Year 12 Mathematics Essentials
2016

Mark

55

Test 1: Data Collection
Chapters 2, 9 and 12

Name: Solutions **Time Allowed:**

Instructions:

- *Answer all questions in the space provided.*
- *Show all working where appropriate to be awarded full marks.*
- *Round all answers to two decimal places unless stated otherwise in the question.*

Question One [4 marks; 1 each]

Classify each of the following data collections according to whether they are a sample or census.

a) Favourite TV programs of secondary school students by surveying students in Year 8.

sample ✓

b) Number of hours computer games played by your class at home each week by surveying each student in your class.

census ✓

c) Survey of first 40 customers at the opening of a new store to gauge customer satisfaction.

sample ✓

d) The most popular colours for cars sold in the last month in Western Australia, taken from the registration board database of all cars registered in the last month.

census ✓

Question Two [3 marks; 1 each]

Give an example of how each of these sampling methods can be conducted.

Random sampling: name out of hat ✓

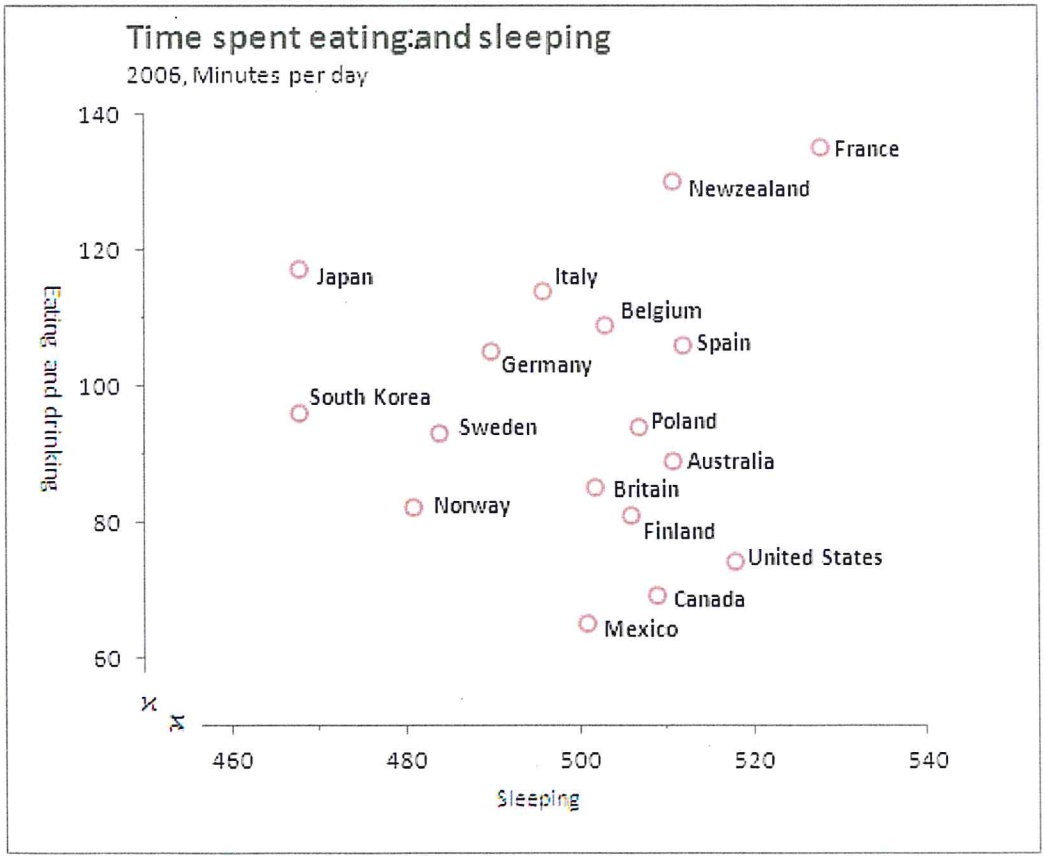
Systematic sampling: every third person entering the building ✓

Self-selecting sampling: taking volunteer to conduct survey ✓

anything relevant

Question Three [6 marks; 1 each]

Answer the following questions about the scatterplot below.



a) Which country spends the least amount of time eating and drinking?

Mexico ✓

b) Which country spends the most time sleeping?

France ✓

c) Which country comes fourth in spending the most time eating and drinking?

Italy ✓

d) After Canada which country is next when it comes to eating and drinking?

United States ✓

e) Which country spends less time than Sweden eating and drinking and less time sleeping than Mexico?

Norway ✓

f) Which country spends less times sleeping and more time eating and drinking than Italy?

Japan ✓

Question Four [4 marks; 1 each]

Give one advantage and disadvantage of conducting a census and sample.

	Census	Sample
Advantage	accurate data ✓	quicker, easier to process ✓
Disadvantage	hard to organise, takes longer ✓	might not be accurate ✓

Question Five [3 marks]

A survey was conducted to find the satisfaction rating of the service provided at a local café. The day of the survey coincided with a hail storm which affected the area. The survey conducted was self-selecting and five people participated. All of the participants were 100% satisfied with the service provided. Explain why this data is misleading and unreliable, give two reasons. Provide an example of how data could be collected in a more reliable manner.

- hail storm impacted on sample size making unreliable ✓
- 100% not reflective of everyone on a normal day, therefor misleading ✓
- any reasonable collection of reliable data ✓

Question Six [1 mark]

A new product is being launched which claims to hide vegetables within its meal while maintaining its appeal to young children. A survey is conducted at these various locations, an accounting office in the city, a retirement village, a primary school and a fine dining restaurant. Explain how to best use the data.

only use primary school, the only appropriate audience ✓

Question Seven [4 marks; 1 mark each]

Circle the independent variable in the following pairs of variables.

a) positive feedback — self confidence

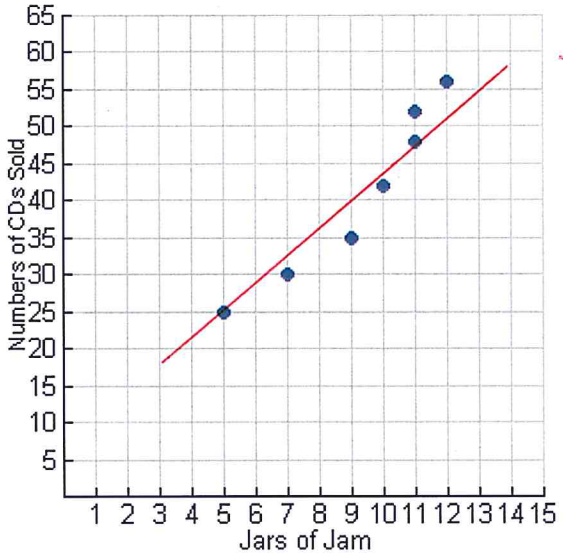
b) headaches suffered — aspirin taken

c) muscle mass — time spent weight-training

d) calcium consumption — bone density

Question Eight [7 marks; 1, 1, 3 and 2 marks]

The number of jam jars sold at a supermarket in a northern suburb of Perth was recorded as well as the number of CDs sold at a record shop in a southern suburb of Perth over a number of days. The scatterplot for the data was produced below.



a) Over how many days was the data collected? 7 days ✓

b) Draw in a line of best fit for the scatterplot.

c) Describe the association between the variables in terms direction, form and strength.

strong ✓ positive ✓ linear ✓

d) What conclusion can you make about the number of jars of jam sold and number of CDs sold at these stores?

no conclusion ✓

not causal ✓

Question Nine [4 marks]

A stratified sample of 75 students is to be selected from the student population of a school. If the school consists of 256 year sevens, 222 year eights, 175 year nines, 198 year tens, 156 year elevens and 98 year twelves, how many of each year group should be in the sample for it to reflect the proportion in each year group?

total : 1105 ✓

year 7 $\frac{256}{1105} \times 75 = 17$ ✓ ½

year 8 $\frac{222}{1105} \times 75 = 15$ ✓ ½

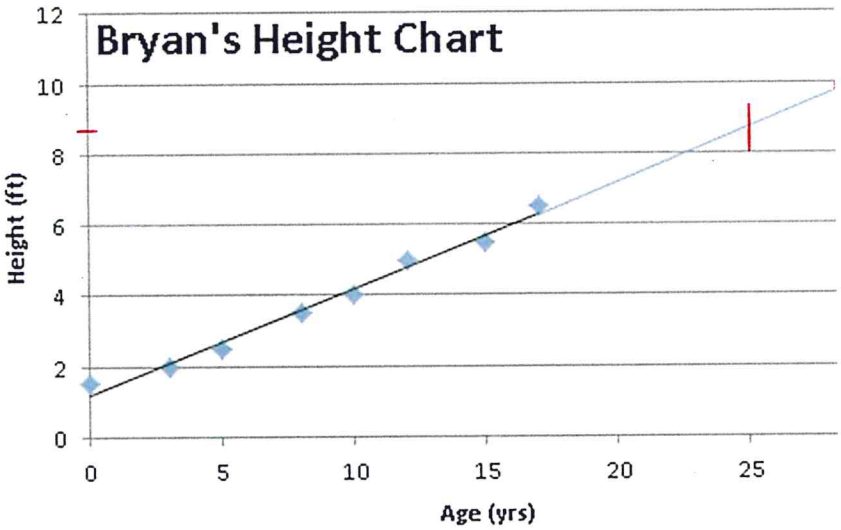
year 9 $\frac{175}{1105} \times 75 = 12$ ✓ ½

year 10 $\frac{198}{1105} \times 75 = 13$ ✓ ½

year 11 $\frac{156}{1105} \times 75 = 11$ ✓ ½

year 12 $\frac{98}{1105} \times 75 = 7$ ✓ ½

Question Ten [5 marks; 1, 1 and 3 marks]



a) For the scatterplot above, use the line of best fit to predict Bryan's height when he is 25 years old.

≈ 8 3/4 feet ✓

b) To convert feet (ft) to centimetres (cm) we multiply the height in feet by 30.48. What would Bryan's height be in centimetres?

266.7 cm ✓

c) Make a comment on the reasonableness of this prediction, giving two reasons why this would be a reasonable/unreasonable prediction.

unreasonable ✓ no one is that tall ✓

extrapolation ✓

not linear trend ✓

any reasonable 2 marks

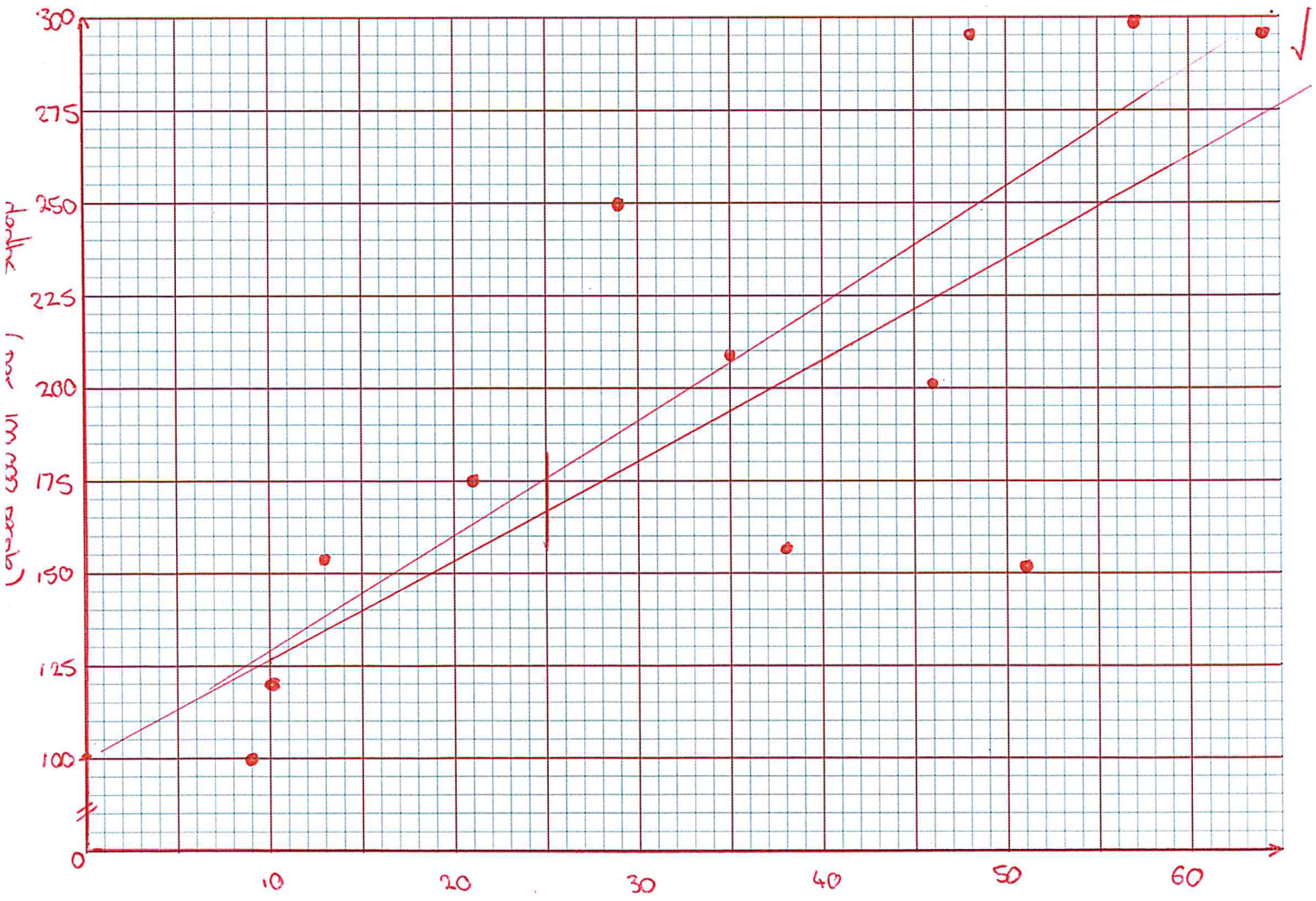
Question Eleven [14 marks; 4, 2, 2, 1, 1 and 4 marks]

A statistician was asked to investigate whether countries with higher levels of consumption of a particular foodstuff tended to have a higher death rate due to heart disease. The following information was obtained for 12 countries.

	CONSUMPTION OF FOODSTUFF kg per person per year	DEATHS DUES TO HEART DISEASE Per year per 100 000 of population
Country A	57	299
Country B	13	154
Country C	51	152
Country D	48	295
Country E	35	208
Country F	9	100
Country G	21	175
Country H	10	120
Country I	46	202
Country J	64	290
Country K	38	156
Country L	29	250

accurate ✓✓ label axis x } ✓ label axis y } appropriate scale x } ✓ " " y }

a) Draw a scatterplot on the axis below.



b) Does there appear to be a relationship between the two variables? If so, describe it.

as the consumption of particular food increases ✓
so do the deaths by heart disease ✓

c) If there is an association, is it a causal relationship or not? If not, suggest other factors that we might need to consider.

not causal, ✓ many factors cause heart disease, eg
lifestyle, exercise, genetics ✓

d) Draw in a line of best fit on your scatterplot.

within region (reasonable line)

e) Give an estimate for what you think the correlation coefficient might be for this data.

≈ 0.75 ✓

f) Use your line of best fit to show how how this can be used to predict the deaths due to heart disease given the consumption of food stuffs is 25 kg per person per year. What is your prediction, make a comment on how reliable this prediction is and why you think so.

≈ 172 ✓

show on graph ✓

reasonable ✓

interpolation ✓