

## Descriptive Statistics - Numerical Measures

### Measures of Central Tendency

The **mean** value of a set of data is:

either

- the sum of the data divided by the number of data
- or
- the weighted average of the data values.

## Sigma Notation

We can use sigma notation to represent a sum of values in a more concise fashion.

$$\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$$

## Sigma Notation

We can use sigma notation to represent a sum of values in a more concise fashion.

$$\sum_{i=1}^n x_i$$

$x_i$  is the expression to be added to itself over and over again

## Sigma Notation

$$\sum 2 = 2 + 2 +$$

But how many times?

$$\sum_{i=1}^n x_i$$

'n' times

$$\sum_{i=1}^3 2 = 2 + 2 + 2 = 6$$

## Sigma Notation

$$\sum_{i=1}^n x_i$$

This indicates the starting value from which we should add on each time

$$\sum_{i=1}^4 i = 1 + 2 + 3 + 4 = 10$$

So add up: four terms, taking values 1 onwards

## Sigma Notation

$$\sum_{i=1}^5 m$$

$$\sum_{i=1}^3 d + 2$$

$$\sum_{i=1}^3 n^2$$

$$\sum_{i=2}^3 n^2$$

### EXAMPLE: Finding the Mean

- Eleven Geography students were asked how much they spent on travel each week (in £).

16 20 24 11 20 15 18 22 10 14 17

- Find the mean travel spend for these students:
- What are the advantages and disadvantages of using the mean to summarise this data?

### Measure of Central Tendency: Median

#### Odd number of data

The middle value when the data is presented in ascending order.

#### Even Number of Data

The mean of the middle two values when presented in ascending order.

### EXAMPLE: Finding the Median

- ▶ Eleven Geography students were asked how much they spent on travel each week (in £).

16 20 24 11 20 15 18 22 10 14 17

- ▶ Find the median travel spend for these students:
- ▶ What are the advantages and disadvantages of using the median to summarise this data?

### Measure of Central Tendency: Mode

The data value(s) that appear(s) most frequently

Every data set has **at least one** modal value

If there are two modal values we say the data set is **bimodal**

### EXAMPLE: Finding the Mode

- ▶ Eleven Geography students were asked how much they spent on travel each week (in £).

16 20 24 11 20 15 18 22 10 14 17

- ▶ Find the mode travel spend for these students:
- ▶ What are the advantages and disadvantages of using the mode to summarise this data?

### Examples

Find the mean, median and mode for the following data sets:

Minutes between received text messages:

10 5 6 2 35 14 12 7 8 10 1 14  
11 5 8 12 2 9 3 4 11 11 5 6

Time in minutes spent studying for a test:

35 60 40 7 100 120 12 50 75 33 50

## Mean, Median and Mode Comparison

### Mean

Advantage: takes into account all data values

Disadvantage: sensitive to extreme values

### Median

Advantage: not sensitive to extreme values

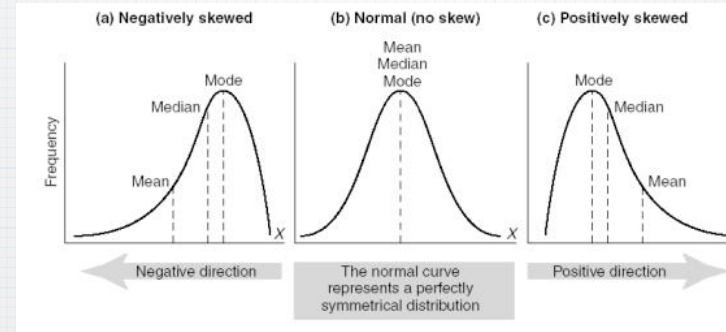
Disadvantage: only considers middle value

### Mode

Advantage: quick and easy to find

Disadvantage: may include all or many values

## Mean, Median and Mode can suggest skew ..



## Mean, Median and Mode from Frequency Distribution

Calculate

No of siblings

## Mean, Median and Mode from Frequency Distribution

Number of students attending 50 lectures			
20	31	40	17
43	31	23	31
18	16	20	31
22	32	20	38
24	30	39	33
25	24	26	39
30	32	26	40
22	24	36	32
41	42	28	18
15	33	28	17
38	33	29	17
23	31	29	33
21	15		

Class	Frequency
15-19	8
20-24	11
25-29	7
30-34	14
35-39	5
40-44	5
<b>Total</b>	<b>50</b>

## Mean, Median and Mode from Frequency Distribution

Classes	Frequency	Relative Frequency	Percentage Frequency	Midpoint	freq*mid
	$f$			$x$	$fx$
15-19	8	0.16	16%	17	136
20-24	11	0.22	22%	22	242
25-29	7	0.14	14%	27	189
30-34	14	0.28	28%	32	448
35-39	5	0.1	10%	37	185
40-44	5	0.1	10%	42	210
	50	1	1		1410
	So mean is	1410	28.2		
		50			
	Most is	30-34			
	Median is at point $50/2 = 25$	25-29			

## Mean, Median and Mode from Frequency Distribution

AA battery life (in minutes)

Classes	Frequencies
	$f$
360 - 369	2
370 - 379	3
380 - 389	5
390 - 399	7
400 - 409	5
410 - 419	4
420 - 429	3
430 - 439	1
	30

Calculate  
 relative frequency  
 midpoints  
 estimated mean  
 median  
 mode

## Using Excels built in frequency distribution function

Classes	frequency	Hints:
15 - 19	8	1 Place classes in different columns
20 - 24	11	2 Highlight ALL the blank frequency cells
25 - 29	7	3 Type in the formula bar the frequency formula
30 - 34	14	4 Select A column for data array
35 - 39	5	5 Select E column for bins array
40 - 44	5	6 Hit ctrl-shift-enter for formula array to be created