

# community project

encouraging academics to share statistics support resources

All stcp resources are released under a Creative Commons licence

stcp-gilchristsamuels-5

The following resources are associated:  
Descriptive Statistics – Measures of Spread worksheet

## Descriptive Statistics – Measures of Middle Value

**Research question type:** Most

**Kind of variables:** Nominal, ordinal and interval/scale

**Common Applications:** Most quantitative studies

Descriptive statistics are a way of summarising a data set into a few values, often for comparative purposes. Commonly they are presented in tables, although appropriate charts are also used. They include measures of middle values (or central tendency) and measures of spread (or dispersion).

The following is a rough guide to appropriate measures for level of data:

Data level	Middle value (central tendency)	Dispersion (spread)
Nominal	Mode	
Ordinal	Mode, median	Range, interquartile range (IQR)
Interval/scale	Median, mean	Variance, standard deviation

Measures of middle value (central tendency)

### 1. Mode

The mode or modal value of a set of data is the most frequently occurring value.

#### Example

Find the mode of: 2, 6, 3, 9, 5, 6, 2, 6.

#### Solution:

The most frequently occurring value is 6 (there are three of them). Hence the mode is 6.



## 2. Median

The median value of a set of data is the middle value of the ordered data. That is, the data must be put in numerical order first.

### Example

Find the medians of the following data sets:

a) 11, 4, 9, 7, 10, 5, 6

b) 1, 3, 0.5, 0.6, 2, 2.5, 3.1, 2.9

### Solution:

a) Ordering the data gives 4, 5, 6, 7, 9, 10, 11 and the middle value is 7. Hence, the median is 7.

b) Ordering the data gives 0.5, 0.6, 1, 2, 2.5, 2.9, 3, 3.1. Here, there is a middle pair: 2 and 2.5.

The median is midpoint between this pair of values. Median =  $\frac{(2+2.5)}{2} = 2.25$ .

In general, if there are  $n$  values, the median is at the  $\frac{(n+1)}{2}$ <sup>th</sup> value when the values are written in ascending order. (Note: for ordinal data, it may not be possible to calculate a midpoint).

## 3. Mean

To calculate the mean of a set of scale data values, we must first add up (sum) all the data values (represented by the variables  $x_1, x_2, \dots, x_n$ ), and then divide the result by the number of values ( $n$ ). A  $\Sigma$  (sigma) symbol is often used to indicate that values are to be summed using a summation variable, such as  $i$ , to represent the numbers from 1 to  $n$ .

The mean of a set of  $n$  data values  $x_1, x_2, \dots, x_n$  is then:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

### Example

Find the mean of: 6, 8, 11, 5, 2, 9, 7, 8.

### Solution:

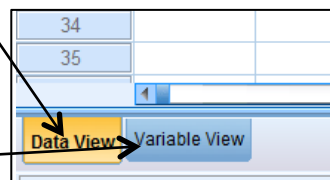
$$\text{Mean} = \bar{x} = \frac{\sum x_i}{n} = \frac{6+8+11+5+2+9+7+8}{8} = \frac{56}{8} = 7$$

## Steps in SPSS

### Data entry

Using the same example above, open SPSS, select Type in Data, select the Data View tab at the bottom of the screen and enter the values into the first column of the Data Editor window as shown on the right.

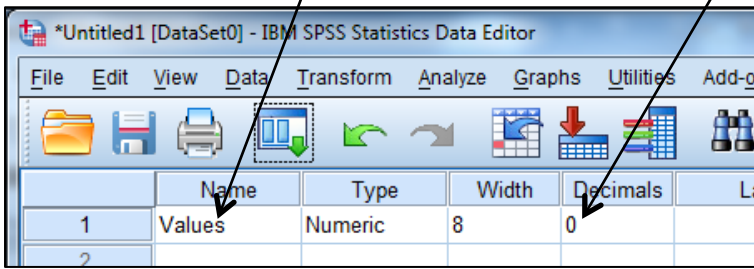
Select the Variable View tab at the bottom of the screen.




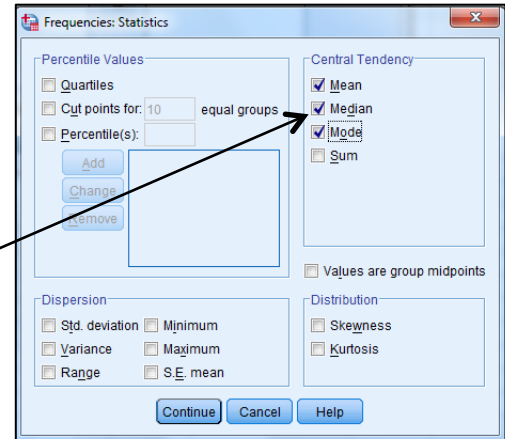
*Untitled1 [DataSet0] - IBM SPSS Data Editor	
9 : VAR00001	
	VAR00001
1	6.00
2	8.00
3	11.00
4	5.00
5	2.00
6	9.00
7	7.00
8	8.00
9	



Rename the variable 'Values', and set the number of decimal places to zero.



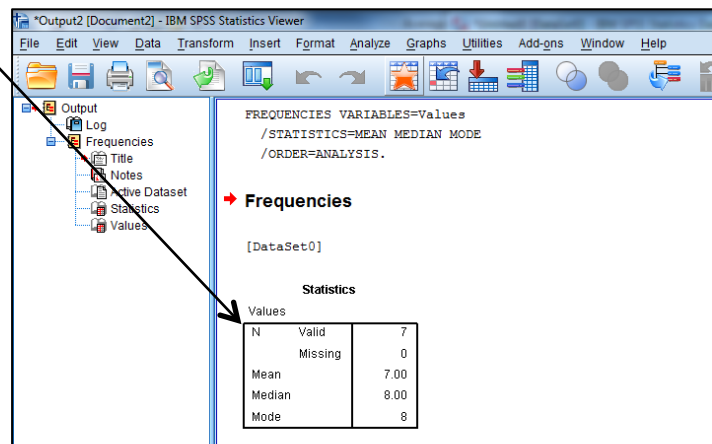
- Select Analyze – Descriptive Statistics – Frequencies... from the menu
- Select the variable Values and use  to move it to the variable(s) list
- Click on the Statistics... button
- Check the boxes next to mean, median and mode
- Click Continue and OK



The results are given in the Output window or SPSS Statistics Viewer window.

The mean, median and mode are 7, 8 and 8 respectively.

**Note:** The Analyze – Descriptive Statistics – Explore... option also provides most of the descriptive statistics that you will need (but not the mode).

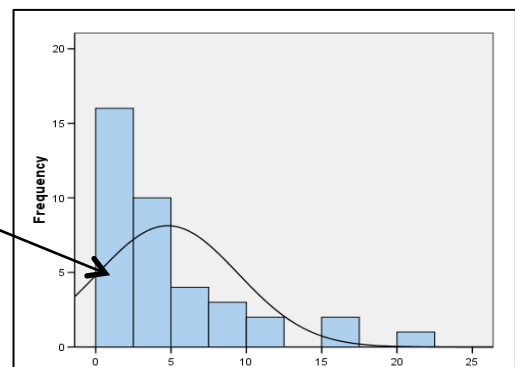


### Choice between mean and median for scale data

If a data set is approximately normally distributed (or at least symmetrical) then the mean and media will be approximately the same.

If it is skewed then the mean and median will be different. For example, if a data set is skewed to the left (see diagram) then the mean (given by the peak in the fitted normal curve) will be higher than the median. The median is then a more meaningful measure of central tendency.

If a data distribution is even more irregular then the mean is probably the best measure to use.



**Exercise**

- Find the mode, median and mean of the following:
  - 3, 12, 11, 7, 5, 5, 6, 4, 10
  - 16, 19, 10, 24, 19
  - 8, 2, 8, 5, 5, 8
  - 28, 39, 42, 29, 39, 40, 36, 46, 41, 30
  - 133, 215, 250, 108, 206, 159, 206, 178
  - 76, 94, 76, 82, 78, 86, 90
  - 52, 61, 49, 52, 49, 52, 41, 58
- The exchange rate for sterling against the US dollar for 15 days in January 2003 is given in the following table. Calculate the mean, median and mode of this data.
 

1.5977	1.6028	1.6108	1.6067	1.5995	1.6064	1.6080	1.6054
1.6098	1.6049	1.6064	1.6179	1.6082	1.6095	1.6134	
- The following data is the hourly pay in pounds offered to shop assistants in a survey of job vacancies for June 2003 based in Reading:
 

5.00	4.21	4.97	5.00	5.00	5.29	5.05	5.50	5.79	5.00	5.40
5.20	5.10	5.06	4.50	4.50	5.50	4.50	5.00	5.00	5.50	

Calculate the mean, median, and mode of this data.

**Answers**

- 5, 6, 7
  - 19, 19, 17.6
  - 8, 6.5, 6
  - 39, 39, 37
  - 206, 192, 181.875
  - 76, 82, 83.14
  - 52, 52, 51.75
- Mean = \$1.6072, median = \$1.6067, mode = \$1.6064
- Mean = £5.05, median = £5.00, mode = £5.00

