OCR

A Level

A Level Mathematics

Understanding the Large Dataset (Answers)

Name:

M M E

Mathsmadeeasy.co.uk

Total Marks:

Understanding the Large Dataset- Answers

OCR

1) **A demographer is investigating the dependency ratio (DR) for a Local Authority (LA) in the East Midlands. An error means that the only information they have is the year, 2001, and the DR, 55.63%. The formula for the dependency ratio is**

$$\boldsymbol{DR = \frac{Number\ of\ People\ Aged\ (0-14) + Aged\ (>65)}{Number\ of\ People\ Aged\ 15-64} \times 100}$$

i) **Determine the name of the Local Authority.**

[2 marks for correct numerator]

[2 marks for correct denominator]

Numerator is 35,332, denominator is 63,513

[1 mark]

Chesterfield.

ii) **Find the DR for the same LA in 2011.**

[2 marks]

Numerator is 36,181, denominator is 67,607.

DR= 53.52%

iii) **For this LA, assuming the DR changes linearly, in what year will the DR = 0%.**

Using the years as *x* and DR as *y*, the following model can be obtained

[2 marks- 1 for intercept, 1 for gradient]

y = 478.423 - 0.211291 x

[1 mark for correct solution]

Solving this, for $y=0$, *year (x)* = 2264.28.

iv) **In 2001 the state pension was £72.50 per person and £102.15 in 2011. Assuming it is given to everyone aged 65 and above, calculated the percentage increase or decrease in pay out for this LA from 2001 to 2011.**

[1 mark]

Cost for 2001 is £1,266,720

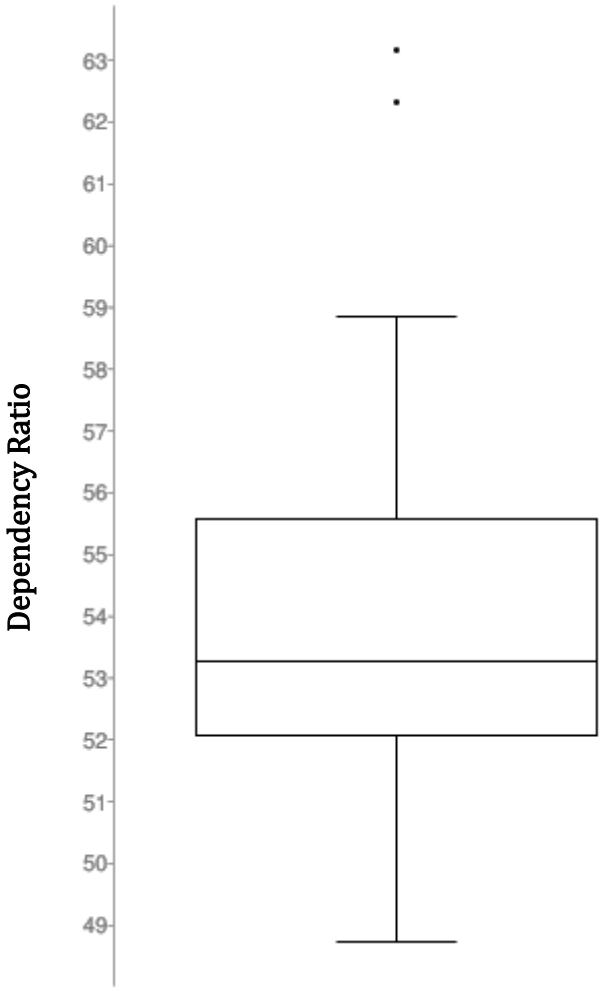
[1 mark]

Cost for 2011 is £1,970,575.65

[1 mark]- Percentage increase- 55.57%

v) **Produce a box-plot showing the 2001 DR for each LA in the East Midlands and comment on outliers.**

[1 mark for correct boxplot]



[1 mark for each correct outlier- 2 max]

[1 mark for suitable comment on each- 2 max]

Both of the outliers are as such, owing to the tight distribution of the rest of the LA.

East Lindsey – mainly rural, no major cities, low population density, several coastal resorts- holiday/retirement homes might account for this.

South Holland- similar to East Lindsey, mainly rural, no major cities, low population density.

vi) **For this region, comment on whether the median or mean would be the most appropriate measure of the average DR.**

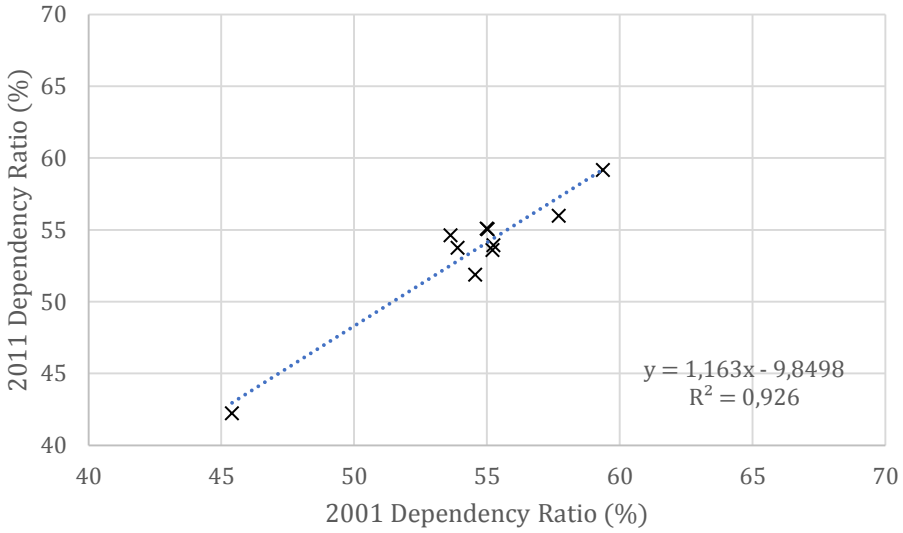
[1 mark]

Either answer is acceptable. The mean will account for the outliers. The median largely is robust to this slight skew. However, the distribution of DR is almost normally distributed with a median of 53.29 and a mean of 53.89 suggesting either measure would be a sufficient method of the average.

vii) **Calculate the mean DR for each region for both 2001 and 2011. Plot these DR on**

a scatter plot and comment on how the distribution has changed over the 10 years.

[5 marks for correct plot – 1 mark for axes, 4 marks for data, which involves calculating a DR for each data entry and then calculating the mean for each region]



[1 mark for any of the following or appropriate comment]

- Adding a trend line, we can see that the dependency ratio for most areas is increasing. i.e more people, with time, are becoming dependent.
- London has the lowest DR for both years, with the South West having the highest for both. Reasons for such are likely to be similar to those outlined in part v).
- The majority of regions have a DR around 55%. Suggesting just over 1 person is dependent on every 2 people of working age.

viii) **For 2011, determine the LA with the highest DR and the LA with the lowest DR. Using your knowledge, or research, comment on why these DR each LA are as such.**

[1 mark]

Highest DR is Christchurch at 79.21%

Wikipedia states - Christchurch is a popular destination for retirees and has one of the oldest populations in the country with 30 per cent of residents aged over 65.

It has a high house prices, rural, coastal.

[1 mark]

Lowest DR is City of London at 28.19%

Urban, London, heart of economy etc.

2) **In your new role as Transport & Travel Coordinator at the Tamworth University you have**

been tasked with investigating people's method of travel. You can remove those who are *Not in employment.*

i) **Produce and comment on two pie-charts showing how methods of travel have changed in the West Midlands from 2001 to 2011.**

[1 mark for each pie chart- 2 max]

[1 mark for comment]

There is very little change in the method of travel for the West Midlands. Car or van is the predominant method for both years, with home working followed by foot being second and third for 2001 and these latter two's order is switched in 2011.

ii) **For Tamworth create a linear regression model, with respect to time, for each of the travel types, modelling how each type has changed.**

[1 mark for ensuring data is in correct order i.e. taxi for 2001 lines up with taxi for 2011]

[1 mark]

The linear model used is

$$y = mx + c$$

where y is the number of people, m is the gradient, x is time (year), c is the intercept.

[1 mark for all correct gradients]

[1 mark for all correct intercepts]

model	Work mainly at or from home	Underground, metro, light rail, tram	Train	Bus, minibus or coach	Motorcycle, scooter or moped	Driving a car or van	Passenger in a car or van	Bicycle	Taxi	On foot	Other method of travel to work
m	2.4	0	26.9	-72.8	-13.3	259.5	-54.1	-10.2	-21.6	4.6	5.5
c	-2037.4	14	-53297.9	147812.8	27009.3	-495802	111500.1	20721.2	44263.6	-6176.6	-10922.5

In 2020, the university will have a population of 10,000 and you may assume that people travelling to it are represented perfectly by the whole population of Tamworth.

iii) **Determine the number of people who will travel to the university by *car or van*.**

[1 mark]

Using the model for car/van

$$\#\ people = 259.5(year) - 495802$$

$$\#\ people = 259.5(2020) - 495802$$

$$\#\ people = 28389\ (rounded\ to\ \%)$$