factsheet 16: t-scores stens and stanines

Standard scores - T-scores, stens and stanines

All of the above are standard score scaling systems. Each type of score enables test scores to be interpreted in relation to a norm group. The relationship between the scoring systems (ie Z-scores, T-scores, stens etc) is like the relationship between centigrade and Fahrenheit, or yards and metres. When measuring temperature: different scales are used by different people in different circumstances, but it is possible to convert from one scale to another. If necessary, a single score can be expressed as a value on each of the different scales.

As you can see from the following formulae, each of the new standard score systems is based on the Z-score:

$$\text{T-score} = (\text{Z-score} \times 10) + 50$$

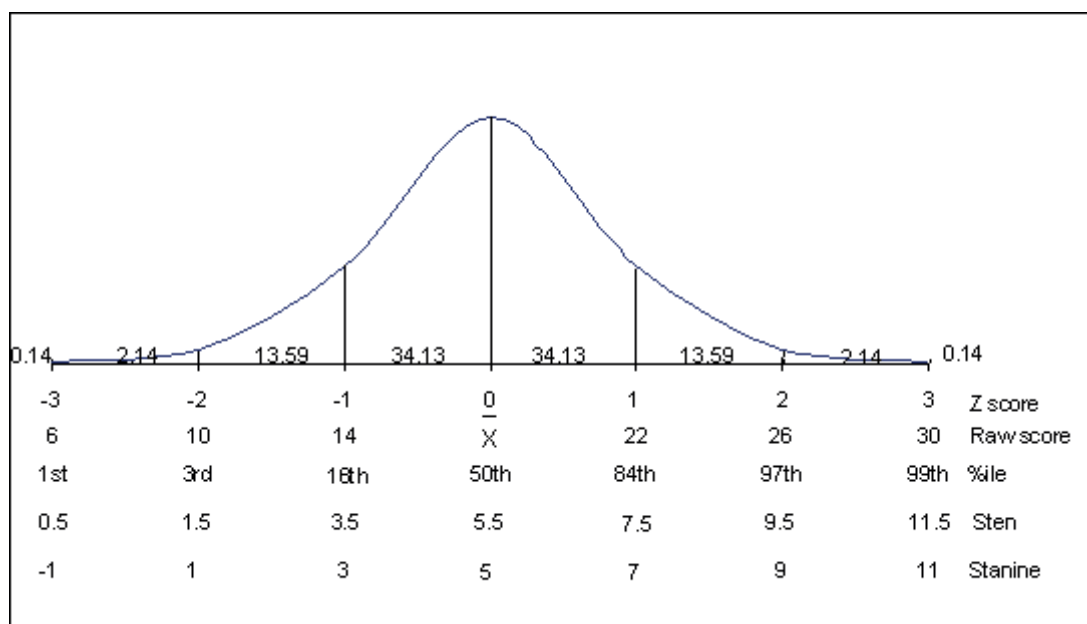$$\text{Sten} = (\text{Z-score} \times 2) + 5.5$$

$$\text{Stanine} = (\text{Z-score} \times 2) + 5$$

If we take some simple Z-scores (such as -2, -1, 0, +1, and +2) we can use the formulae to calculate other standard score equivalents as follows:

Z-Score	T-Score $(Z \times 10) + 50$	Sten $(Z \times 2) + 5.5$	Stanine $(Z \times 2) + 5$
-2	(-2 x 10) + 50 = (-20 + 50) = **30**	(-2 x 2) + 5.5 = (-4 + 5.5) = **1.5**	(-2 x 2) + 5 = (-4 + 5) = **1**
-1	(-1 x 10) + 50 = (-10 + 50) = **40**	(-1 x 2) + 5.5 = (-2 + 5.5) = **3.5**	(-1 x 2) +5 = (-2 + 5) = **3**
0	(0 x 10) +50 = (0 + 50) = **50**	(0 x 2) + 5.5 = (0 + 5.5) = **5.5**	(0 x 2) + 5 = (0 + 5) = **5**
+1	(1 x 10) + 50 = (10 + 50) = **60**	(1 x 2) + 5.5 = (2 + 5.5) = **7.5**	(1 x 2) + 5 = (2 + 5) = **7**
+2	(2 x 10) + 50 = (20 + 50) = **70**	(2 x 2) + 5.5 = (4 + 5.5) = **9.5**	(2 x 2) + 5 = (4 + 5) = **9**

0.14	2.14	13.59	34.13	34.13	13.59	2.14	0.14
-3	-2	-1	0	1	2	3	Z score
6	10	14	X̄	22	26	30	Raw score
1st	3rd	16th	50th	84th	97th	99th	%ile
0.5	1.5	3.5	5.5	7.5	9.5	11.5	Sten
-1	1	3	5	7	9	11	Stanine

Although some decimals are shown in the tables above, T-scores, stens and stanines are all usually rounded to the **nearest whole number**. A sten calculated to be 7.78 would therefore be rounded up to 8.

If the calculated standard score falls precisely between two whole numbers (eg 7.5), there are (unfortunately) different views about how this value should be rounded. If it is a score from an ability test, the candidate may be given "the benefit of the doubt" by rounding up to 8. Another convention used by some practitioners is always to round away from the mean, so a sten of 7.5 would be rounded to 8, but a sten of 3.5 would be rounded down to 3.

The name "T-score" is derived from "**transformed score**" - all standard score are based on linear transformations, first from raw scores to Z-scores, and then from Z-scores to another standard scale.

The name "**sten**" comes from "standard ten", because stens are expressed as a whole number ranging from 1 to 10 (with 5.5 representing the mean). Stens calculated to be lower than 1 (eg 0.5, -1.2) are rounded up to 1, the bottom value for the scale. Stens calculated to be greater than 10 (eg 11.4, 10.9) are rounded down to 10, the highest value for the scale.

Stanines (from "**standard nine**") are similar to stens except that the mean is 5.0, and the scale ranges from 1 to 9 (ie stanines calculated to be greater than 9 are rounded down to 9).

In the same way that we can state, for any normally distributed group, that the Z-score mean will be 0 and the Z-score SD will be 1, we can identify the mean and SD for each of the other standard score scales.

Again, these will be consistent for any group regardless of the raw score mean and SD.

	Mean	SD
Z-Score scale	0	1
T-Score scale	50	10
Sten score scale	5.5	2
Stanine score scale	5	2

T-scores are often used (as an alternative to percentiles) in the interpretation of ability tests. This is because they indicate each candidate's score with a fair degree of precision, usually in the range 20 to 80. Personality measures tend to utilise stens or stanines: this is because these scales are a little less precise, rounding to whole numbers on a scale of 1-10 or 1-9. Since personality measurement tends to be less precise than ability measurement, stens or stanines are more appropriate.