Statistical investigations

Extended investigation **Part 1: Preparation activities**
Comments for teachers

General

In some of the problems provided the investigation will be based on increasing age. It is suggested that the numbers given are more finely graded than just in years; so in months or to the nearest quarter.

Activity 1

Students are to be encouraged to thoroughly analyse the statement provided and to engage in discussion to confirm that they have a common understanding of the problem posed. These questions are designed to guide this discussion and to challenge the students' understandings of the two variables indicated (time for paid work and time for study).

Activity 2

The final question posed for the study should be ratified by the teacher before the collection of data proceeds. A suitable question would be

Is there a relationship between the number of hours that students study on the weekend and the number of hours that students work on the weekend?

Activity 3

Students should be encouraged to collect data from only one year group and to consider the influence of other factors if more than one year group is selected. Collection of data for time needs to be consistent, e.g. to the nearest half hour. Whatever decisions are made in the defining of study and paid work, students need to find ways to communicate their meanings to those that are surveyed.

In designing the data collection process, students may discover that the question for the investigation will not be suitable and needs to be refined. All of the problems posed provide opportunities for the collection and analysis of **bivariate** data and students should be guided in that direction so that their preparation for the in-class component of this assessment is relevant.

Activity 5

In this situation the explanatory variable would be time spent doing paid work and the response variable, time spent on study.

Activity 6

Students should look for a linear relationship between the two variables. As a result of doing this investigation, students may have further questions e.g. what other factors might influence the amount of time students spend studying on the weekend? Does time spent using social media influence the amount of time students spent studying on the weekend?

Statistical investigations

Extended investigation **Part 2: In-class validation**
Solutions and marking key

Part A: Question 1

	Solution	Mathematical behaviours	Marks
(a)	It assumes pensioners are all old and students are all young	• identifies problem with terminology	1
(b)	Television shows have different lengths and the number of shows may not reflect the time spent.	• identifies problem with definition of time watched.	1
(c)	Adult – needs to be defined as being 18 onwards. Watching television – does this include having it on while you work? Does it include watching a TV show on your computer or Ipad?	• identifies word/phrase requiring clarification.	1

Part A: Question 2

	Solution	Mathematical behaviours	Marks
(a)	Needs to have a defined period eg over the last week	• identifies problem with terminology	1
(b)	*time spent watching television* Time spent watching television is likely to depend on age and not age depend on time spent watching television	• identifies response variable • explains role of response variable	1 1
(c)	To the nearest year	• identifies appropriate rounding	1
(d)	Not easy to remember Not easy to calculate	• identifies reason for difficulty • identifies reason for difficulty	1 1
(e)	Age It is the explanatory / independent variable	• identifies variable • justifies choice	1 1

SCENARIO 2 (3)

Part B: Question 1

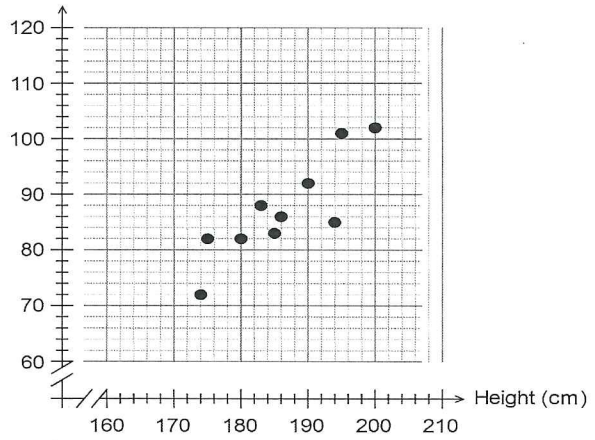
Solution
Suggests at least 50 students
Selects students at random
Describes the random process
Nominates suitable time and date to fit local situation
Nominates suitable place for the survey to occur
Students complete written survey / data recording sheets kept secure and backed up

Mathematical behaviours	Marks
• selects suitable number of respondents	1
• identifies need for random selection	1
• describes random collection of data	1
• nominates appropriate time for survey	1
• nominates appropriate place for survey	1
• describes data collection process	1

SCENARIO 3 (5) → SEE AMENDMENTS (sep. sheet)

Part B: Question 2

Solution
Footballers' heights and weights

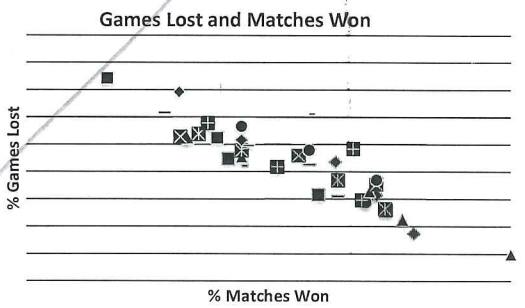


Mathematical behaviours	Marks
• selects correct type of graph (scatter graph)	1
• places variables on correct axes	1
• labels axes with variable names	1
• labels variables on axes with correct units	1
• selects suitable scales for both axes	1
• plots 4 points correctly	1
• plots 3 points correctly	1
• plots 3 points correctly	1

SCENARIO 3 → SEE AMENDMENTS (sep. sheet)

Part C

Question 1 (6)

	Solution	Mathematical behaviours	Marks
(a)	(i) points are close together (ii) points head downwards (iii) points nearly in a line	• identifies strength from graphs • identifies negative slope • identifies linear pattern	1 1 1
(b)	-0.88 Needs to be negative Is not perfect (-1)	• selects correct correlation • identifies negative value needed • identifies relationship not perfect	1 1 1
(c)	No r would need to be 1 (it is not)	• identifies value incorrect • justifies decisions	1 1
(d)	(i) %games lost = -0.3 x 100 + 65 = 35 (iii) Not reliable – extrapolated beyond the data provided. The trend may not continue.	• substitutes into given equation • determines % games lost • plots point on scatter graph • identifies lack of reliability • explains lack of reliability	1 1 1 1 1
(d) (ii)			

Part D SCENARIO 4 Q7

Solution
Note: These will vary

Mathematical behaviours	Marks
• identifies all components of the statistical investigation process	1
• describes the data to be collected – age + # of accidents	1
• describes how the data will be collected – any reasonable answer	1
• describes the graph to display the data collected – scatter	1
• describes the process to identify the trend – graphic analysis	1
• identifies the need to determine the linear relationship and the correlation coefficient	2

Q5 Scatter plot ✓
All dots correct ✓✓ (6 marks)
Appropriate scale ✓
Height explanatory ✓
Ht & Wt correct ✓

Q6 $r = 0.8601$ ✓ ✓ v. strong positive linear ✓

$r^2 = 0.7397857$ ✓

∴ 74% of assoc due to Height ✓
(variation in wt is explained by Ht)

$Wt = 0.907\,Ht - 81.55$ ✓✓

Residuals

1	−9.373
2	2.1857
3	3.6019
4	4.8565
5	1.254
6	5.7199
7	−3.212
8	0.3223
9	−4.237
10	−1.119

✓

Random pattern in residuals ✓
∴ suited to linear fitting

HOWEVER Very small sample size means association should be treated with caution ✓