louise’s plan.

Include details of how she could use her diagrams to compare.

**(Total for Question 7 is 7 marks)**

**8\*** Lucas wants to investigate how an athlete’s performance in one field event is related to their performance in a different field event.

He wants to collect data on the shot put distances and javelin throw distances for athletes by using the internet.

Write down **one** thing that he should include in his plan for each of:

* data collection
* analysing and presenting the data
* interpreting his diagrams and / or calculations

Explain why each of these things is appropriate.

**(Total for Question 8 is 6 marks)**

Higher tier

**1\*** Mr Jones is investigating the amount of time that students at his school spend on their homework each week.

He writes the following hypothesis for his investigation:

“Older students will spend more time on their homework each week than younger students.”

Mr Jones decides to use a census of students from his school.

He is going to ask each student to record the amount of time that they spent on their homework in the last week.

Mr Jones then collects this information from each student by asking them to complete an online database during their registration time.

Some of the database is shown below.

|  |  |  |
| --- | --- | --- |
|  | **Age** | **Time spent on homework in the last week** |
| 1 | 14 | 50 min |
| 2 | 15 years old | 10 |
| 3 | 11 ½ | 90 minutes |
| 4 | 13 | Don’t know |
| 5 | 13years 4months | 3 hours |
| 6 | 12 | 40 |
| 7 | 14y 7m | * 1. hours |

1. Give two reasons why Mr Jones must clean the data before processing it. (2)
2. Discuss how Mr Jones’ data collection plan could affect the reliability of his conclusions. (2)

**(Total for Question 1 is 4 marks)**

**2\*** Anna is investigating whether there is a difference between the amount of money spent on sport by boys and the amount of money spent on sport by girls.

She writes the following hypothesis for her investigation:

“The amount of money spent on sport by boys is greater than the amount of money spent on sport by girls.”

Anna is planning to take a quota sample of 10 boys and 10 girls at the local leisure centre.

Comment on whether Anna’s plan is appropriate.

**(Total for Question 2 is 3 marks)**

**3\*** Mike and Nick are investigating the number of concerts people attended in the last year.

The table shows information about the number of concerts attended by 100 people.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Number of concerts** | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| **Frequency** | 15 | 14 | 12 | 14 | 9 | 11 | 9 | 8 | 5 | 3 |

Mike and Nick want to work out the average number of concerts attended by these 100 people in the last year.

Mike thinks that they should find the mean of 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9

Nick thinks that they should multiply each number of concerts by its frequency, add these up and divide by 100

Which one of these two averages should they use?

Give a reason for your answer.

**(Total for Question 3 is 2 marks)**

**4\*** Dominique is investigating to see if there is an association between the tail length and wing   
length of female hook-billed kites.

Her hypothesis is

“Kites with a high tail length will have a high wing length”

Dominique finds the tail length (in millimetres) and the wing length (in millimetres) for each of 10 female hook-billed kites.

This information is shown in the table below.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Tail length (mm)** | 191 | 197 | 208 | 180 | 188 | 210 | 196 | 191 | 179 | 216 |
| **Wing length**  **(mm)** | 284 | 285 | 288 | 273 | 280 | 283 | 288 | 271 | 257 | 305 |

(Source: *public.iastate.edu*)

Some of this information is shown on the scatter diagram below.

170

190

180

200

210

220

290

280

270

260

300

Wing length (mm)

Tail length (mm)

320

310

250

**x**

**x**

**x**

**x**

**x**

**x**

**x**

**x**

The last two pieces of data in the table have not been plotted on the scatter diagram.

1. Plot the last two points on the scatter diagram. (1)

(b) Explain, giving a statistical reason, whether or not this scatter diagram supports Dominique’s hypothesis. (2)

(c) Draw a line of best fit on the scatter diagram. (1)

Dominique wants to use the line of best fit to predict the wing length of a female hook-billed kite with tail length 240 mm.

(d) Is it appropriate for Dominique to use the line of best fit for this prediction?

Give a reason for your answer. (2)

Dominique wants to use the line of best fit to predict the tail length of a male hook-billed kite with wing length 285 millimetres.

(e) Explain why it is **not** appropriate for Dominique to use the line of best fit for this prediction. (1)

**(Total for Question 4 is 7 marks)**

**5\*** Michele wants to compare the heights of football players with the heights of rugby players.

She has written a plan for her investigation.

**Data collection**

Collect data from the internet.

Collect data on the heights of 20 football players and the heights of 20 rugby players.

**Diagrams and calculations**

Draw a scatter graph of the heights of football players and the heights of rugby players.

Calculate the Spearman’s rank correlation coefficient.

Discuss whether Michele’s plan for her investigation is appropriate.

**(Total for Question 5 is 6 marks)**

1. Timo works for a holiday company.

He has been asked to produce a report on any trends in quarterly profits over the last two years.

He has been provided with the quarterly profit figures for the last two years.

He is going to write a plan for his report.

Which diagrams and calculations should Timo include in his report?

Explain how these diagrams and calculations will be interpreted.

Indicate any limitations on the conclusions that Timo can draw.

**(Total for Question 6 is 6 marks)**

1. Munira is investigating whether the distances jumped in the Olympic long jump competition are normally distributed.

She is going to write a plan for her investigation.

Write down **one** thing that she should include in her plan for each of:

* data collection
* analysing and representing data
* interpreting the diagrams and / or calculations

Explain why each of these things is appropriate.

**(Total for Question 7 is 6 marks)**

Mark scheme – Foundation

|  |  |  |  |
| --- | --- | --- | --- |
| **Question number** | **Answer** | **Additional guidance** | **Mark** |
| **1(a)** | B2 for not appropriate AND a correct reason, e.g.   * reference to sampling people who use the garden centre * potentially low response rate   OR if B2 not earned…  B1 for not appropriate with an attempt at a reason | B2 for a complete answer assessing the appropriateness of the method of data collection  OR if B2 not earned…  B1 for an incomplete answer assessing the appropriateness of the method of data collection | (2) |
| **1(b)** | B1B1 for each correct comment from:   * biased / leading question * options vague * two positive (agree) options, only one negative (disagree) option | B1B1 for assessing the appropriateness of the given question | (2) |
| **1(c)** | B2 for not suitable AND a correct reason, e.g.   * no option for 0 / no option for more than £50 * overlapping response boxes * no time frame   OR if B2 not earned…  B1 for not suitable with an attempt at a reason | B2 for assessing the suitability of the given question and giving a supporting reason  OR if B2 not earned…  B1 for an incomplete answer assessing the suitability of the given question | (2) |

|  |  |  |  |
| --- | --- | --- | --- |
| **2(a)** | B1 for Rap |  | (1) |
| **2(b)** | B1 for Rock |  | (1) |
| **2(c)** | B1B1 for each of two limitations of Ricardo’s conclusion,  e.g.   * the data only relates to people who attend the youth club * the data only relates to one age group * this is only a small sample | B1B1 for each comment assessing the limitation of the stated conclusion | (2) |
| **3(a)** | B1 for a correct reason, e.g.   * reference to the extreme value in the female distances thrown * reference to the mean being affected by extreme values | B1 for assessing the appropriateness of the mean as a choice of average | (1) |
| **3(b)** | B2 Not suitable as the data is not bivariate / in related pairs oe  OR if B2 not earned...  B1 Not suitable with an attempt at a reason | B2 for a complete answer assessing the appropriateness of the choice of diagram  OR if B2 not earned...  B1 for an incomplete answer assessing the appropriateness of the choice of diagram | (2) |
| **4(a)\*** | B1 e.g. some people may not want to say how much money they spend on shopping (personal question)  OR  B1 e.g. people may interpret ‘shopping’ in different ways / may not be clear on what they should include in the total |  | (1) |
| **4(b)\*** | B1 Advantage:   * e.g. convenient * e.g. equal numbers of men and women sampled   B1 Disadvantage:   * e.g. not representative / people shopping on a Monday morning unlikely to be representative of everyone who lives in Boxhill * e.g. biased / asking people who are going shopping | 1st B1 for any correct advantage  2nd B1 for any correct disadvantage | (2) |
| **4(c)\*** | B1B1 for each appropriate comment, e.g.   * a spreadsheet makes data easy to sort * the spreadsheet can be used to perform statistical calculations quickly on large amounts of data * the spreadsheet can be used to produce graphs containing large amounts of data faster than doing it by hand | B1B1 for assessing the use of technology to analyse data | (2) |
| **5(a)\*** | B2 for scatter diagram with a correct reason, e.g.   * the data is bivariate * looking for a relationship / correlation / association between the two variables   OR if B2 not earned...  B1 for scatter diagram with an attempt at a reason | B2 for a complete answer assessing the appropriateness of the choice of diagram  OR if B2 not earned...  B1 for an incomplete answer assessing the appropriateness of the choice of diagram | (2) |
| **5(b)\*** | B2 for not suitable with an appropriate reason, e.g.   * the target audience will not understand the correlation coefficient   OR  B1 for suitable with an appropriate reason, e.g.   * Spearman’s rank correlation coefficient will give a numerical measure of the correlation | B2 for a complete answer assessing the appropriateness of the choice of calculation considering target audience  OR if B2 not earned...  B1 for an incomplete answer assessing the appropriateness of the choice of calculation without consideration of target audience | (2) |
| **6(a)\*** | B1 for an appropriate reason, e.g.   * it is not possible to collect primary data for rainfall over the last 100 years * the large sample size (100 years of data) would increase the reliability of the conclusions | B1 for a complete answer assessing the appropriateness of the method of data collection | (1) |
| **6(b)\*** | B1 for a possible problem with data collection, e.g.   * data for some years may be missing * data may not be recorded in a consistent way over the last 100 years | B1 for any appropriate potential problem identified | (1) |
| **6(c)\*** | B2 for time series graph with an appropriate reason, e.g.   * this will show the trends in rainfall over the years   OR if B2 not earned  B1 for time series with an attempt at reason | B2 for a complete answer including assessment of appropriateness of the choice of diagram  OR if B2 not earned...  B1 for an incomplete answer assessing the appropriateness of the choice of diagram | (2) |
| **7\*** | B1 for hypothesis not suitable with an appropriate reason, e.g.   * this is a question rather than a hypothesis   B1B1 for each correct comment from   * sample size is small * random sampling means that all songs have an equal chance of being selected * specifies collecting song lengths in minutes   B2 for e.g. she can use the box plots to compare medians which show comparison of average length of songs  OR if B2 not scored…  B1 for can compare medians  AND  B2 for e.g. she can use the box plots to compare interquartile ranges / ranges which show comparison of spread of / consistency of song lengths  OR if B2 not scored…  B1 for can compare interquartile ranges / ranges | B1 for a complete answer assessing the appropriateness of the proposed hypothesis  B1B1 for assessing appropriateness of approach to data collection  B2B2 for assessing appropriateness of approach to presenting data | (7) |
| **8\*** | **Data collection**  B1 for identifying one thing that should be included in the plan for data collection  **and**  B1 for explaining why this aspect is appropriate  B1 for deciding what data to collect and / or how to collect and record it **and**  B1 for an appropriate reason  OR  B1 for designing a collection method for primary / secondary data **and**  B1 for an appropriate reason  OR  B1 for appreciating the importance of acknowledging sources **and**  B1 for an appropriate reason  **Analysing and representing data**  B1 for identifying one thing that should be included in the plan for analysing and representing data  **and**  B1 for explaining why this aspect is appropriate  B1 for planning to generate diagrams and / or visualisations to represent the data  B1 for an appropriate reason  OR  B1 for planning to generate statistical measures to compare data  B1 for an appropriate reason  **Interpreting your diagrams and / or calculations**  B1 for identifying one thing that should be included in the plan for interpreting diagrams and / or calculations  **and**  B1 for explaining why this aspect is appropriate  B1 for planning to interpret diagrams and / or calculations / measures **and**  B1 for an appropriate reason  OR  B1 for planning to reach conclusions that relate to their hypothesis / question **and**  B1 for an appropriate reason | B1B1B1 for each of three planned elements **and** B1B1B1 for each of three appropriate reasons  B1 for e.g. using distances measured to the nearest centimetre **and**  B1 for e.g. as this will be sufficient to distinguish between the distances  B1 for e.g. use results from competitions where athletes throw both shot put and javelin **and**  B1 for e.g. this means that there will be paired data  B1 for identifying the source of the data, e.g. an official sporting event website **and**  B1 for e.g. it allows other people to judge the reliability of your source  B1 for e.g. use a scatter diagram for the data **and**  B1 for e.g. as this will allow you to see if there is a correlation (between the shot put throw distance and the javelin throw distance)  B1 for e.g. calculate the mean point **and**  B1 for e.g. to be able to plot an accurate line of best fit  OR  B1 for e.g. calculate the equation of the line of best fit **and**  B1 for e.g. to be able to see the relationship between shot put throw and javelin throw  B1 for e.g. interpret results for men and women separately **and**  B1 for e.g. as results for men and women may not be the same  B1 for e.g. looking for correlation on a scatter diagram **and**  B1 for e.g. you can see whether, as the distance thrown in the shot put increases, the distance thrown in the javelin increases. | (6) |

Mark scheme – Higher

|  |  |  |  |
| --- | --- | --- | --- |
| **Question number** | **Answer** | **Additional guidance** | **Mark** |
| **1(a)\*** | B1B1 for two correct reasons, e.g.   * ages given in different formats (nearest year / year and month, minutes / hours) * remove extraneous symbols * remove missing data (‘don’t know’) | B1 for each correct reason for the need to clean data on the database prior to processing it | (2) |
| **1(b)\*** | B1B1 for two correct aspects, e.g.   * large sample size increases reliability * issues with consistency in recording time taken for homework * issues with how the data collection is carried out (may be errors in data entry) * potential for students to lie about how much time they spend on homework * non-response decreases reliability | B1B1 for two correct comparisons assessing the reliability of the conclusions drawn | (2) |
| **2\*** | B1B1B1 for three correct aspects, e.g.   * sample size small * biased – not appropriate to sample at leisure centre * boys / girls selected for her sample may be in a group * both boys and girls included in sample | B1 for each correct comment assessing the appropriateness of the plan for data collection | (3) |
| **3\*** | B2 for Nick’s mean AND reference to there being different proportions of people going to different numbers of concerts  OR  B1 for Nick’s mean with attempt at reason | B2 for a complete assessment of the appropriate choice with reason  OR  B1 for an incomplete assessment of the appropriate choice | (2) |
| **4(a)\*** | B1 for two correct points plotted at (179, 257) and (216, 305) |  | (1) |
| **4(b)\*** | B1 for graph supports the hypothesis and reason  B1 for scatter graph shows **positive correlation** | B1 for conclusion supported by sensible reason  B1 for statistical reasoning using words in bold | (2) |
| **4(c)\*** | B1 for an appropriate line of best fit |  | (1) |
| **4(d)\*** | B2 for not appropriate AND reference to extrapolation  OR if B2 not earned…  B1 for not appropriate with an attempt at a reason | B2 for assessing the appropriateness of the statistical methodology with reason  OR if B2 not earned…  B1 for an incomplete attempt to assess the appropriateness of the statistical methodology | (2) |
| **4(e)\*** | B1 for reference to the scatter graph being for female kites and not male kites | B1 for correct explanation showing understanding of the limitations of the conclusions that can be drawn from the data provided | (1) |
| **5\*** | **Data collection**  B1B1B1 for correct comments relating to the method of data collection, e.g.   * source of data is vague * need to choose reliable internet sites * sample size is acceptable * need to be more specific about which football players / rugby players will be selected (from which leagues) * need to describe how they will sample the 20 football players and 20 rugby players   **Diagrams and calculations**  B1B1B1 for correct comments relating to the method of processing and presenting, e.g.   * scatter graph / Spearman’s rank correlation coefficient not appropriate * not bivariate data * we want to compare rather than look for correlation * comparative box plots would be more appropriate | B1B1B1 for comments assessing the appropriateness of the approach to data collection  B1B1B1 for comments assessing the appropriateness of the diagrams and calculations | (6) |
| **6** | **Diagrams**  B2 for time series AND appropriate reason, e.g. allows trends to be seen over time  OR if B2 not scored…  B1 for time series with an attempt at a reason  **Calculations**  B2 for reference to an appropriate calculation AND reason, e.g.   * calculate (4-point) moving averages AND helps to see trends * calculate seasonal variation AND to see how different seasons affect the profit   OR if B2 not scored…  B1 for moving averages or seasonal variation with an attempt at a reason  **Limitations**  B1B1 for correct comments relating to limitations, e.g.   * conclusions are only valid for these two years of data * reference to any identified trends may not continue * reference to only applying to this company | B2 for assessing the appropriateness of statistical methods for presenting data  B2 for assessing the appropriateness of statistical methods for processing / analysing data  B1B1 for identifying limitations of any conclusions | (6) |
| **7** | **Data collection**  B1 for identifying one thing that should be included in the plan for data collection  **and**  B1 for explaining why this aspect is appropriate  B1 for deciding what data to collect and / or how to collect and record it **and**  B1 for an appropriate reason  OR  B1 for designing a collection method for primary / secondary data **and**  B1 for an appropriate reason  OR  B1 for appreciating the importance of acknowledging sources **and**  B1 for an appropriate reason  **Analysing and representing data**  B1 for identifying one thing that should be included in the plan for analysing and representing data  **and**  B1 for explaining why this aspect is appropriate  B1 for planning to generate diagrams and / or visualisations to represent the data  B1 for an appropriate reason  **OR**  B1 for planning to generate statistical measures to compare data  B1 for an appropriate reason  **Interpreting your diagrams and / or calculations**  B1 for identifying one thing that should be included in the plan for interpreting diagrams and / or calculations  **and**  B1 for explaining why this aspect is appropriate  B1 for planning to interpret diagrams and / or calculations / measures **and**  B1 for an appropriate reason  OR  B1 for planning to reach conclusions that relate to their hypothesis / question **and**  B1 for an appropriate reason | B1B1B1 for each of three planned elements **and** B1B1B1 for each of three appropriate reasons.  B1 for e.g. use distance for jump to the nearest centimetre **and**  B1 for e.g. this is sufficient as there will be a large range of distances / this is an appropriate measurement accuracy for jumps  B1 for e.g. use census of all Olympic finals for last four Olympics **and**  B1 for e.g. this ensures that you have enough data / using data from heats and finals would include the same athletes more than once  OR  B1 for e.g. collecting data for male athletes **and**  B1 for e.g. male and female athletes are likely to jump different distances  B1 for identifying the source of the data, e.g. official website of the Olympics **and**  B1 for e.g. it allows other people to judge the reliability of your source  B1 for e.g. use a histogram for the data **and**  B1 for e.g. as this will allow you to see the shape of the distribution  B1 for e.g. calculate the mean and standard deviation **and**  B1 for e.g. to be able to see how much of the data lies within 2 standard deviations of the mean  B1 for e.g. looking for a bell shape to the histogram **and**  B1 for e.g. normally distributed data would fit a bell curve  B1 for e.g. considering the amount of data within 2 standard deviations of the mean **and**  B1 for e.g. if data is normally distributed then 95% would be expected to lie within + / - 2 standard deviations of the mean | (6) |

Appendix: Further themes for investigation

Overview

The themes in this appendix comprise the remainder of the past controlled assessment briefs (from specification 2ST01) not included in section 6. They are presented in reverse order of date, from the most recent, in 2017, going back to 2010 and the original sample assessment materials.

Data sources, where given, have been checked as of the date of publication of this document. Teaching activity ideas have been given in some instances.

The following table lists *all* the controlled assessments briefs from specification 2ST01. Those given in bold can be found in section 6 of this document. The remainder are given in this appendix.

|  |  |  |  |
| --- | --- | --- | --- |
| **Series** | **Controlled assessment themes** | | |
| 2017 | Employment | **Music (Theme 6.5)** | Weight |
| 2016 | Books | Holidays | **Team sports and team competitions (Theme 6.3)** |
| 2015 | Climate | Journeys | Puzzles and games |
| 2014 | Cars | Earthquakes | Time |
| 2013 | Angles and lines | Films | Food |
| 2012 | **Money (Theme 6.1)** | Paper and pencil | **Track and field (Theme 6.4)** |
| 2011 | Reaction times | Road transport | Trees |
| 2010 | Houses | Memory | **Weather (Theme 6.2)** |
| SAMs | Estimation | Human body |  |

Theme: Employment

Ideas for investigation

The following are some possible ideas to investigate relating to employment but there are many others that you might choose.

You may want to collect information about:

• the types of job that people have

• people’s feelings about their jobs

• the jobs students would like to do in the future

• job vacancies in different parts of the country

• rates of pay

• retirement ages in different countries

• the average number of hours worked in different countries

• employment rates

• levels of trade union membership over time

• height or weight of people working in a particular job.

You can ask a variety of questions that can be investigated statistically. For example:

• Which jobs would students like to do in the future?

• How do the types of jobs done by men and women compare?

• Is there an association between ratings for job satisfaction, work stress, job meaning, job flexibility and/or pay?

• How do the numbers of job vacancies for a particular job in different parts of the country compare?

• How do the rates of employment in different categories of jobs, e.g. agriculture, manufacturing or financial services, vary by country?

• How do the rates of pay for men and women compare for different parts of the country, for different employment types and/or over time?

• How do average numbers of hours worked each year vary over time?

• Is there an association between average number of hours worked, average earnings and/or retirement age for different countries?

• How do the employment rates for different age groups, for men and women, for different levels of education and/or over time for countries around the world compare?

• Is there an association between rate of full-time employment and rate of part-time employment in different countries worldwide?

• How has trade union membership varied over time?

• How do worldwide retirement ages for men and women compare?

• How do the heights and/or weights of people in different types of employment compare?

Data sources

<http://www.jobsite.co.uk/search>   
The jobsite website provides details on jobs available with rates of pay. The data can be sorted by location and by sector of employment.

<http://www.theguardian.com/news/datablog/2011/nov/24/wages-britain-ashe-mapped>   
Data from The Guardian newspaper giving median salary for different locations in the United Kingdom.

<http://www.oecd.org/employment/emp/onlineoecdemploymentdatabase.htm>   
OECD website providing numerous data sets relating to employment.

<https://data.oecd.org/emp/hours-worked.htm>   
OECD data on hours worked and employment rates (grouped by a number of factors) for a number of countries.

<http://www.dailyinfographic.com/average-hours-of-work>   
This website provides data on number of hours worked per year and per week for a range of countries around the world.

<http://www.thisismoney.co.uk/money/article-2067258/Best-paid-jobs-2011-Tables-officialfigures-UK-salaries.html> <http://www.thisismoney.co.uk/money/article-1709280/Best-paid-jobs-A-guide-UK-salarieswages-2010.html>   
These webpages provide tables of average pay for workers in over 300 careers in the UK for the years 2010 and 2011.

<http://unstats.un.org/unsd/demographic/products/indwm/default.htm>   
A United Nations website which provides datasets relating to women and men in six fields of concern including work.

<http://www.texaswages.com/>   
This website provides information on mean, median, entry level and experienced employee rates of pay (hourly or annual) for the state of Texas. The data is divided by occupation and area of Texas.

[http://www.stltoday.com/news/local/stl-info/public-workers-public-salaries/collection\_ 7d571a8c-6e62-11df-98c9-00127992bc8b.html](http://www.stltoday.com/news/local/stl-info/public-workers-public-salaries/collection_%207d571a8c-6e62-11df-98c9-00127992bc8b.html)  
This website provides links to searchable databases of salaries of public employees for a variety of places in the USA.

<https://www.statcrunch.com/5.0/shareddata.php?keywords=SALARIES&startlimit=0>   
This website allows users to share datasets. The salaries category contains multiple datasets which may be of relevance.

<http://chartsbin.com/view/2468>   
This website provides an interactive map with retirement ages for men and women worldwide.

<http://www.payscale.com/top-tech-employers-compared-2012/job-satisfaction-survey-data>   
This website provides data on employee opinions at technology companies: employee job satisfaction rating, stress rating, job meaning rating and job flexibility rating. There is also data on median starting salary, mid-point salary and weeks of holiday.

<http://www.iisg.nl/hpw/data.php>   
This website has multiple datasets which include data on historical wages from around the world. Some of the datasets include information on rural and urban wages, some include skilled and unskilled wages, some are divided by employment type and some show the wages divided by gender. The data is generally from pre-1950.

<http://www.stat.ufl.edu/~winner/datasets.html>   
This website contains many datasets. Towards the bottom of the page is a dataset for NBA player heights and weights for the year 2013.

Teaching activity

Pay rates for males and females:

<http://www.nuffieldfoundation.org/fsmqs/level-3-data-analysis#pay%20rates>

Theme: Weight

Ideas for investigation

The following are some possible ideas to investigate about weight but there are many others that you might choose.

You may want to collect information about:

• the weights of things, e.g. fruit, vegetables, animals, mobile phones/laptop computers, etc

• people’s estimates of the weights of things

• the factors that may influence people’s ability to estimate weight, e.g. age, gender, time of day, distractions, units of weight, etc.

You can ask a variety of questions that can be investigated statistically. For example:

• How do the weights of things compare, e.g. the weights of vegetables/fruit/mobile phones/laptop computers/pets?

• How do the weights of things change with time, e.g. the growth of vegetables/pets?

• How do the weights and prices of things compare, e.g. laptop computers/mobile phones?

• How do the weights of animals compare, e.g. different genders of the same species?

• Are people better at estimating weights in imperial units than in metric units?

• Are people better at estimating smaller/larger weights?

• How do different factors affect people’s ability to estimate weights, e.g. time of day, gender, age?

• Is there an association between people’s ability to estimate weights using their left hand and using their right hand?

• Is there an association between the body weight and the brain weight of animals?

• Can the weights of things be modelled by a known distribution, e.g. the weights of vegetables/fruit/loaves of bread?

• Can estimates of weight be modelled by a known distribution?

Data sources

[www.egginfo.co.uk/egg-sizes](http://www.egginfo.co.uk/egg-sizes)   
This website gives the weight ranges for different sizes of eggs.

[www.metric-conversions.org/weight-conversion.htm](http://www.metric-conversions.org/weight-conversion.htm)   
This website enables you to convert between metric and imperial units of weight.

<http://thewebsiteofeverything.com/animals/mammals/adult-weight.html>   
This website gives information about the weights of a variety of mammals including weights at birth.

[www.rhs.org.uk/advice/grow-your-own/vegetables](http://www.rhs.org.uk/advice/grow-your-own/vegetables)   
This website gives information about the planting and growing of vegetables from seed.

[www.recombu.com/mobile/compare](http://www.recombu.com/mobile/compare)   
This website gives information about the weights and prices (and other technical information) for a variety of mobile phones.

[www.pcworld.com/category/laptop-computers/](http://www.pcworld.com/category/laptop-computers/)  
This website gives information about the weights and prices (and other technical information) for a variety of laptop computers.

<http://mste.illinois.edu/malcz/DATA/BIOLOGY/Animals.html>   
This website gives the mean body weight and the mean brain weight for a variety of animals.

Theme: Books

Ideas for investigation

The following are some possible ideas to investigate about books but there are many others that you might choose.

You may want to collect information about:

* the number of words in the sentences, paragraphs and chapters of books
* the time taken to read the page of a book
* the amount of punctuation used in books
* the types of books people like to read, e.g. Fantasy, Horror, Crime etc
* the times taken for people to find a word in a dictionary
* the authors and publishers of books
* the weights of books.

You can ask a variety of questions that can be investigated statistically. For example:

* Do boys and girls prefer to read different types of books?
* How do the books written for children differ from the books written for adults?
* How do the books written for girls differ from the books written for boys?
* How do the books written by different authors compare?
* How long does it take to find a word in a dictionary?
* How do the times taken for people to read a page of a book compare?
* How do the books written in the last 10 years compare to books written 100 years ago?
* Does an author’s writing style change over time?
* How do the numbers of pages and the weights of books compare?

Teaching activity

Investigate the nation’s reading habits:

<http://www.icse.xyz/stats4schools/large_datasets/reading/default.html>

Theme: Holidays

Ideas for investigation

The following are some possible ideas to investigate relating to holidays but there are many others that you might choose.

You may want to collect information about:

* people’s preferred types of holiday and holiday destination
* costs of holidays in the UK and abroad
* most popular tourist destinations
* the income generated by tourism
* the variation in the costs of holidays through the year
* costs of different types of holidays
* costs of holidays purchased from price comparison websites and from tour operator websites
* the ratings given to holidays by holidaymakers.

You can ask a variety of questions that can be investigated statistically. For example:

* What types of holidays do males and females prefer to go on?
* Which is the most popular holiday destination? Has this changed over time?
* Do people spend longer on holiday when they holiday in the UK than when they
* holiday abroad?
* How do the prices of holidays from price comparison websites compare to the prices from tour operator websites?
* How do the prices of holidays vary during the year?
* How do the prices of different types of holidays compare?
* How do the prices of holidays in the UK compare to prices of non-UK holidays?
* Is there a relationship between the distance to a holiday destination and the price of a holiday?
* Is there a relationship between the adult price and the child price for holidays?
* How do the prices of holidays aimed at specific age groups compare?
* How has the number of people going on international holidays changed over time?
* How has national income from tourism changed over time?
* What factors affect the price of a holiday?
* Is there a relationship between popularity of a destination, ratings given by holiday makers and/or cost of holidays?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

<http://www.gocompare.com/holidays/>  
The holiday section of the Go Compare price comparison website.

<http://mkt.unwto.org/en/publication/unwto-tourism-highlights-2013-edition>   
Part of the United Nations World Tourism Organisation (UNWTO) website. The Facts and Figures tab offers a choice of different options which lead to pages with free downloadable PDF files of data.

<http://www.bbc.co.uk/news/magazine-23433149>   
A BBC news article which contains data on the most popular tourist destinations and spending (total and per head).

<http://data.gov.uk/dataset/overseas_travel_and_tourism>   
The reports on this site provide quarterly figures relating to UK holidaymakers going abroad and non-UK holidaymakers visiting the UK.

<http://www.visitengland.org/insight-statistics/major-tourism-surveys/overnightvisitors/UKTS2010/Online_Data_Browser_2010.aspx>   
There is a link on this page which allows access to the data browser for Visit England statistics relating to UK trips by UK residents from 2006 onwards.

<http://wales.gov.uk/topics/tourism/researchl1/tourisminwales/?lang=en>   
The tourism section of the Welsh Government website. There are a variety of PDF reports containing figures under each of the sections within this area of the website.

Teaching activity

Tourism in London:

<http://www.icse.xyz/stats4schools/lesson_ideas/tourism_london/default.html>

Theme: Climate

Ideas for investigation

The following are some possible ideas to investigate relating to climate but there are many others that you might choose.

You may want to collect information about:

* the number of days of rain and amounts of rainfall
* the numbers of storms and the wind speeds of the storms
* white Christmases
* land and sea temperatures
* the size of the hole in the ozone
* amounts of snow coverage
* numbers of hours of sunshine
* climate change.

You can ask a variety of questions that can be investigated statistically. For example:

* Where and/or when would you go if you wanted to do something that is dependent on the weather, such as wanting wind for windsurfing, sun for a holiday, or rain and sun for growing vines?
* How does the amount of rainfall compare for different places in the UK?
* Is there an association/relationship between hours of sunshine, amounts of rainfall and/or temperatures in a region?
* Is the sea getting warmer?
* Are the polar ice caps getting smaller?
* Are there more storms now than in the past?
* What will the Earth’s climate be like in 10 years’ time?
* To what extent is the climate affected by volcanoes?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

<http://www.weather2travel.com/climate-guides/>   
Gives a variety of monthly climate data for many global holiday destinations.

<http://www.weatherbase.com>   
Gives a variety of monthly climate data for many places throughout the world.

<http://data.giss.nasa.gov/gistemp/>   
Database of regional land/sea temperature anomalies since 1880.

<http://www.metoffice.gov.uk/public/weather>  
Database of UK climate indicators such as maximum/minimum temperatures, sunshine hours, rainfall and frost days since 1910.

<http://sharaku.eorc.jaxa.jp/TYP_DB/index_e.shtml>   
Database of tropical cyclones since 1997.

<http://solarscience.msfc.nasa.gov/greenwch/spot_num.txt>   
Database of monthly average numbers of sunspots since 1749.

<http://ozonewatch.gsfc.nasa.gov/meteorology/annual_data.txt>   
Database of mean ozone hole area and minimum ozone column height since 1979.

<http://www.climate4you.com>   
Website that discusses various climate change indicators with links to useful databases

Teaching activity

Weather investigation with database covering many years: <https://nrich.maths.org/10470>

Theme: Journeys

Ideas for investigation

The following are some possible ideas to investigate relating to journeys but there are many others that you might choose.

You may want to collect information about:

* the method of travel to school/work of students/teachers/parents
* whether age/gender affects method of travel
* the distances travelled to school/work or any other activity
* the time it takes to get to school/work on different days of the week for a particular method of travel
* the time it takes to travel to school/work compared with how long it takes to travel home
* whether there is an association between time and distance travelled for a particular method of travel
* the interrelationships between cost, distance and journey time
* the effect weather has on methods of travel or journey times
* whether the times of a particular journey can be modelled by a normal distribution for a particular method of travel.

You can ask a variety of questions that can be investigated statistically. For example:

* How do students travel to school?
* Is there any difference between the way boys and girls travel to school?
* Do most teachers use a car to travel to school?
* Is there a difference between the mean time it takes to get to school in the morning and the mean time it takes to get home in the evening?
* Are boys’ average journey times similar to girls’ average journey times for the same method of travel?
* How do journey times to school and journey times going home compare?
* Do the different days of the week affect journey times to school/work or any other activity?
* Is there any correlation between the time it takes to get to school in the morning and the time it takes to get home in the evening?
* Is there an association between the distance to school and the journey time for a particular method of travel?
* What are the interrelationships between cost, distance and journey time?
* How do the costs for different methods of travel compare?
* How does the punctuality of different train companies compare over time?
* Can journey times be modelled by a normal distribution?

Data sources

The following are some data sources that might prove helpful. There are many others.

<http://www.censusatschool.com/resources.html>  
This site gives some data on students’ journeys to school (distance/time/mode of travel). Look for the ”Travelling to School” task.

[www.passengerfocus.org.uk](http://www.passengerfocus.org.uk)   
This site gives a lot of data about train journeys, including punctuality.

[www.nationalrail.co.uk](http://www.nationalrail.co.uk)   
[www.thetrainline.com](http://www.thetrainline.com)   
These and many other sites give information on costs and times of train journeys.

Teaching activity

Travel to school task with data:  
<http://www.icse.xyz/stats4schools/lesson_ideas/travel_school/default.html>

Theme: Puzzles and games

Ideas for investigation

The following are some possible ideas to investigate relating to puzzles and games but there are many others that you might choose.

You may want to collect information about:

* male and female scores when doing word, number and/or logic puzzles on paper
* male and female scores from computer-based games
* the time it takes for males and females to complete word, number, logic (e.g. Frogs, coin rearrangement) or jigsaw puzzles
* the number of steps taken to complete logic puzzles (e.g. Towers of Hanoi)
* the time taken to complete word, number, logic and/or jigsaw puzzles with and without distractions.

You can ask a variety of questions that can be investigated statistically. For example:

* Are males or females better at a particular type of puzzle or computer-based game?
* Do males improve their performance in puzzles or computer-based games more quickly than females?
* Is there any correlation between abilities in different types of puzzle or game?
* Is performance in simple computer games more strongly correlated with performance in logic puzzles than with performance in word puzzles?
* Does age affect the time taken to complete puzzles?
* How do the times taken to complete word, number and/or logic puzzles differ with and without distractions?
* How does performance on a particular puzzle or game type improve with practice?
* How does performance on a puzzle or game vary between type of game, between gender and between different age groups?
* What distribution does the time taken to complete a particular puzzle or game have? How do distractions affect the distribution?
* Does the difference in time taken between first and second attempts at a particular puzzle or game show a known distribution? How does this compare with other puzzles or games?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

<http://www.printable-puzzles.com/>   
Wide variety of different puzzles – word, number and logic.

<http://puzzles.about.com/od/toppicks/u/FreePuzzles.htm>   
Wide variety of different puzzles – word, number and logic.

<http://www.puzzles.com/index.htm>   
A variety of different types of puzzle which includes examples of paper-based, interactive computer puzzles and those utilising physical manipulation of props (e.g. coin rearrangement puzzles).

<http://www.jigzone.com/>   
A variety of different images which can be split into jigsaws starting at six pieces, including a timer for its interactive jigsaw puzzles.

Theme: Cars

Ideas for investigation

The following are some possible ideas to investigate relating to cars but there are many others that you might choose.

You may want to collect information about:

• the prices of used cars locally and/or nationally

• the mileages and ages of used cars

• the colours of cars

• the numbers of passengers in cars

• the numbers of registered cars on UK roads.

You can ask a variety of questions that can be investigated statistically. For example:

• What are the common makes of car on UK roads?

• What are the factors that affect the price of a used car?

• How do the ages of cars in your area compare to the ages of cars nationally?

• How do the prices of used cars in your area compare to the prices of used cars in other areas?

• How has the numbers of cars on UK roads changed over time?

• How are the mileages of used cars distributed?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

[www.autotrader.co.uk](http://www.autotrader.co.uk)   
Provides a variety of information about the purchase of used cars (large database).

[www.carshop.co.uk](http://www.carshop.co.uk)   
An alternative source to AutoTrader (smaller database).

<http://webarchive.nationalarchives.gov.uk/20160105185815/http://www.ons.gov.uk/ons/rel/abs/annual-business-survey/car-production/sty-car.html>  
Information on car registrations among other items.

<https://www.gov.uk/government/statistical-data-sets/vehicles-statistical-tables-index>  
Various data on vehicles.

Theme: Earthquakes

Ideas for investigation

The following are some possible ideas to investigate relating to earthquakes but there are many others that you might choose.

You may want to collect information about:

• the locations of earthquakes

• the numbers of casualties in earthquakes

• the times/dates of earthquakes

• the numbers of earthquakes

• the magnitudes of earthquakes (as measured on the Richter scale)

• the depths of earthquakes.

You can ask a variety of questions that can be investigated statistically. For example:

• Have the numbers of earthquakes changed with time?

• Have the magnitudes of earthquakes changed in the last 10 years?

• Do deeper earthquakes have greater magnitudes?

• Have the locations of earthquakes changed with time?

• How do the magnitudes and depths of earthquakes in different regions of the world compare?

• Can we predict when the next earthquake is likely to occur in a region?

• Are large primary earthquakes followed by large secondary earthquakes?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

[www.world-earthquakes.com](http://www.world-earthquakes.com)   
Provides a database to find earthquakes by geographic region, e.g. continent/hemisphere, and time, e.g. last hour/day/week/year.

<http://www.usgs.gov/>   
Provides a database to search for significant earthquakes globally and in the US from 1971 to the present.

<http://ncedc.org/maps/>   
Provides a database to search for earthquakes by date/time, magnitude and depth.

Theme: Time

Ideas for investigation

The following are some possible ideas to investigate relating to time but there are many others that you might choose.

You may want to collect information about:

* the time taken to complete an activity (e.g. a number puzzle, a crossword puzzle, a 100 m run/swim, or a run up steps)
* the time taken by an object (e.g. a ball) to fall from different heights
* how accurate people are when asked to estimate 30 seconds (or one minute or five minutes)
* how age/gender affects the time taken to complete an activity.
* how age/gender affects the ability to estimate lengths of time accurately
* the factors that affect the time taken to complete an activity (e.g. time of day, noise).

You can ask a variety of questions that can be investigated statistically. For example:

• What is the average time taken for one or more people to complete an activity?

• Are boys quicker than girls at completing an activity?

• Does age affect how long it takes to complete an activity?

• Are boys better than girls at estimating times?

• Are people as good at estimating 30 seconds as they are at estimating five minutes?

• How does age/gender affect the time taken to complete an activity or estimate a length of time?

• Is there an association between the times taken to complete two different activities (e.g. a number puzzle and a crossword)?

• Is there an association between the errors in estimating a short length of time and the errors in estimating a long length of time?

• How do different factors (e.g. noise, time of day) affect the time taken to complete an activity?

• Can the times taken to complete an activity be modelled by a normal distribution?

• If an object (e.g. paper) is dropped from different heights, is there an association between the height and the time taken to reach the ground?

• What is the equation of the curve/line of best fit for the times taken for an object to fall from different heights?

Data sources

The following is a data source found at the time of publication that might prove helpful. There are many others.

[www.puzzlechoice.com](http://www.puzzlechoice.com)   
This site has crossword, word and number puzzles of various types.

There are many websites for puzzles – just put ‘number puzzles’ or ‘crossword puzzles’ into a search engine.

Teaching activity

Gender differences in household chores:  
<http://www.icse.xyz/stats4schools/lesson_ideas/household_chores/default.html>

Theme: Angles and lines

Ideas for investigation

The following are some possible ideas to investigate relating to angles and lines but there are many others that you might choose.

You may want to collect information about:

• estimates of the sizes of angles of different types (acute/reflex/obtuse) by different age groups and genders

• estimates of the lengths of lines/curves for different age groups and genders

• estimates of the lengths of lines drawn as spirals

• using your left eye or right eye when estimating the sizes of angles/lengths of lines

• estimates of the size of an angle/length of a line when there is a distraction

(e.g. loud music)

• estimates of the size of an angle/length of a line from different distances.

You can ask a variety of questions that can be investigated statistically. For example:

• Are males better than females at estimating the lengths of lines?

• Are females better than males at estimating the sizes of angles?

• Are people better at estimating the sizes of acute, obtuse or reflex angles?

• Are people better at estimating the lengths of straight lines or the lengths of curves?

• Does age affect your ability to estimate the sizes of angles/lengths of lines?

• Is there a correlation between the ability to estimate the sizes of angles and the ability to estimate lengths of lines?

• Can the errors in estimating the sizes of angles/lengths of lines be modelled by a normal distribution?

• Are people better at estimating the sizes of angles/lengths of lines using their right eye or their left eye?

• Do distractions, like loud music, affect the accuracy of estimates?

• Are people better at estimating the lengths of straight lines or at drawing lines of a given length without measuring instruments?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

<https://www.stem.org.uk/resources/collection/3834/experimentsatschool>  
Free registration/Login required.

<https://nrich.maths.org/1235>  
Estimating angles task.

Theme: Films

Ideas for investigation

The following are some possible ideas to investigate relating to films but there are many others that you might choose.

You may want to collect information about:

• the lengths of films

• the year in which a film is made

• the box office takings of films

• the categories of films (e.g. horror)

• the directors/producers of films

• the actors/actresses in films

• the country of production/studio of films and their language (e.g. Japanese)

• whether the film is black and white or colour.

You can ask a variety of questions that can be investigated statistically. For example:

• What categories of films do males and females prefer to watch?

• Are action films longer than comedy films?

• Have film-watching tastes changed over time?

• Is there a relationship between the production costs of films and the box office takings of films?

• How do the lengths of films of directors/producers compare in different periods of their careers?

• How do the box office takings of films in different categories/genres and/or countries compare?

• How do the lengths of films made in different countries compare?

• Is there a relationship between the box office takings of films and when they are released?

• Is there a relationship between the production costs, the box office takings, and the lengths of films?

• Is there a significant difference between the lengths of films made in the last 5 years and those made in earlier decades?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

[www.IMDb.com](file:///C:\Users\helemars\Downloads\www.IMDb.com)  
The Internet Movie Database (IMDb) provides an extensive data base for a considerable number of films and is considered by some to be the primary source for film information.

<http://en.wikipedia.org/wiki/lists_of_films>   
Wikipedia gives a considerable database for films that is relatively easy to access.

[www.the-numbers.com/market/](http://www.the-numbers.com/market/)   
The Numbers® provides year on year market information broken down into year, genre and distributor, and provides inflation-adjusted box office takings from 1995 to the present.

Theme: Food

Ideas for investigation

The following are some possible ideas to investigate relating to food but there are many others that you might choose.

You may want to collect information about:

• the food preferences of males and females

• whether food preferences are affected by age

• the choices made in the school canteen

• the total/average amount of money spent per day/week in the school canteen by each canteen user

• the amount of money spent per day/week on food by a student

• the weights/lengths of some type of fruit/vegetable, e.g. potatoes, apples, tomatoes

• how food prices have changed over time and whether there are seasonal variations in prices

• the nutritional contents of a type of food, e.g. cereal, chocolate

• the rise in food prices compared with inflation (RPI/CPI)

You can ask a variety of questions that can be investigated statistically. For example:

• Do boys choose different types of food compared with girls?

• Does age affect choice of meals?

• The probability of choosing a starter and a sweet from the canteen and how this can be used to predict what should be offered.

• Do boys eat different numbers of portions of fruit, vegetables, sweets or crisps per day than girls?

• What are the favourite food types of students, e.g. dairy, protein and fruit/vegetables?

• In any day/week do boys spend more money in the canteen than girls?

• For one particular type of fruit/vegetable, e.g. potato, apple, tomato, is there a relationship between the length/weight?

• Are fruits/vegetables of different varieties, e.g. baking/boiling potatoes or green/red apples, similar in weight?

• How do the nutritional contents of different foods, e.g. cereals, biscuits, compare?

• Are the colours/varieties in assorted packets of sweets uniformly distributed?

• Are the measured weights of individual items of a particular fruit/vegetable normally distributed?

• Are food price indices the same, e.g. for potatoes and rice?

• How have food prices/indices changed over time and do they show seasonal variation?

• Does the price of a particular food item vary with the seasons?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

[www.censusatschool.org.uk](http://www.censusatschool.org.uk)   
On this site, there is data on fast foods and portions of types of food eaten for different UK years.

[www.defra.gov.uk/](http://www.defra.gov.uk/)   
This site has several areas concerning food.

<https://www.gov.uk/government/collections/family-food-statistics>  
This section of Defra gives information on household purchase and consumption of food.

<https://www.gov.uk/government/collections/food-statistics-pocketbook>  
This has various data lists on food – look at the ‘Food Statistics Pocket Books’.

Other sources:

• Supermarkets/shops

• Internet shopping sites

Theme: Paper and pencil

Ideas for investigation

The following are some possible ideas to investigate relating to paper and pencil but there are many others that you might choose.

You may want to collect information about:

* male and female scores when doing spelling, number tests and/or other paper-based puzzles
* the time it takes for males and females to complete a paper-based puzzle
* the time it takes to make a paper model (e.g. paper aeroplane/boat)
* the distance model aeroplanes fly
* estimating the areas of regular shapes drawn on paper
* estimating the lengths of lines/curves drawn on paper.

You can ask a variety of questions that can be investigated statistically. For example:

* Are males better/quicker than females at spelling/number problems?
* Is there any correlation between ability to do spelling and the ability to use numbers?
* How do scores for males differ from those of females?
* Does age affect your ability at number problems/spelling?
* Are females better at estimating than males?
* How do the times to do a paper puzzle/make a paper model differ between males and females?
* Is there any correlation between the time it takes to make a paper aeroplane and the distance it flies?
* How do the estimates for the areas of the pieces of paper vary between gender and between different age groups?
* Can the time taken to complete a puzzle be modelled by a normal distribution?
* Can the time taken to make a paper model or the distance it flies be modelled by a normal distribution?
* Can the errors in estimating areas/lengths be modelled by a normal distribution?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

[www.printable-puzzles.com/](http://www.printable-puzzles.com/)   
Crosswords, word searches, logic.

[www.bestpaperairplanes.com/](http://www.bestpaperairplanes.com/)   
Aeroplane templates.

[www.funpaperairplanes.com/](http://www.funpaperairplanes.com/)   
Aeroplane templates.

<https://nrich.maths.org/10629>   
Estimating time activity to generate data.

<https://nrich.maths.org/10999>  
Estimating 30 seconds with data set.

Theme: Reaction times

Ideas for investigation

A quick reaction time is very useful. In athletics, the sprinter with the quickest reaction time is the first out of the blocks. In football, a goalkeeper must react quickly to prevent a goal being scored. On the roads, reacting to a car horn can save your life!

The following are some possible ideas to investigate relating to reaction times but there are many others that you might choose.

You may want to collect information about:

* age
* gender
* time of day
* left- or right-handedness
* the time taken to react to a visual, touch or sound stimulus.

You can ask a variety of questions that can be investigated statistically. For example:

* Do females have quicker reaction times than males?
* Do younger people have quicker reaction times than older people?
* Does the time of day affect reaction time?
* Can reaction time be improved with practice?
* Do people react quicker to touch or sound or visual stimuli?
* Do left-handed people have quicker reactions than right-handed people?
* How do distractions, e.g. loud music, affect reaction times?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

<http://mypages.iit.edu/~smile/ph92rc.html>   
This website gives a description of how an experiment may be set up. It also gives some secondary data.

<http://www.humanbenchmark.com/tests/reactiontime/index.php>   
This site gives a simple on-screen test with the top 100 results recorded.

<http://www.sciencejoywagon.com/explrsci/media/reflex.htm>   
This site provides a simple test for sight and sound.

<http://getyourwebsitehere.com/jswb/rttest01.html>   
Online reaction timer, traffic light test. Top 5 scores are available for comparison.

<http://faculty.washington.edu/chudler/java/dottime.html>   
Hitting dots as they appear.

Theme: Road transport

Ideas for investigation

The following are some possible ideas to investigate relating to road transport but there are many others that you might choose.

You may want to collect information about:

* the colours and makes of various forms of road transport
* the methods of travelling to school
* second hand car/bike prices
* fuel prices over time
* sales of new car/lorry/bike/buses
* levels of greenhouse gas (CO2) produced by different types of road transport
* CO2 emissions and engine size
* seasonal sales.

You can ask a variety of questions that can be investigated statistically. For example:

• What is the most popular colour/make of car in the school car park?

• How do students travel to school?

• How do the ways parents of pupils in my school travel to work compare with the

national statistics?

• How do the numbers of sales of certain types/makes of car compare with sales of

other types/makes over time?

• How do second hand prices change over time?

• How does the fuel consumption change with engine size?

• How do the prices of diesel and petrol compare over time?

• Is there a relationship between CO2 emissions and engine size?

• Are CO2 emissions for cars of similar engine size and fuel type normally distributed?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

Secondary Data Sources:

• Local Newspapers

• Magazines: many magazine publications give data suitable for sampling, e.g. What Car: This gives prices, ages, engine sizes, type, CO2 emissions etc.  
What Bike:This gives many of the same pieces of information as What Car.

• Web Sites: There are many websites that give transport information but many give data that is in an unsuitable form and difficult to follow. The following are sites that are easy to use.

[www.theaa.com/motoring\_advice/running\_costs/archive.html](http://www.theaa.com/motoring_advice/running_costs/archive.html)   
This gives archive information on running costs of cars.

[www.scotland.gov.uk/Publications/2008/12/22091243/2](http://www.scotland.gov.uk/Publications/2008/12/22091243/2)   
This is an easy to use site. It gives lists of transport-related statistical tables and suitable ones can be selected for use. For example, table 6.13 is a table giving CO2 emissions for various methods of transport over several years.

[www.dft.gov.uk/pgr/statistics/datatablespublications/](http://www.dft.gov.uk/pgr/statistics/datatablespublications/)   
This gives a list of tables for various methods of transport. The required ones are easily accessed.

[www.whatcar.com](http://www.whatcar.com)   
This site gives the information found in the magazine, plus other interesting facts. The magazine is often easier to use for lower ability students.

[www.statistics.gov.uk/cci/nscl.asp?ID=7627](http://www.statistics.gov.uk/cci/nscl.asp?ID=7627)   
This site gives percentages and numbers of people using different methods of transport to go to work.

[www.am-online.com/NewCarSalesFigures/](http://www.am-online.com/NewCarSalesFigures/)   
This site has an archive of new car sales figures.

[www.gov.uk/government/statistics/announcements?utf8=%E2%9C%93&keywords=transport&topics%5B%5D=transport](http://www.gov.uk/government/statistics/announcements?utf8=%E2%9C%93&keywords=transport&topics%5B%5D=transport)  
This is a government site giving a variety of data on transport.

Theme: Trees

Ideas for investigation

The following are some possible ideas to investigate relating to trees but there are many others that you might choose.

You may want to collect information about:

* the different types of tree in a wood, orchard or park
* the lengths and widths of leaves on trees
* the cost of buying trees.

There is other information that you may choose to collect.

You can ask a variety of questions that can be investigated statistically. For example:

* What types of trees are there in my area?
* What are the average lengths of leaves on different trees?
* What is the distribution of the lengths of leaves on trees?
* What is the proportion of different trees in a park in my area?
* How are trees affected by their location?
* How does the weather affect trees?

Theme: Houses

Ideas for investigation

The following are some possible ideas to investigate about houses but there are many others that you might choose.

You may want to collect information about:

• different types of house

• the cost of different types of house

• the change in house prices over time

• the cost of renting houses

• the cost of buying houses

• the types of housing in urban and rural areas.

You can ask a variety of questions that can be investigated statistically. For example:

• What are the types of house in my area?

• How has the price of houses changed over time?

• What are the factors that influence the price of a house?

• Is it better to buy or rent a house?

• How does location affect the price of a house?

• How have the types of housing in my area changed over time?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

[www.rightmove.co.uk](http://www.rightmove.co.uk)   
Provides information about property prices and rental fees nationally.

[www.findaproperty.co.uk](http://www.findaproperty.co.uk)   
Provides information about property prices locally and nationally.

[www.landregistry.gov.uk](http://www.landregistry.gov.uk)   
Provides information about how house prices have changed nationally over time, and other useful information.

[www.home.co.uk](http://www.home.co.uk)

provides information about how house prices have changed over time by postcode, and other useful information.

Theme: Memory

Ideas for investigation

Can you remember names? Places? Colours? Numbers? Words? Song titles?

At the World Memory Championships last year, Johannes Mallow from Germany memorised 110 historic dates in 15 minutes; this is more than the average student remembers in a year of studying history. In the Random Words discipline, Boris Konrad from Germany memorised 255 words in 15 minutes.

The following are some possible ideas to investigate relating to memory but there are many others that you might choose.

You may want to collect information about:

* age
* gender
* the time it takes to remember something
* number of items remembered
* type of items remembered.

You can ask a variety of questions that can be investigated statistically. For example:

* Do girls have better memories than boys?
* Do older people have better memories?
* How many numbers can be remembered in 5 minutes?
* Are people better at remembering numbers, words or pictures?
* Can we predict how many numbers can be remembered if we know how many objects can be remembered?

Data sources

The following are some data sources found at the time of publication that might prove helpful. There are many others.

<http://world-memory-statistics.com/c_world.php>  
An archive of the results of the World Memory Championships.

<http://www.lumosity.com/brain_attribute_categories/memory>   
This site offers brain training games.

Theme: Estimation

Ideas for investigation

A valuable life skill is the ability to estimate. For example, you may be asked to estimate how long a task is going to take or how much it is going to cost.

An estimate is an approximate idea of length, weight, time etc that is given without actually taking measurements, but is based upon your previous experience of such things.

All sorts of factors may affect our ability to estimate.

For example:

• age of the person doing the estimation

• size of the object being estimated

• orientation of the object being estimated

• colour of the object being estimated.

Theme: Human body

Ideas for investigation

When you look round a room full of people you will see that the human body has many shapes and sizes. You can investigate the human body statistically in many different ways.

You may want to collect information about:

* eye/hair colours
* left- or right-handedness
* lengths of arms/legs
* height
* weight
* body mass index.
* There are many questions that can be investigated statistically. For example:
* Do you expect there to be more people with brown eyes than any other colour?
* Do you expect people with long arms to be tall?
* Do you expect tall people to weigh more than short people?
* Do you expect people to cluster around an average height or are heights distributed evenly?

Teaching activities

1. Stature – This activity looks at simulated stature data for adults in different countries. This is an opportunity to construct histograms and consider the Normal distribution model.  
<http://www.nuffieldfoundation.org/fsmqs/level-3-data-analysis#Stature>

**2. Anthropometric data – Students investigate relationships between anthropometric variables and write a report on their findings.** <http://www.nuffieldfoundation.org/fsmqs/level-3-data-analysis>