# STAY AHEAD OF THE CURVE





## TEACHING STATISTICS USING LARGE DATA SETS



y= 4 -1

3

### LARGE DATA SET REQUIREMENT

AS and A level mathematics specifications must require students to:

- become familiar with one or more specific large data set(s) in advance of the final assessment (these data must be real and sufficiently rich to enable the concepts and skills of data presentation and interpretation in the specification to be explored)
- use technology such as spreadsheets or specialist statistical packages to explore the data set(s)
- interpret real data presented in summary or graphical form
- use data to investigate questions arising in real contexts



#### **IMPLICATIONS**

- Students should work with the data
- Students will not have access to the large data set in the exam
- Summary data or samples of data may be provided within the exam paper
- The data will be available in a spreadsheet this makes it easy to copy into other software
- Some of the questions on the statistics content will be set in the context of the large data set, students who have spent time exploring the data will have an advantage



### **OCR SPECIFICATION A DATA SET**

- LDS will remain for the life of the specification, unless the review process identifies a necessary change.
- Any change to the data set will be made before the beginning of any given two year course and centres will be notified a year in advance.
- The Large Data Set (LDS) consists of four sets of data: two each from the censuses of 2001 and 2011
  - method of travel to work
  - age structure of the population
- The data are aggregated by local authority.



### **OCR (MEI) SPECIFICATION B DATA SETS**

- Three data sets will be issued in June 2017 each will be of similar size to the specimen data set.
- Let's call the first three data sets LDS1, LDS2, LDS3.
  - LDS1 is for AS 2018, A level 2018 and A level 2019
  - LDS2 is for AS 2019 and A level 2020
  - LDS3 is for AS 2020 and A level 2021.
- The intention is that teachers can use all three data sets in their teaching perhaps concentrating more on the exam data set for a class as examinations get nearer.
- Each data set will be labelled with exam years.



#### AS SPECIFICATION A SAMS

	Method of travel to work Individual motorised transport		Shared motorised transport	Public transport	No motorised transport	
	Carbon emissions category	High	Medium	Low	None	Total
Local Authority A	Number of workers	174374	42 112	<mark>61 4</mark> 83	76024	353 993
	Percentage of workers	49.3	11.9	17.4	21.5	100
Local Authority B	Number of workers	39 433	9944	4614	16232	70223
	Percentage of workers	56.2	14.2	6.6	23.1	100

 (i) Clara calculated the values for the column headed "shared motorised transport" by doubling the value in the "passenger in a car or van" column of the original data set. Explain what assumption she has made and what other adjustment would need to be made to the data to take account of this.
[2]

#### AS SPECIFICATION A SAMS ANSWER

(	Questi	on	Answer	Marks	AO	Guidance
13	(i)		She has assumed that any car has exactly two people in it: one passenger and the driver. Subtract the value in "Passenger in a car or van" from the value in "Driving a car or van" to get the	B1 B1	2.2b 2.2a	Must refer to "Driving a car or van", or equivalent
			number of people driving alone.	[2]		

#### AS SPECIFICATION B SAMS

(ii) The birth rate for all the countries in Australasia are shown below.

Country	Birth rate per 1000			
Australia	12.19			
New Zealand	13.4			
Papua New Guinea	24.89			

(A) Explain why the calculation below is not a correct method for finding the birth rate per 1000 for Australasia as a whole.

$$\frac{12.19 + 13.4 + 24.89}{3} \approx 16.83$$

(B) Without doing any calculations, explain whether the birth rate per 1000 for Australasia as a whole is higher or lower than 16.83.
[1]



[1]

#### AS SPECIFICATION B SAMS ANSWER

9	(ii)	(A)	E.g. The calculation doesn't use the populations as weights	E1	2.3
			L.g. Does not take the populations into account	[1]	
		(B)	E.g. Lower because Australia has the highest population but the lowest birth rate oe	E1	2.2a
			E.g. answer given is too high as too much weight is given to Papua New Guinea		
				[1]	



#### **IDEAS FOR TEACHING WITH DATA SETS 1**

Found on the internet:

Eastbourne has the oldest population in Britain with a median age of 71.5

Could it be true?



#### **IDEAS FOR TEACHING WITH DATA SETS 2**

Which local authority in England and Wales has the highest percentage of the population employed?

Which has the lowest?

#### WHAT DO YOU NOTICE?







#### **IDEAS FOR TEACHING WITH DATA SETS 3**

Are adult heights Normally distributed?



#### WOMEN AND MEN SEPARATELY





#### **IDEAS FOR TEACHING WITH LARGE DATA SETS 4**

Physician density (physicians/1000 population)

What will the scatter diagram look like?

Health expenditure (% of GDP)



#### **IDEAS FOR TEACHING WITH LARGE DATA SETS 5**





### AGES IN 6 GROUPS (IDEA 6)

The groups are:

Divorced, Living with partner, Married, Never married, Separated, Widowed. Which box and whisker diagram goes with which group?



OXFOR CAMBRIDGE AND REA

ocr.org.uk/alevelmaths #OCRbig4

#### **EXAMBUILDER**



## **PRACTICE MAKES**

# PERFECT

Sign up to our new, free online mock assessment service!

ocr.org.uk/exambuilder



#### **NEXT STEPS**

 Download the specifications
Find out more about OCR and MEI's CPD and Teacher Networks
Talk to us

#### OCR

ocr.org.uk/alevelmaths maths@ocr.org.uk @OCR\_Maths 01223 553998 MEI office@mei.org.uk mei.org.uk/2017 @MEIMaths 01225 776776



