

Averages from lists R

The Mean

A measure of average to find the central tendency... a typical value that represents the data

24, 8, 4, 11, 8

Find the sum of the data (add the values)

55

Divide the overall total by how many pieces of data you have

$55 \div 5$

Mean = 11

The Mode (The modal value)

This is the number OR the item that occurs the most (it does not have to be numerical)

24, 8, 4, 11, 8

This can still be easier if the data is ordered first

Mode = 8

The Median

The value in the center (in the middle) of the data

24, 8, 4, 11, 8

Put the data in order

4, 8, 8, 11, 24

Find the value in the middle

4, 8, 8, 11, 24

Median = 8

NOTE: If there is no single middle value find the mean of the two numbers left

For Grouped Data

The modal group – which group has the highest frequency

Choosing the appropriate average

The average should be a representative of the data set – so it should be compared to the set as a whole – to check if it is an appropriate average

Here are the weekly wages of a small firm

£240 £240 £240 £240 £240
£260 £260 £300 £350 £700

Which average best represents the weekly wage?

The Mean = £307

The Median = £250

The Mode = £240

Put the data back into context

Mean/Median – too high (most of this company earn £240)
Mode is the best average that represents this wage

It is likely that the salaries above £240 are more senior staff members – their salary doesn't represent the average weekly wage of the majority of employees

Comparing distributions

Comparisons should include a statement of average and central tendency, as well as a statement about spread and consistency

Here are the number of runs scored last month by Lucy and James in cricket matches

Lucy: 45, 32, 37, 41, 48, 35

James: 60, 90, 41, 23, 14, 23

Lucy

Mean: 39.6 (1dp), Median: 38, Mode: no mode, Range: 16

James

Mean: 41.8 (1dp), Median: 32, Mode: 23, Range: 76

James has two extreme values that have a big impact on the range

James is less consistent than Lucy because his scores have a greater range. Lucy performed better on average because her scores have a similar mean and a higher median

Averages from a table R

Non-grouped data

Number of Siblings	0	1	2
Frequency	6	8	6
Subtotal	0	8	12

Overall Frequency: 20

Total number of siblings: 20

The data in a list: 0,0,0,0,0,1,1,1,1,1,1,1,2,2,2,2,2,2

Mean: $\frac{\text{total number of siblings}}{\text{Total frequency}} = 1$

Grouped data

x	Frequency	Mid Point	MP x Freq
40 < x ≤ 50	1	45	45
50 < x ≤ 60	3	65	195
60 < x ≤ 70	5	65	325

Overall Frequency: 9

Overall Total: 565

Mean: 62.8g

The data in a list: 45, 55, 55, 55, 65, 65, 65, 65, 65

