**AQA**

# A Level

# A Level Mathematics

Understanding the Large Dataset (Answers)

Name:

**MME**

Mathsmadeeasy.co.uk

Total Marks:

1) Emma lives in Newcastle. She likes to eat a cheddar and pickle sandwich on standard sliced white bread.

    i)      If Emma, is representative of the population of her local area, in which year did she eat the most pickle?

[1 mark]

2001. 24g.

    ii)      State the modal year(s) for consumption of cheese like cheddar.

[1 mark]

2004. 55g.

Based on bread consumption, Emma thinks that standard sliced white $sws$ will still be the most popular in the year 2020. Her friend, Mo, thinks that wholemeal and granary bread $wgb$, sliced and unsliced will be more popular.

    iii)      Assuming linear trends calculate whether Emma or Mo will be correct.

[2 marks for each correct intercept and gradient for each model – 4 max]

For both models $t$ is time in years.

For Emma's bread we get the model

$$sws = -11.45(t) + 23275.97$$

For Mo's bread we get the model

$$wgb = -2.34(t) - 4583.44$$

For $t = 2020$ we get

$$sws = 150.70 \ g$$
$$wgb = 149.55 \ g$$

Which means Emma is correct, standard white sliced will still be more popular in 2020, just.

    iv)      Why is assuming linear trends, without revising the model, unlikely to be a useful assumption to make. Use the fact that Emma intends to be eating sliced white bread in the year 2040.

[1 mark for any comment that refers to negative values]

Using the current model Emma will be eating -78.26g of bread a week in 2040. This means that a linear trend is not likely for the next 20 years and the model needs revising to be polynomial.

2) In the East Midlands from 2002 to 2004 (inclusive) the total consumption of fats remained the same at 186g per person. However, the different types of fats consumed changed.

    i) Produce one radar demonstrating the changing tastes in fat consumption over the three years.

This chart must not include the cumulative totals which are entitled vegetable and salad oils, all other fats and margarine, otherwise the total consumption does not add up to 186g.

[1 mark for each of the graphs draw correctly- 3 max]

[1 mark for 10 correct labels]

[1 mark for key]



······ 2002 ── 2003 ── ·2004
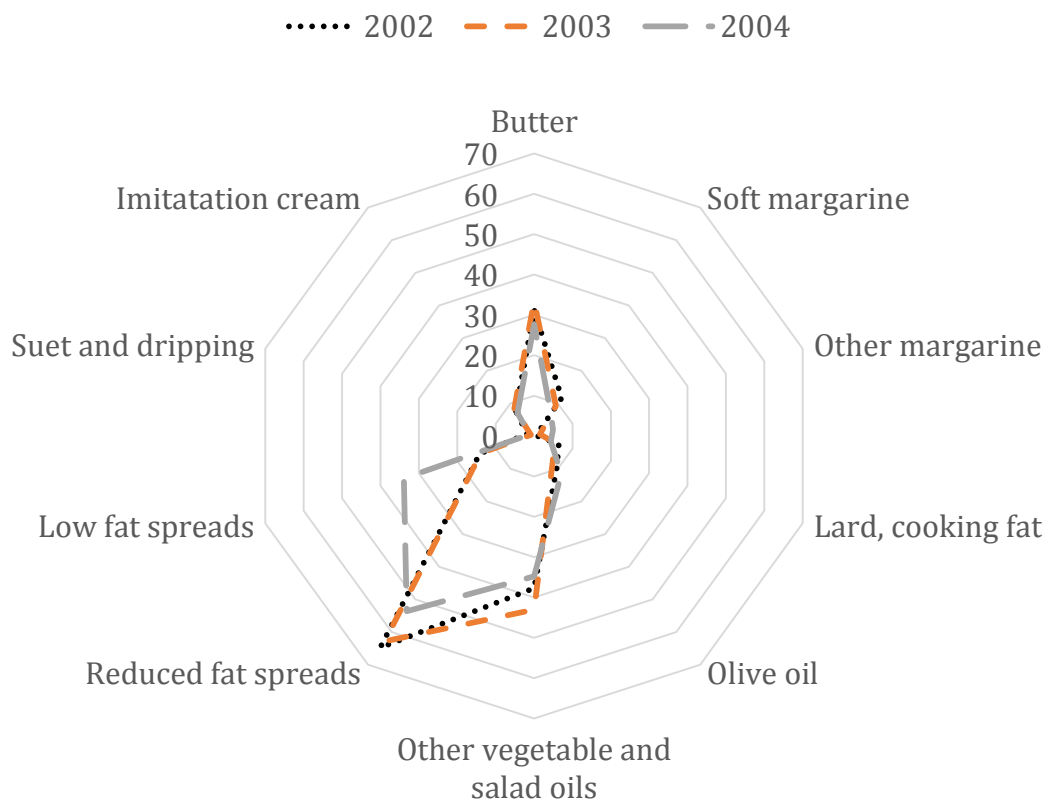
    ii) Comment on the changing tastes.

[1 mark for each of the following comments – 2 max]

- Largely the types of fats consumed remained the same, with reduced fat spreads being the favourite across all three years.
- Low fat spreads were consumed in almost 3x as many grams in 2004 than in 2002 and 2003, with a reduction in reduced fat spread consumption.
- Lard, although consumed in small proportions, appears to be reducing each year and conversely other margarines appear to be increasing in consumption.
- Other vegetable and salad oils remain as the second most consumed subcategory across the three years.

3) The consumption of jams and fruit curds in England was 146g per person per week in 2001 and 14g per person per week in 2014.

   i) Comment on the significance on pattern/conclusion/hypothesis you can draw from the data spanning this period.

[1 mark for correlation]

[1 mark for acceptance and significance of this finding]

Correlation is significant., with a value of $\rho = -0.73$. The P-Value is 0.003348. The result is significant at $p < 0.05$. Therefore, we can reject the null hypothesis of $H_0: \rho = 0$ in favour of $H_1: \rho \neq 0$. I.e. there is strong evidence of correlation. In this instance, we can say as time increases, the consumptions of jams decrease.

   ii) Find the region in the England which, for jams and fruit curds, is most like the national trend.

[1 mark for South West]

South west is correlated to England with a correlation of 0.68, with an associated p-value of 0.003642, thus we can conclude there is a correlation.

   iii) What sampling errors might be occurring to account for the results you have?

[1 mark for any of the following – 2 max]

- using self-reported diaries, this reporting method might not be consistent across individuals
- over a two-week period, this has a strong potential to not be representative of the entire year
- where possible quantities are recorded in the diaries but otherwise estimated, the method of estimation is not reported with the dataset, assumptions for the estimation may be biased.