

Σ+ SPSS TUTORIALS

[BASICS](#) [DATA ANALYSIS](#) [T-TEST](#) [ANOVA](#) [CHI-SQUARE TEST](#)

Median – Simple Tutorial & Examples

- [Median - Simple Data Examples](#)



[Relation Median and Mean](#)

- [Median - Strengths & Weaknesses](#)
- [Finding Medians in SPSS](#)
- [Statistical Significance for Medians - Sign Tests](#)

The median is the middle value after sorting all values

for an odd number of values. For an even number of values, it's the

You are here: [Home](#) → [Basics](#) → [Must-Know Statistics](#) → [Median – Simple Tutorial & Examples](#)

average of the 2 middle values after sorting all values. The examples below from [this GoogleSheet](#) (read only) will make this perfectly clear.

[Median – Simple Tutorial & Examples](#)

Median - Simple Data Examples



Median ☆ www.spss-tutorials.com

File Edit View Insert Format Data Tools Add-ons Help All changes saved in Drive

100% \$ % .0 .00 123 Default (Ari... 16 B I S A

fx =MEDIAN(B2:B7)

	A	B	C	D	E	F	G
1		V1	V2	V3	V4	V5	
2		1	1	1	4	1	
3		2	2	2	2	1	
4		3	3	3	100	1	
5		4	4	4	5	1	
6		5	5	5	3	1	
7			6	100	1	2	
8							
9	MEDIAN	3	3.5	3.5	3.5	1	
10	MEAN	3	3.5	19.17	19.17	1.17	
11							
12							

- **V1** holds values 1 through 5 sorted ascendingly. The median -middle value- is 3.
- **V2** holds values 1 through 6 sorted ascendingly. The median is 3.5. It is the average of the 2 middle values 3 and 4.
- **V3** is V2 with 6 replaced by 100. This greatly affects the mean but the 2 middle values -and hence the median- stay the same.
- **V4** holds the values of V3 in random order. The median is *not* the average of the 2 middle values unless we first sort them.
- **V5** contains ties: the value 1 occurs 5 times. Since the values are sorted, the median is the average of the 2 middle values (1 and 1).

Note that for V2 through V4,

the median is the value that separates the 50% highest values from the 50% lowest values.

This turns out to hold for most (semi)continuous variables that we find in real-world data such as



- exact monthly incomes in dollars,
- body weight in grams or
- age in days.

However, it may not hold at all for heavily tied data (such as V5) or small numbers of observations.

[▶ AdChoices](#)[Histogram Chart](#)[Analyzing Data](#)[Probability Calculator](#)

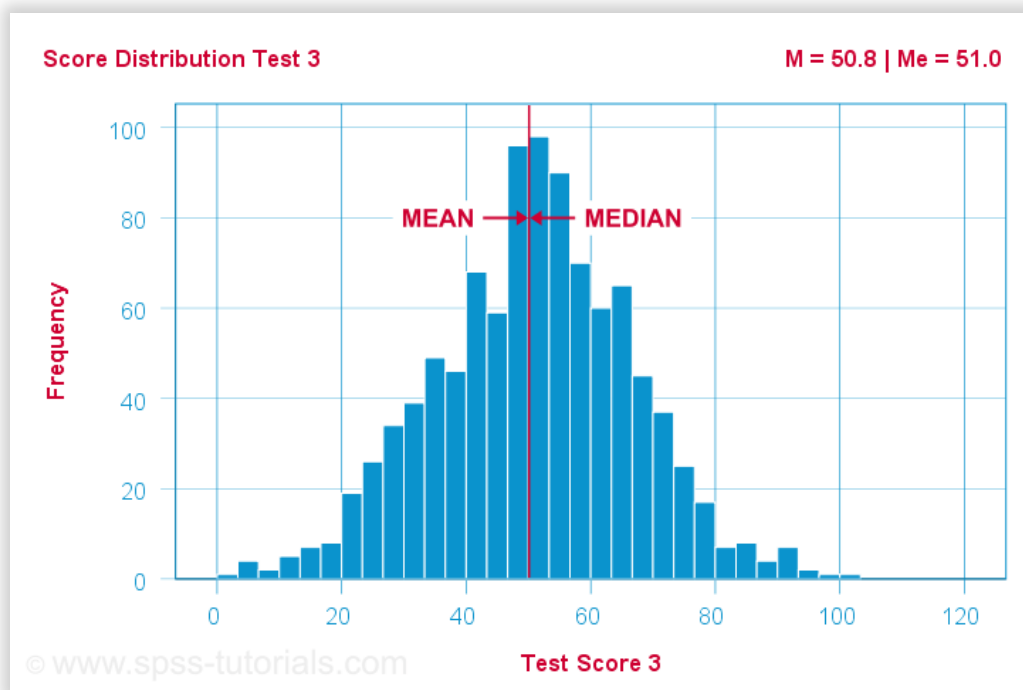
Relation Median and Mean

We'll discuss the pros and cons of medians versus means in a minute. Let's first see how they relate in the first place. This depends mostly on the skewness of the frequency distribution of some variable:

**the median is *equal* to the mean
for symmetrically distributed variables**

which implies skewness = 0. The histogram shown below illustrates this point.





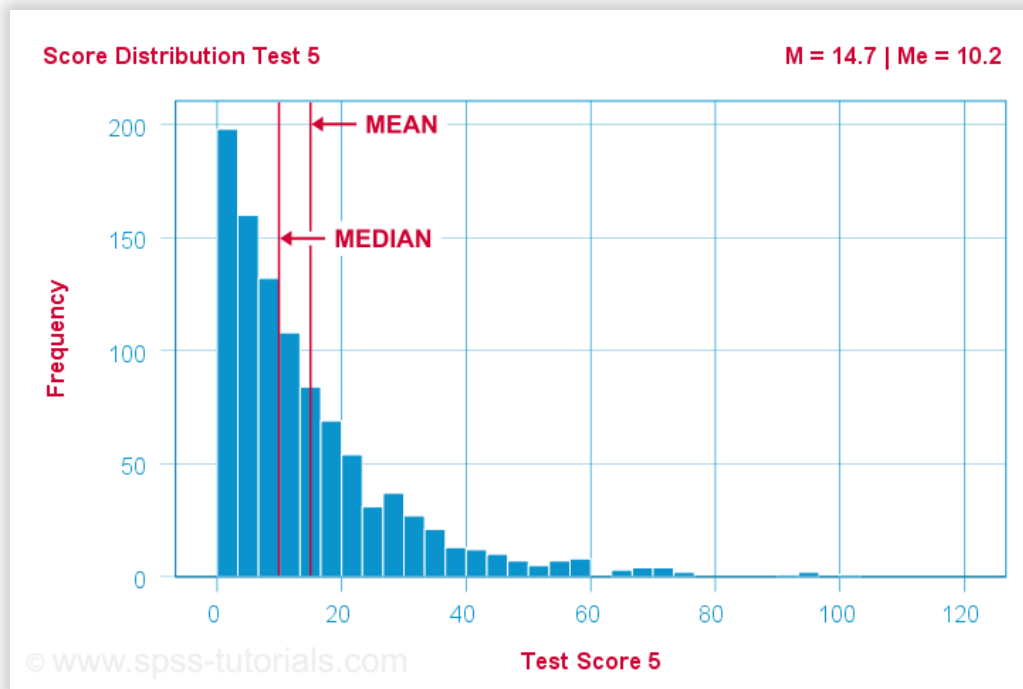
Skewness is basically zero for these 1,000 test scores. The sample mean (M) = 50.8 while the median (Me) = 51.0. The red lines indicating them on the x-axis are indistinguishable.

Different patterns occur when skewness is substantial. First off,

**the median is *smaller* than the mean
for positively skewed variables**

as shown below.





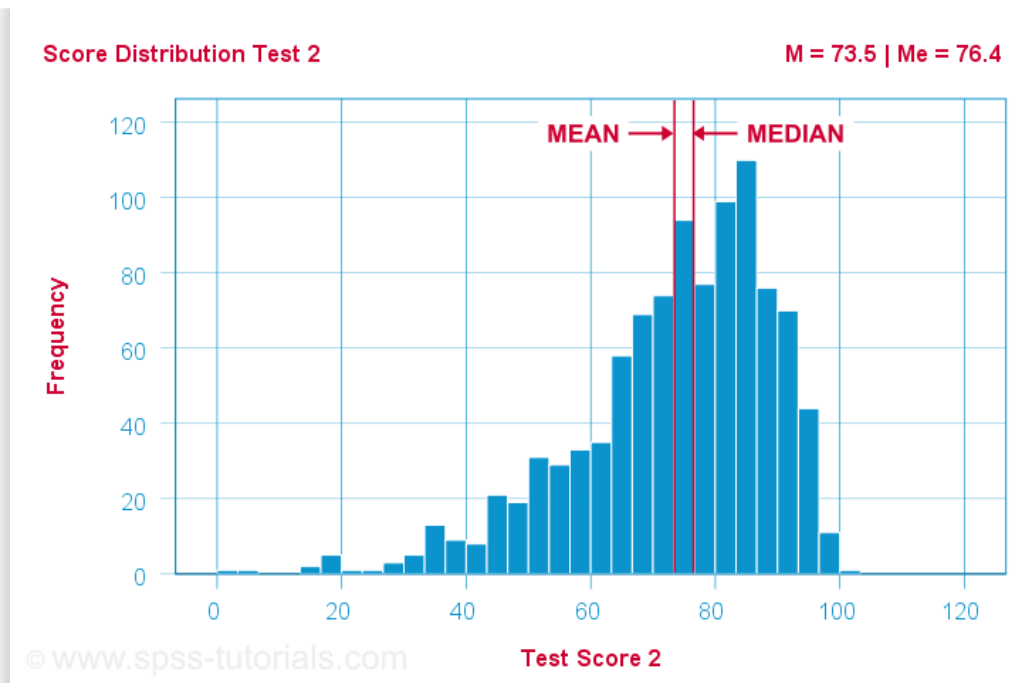
What basically happens here is that some very high scores affect the mean but not the median. We already saw this in our initial examples: changing {1,2,3,4,5,6} to {1,2,3,4,5,100} greatly affects the mean but the median is 3.5 for both variables. The histogram above shows the exact same phenomenon but it uses more realistic data.

As you can probably guess by now, the opposite also holds:

**the median is *larger* than the mean
for negatively skewed variables**

as illustrated by the histogram below.



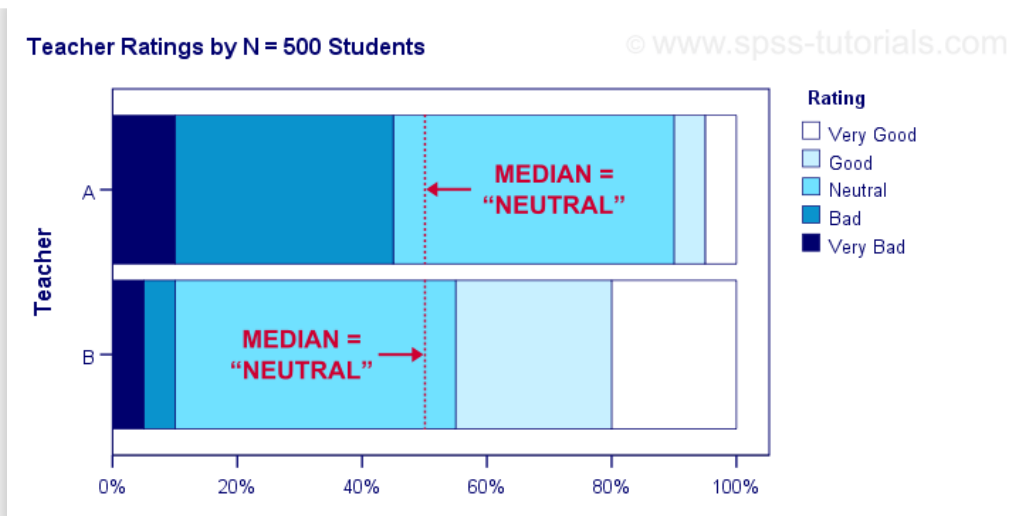


What basically happens here is that the very low scores “drag down” the mean. The median, however, is unaffected by these.

Median - Strengths & Weaknesses

Thus far, this introduction implicitly pointed out some **strengths** of the median compared to the mean:

- The median is not sensitive to outliers. So perhaps the *mean* salary for some people is high due to a single billionaire. I'd rather know the *median* salary in this case. This'll tell me (roughly) which salary separates the 50% lowest from the 50% highest incomes. It's a more realistic estimate of what these people *tend to* earn.
- Means are applicable only to quantitative variables. Medians are also suitable for **ordinal variables**. However, ordinal variables typically have tons of ties (values that occur more than once). For such variables, medians may be misleading as shown below.



Although teacher B is rated much more favorably than teacher A, their median ratings are identical.

Apart from these strengths, medians have some **weaknesses** too:

- Medians are unsuitable for numeric calculations. For instance, sums can be calculated from means and sample sizes but not from medians. The difference between 2 means is easily interpretable but the difference between 2 medians hardly so.
- In the presence of ties, very different variables may have similar medians.
- The median may not actually exist. For instance, if two people have 0 and 1 children, their median is 0.5 children.
- The median is said to fluctuate more from sample to sample than the mean. That is, it is less stable and has a larger standard error.

Finding Medians in Googlesheets

Finding medians is super easy with Googlesheets. For example, typing

```
=MEDIAN (B2 : B7)
```

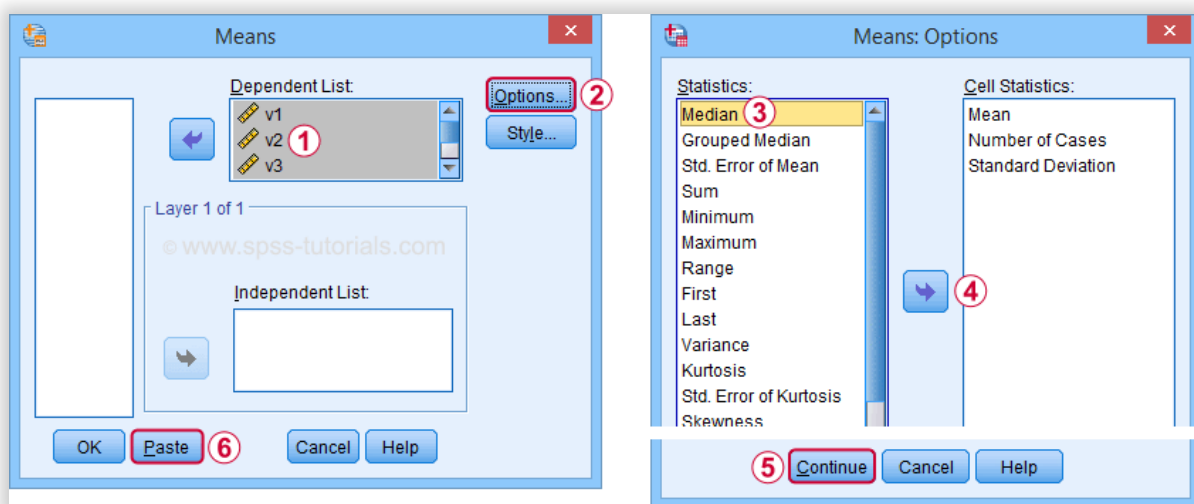
into any cell results in the median of cells B2 through B7 (assuming all non empty cells contain numbers). Some more examples are shown in

this [Googlesheet](#) (read only).

Finding Medians in SPSS

In **SPSS**, the best option to find medians is from **Analyze** ► **Compare Means** ► **Means**

Use this dialog to create a table showing a wide variety of descriptive statistics including the mean, **standard deviation**, skewness, kurtosis and more. Optionally, these are reported for separate groups defined by "Independent List".



An even faster option is typing and running the resulting **syntax** -a simple **MEANS** command- such as

```
means v1 to v5  
/cells count mean median.
```

An example of the resulting table -after some adjustments- is shown below.



Descriptive Statistics

© www.spss-tutorials.com

	N	Minimum	Maximum	Mean	Median	SD	Skewness	Kurtosis
Test Score 1	1000	.0	100.0	86.9	90.8	12.9	-2.1	6.0
Test Score 2	1000	.0	100.0	73.5	76.4	16.0	-1.0	1.3
Test Score 3	1000	.0	100.0	50.8	51.0	16.0	-.1	.1
Test Score 4	1000	.0	100.0	24.5	22.8	13.0	1.0	2.2
Test Score 5	1000	.0	100.0	14.7	10.2	14.8	2.0	5.1

Notice the huge positive **correlation** between skewness and (mean - median): the median is larger than the mean insofar as a variable is more negatively (left) skewed. The opposite pattern -mean larger than median- occurs for positively (right) skewed variables. This was previously illustrated with some histograms based on the same data file as this table.

Statistical Significance for Medians - Sign Tests

Among the most popular statistical techniques are **t-tests**. These test if the difference between 2 means is **statistically significant**. But what if we want to test for medians instead of means? In this case we'll end up with one of 3 median tests, sometimes called **sign tests**:

- A **sign test for 1 median** is similar to a **one sample t-test** for a median: it compares the sample median to a hypothesized value.
- The **sign test for independent medians** is similar to an **independent samples t-test** or a **one-way ANOVA** for medians: it tests if 2 or more populations have equal medians.
- A **sign test for related medians** is similar to a **paired samples t-test** for medians: it tests if 2 variables measured on the same people or other observations have equal medians.



A sign test for 1 median basically works like this:

- each value *smaller* than the hypothesized median is replaced by a minus (-) sign;
- values larger than the hypothesized median are replaced by plus (+) signs;
- if the hypothesized median is correct, then some 50% of all signs should be plusses;
- a **binomial test** examines if the sample proportion of plusses is significantly different from 0.5.

© www.spss-tutorials.com

SIGN TEST FOR 1 MEDIAN

	id	gender	ad1	me0	t1
4	5	Female	93	79.50	+ (above median)
5	8	Male	58	79.50	- (below median)
6	10	Female	89	79.50	+ (above median)
7	4	Female	92	79.50	+ (above median)
8	6	Male	49	79.50	- (below median)
9	17	Female	86	79.50	+ (above median)

AD1: TEST VARIABLE

ME0: HYPOTHESIZED MEDIAN

T1: SIGN (50% SHOULD BE +)

The other sign tests follow the same basic reasoning. Sign tests are not very popular because ties are problematic for them and they tend to have low statistical power.



Thanks for reading!



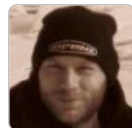
Let me know what you think!

Done!

**Required field. Your comment will show up after approval from a moderator.*

This tutorial has 6 comments

By **Ruben Geert van den Berg** on November 5th, 2019



Hi Mariam!

There's no .pdf versions of any of our tutorials. However, you can try and **print** any page to a pdf file from your web browser.

Kind regards

Expand [comment](#) | [all comments](#)



Get In Touch!

SPSS Help (Netherlands)

Ruben Geert van den Berg

Sigma Plus Statistiek

 LinkedIn

www.sigma-plus-statistiek.nl

 Facebook

info@sigma-plus-statistiek.nl

SPSS Help (International)

SPSS tutorials

www.spss-tutorials.com

info@spss-tutorials.com

 **SPSS tutorials**

© Copyright Protected 2019 by **Sigma Plus Statistiek**

Disclaim

