

Stanscore 5: Normalized - Standardized Scores Norm Table Generation



Paul Barrett

Advanced Projects R&D Ltd.
New Zealand

Web: <https://www.pbarrett.net>

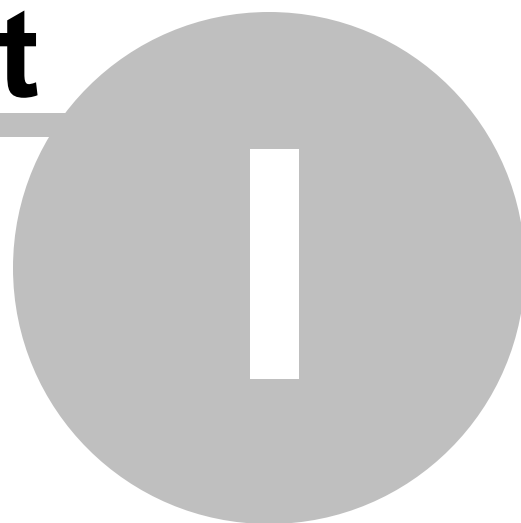
Email: paul@pbarrett.net

December 1st, 2018

Table of Contents

Part I Introduction	4
1 New in Stanscore 5.....	5
2 New in Stanscore 4.....	5
3 StanScore 5 Basics and Information.....	7
4 Using the program.....	9
Part II Getting Data into the Program	17
1 Access Database Tables.....	18
2 ASCII Files.....	19
3 Excel Files.....	20
4 PsWin v2 & above Scale Score files.....	21
5 STATISTICA Text Export files.....	21
Part III Saving Tables	26
1 Saving Excel Tables.....	27
Part IV The Results	29
1 Initial Data-Check Table.....	30
2 Detailed Score Distribution.....	32
3 Lookup Tables.....	36
Part V The Stanscore Algorithms	38
1 Stanscore 1 and 2 algorithm.....	39
2 Stanscore 3 algorithm.....	40
3 Stanscore 4 algorithm.....	41
Part VI Definitions	42
1 normalised scores.....	43
2 percentiles.....	43
3 stanines.....	45
4 standardised scores.....	47
5 stens.....	47
6 T-scores.....	49
7 z-scores.....	50

Part



1 Introduction

1.1 New in Stanscore 5

1. This version now uses Excel 2007 and upwards format files for all input/output (.xlsx) files as the default. I've left the older format .xls file input/output as 'selectable', in order to maintain compatibility for those using .xls files.
2. As a result of #1, the **number of cases** can now extend to **n = 1,048,570** instead of the 65,530 limitation of the older .xls file-format.
3. The maximum **score range** for any scale is now dynamically allocated (in essence, an unlimited range).
4. The Microsoft Access input file routine only reads older .mdb format files - but I fixed the problem loading such files in the previous version.
5. The Raw data table output file and Norm-scored table output have now been nicely formatted in Excel (.xlsx) output.
6. Some other cosmetic changes were made.

1.2 New in Stanscore 4

2011 Version 4

Stanscore version 3, in line with many authors, computed percentile ranks assuming continuous scores. This used a formula that assumes a score is actually a mid-point between two integer scores. The upshot of this is that the percentiles created using this formula are usually (**but not always**) interpreted as the values "BELOW WHICH" n% of individuals score. Another way of putting this is to say person X with percentile rank of say 80 scores higher than 80% of the group.

The problem is this is just plain wrong for integer-scored data where percentiles need to be computed for each integer score of a psychological test. There is no continuity; there are just the numbers 0 to some maximum-possible score, with the intervals between scores as an indivisible single number of 1.

I tried to make some sense of the competing definitions in one of my Technical Whitepapers entitled "**Percentiles and Percentile Ranks - confused or what?**" which can be downloaded from: <http://www.pbarrett.net/#whitepapers>, Technical whitepaper #3. This contains about 15 different definitions of percentiles - and unravels the likely reason why the two classes of definitions differ from one another.

But, on reflection this week, I think all these authors' have not properly understood themselves how the calculation of the cumulative proportion used to construct the percentiles makes a mockery of both definitions, when dealing with unit-change test scores (i.e. scores which increase a unit at a time 0, 1, 2, 3, 4, 5, 6 etc). These scores are discrete; they are not real-valued, continuous numbers.

The "[continuous](#)" [assumption calculation](#) only uses half the frequencies at an observed score. So for the maximum possible score, only half the observed frequencies are used to express the final percentile. This is crazy. It is factually incorrect to state the percentile for this score is **the value below which n%** of the norm-group score, because half the people scoring the maximum observed score are included in the cumulative proportion. Alternatively, it is also wrong to state the percentile is the score **at or below which** n% of the group score, because half the people scoring at the percentile score are missing from the calculation.

So, Stanscore 4 now computes the cumulative frequency distribution from the observed frequencies at each integer score, assuming an integer score is a discrete entity (no continuity assumption). This distribution is what you find if you computed a standard frequency distribution for a set of scores in say SPSS, STATISTICA, or another stats package. The percentiles of this distribution range between a potential minimum value of 0 (if there are no observed frequencies at 0), and **ALWAYS** 100.

The interpretation of these percentiles is very simple, accurate, and very clear.

It is that that person X with percentile rank of say 80 scores **at or above** n% of the group.

Or

n% of the group score **at or below** a percentile rank of 80.

A person receiving the maximum score would always have a percentile rank of 100, which would mean they score at or above 100% of the norm group.

So, you'll now find that two frequency distributions for percentiles are now presented, one using the original Stanscore 3 continuous-score cumulative proportions (for legacy comparisons with "textbook solutions"), the other using the correct discrete score cumulative proportions.

Normalised Sten-Stanine tables are computed for both kinds of proportions, although you will see little difference between them as stens/stanines are based upon indexing proportions of scores within a score distribution ... but, I calculated them both here in order that those who choose discrete score frequencies might wish to remain consistent and use discrete-score proportion stens/stanines.

Sometimes, you will see huge discrepancies between percentiles computed using each approach - IF the data are severely right or left skewed. Within a "continuous" perspective, the frequencies are being truncated at a terminal value (the lowest or highest score); within the "discrete" perspective, no such truncation is possible - frequencies are "just as they come".

Anyway, providing the two "perspective" results has doubled the quantity of matrices output for printing, saving.

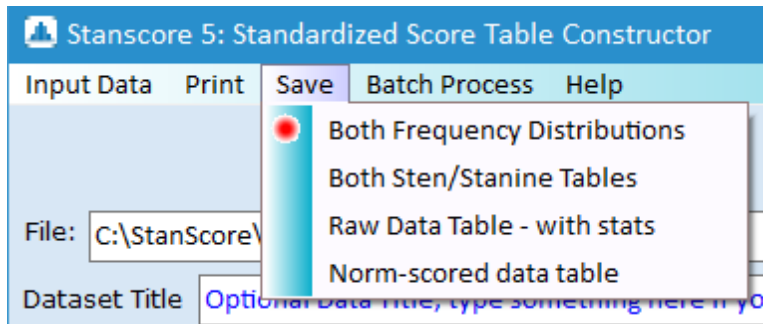
I've updated screen images and the relevant section help accordingly (e.g. [percentiles](#))

Specific Changes in Version 4.2 (December 30th, 2011)

The big update is #1 below. As an option, the program now produces normed scores for all raw scores from which the norms were created.

1. There is now a new option available to save the raw scores for all scales along with their norm-score counterparts (Percentiles, Stens, Stanines, and T-Scores) computed using the lookup-table for each scale. Two sets of tables are created as .xls (Office 2003 version) files, once using Continuous score assumption norms, the other using Discrete score norms.

This option is found under the Save menu option:



2. The program no longer uses a read/write Excel file component which relies upon a user having Excel installed on their computer. This speeds up reading and writing of such files considerably. There is no difference in file formats from previous versions of Stanscore.

3. Batch printing has been better automated with more helpful status messaging.

4. Batch output of norm tables into Excel Workbooks (one sheet per scale) now includes a separate summary percentile lookup sheet for all scales.

5. Single scale output files now contain the scale name as part of the Excel filename.

6. Individual help pages can now be printed using the print button located at the top right of the help



screen ...

Paul Barrett (2011) paul@pbarrett.net

1.3 StanScore 5 Basics and Information

StanScore 4 is a simple but useful utility program that converts one or more sets of raw test scores into five transformed scores:

[normalized-z-scores](#)

[stens](#)

[stanines](#)

[percentiles](#)

[T-scores](#)

The program accepts data in several convenient formats such as from Microsoft Excel Spreadsheets, Access database tables, and standard comma-separated free text files. This means that data contained in such statistical packages as SPSS, STATISTICA, SYSTAT etc. can be easily exported into say an Excel file, then read straight into Stanscore "as is" for processing.

Missing Data are recognized for each variable separately by the program - recognizing a -9999 code, a null field (two commas together in a comma-separated "[csv](#)" or "[txt](#)" file), or the usual blank entries in SPSS, STATISTICA, Excel, or ACCESS

StanScore currently calculates the 5 transformed scores in an on-screen, scrollable, detailed score distribution table (which can be saved as an Excel spreadsheet - see the [Detailed Score Distribution](#) help section). The detailed on-screen score histogram can be printed in tabular form, and saved as an Excel file for further formatting.

The program also provides summary raw-score to [stanine](#) and [sten](#) look-up tables that are displayed on-screen, can be printed in tabular form, and can be saved to an Excel file for other uses or re-formatting for presentation-documentation purposes.

The program accepts anywhere from 1 to 50 scale or test score variables. This means that using the "Print-All" facility, percentile and raw-score conversion lookup tables can be printed/saved for up to 50 variables as one command.

Very Important

Stanscore 5 requires that input data file contains only the variables to be analyzed by the routine. That is, if you have say an SPSS, Excel, or ACCESS file which contains scale scores as well as other variables, then you must create a subset dataset - which will then be exported into Excel and into Stanscore. I simply did not have the time to program a variable selection interface!

The PsWin files have a pre-specified structure - I've built this in automatically into the code so that importing of these files is totally automated (in line with the PsWin design). **[PB 2011]:** *the PsWin format is merely legacy value now - as the package is totally obsolete).*

Some Important Constraints for Version 5

- It requires the raw test scores/variables to be integer (whole number) and non-negative values.
- The maximum score range is between 0 and 2000.
- This version now uses Excel 2007 and upwards format files for all input/output (.xlsx) files as the default. I've left the older format .xls file input/output as 'selectable', in order to maintain compatibility for those using .xls files.
- The **number of cases** can now extend to **n = 1,048,570** instead of the 65,530 limitation of the older .xls file-format.

Computer Stuff

The code is written in Delphi Seattle - using double-precision for all real-valued variables. Help written using [Help&Manual](#) Professional version 7.4

Paul Barrett (2011) paul@pbarrett.net

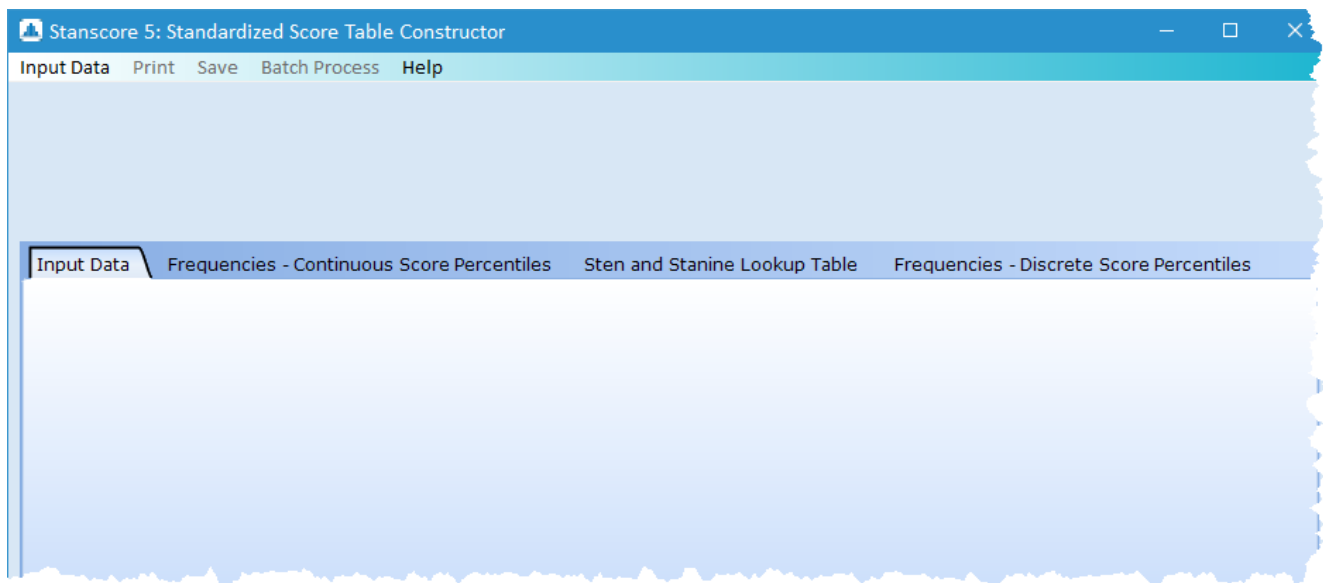
1.4 Using the program

Basically, the program functions in three steps:

1. Select a file from which to read in your data
2. Compute the Norms from the data
3. View/print the frequency table and norm look-up tables.
4. Save the Norm look-up tables in Excel that are ready for immediate use, as well as the input raw scores and their normed equivalents.

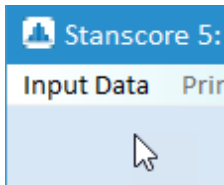
Step 1: Select a file from which to read in your data

When you start the program, you'll see the following screen ... the menu bar and the tabbed area below this - with the Input Data Tab selected (but nothing to show as yet).



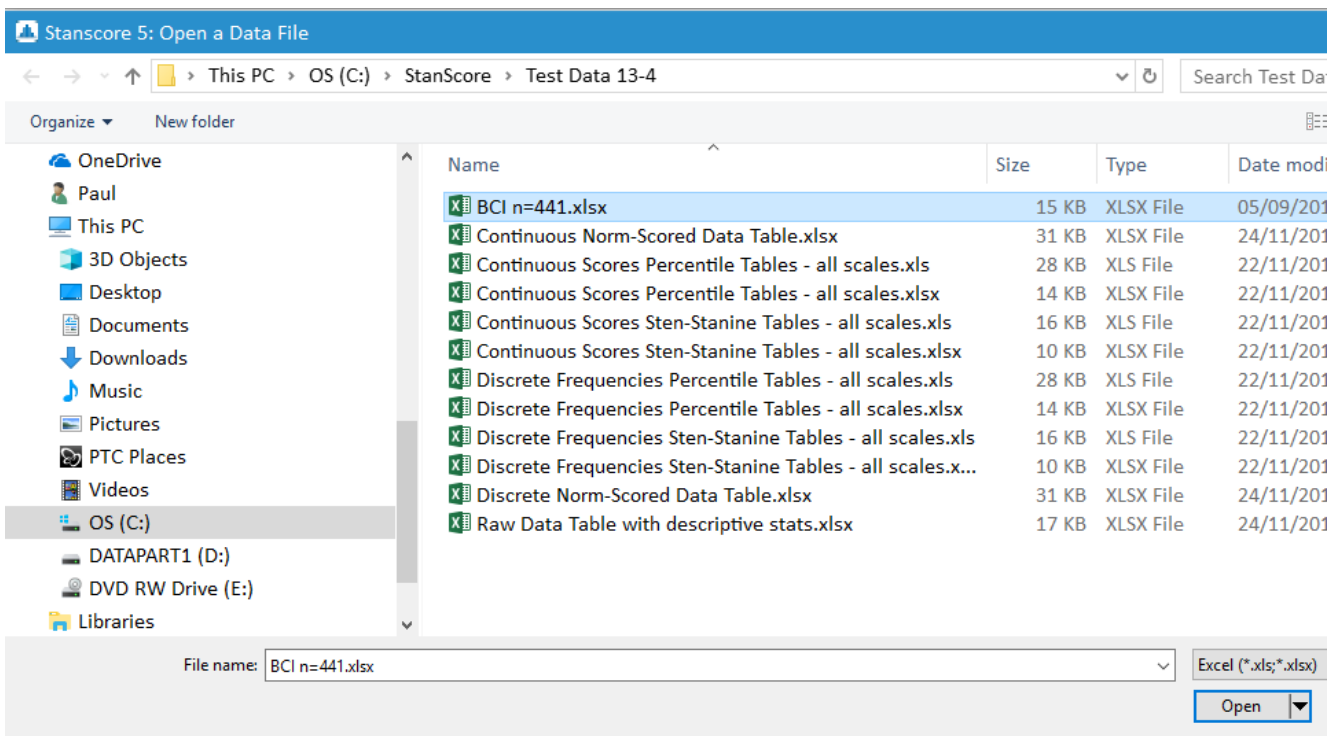
Only two menu options are available - Input Data and Help.

To begin, click on the menu option



Clicking this menu option displays a standard file selection dialog. The program will default to the subdirectory `C:\stanscore\input data files` when you first run it. From then on, it "remembers" the last subdirectory from which you opened a file and automatically opens the file selection dialog in that directory.

The program is defaulted to the Excel file type as its initial file-selection routine. However, clicking on the "Files of Type" field combobox, you can choose from all the available file-types:



So, let's say I wanted to construct norms from the file "test.xls", I would simply select the file by clicking on the filename (I'm using Windows 10) - which then immediately opens the file... displaying the data along with some summary statistical information. Note, I also display the missing cases for you - along with the counts., as well as an a summary indication of how many variables were found in the file, and the number of cases.

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

Compute Norm Tables

File: C:\StanScore\input data files\test.xls No. of Scales = 3

Dataset Title Optional Data Title, type something here if you wish Cases = 26

Input Data Frequencies - Continuous Score Percentiles Sten and Stanine Lookup Table Frequencies - Discrete Score Percentiles

Cases	Variable 1 Long Name	Variable 2	Variable 3
13	12	16	9
14	3	23	17
15	15	26	18
16	32	33	23
17	19	26	19
18	9	23	22
19	10	14	17
20	7	19	6
21	2	8	12
22	6	22	17
23	12	4	12
24	23	8	15
25	33	9	11
26	-9999	2	3
Valid N	24	26	26
Missing	2	0	0
Min Val	2	2	3
Max Val	34	33	35
Mean	12.33333	14.73077	16.03846
Std.Dev	9.90242	8.90194	8.41180

You can type in a new dataset title if you wish, overwriting the default blue text.

Note, a new button has appeared for you as well -

Compute Norm Tables

Click on this to compute all norm tables for each variable. The program displays these results in the three tabbed pages:

Input Data Frequencies - Continuous Score Percentiles Sten and Stanine Lookup Table Frequencies - Discrete Score Percentiles

For reasons given in the [Percentiles Definitions](#) help, use the percentiles, stens, and stanines from this table {Frequencies - Discrete Score Percentiles}.

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable Variable 1 Long Name

File: C:\StanScore\input data files\test.xls No. of Scales = 3

Dataset Title Optional Data Title, type something here if you wish Cases = 26

Raw Score	Frequency	Cumulative Freq.	Proportion	Cumulative Propn.	z-score	stanine	sten	T-Score	Percentile
0	0	0	0.000	0.00000	-4.00000	1	1	10	0
1	0	0	0.000	0.00000	-4.00000	1	1	10	0
2	2	2	0.083	0.08333	-1.38187	2	3	36	8
3	2	4	0.083	0.16667	-0.96515	3	4	40	17
4	0	4	0.000	0.16667	-0.96515	3	4	40	17
5	3	7	0.125	0.29167	-0.54574	4	4	45	29
6	1	8	0.042	0.33333	-0.42808	4	5	46	33
7	2	10	0.083	0.41667	-0.20861	5	5	48	42
8	1	11	0.042	0.45833	-0.10357	5	5	49	46
9	2	13	0.083	0.54167	0.10357	5	6	51	54
10	1	14	0.042	0.58333	0.20861	5	6	52	58
11	0	14	0.000	0.58333	0.20861	5	6	52	58
12	3	17	0.125	0.70833	0.54574	6	7	55	71
13	0	17	0.000	0.70833	0.54574	6	7	55	71
14	0	17	0.000	0.70833	0.54574	6	7	55	71
15	1	18	0.042	0.75000	0.67173	6	7	57	75
16	0	18	0.000	0.75000	0.67173	6	7	57	75
17	0	18	0.000	0.75000	0.67173	6	7	57	75
18	0	18	0.000	0.75000	0.67173	6	7	57	75
19	1	19	0.042	0.79167	0.80963	7	7	58	79
20	0	19	0.000	0.79167	0.80963	7	7	58	79
21	0	19	0.000	0.79167	0.80963	7	7	58	79
22	0	19	0.000	0.79167	0.80963	7	7	58	79

For reasons given in the [Percentiles Definitions](#) help, DO NOT use the percentiles, stens, and stanines from the below table {Frequencies, Continuous Score Percentiles}. They are reported here as "textbook" values, with really no more value than that.

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable Variable 1 Long Name

File: C:\StanScore\input data files\test.xls No. of Scales = 3

Dataset Title Optional Data Title, type something here if you wish Cases = 26

Input Data Frequencies - Continuous Score Percentiles Sten and Stanine Lookup Table Frequencies - Discrete Score Percentiles

Raw Score	Frequency	Cumulative Freq	Percentile z-score	Exact Proportion	Exact z-score	stanine	sten	T-Score	Percentile
0	0	0	-4.000	0.00000	-4.00000	1	1	10	0
1	0	0	-4.000	0.00000	-4.00000	1	1	10	0
2	2	2	-1.751	0.04167	-1.73155	2	2	33	4
3	2	4	-1.173	0.12500	-1.14855	3	3	39	12
4	0	4	-0.952	0.16667	-0.96515	3	4	40	17
5	3	7	-0.736	0.22917	-0.73890	4	4	43	23
6	1	8	-0.493	0.31250	-0.48604	4	5	45	31
7	2	10	-0.303	0.37500	-0.31630	4	5	47	38
8	1	11	-0.150	0.43750	-0.15584	5	5	48	44
9	2	13	0.000	0.50000	0.00000	5	6	50	50
10	1	14	0.150	0.56250	0.15584	5	6	52	56
11	0	14	0.200	0.58333	0.20861	5	6	52	58
12	3	17	0.383	0.64583	0.37158	6	6	54	65
13	0	17	0.551	0.70833	0.54574	6	7	55	71
14	0	17	0.551	0.70833	0.54574	6	7	55	71
15	1	18	0.610	0.72917	0.60751	6	7	56	73
16	0	18	0.672	0.75000	0.67173	6	7	57	75
17	0	18	0.672	0.75000	0.67173	6	7	57	75
18	0	18	0.672	0.75000	0.67173	6	7	57	75
19	1	19	0.736	0.77083	0.73890	6	7	57	77
20	0	19	0.804	0.79167	0.80963	7	7	58	79
21	0	19	0.804	0.79167	0.80963	7	7	58	79
22	0	19	0.804	0.79167	0.80963	7	7	58	79

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable Variable 1 Long Name

File: C:\StanScore\input data files\test.xls No. of Scales = 3

Dataset Title Optional Data Title, type something here if you wish Cases = 26

Input Data Frequencies - Continuous Score Percentiles **Sten and Stanine Lookup Table** Frequencies - Discrete Score Percentiles

Raw Score Range	Stanine Equivalent	---	Raw Score Range	Sten Equivalent
0 to 1	1	---	0 to 1	1
2 to 2	2	---	2 to 2	2
3 to 4	3	---	3 to 3	3
5 to 7	4	---	4 to 5	4
8 to 11	5	---	6 to 8	5
12 to 19	6	---	9 to 12	6
20 to 31	7	---	13 to 23	7
32 to 33	8	---	24 to 32	8
34 to 34	9	---	33 to 33	9
		---	34 to 34	10

Variable 1 Long Name

Lookup Table
Stens and Stanines computed using the Continuous Score Percentile Rank Formula

Raw Score Range	Stanine Equivalent	---	Raw Score Range	Sten Equivalent
0 to 1	1	---	0 to 1	1
2 to 2	2	---	2 to 2	3
3 to 4	3	---	3 to 5	4
5 to 6	4	---	6 to 8	5
7 to 11	5	---	9 to 11	6
12 to 18	6	---	12 to 22	7
19 to 31	7	---	23 to 32	8
32 to 33	8	---	33 to 33	9
34 to 34	9	---	34 to 34	10

Lookup Table
Stens and Stanines computed using the Discrete Score Frequencies

If you choose another variable using the variable selection combobox "View Results for which variable", the results are automatically updated to reflect the chosen variable in both tabbed areas.

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable Variable 1 Long Name

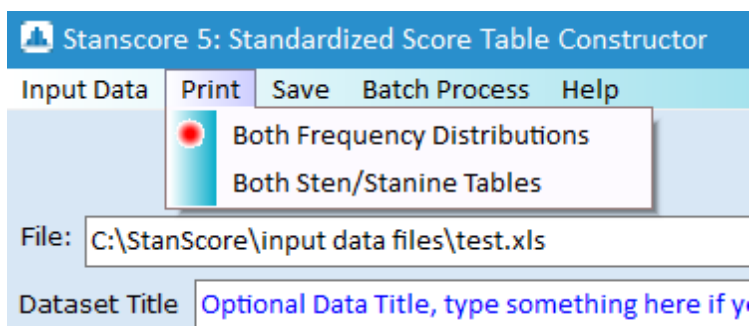
File: C:\StanScore\input data files\test.xls

Dataset Title Optional Data Title, type something here if you wish Cases = 26

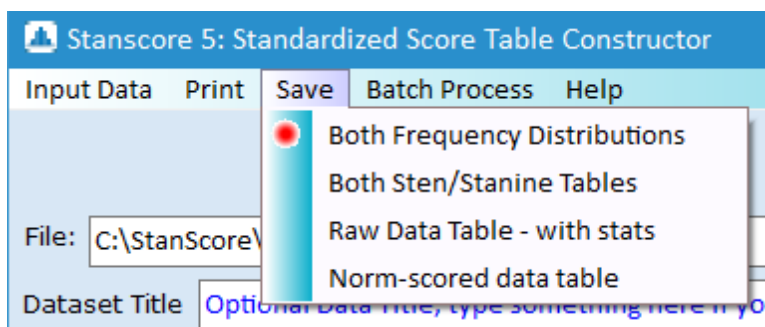
Input Data Frequencies - Continuous Score Percentiles **Sten and Stanine Lookup Table** Frequencies - Discrete Score Percentiles

Variable 1 Long Name
Variable 2
Variable 3

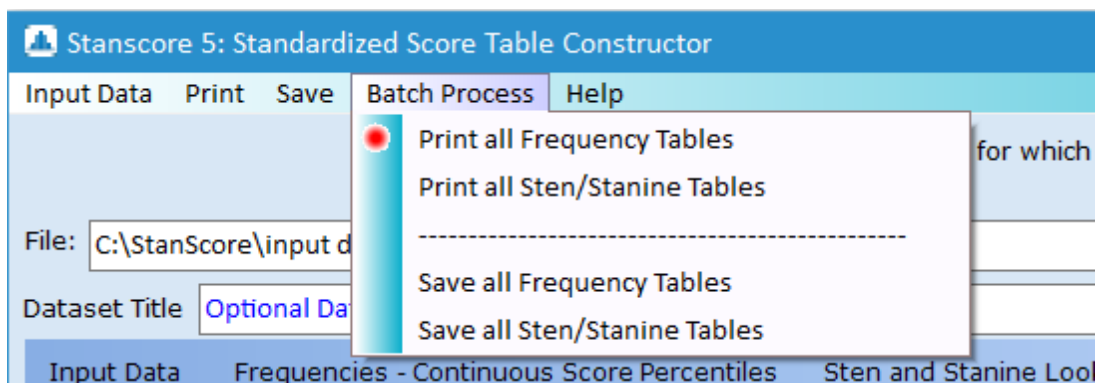
You can print any of the displayed results tables using the menu option "Print" ...



and save them as excel files using the menu option "Save":

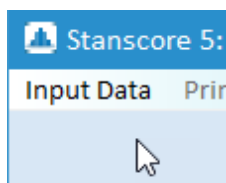


or print/save all variable tables with one menu option choice using the "Batch" menu option:



Note, the excel files are saved as a multiple-tabbed-sheet file. See the section entitled [Saving Excel Tables](#).

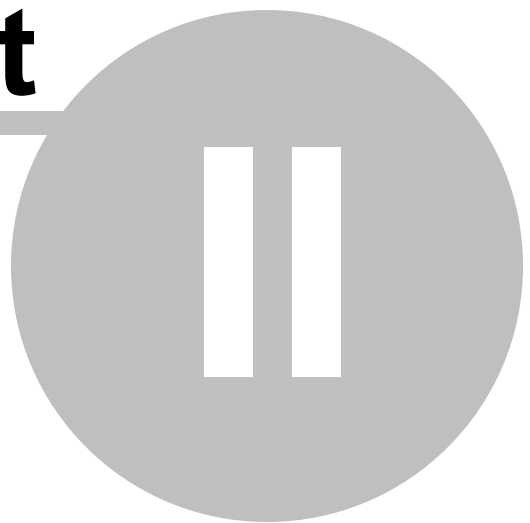
If you wish to open a new file, you don't have to close the program, just click on the menu option "Input Data" and you will be prompted for a new filename ...



OK - that's all there is to the program ... enjoy! If you are thinking of using a file format other than Excel, make sure you take a quick look at each of the file format help sections under the chapter title "Getting Data into the Program".

Paul Barrett (2011) paul@pbarrett.net

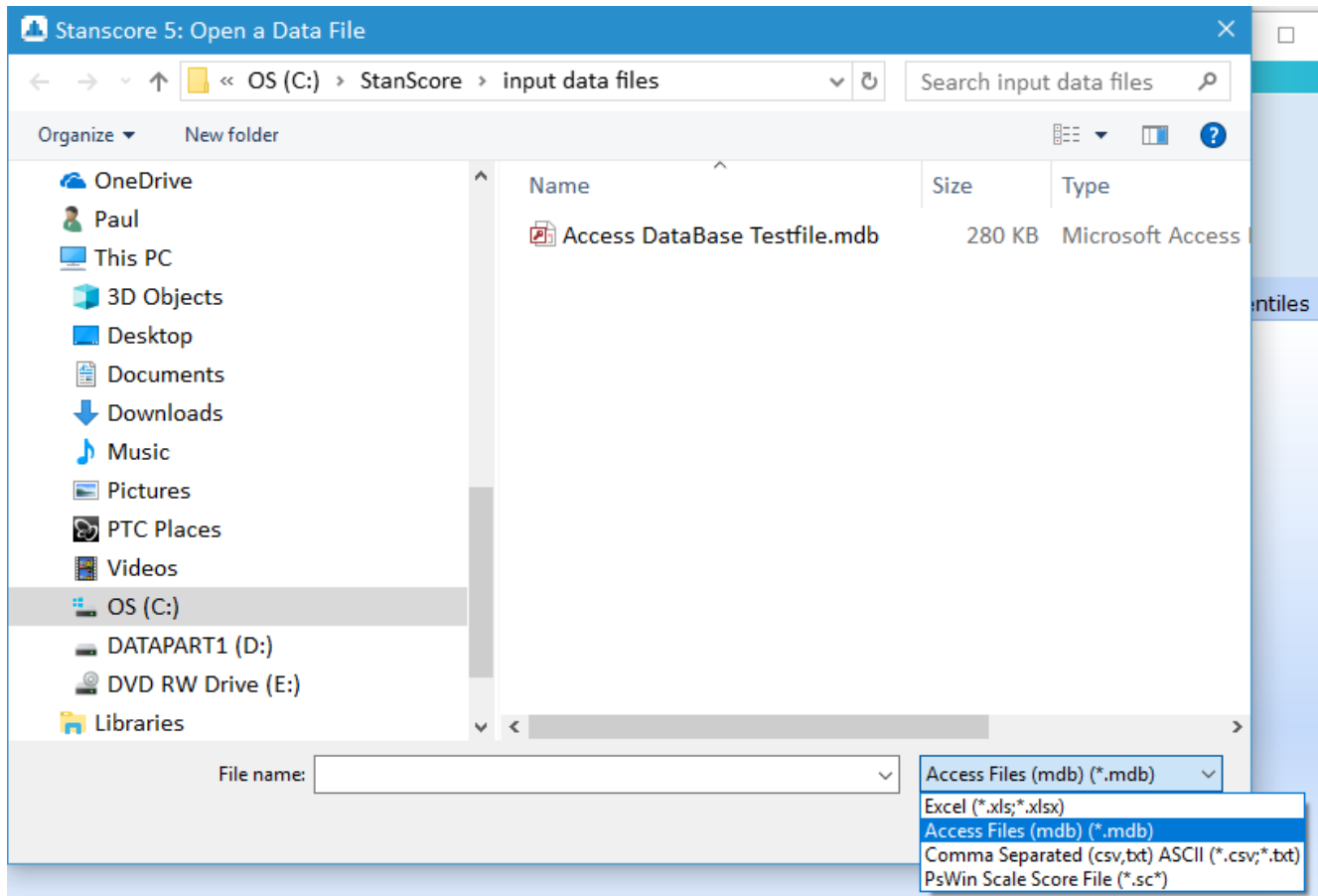
Part



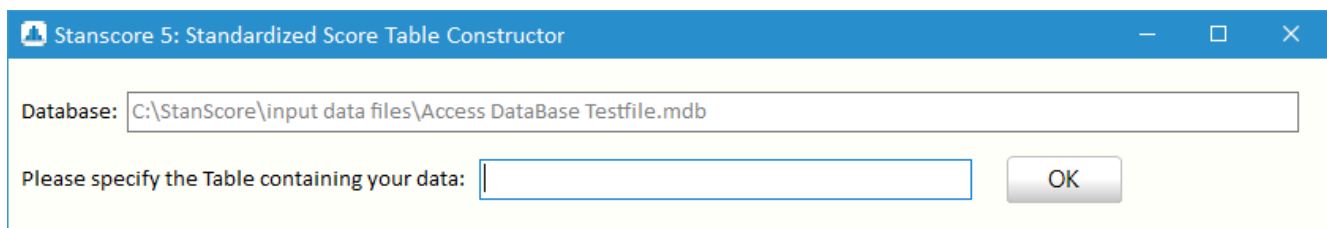
2 Getting Data into the Program

2.1 Access Database Tables

When reading in an ACCESS file, you not only have to know the ACCESS database name, but also the **Table name** in which the data are present. Remember, the only variables to appear in that table are those for which you wish to create norm tables. So - the input procedure is:



Selecting the file and clicking on open shows another window, which asks for the Table you wish to input into the routine..



Whatever name you have given the table in your database can be input here. Then, things proceed as outlined in the section entitled "[Using the Program](#)"

IMPORTANT

Currently, only older .mdb format Access files can be opened with Stanscore.

Paul Barrett (2011) paul@pbarrett.net

2.2 ASCII Files

Here you can read in comma-separated textfiles - that use the fixed suffixes **.csv** and **.txt**. These are the default filenames created by Excel and by other programs such as SPSS and STATISTICA. The file format should be like this ...

File = Test.csv

```
Variable1,Variable2,Variable3
,2,14
2,33,5
3,12,22
5,13,13
5,12,15
5,10,12
7,19,17
8,9,33
9,8,34
34,3,35
23,12,7
12,17,9
12,16,9
3,23,17
15,26,18
32,33,23
19,26,19
9,23,22
10,14,17
7,19,6
2,8,12
6,22,17
12,4,12
23,8,15
33,9,11
,2,3
```

The three variable names appear on the first line, followed by the data, case by case. Note the missing item for case 1, variable 1, and the same for the last case. However, unless you are constructing files like this yourself, just leave it to the relevant export program to set files up like this.

When exporting ASCII text from a program, make sure you request variable names on the first line of the file, and a **comma** as the data separator character.

Paul Barrett (2011) paul@pbarrett.net

2.3 Excel Files

Well, nothing to it. Just provide the filename and off you go!

The program assumes that the data will be in the first sheet of the Excel table. If you are exporting data as an Excel file from some other program (say SPSS), make sure you request variable names on the first line of the file.

An Excel input file looks like:

	A	B	C
1	Psychological Health	Job Satisfaction	Workload
2	40	35	30
3	58	63	36
4	54	58	23
5	61	71	16
6	48	63	35
7	54	42	34
8	68	49	31
9	66	20	20
10	55	53	26
11	36	36	19
12	59	50	28
13	38	14	33
14	59	62	42
15	68	62	34
16	37	64	25
17	64	71	27
18	59	56	25
19	60	67	35
20	45	63	39
21	55	44	33
22	44	62	47
23	47	59	38
24	39	47	43
25	37	49	36
26	48	67	34
27	68	66	29
28	54	33	39

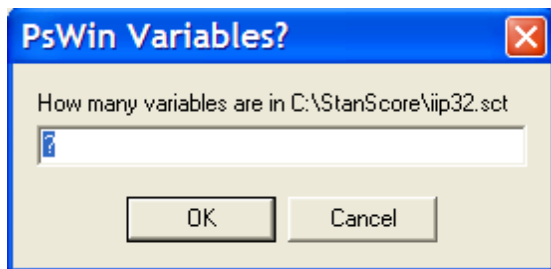
IMPORTANT

Stanscore 5 now uses Excel 2007 and upwards format files for all input/output (**.xlsx**) files as the default. I've left the older format .xls file input/output as 'selectable', in order to maintain compatibility for those using .xls files.

Paul Barrett (2011) paul@pbarrett.net

2.4 PsWin v2 & above Scale Score files

This utility is specifically for PsWin scale score files - which have a special format. The only special thing about reading in such files is that you will be prompted for the number of variables to read from the file. E.g. ...



PsWin scale score files all end with the suffixesscl, .sct, .scf, .scm. These are recognized automatically by the software,

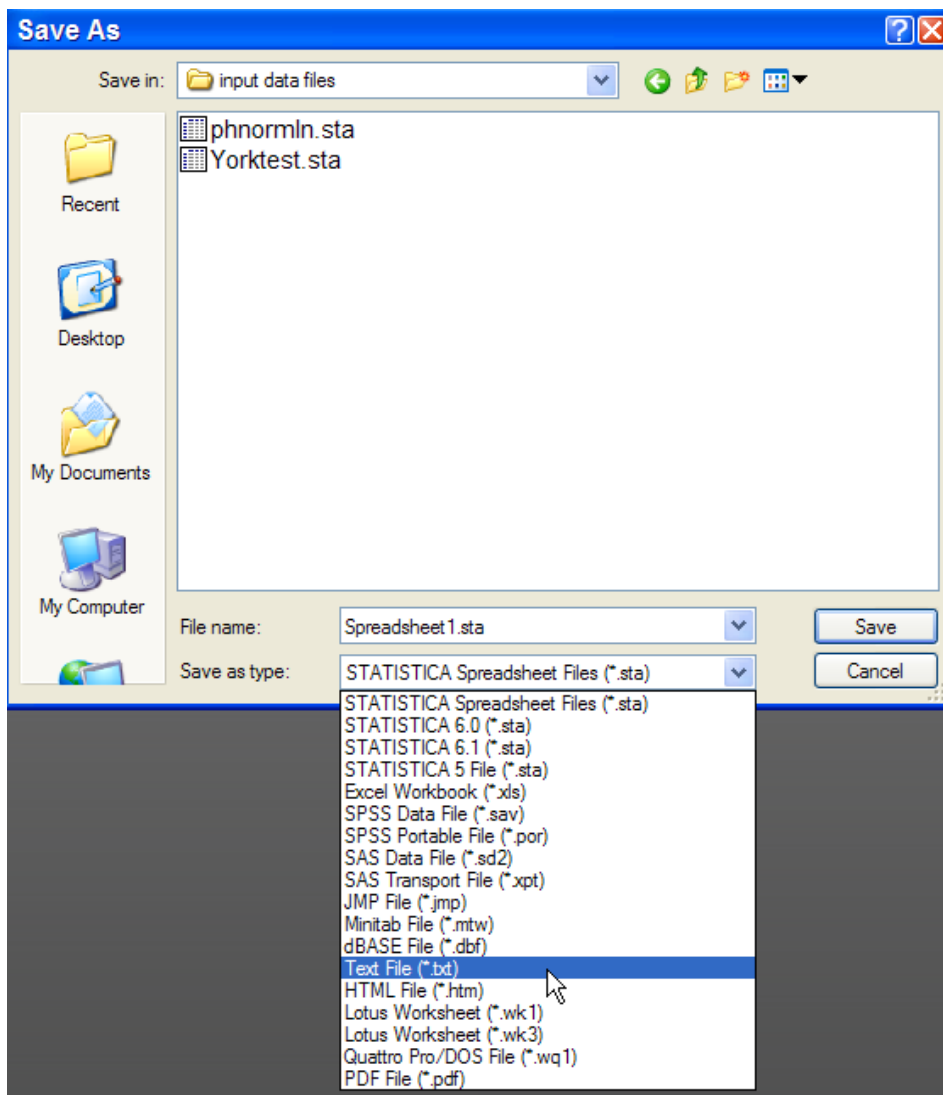
However, this input type is merely legacy now as the PsWin package is obsolete.

Paul Barrett (2011) paul@pbarrett.net

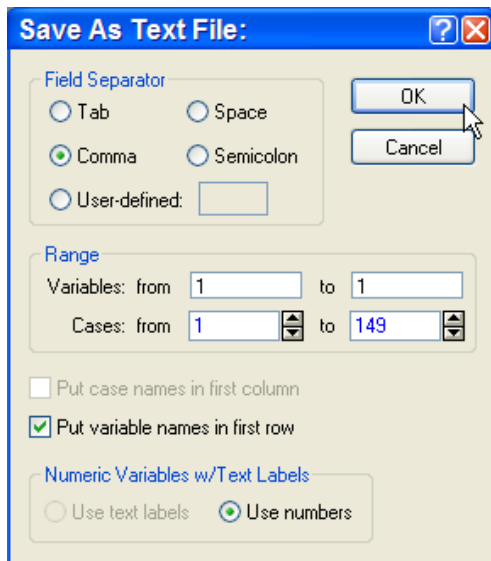
2.5 STATISTICA Text Export files

This is no different from the [ASCII-csv](#) file input instructions - except that the older Statistica export text files could end in in .fre. Really, this is just for backward compatibility with version 5 STATISTICA free format text output files.

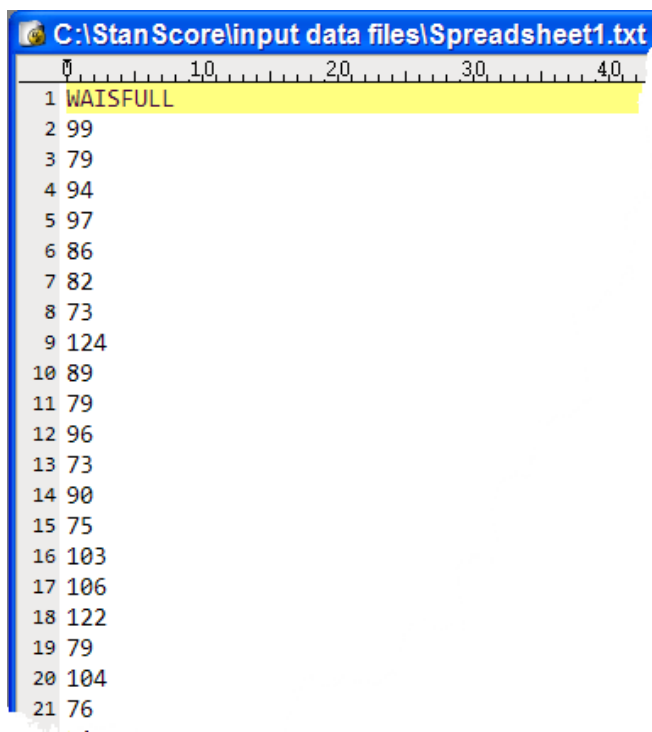
Saving a STATISTICA 9 file as a text file - ready for input into Stanscore looks like this:



clicking on Save takes you to an options screen:



Note, I have selected "comma" as the Field Separator, and ticked the box "Put variable names in first row". I click on OK ... the file created - Spreadsheet1.Txt looks like:



Which reads directly into Stanscore "as is" using the ASCII text file input option ...

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

Compute Norm Tables

File: C:\StanScore\input data files\Spreadsheet1.txt No. of Scales = 1

Dataset Title: Optional Data Title, type something here if you wish Cases = 149

Input Data Frequencies - Continuous Score Percentiles Sten and Stanine Lookup Table Frequencies - Discrete Score Percentiles

135	102
136	118
137	132
138	111
139	112
140	125
141	123
142	72
143	73
144	134
145	98
146	91
147	83
148	91
149	78
Valid N	149
Missing	0
Min Val	61
Max Val	134
Mean	103.97987
Std.Dev	16.79787

and norm table:

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable: WAISFULL

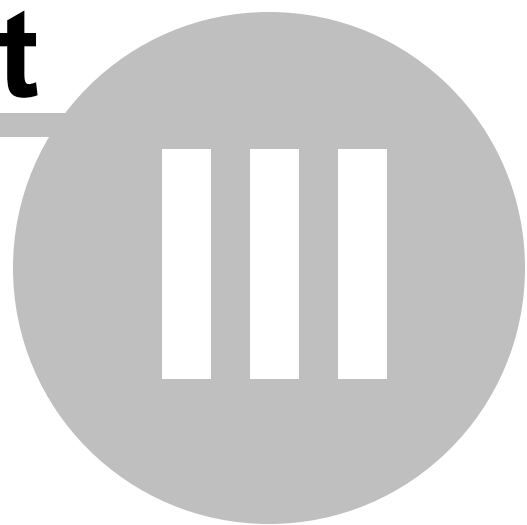
File: C:\StanScore\input data files\Spreadsheet1.txt No. of Scales = 1

Dataset Title: Optional Data Title, type something here if you wish Cases = 149

Raw Score	Frequency	Cumulative Freq.	Proportion	Cumulative Propn.	z-score	stanine	sten	T-Score	Percentile
115	2	104	0.013	0.69799	0.51585	6	7	55	70
116	5	109	0.034	0.73154	0.61470	6	7	56	73
117	3	112	0.020	0.75168	0.67702	6	7	57	75
118	5	117	0.034	0.78523	0.78737	7	7	58	79
119	4	121	0.027	0.81208	0.88314	7	7	59	81
120	0	121	0.000	0.81208	0.88314	7	7	59	81
121	0	121	0.000	0.81208	0.88314	7	7	59	81
122	5	126	0.034	0.84564	1.01575	7	8	60	85
123	4	130	0.027	0.87248	1.13637	7	8	61	87
124	3	133	0.020	0.89262	1.23902	7	8	62	89
125	3	136	0.020	0.91275	1.35670	8	8	64	91
126	1	137	0.007	0.91946	1.40040	8	8	64	92
127	1	138	0.007	0.92617	1.44694	8	8	64	93
128	1	139	0.007	0.93289	1.49685	8	8	65	93
129	0	139	0.000	0.93289	1.49685	8	8	65	93
130	2	141	0.013	0.94631	1.60961	8	9	66	95
131	1	142	0.007	0.95302	1.67459	8	9	67	95
132	3	145	0.020	0.97315	1.92973	9	9	69	97
133	1	146	0.007	0.97987	2.05169	9	10	71	98
134	3	149	0.020	1.00000	4.00000	9	10	90	100
Total Freq	149								

Paul Barrett (2011) paul@pbarrett.net

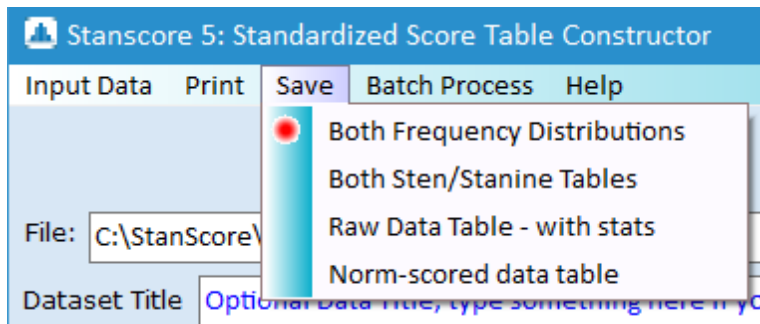
Part



3 Saving Tables

3.1 Saving Excel Tables

The results tables, normed raw score tables, and data table with descriptive statistics can be saved as Excel files.



Stanscore 5 now uses Excel 2007 and upwards format files for all input/output (.xlsx) files as the default. I've left the older format .xls file input/output as 'selectable', in order to maintain compatibility for those using .xls files.

All Save options produce a Continuous-score and Discrete score normative table. Default filenames are allocated but you can choose to overwrite or call them any another convenient filenames before saving.

The **Norm-scored data table** option saves the raw data, alongside all normalised score equivalents. The Excel files (for continuous and discrete norms) look like:

	A	B	C	D	E
1	Rule Applying_Score	Rule Applying_D_Percentile	Rule Applying_D_Sten	Rule Applying_D_Stanine	Rule Applying_D_T-Score
2	7	69	6	6	55
3	8	87	8	7	61
4	8	87	8	7	61
5	8	87	8	7	61
6	6	50	6	5	50
7	7	69	6	6	55
8	9	96	9	9	68
9	6	50	6	5	50
10	5	34	5	4	46
11	6	50	6	5	50
12	7	69	6	6	55
13	7	69	6	6	55
14	8	87	8	7	61
15	9	96	9	9	68

Note that for multiscale tests, all scales are listed block-by-block in the file. Converting the file above (which contains the data for three scales) into a Statistica file, we can then more easily view the full variable list for the file:

Normed raw data save-file va	
1	
Variable Name	
1	Rule Applying_Score
2	Rule Applying_D_Percentile
3	Rule Applying_D_Sten
4	Rule Applying_D_Stanine
5	Rule Applying_D_T-Score
6	Rule Finding_Score
7	Rule Finding_D_Percentile
8	Rule Finding_D_Sten
9	Rule Finding_D_Stanine
10	Rule Finding_D_T-Score
11	Total Score_Score
12	Total Score_D_Percentile
13	Total Score_D_Sten
14	Total Score_D_Stanine
15	Total Score_D_T-Score

Paul Barrett (2011) paul@pbarrett.net

Part

IV

4 The Results

4.1 Initial Data-Check Table

This is the table that simple displays the imported scale score data - along with some summary statistics for each variable ...

For example, when using [Spreadsheet1.txt](#) (included with the program setup - in [c:\stanscore\input data files\](#) .. the data table shows these summary statistics for the variable WAISFULL ...

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable **WAISFU.**

File:

Dataset Title

Input Data Frequencies - Continuous Score Percentiles Sten and Stanine Lookup Table Fre

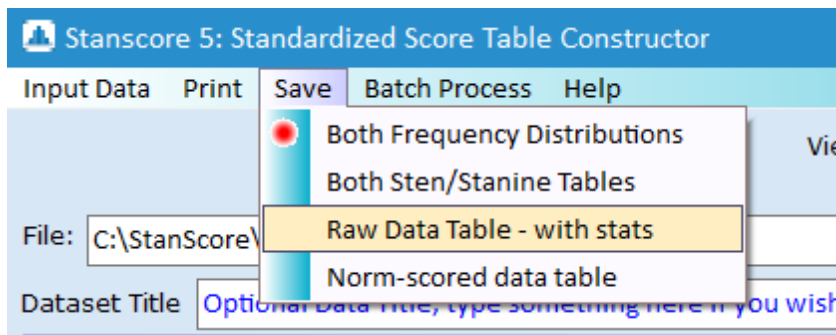
135	102
136	118
137	132
138	111
139	112
140	125
141	123
142	72
143	73
144	134
145	98
146	91
147	83
148	91
149	78
Valid N	149
Missing	0
Min Val	61
Max Val	134
Mean	103.97987
Std.Dev	16.79787

You can scroll through the variables using the scroll-bars that will appear on the table if there are more variables than can be fit into the available table area. The standard deviation calculation is made using the "sample" formula (dividing the squared deviations from the mean by N-1 instead of N).

You can save this displayed table as an excel spreadsheet (with headers) by accessing the Save menu option, then "Raw Data Table" once a norm table has been generated using the

Compute Norm Tables

button.



Paul Barrett (2011) paul@pbarrett.net

4.2 Detailed Score Distribution

This is the listing of all possible scores along with the relevant frequencies, proportions and z-values which are used to produce the final transformed scale scores ... it corresponds to the "textbook" percentile formula which assumes a continuous, real-valued score range underlies the actual integer scores in any test. **I no longer recommend it is used for test-score ranges which Stanscore works with (integer, unit-intervals).**

Continuous Score Formula Percentiles

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable Variable 1 Long Name

File: C:\StanScore\input data files\test.xls No. of Scales = 3

Dataset Title Optional Data Title, type something here if you wish Cases = 26

Input Data Frequencies - Continuous Score Percentiles Sten and Stanine Lookup Table Frequencies - Discrete Score Percentiles

Raw Score	Frequency	Cumulative Freq	Percentile z-score	Exact Proportion	Exact z-score	stanine	sten	T-Score	Percentile
0	0	0	-4.000	0.00000	-4.00000	1	1	10	0
1	0	0	-4.000	0.00000	-4.00000	1	1	10	0
2	2	2	-1.751	0.04167	-1.73155	2	2	33	4
3	2	4	-1.173	0.12500	-1.14855	3	3	39	12
4	0	4	-0.952	0.16667	-0.96515	3	4	40	17
5	3	7	-0.736	0.22917	-0.73890	4	4	43	23
6	1	8	-0.493	0.31250	-0.48604	4	5	45	31
7	2	10	-0.303	0.37500	-0.31630	4	5	47	38
8	1	11	-0.150	0.43750	-0.15584	5	5	48	44
9	2	13	0.000	0.50000	0.00000	5	6	50	50
10	1	14	0.150	0.56250	0.15584	5	6	52	56
11	0	14	0.200	0.58333	0.20861	5	6	52	58
12	3	17	0.383	0.64583	0.37158	6	6	54	65
13	0	17	0.551	0.70833	0.54574	6	7	55	71
14	0	17	0.551	0.70833	0.54574	6	7	55	71
15	1	18	0.610	0.72917	0.60751	6	7	56	73
16	0	18	0.672	0.75000	0.67173	6	7	57	75
17	0	18	0.672	0.75000	0.67173	6	7	57	75
18	0	18	0.672	0.75000	0.67173	6	7	57	75
19	1	19	0.736	0.77083	0.73890	6	7	57	77
20	0	19	0.804	0.79167	0.80963	7	7	58	79
21	0	19	0.804	0.79167	0.80963	7	7	58	79
22	0	19	0.804	0.79167	0.80963	7	7	58	79

The columns are defined in detail as:

Raw Score: the range (in unit steps) of all possible scores between an assumed minimum value of 0 and a maximum value observed in the data provided to the program.

Note, the program has no concept of what an absolute maximum might be for any scale - it can only work with the data presented to it.

Frequency: the number of cases with a particular score

Cumulative Freq: the cumulative frequency at each score level. This is computed as the sum of the frequencies observed **up to and including the "current score"**. For example, see the image just above. In Stanscore 3, I now follow the method used by SPSS, SYSTAT, and STATISTICA. In Stanscore 1 and 2, these frequencies were computed using frequency values up to but NOT including the current score, following the transformed score algorithm of Kline (1986). Specific details about how each produced transformed scores, and the differences between [Stanscore 1 and 2](#), and [Stanscore 3](#) are given in the help chapter entitled "The Stanscore Algorithms".

Percentile z-score: This is the z-score computed from the integer-valued percentile (the last column of the table). This is how Crocker and Algina (Crocker, L., & Algina, J. (1986) Introduction to Classical and Modern Test Theory. Harcourt Brace Jovanovich Publishers. ISBN: 0-03-061634-4, p. 439-440, see especially Table 19.2) compute their z-scores. I've included their data as an example file - the results from Stanscore are:

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable: Raw Score

File: C:\StanScore\input data files\Crocker_Algina_Table19-2_p440.xls

No. of Scales = 1

Dataset Title: Optional Data Title, type something here if you wish

Cases = 150

Raw Score	Frequency	Cumulative Freq	Percentile z-score	Exact Proportion	Exact z-score	stanine	sten	T-Score	Percentile
5	0	0	-4.000	0.00000	-4.00000	1	1	10	0
6	0	0	-4.000	0.00000	-4.00000	1	1	10	0
7	0	0	-4.000	0.00000	-4.00000	1	1	10	0
8	0	0	-4.000	0.00000	-4.00000	1	1	10	0
9	0	0	-4.000	0.00000	-4.00000	1	1	10	0
10	0	0	-4.000	0.00000	-4.00000	1	1	10	0
11	2	2	-2.328	0.00667	-2.47632	1	1	25	1
12	1	3	-2.054	0.01667	-2.12893	1	1	29	2
13	6	9	-1.751	0.04000	-1.75062	1	2	32	4
14	5	14	-1.404	0.07667	-1.42686	2	3	36	8
15	12	26	-1.125	0.13333	-1.10886	3	3	39	13
16	17	43	-0.736	0.23000	-0.73615	4	4	43	23
17	21	64	-0.356	0.35667	-0.36488	4	5	46	36
18	28	92	0.050	0.52000	0.04960	5	6	50	52
19	19	111	0.465	0.67667	0.45570	6	6	55	68
20	15	126	0.804	0.79000	0.80382	7	7	58	79
21	10	136	1.125	0.87333	1.14047	7	8	61	87
22	5	141	1.404	0.92333	1.42686	8	8	64	92
23	3	144	1.644	0.95000	1.64449	8	9	66	95
24	4	148	1.881	0.97333	1.93263	9	9	69	97
25	2	150	2.328	0.99333	2.47632	9	10	75	99
Total Freq	150								

which agree to within rounding error to the printed example. However, it's a moot point whether to avoid utilizing real-valued percentiles as the "proportion" of the normal curve from which the z-score is computed. So, in Stanscore 3, I also showed the Exact Proportion as well as the Exact z-score, using these values to construct the integer stanine, sten, T-scores, and percentiles.

Exact Proportion: this is the cumulative proportion of the scores, computed using the cumulative frequency for the score immediately below the score of interest added to one half of the frequencies at the current score, all divided by the total number of cases (total frequency). See the section [Stanscore 3 algorithm](#) for specific computational details. For a deeper understanding of percentiles and percentile ranks, see the document on my website entitled "[Percentiles and Percentile Ranks - confused or what?](#)" which can be downloaded from: <https://www.pbarrett.net/techpapers.html> ... #3 in the series. For very specific information on how to interpret percentiles, take a look at the [Percentiles definitions page](#).

Exact z-score: This is the **standardized z-score for the Exact Proportion**, with reference to a standard normal distribution. That is, we calculate the exact z-score which subtends the observed proportion of area under the curve within a normal distribution. In effect, an "inverse-normal" distribution calculation.

The Stanine, Sten, T-scores, and Percentiles: just that - the transformed raw scores expressed in integer [sten](#), [stanine](#), [T-score](#), and [Percentile](#) form.

Discrete Score Frequency Distribution Percentiles

The Discrete Frequencies table for the same Crocker and Algina table data shows:

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable: Raw Score

File: C:\StanScore\input data files\Crocker_Algina_Table19-2_p440.xls

Dataset Title: Optional Data Title, type something here if you wish

No. of Scales = 1

Cases = 150

Raw Score	Frequency	Cumulative Freq.	Proportion	Cumulative Propn.	z-score	stanine	sten	T-Score	Percentile
5	0	0	0.000	0.00000	-4.00000	1	1	10	0
6	0	0	0.000	0.00000	-4.00000	1	1	10	0
7	0	0	0.000	0.00000	-4.00000	1	1	10	0
8	0	0	0.000	0.00000	-4.00000	1	1	10	0
9	0	0	0.000	0.00000	-4.00000	1	1	10	0
10	0	0	0.000	0.00000	-4.00000	1	1	10	0
11	2	2	0.013	0.01333	-2.21744	1	1	28	1
12	1	3	0.007	0.02000	-2.05446	1	1	29	2
13	6	9	0.040	0.06000	-1.55415	2	2	34	6
14	5	14	0.033	0.09333	-1.31919	2	3	37	9
15	12	26	0.080	0.17333	-0.93874	3	4	41	17
16	17	43	0.113	0.28667	-0.56036	4	4	44	29
17	21	64	0.140	0.42667	-0.18320	5	5	48	43
18	28	92	0.187	0.61333	0.28580	6	6	53	61
19	19	111	0.127	0.74000	0.64057	6	7	56	74
20	15	126	0.100	0.84000	0.99225	7	7	60	84
21	10	136	0.067	0.90667	1.31919	8	8	63	91
22	5	141	0.033	0.94000	1.55415	8	9	66	94
23	3	144	0.020	0.96000	1.75062	9	9	68	96
24	4	148	0.027	0.98667	2.21744	9	10	72	99
25	2	150	0.013	1.00000	4.00000	9	10	90	100
Total Freq	150								

Raw Score: the range (in unit steps) of all possible scores between an assumed minimum value of 0 and a maximum value observed in the data provided to the program.

Note, the program has no concept of what an absolute maximum might be for any scale - it can only work with the data presented to it.

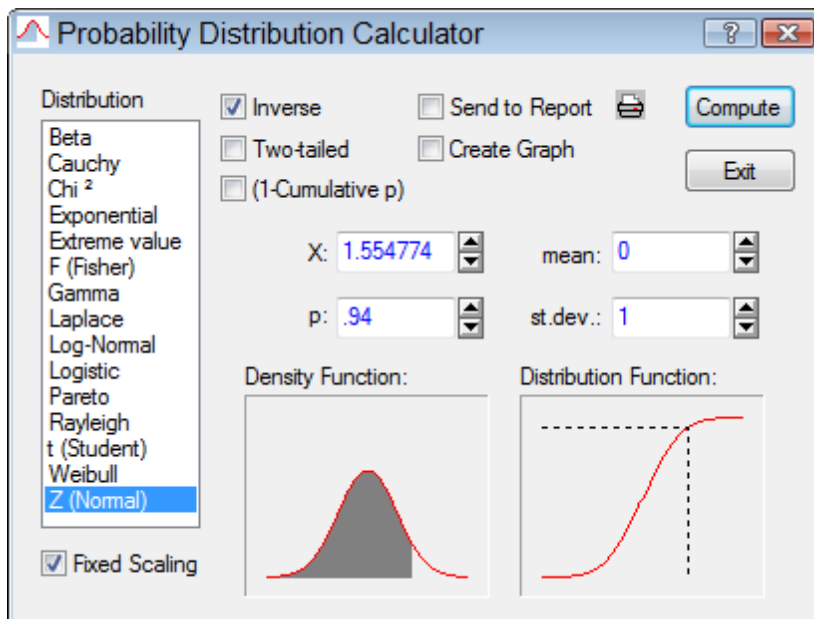
Frequency: the number of cases with a particular score

Cumulative Freq: the cumulative frequency at each score level. This is computed as the sum of the frequencies observed **up to and including the "current score"**.

Proportion: the proportion of frequencies for a particular score (score frequency divided by the Total Frequencies).

Cumulative Propn.: the cumulative proportion of frequencies ... which will always add up to 1.0 in this table.

z-score: the normal distribution z-value associated with the cumulative proportion. For example, the z-score value for a raw score of 22 in the above table is computed as 1.55415. If I use the Statistica 10 probability function and it's inverse-z algorithm, I obtain:



Which indicates reasonably close "good enough for purpose" agreement between the algorithm I'm using and that in Statistica 10.

The Stanine, Sten, T-scores, and Percentiles: just that - the transformed raw scores expressed in integer [sten](#), [stanine](#), [T-score](#) form, and [Discrete-Formula Percentile](#) form (this is nothing more than the rounded integer (**Cumulative Propn.** value x 100.0)!

Paul Barrett (2011) paul@pbarrett.net

4.3 Lookup Tables

This is the whole point of the program - the creation of a simple raw-score to normative score lookup table for [stens](#) and [stanines](#).

Stanscore 5: Standardized Score Table Constructor

Input Data Print Save Batch Process Help

View Results for which Variable: Raw Score

File: C:\StanScore\input data files\Crocker_Algina_Table19-2_p440.xls No. of Scales = 1

Dataset Title: Optional Data Title, type something here if you wish Cases = 150

Input Data Frequencies - Continuous Score Percentiles **Sten and Stanine Lookup Table** Frequencies - Discrete Score Percentiles

Raw Score Range	Stanine Equivalent	---	Raw Score Range	Sten Equivalent
0 to 13	1	---	0 to 12	1
14 to 14	2	---	13 to 13	2
15 to 15	3	---	14 to 15	3
16 to 17	4	---	16 to 16	4
18 to 18	5	---	17 to 17	5
19 to 19	6	---	18 to 19	6
20 to 21	7	---	20 to 20	7
22 to 23	8	---	21 to 22	8
24 to 25	9	---	23 to 24	9
		---	25 to 25	10

Raw Score Range	Stanine Equivalent	---	Raw Score Range	Sten Equivalent
0 to 12	1	---	0 to 12	1
13 to 14	2	---	13 to 13	2
15 to 15	3	---	14 to 14	3
16 to 16	4	---	15 to 16	4
17 to 17	5	---	17 to 17	5
18 to 19	6	---	18 to 18	6
20 to 20	7	---	19 to 20	7
21 to 22	8	---	21 to 21	8
23 to 25	9	---	22 to 23	9
		---	24 to 25	10

Raw Score

Lookup Table
Stens and Stanines computed using the Continuous Score Percentile Rank Formula

Lookup Table
Stens and Stanines computed using the Discrete Score Frequencies

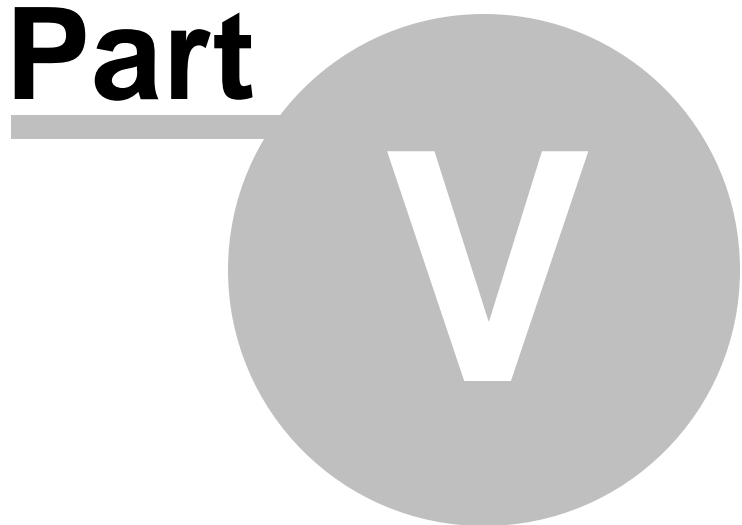
In Stanscore 4 (and 5), two lookup tables are produced, one based upon using the Continuous-Score cumulative proportions for stens and stanines, the other based upon the discrete score cumulative proportions. There won't be much difference between these two tables because really we are only playing around with halving frequencies at a particular score vs using all them to calculate a z-score and percentile. They'll be a bit of "jockeying around" with boundary values, as above ... (see a Sten of 6 in both tables above, one of for scores between 18 to 19, the other for scores 18 only), but it is marginal.

My recommendation is that you use the Discrete Score Frequency Sten-Stanine Lookup table.

It can be printed and saved as an Excel file for inclusion in other documents.

Paul Barrett (2011) paul@pbarrett.net

Part



5 The Stanscore Algorithms

5.1 Stanscore 1 and 2 algorithm

Versions 1 and 2 of this software were constructed using the procedure outlined by Kline, P. (1986) [A Handbook of Test Construction](#). London: Methuen and Co.. ISBN: 0-416-39430-2, pp. 167-168 .

1. Work out the cumulative proportion (CP) for each raw score – doing steps 2-5 ...
2. Prepare a frequency distribution of the scores
3. Compute the Cumulative Frequency (CF) per score, where CF = the sum of all frequencies below a score.
4. Compute the CF to the mid-point of each score interval by adding the CF at a score to half the numbers of cases observed at the same score.
5. Divide the mid-point CF (from #5) by the total number of cases (N). This is the CP (Cumulative proportion).
6. Obtain the theoretical Normal Distribution z-score for each CP (the inverse normal calculation) – here we find the z-score from a given proportion of the area under the curve.
7. Use the conventional transformation formula to convert the z-score into a sten, stanine, or T-score.
e.g. a sten score transformation would be:

$$\text{stenscore} = \text{round}(5.5 + \text{zscore} * 2.0)$$

So, there is no interpolation required – but, here, the expected cumulative proportions for a **perfect** sten score distribution may not be met in practice ...

i.e.

Sten Score	Expected Proportion
1	.02
2	.05
3	.09
4	.15
5	.19
6	.19
7	.15
8	.09
9	.05
10	.02

This is where another approach to constructing a sten/stanine/percentile score distribution could be attempted, which would involve interpolation. That is, you interpolate “empty, no-occurrences” scores to construct a more finely graded cumulative proportion distribution – then construct the sten score

distribution “algorithmically” - to better approximate (if possible) the expected proportions. This would only work better if there were great gaps between scores.

Given the whole purpose of norms is, however, to be representative of a group, it’s unlikely that there would be huge gaps between scores except at the tails (nobody scoring say below 8 on a 30-score-point test). But it can happen with very small norm groups – but very small norm groups have a much greater problem than this! Their lack of representativeness. And yes, it can happen where a test is badly constructed (“lumpy” blocks of nearly equal item difficulties).

There will be differences between the normative scores computed using this algorithm and that used in Stanscore 3 - but these will be trivial. In the end, I chose the different method as this brings the computation of cumulative frequencies in line with all statistical software packages.

Paul Barrett (2011) paul@pbarrett.net

5.2 Stanscore 3 algorithm

The main change in this algorithm is how the cumulative frequencies are computed. A cumulative frequency for a score is computed as the sum of the frequencies observed **up to and including the "current score"**.

Then, this quantity is used in the formula for percentiles and proportions ...

$$PR_i = \text{round} \left(\left[\frac{(cf_{i-1} + 0.5 \cdot f_i)}{N} \right] \cdot 100 \right)$$

where

PR_i = integer percentile rank (the percentile) for score i

round = round the calculation to the nearest integer

cf_{i-1} = the cumulative frequency for immediately preceding score $i - 1$

f_i = the frequency at score i

N = the sum of all frequencies across all scores.

The exact proportion associated with a score (allowing for the hypothetical continuity between integer scores) is

$$\text{exact_proportion}_i = \left(\left[\frac{(cf_{i-1} + 0.5 \cdot f_i)}{N} \right] \right)$$

This exact proportion is then used to create the z-score, and then [sten](#), [stanine](#), and [T-scores](#).

Because of the way the cumulative frequencies are formed, there will be some slight changes between Stanscore 3 and versions 1 & 2 - seen in the percentiles and other transformed scores. But, these will always likely be trivial.

In the end, I chose this different methodology as this brings the computation of cumulative frequencies in line with all statistical software packages.

Paul Barrett (2011) paul@pbarrett.net

5.3 Stanscore 4 algorithm

For Continuous score percentiles, the cumulative proportions are computed as per [Stanscore 3](#) algorithm.

For Discrete score percentiles, the cumulative proportions are quite literally the cumulative proportions computed by adding the observed proportion at a particular score to those proportions of scores immediately below it.

E.g.

Input Data		Frequencies - Continuous Score Percentiles		Sten and Stanine Lc
Raw Score	Frequency	Cumulative Freq.	Proportion	Cumulative Propn.
0	3	3	0.022	0.02206
1	1	4	0.007	0.02941
2	0	4	0.000	0.02941
3	0	4	0.000	0.02941
4	1	5	0.007	0.03676
5	0	5	0.000	0.03676
6	0	5	0.000	0.03676
7	0	5	0.000	0.03676
8	0	5	0.000	0.03676
9	1	6	0.007	0.04412
10	1	7	0.007	0.05147
11	0	7	0.000	0.05147

The data above show the proportions for each frequency of occurrence of a particular raw score. The total number of frequencies for these data is 136. The cumulative proportion for any score is simply the addition of all proportions below a raw score plus added to the proportion at that score. E.g. for a raw score of 4, we add: 0.022, 0.007, 0.000, 0.000, 0.007 = 0.03676 (using 15-decimal place precision, the numbers in the table are rounded versions); for example 3/136 to 9 places of precision is 0.022058824, which is reported as 0.022 in the above table).

Paul Barrett (2011) paul@pbarrett.net

Part

VI

6 Definitions

6.1 normalised scores

Normalized scores are specifically constructed to provide proportions of cases at each score level according to the expected frequencies of a perfect normal distribution. The mean and standard deviation of the raw data are the key parameters which enable standard scores and the normalization to be created. In this way, one can refer to the transformed integer score, with a very good idea just how many people obtained that score. Obviously, this can still end up with some inaccuracies where observed score ranges have big gaps between them, and frequencies are not spread very well across the whole score range. In other words, it's not a magic bullet to fix up badly skewed distributions of scores!

[z-scores](#) are defined as the raw score expressed in standard units, where each raw score is subtracted from the mean of all scores, and divided by the standard deviation of the raw score distribution.

For more information on norms, standardisation and normalisation, and the advantages/disadvantages of each I recommend:

Kline, P. (1986) *A Handbook of Test Construction*. London: Methuen and Co.. ISBN: 0-416-39430-2 pages ... 159-174 (Chapter 8) as an excellent explanatory chapter. Note that Colin Cooper provided the algorithms and software in the book (written in BASIC). I used the inverse-normal routine from his program P.SCALING.

also very good are ...

Bartram, D. and Lindley, P.A. (1994) *Scaling, Norms, and Standardization. Module 2, Part 2 of the British Psychological Society Level A Open Learning Programme*. Leicester UK: BPS Books. **Total Module Set ISBN: 1-85433-075-6**. There is no ISBN for the constituent modules - but you should be able to request a library copy of the specific module.

Kline, P. (2000) *Handbook of Psychological Testing*. London: Routledge. ISBN: 0-415-21158-1. Pages ... 59-64 (Chapter 4).

Crocker, L., and Algina, J. (1986) *Introduction to Classical and Modern Test Theory*. Harcourt Brace Jovanovich Publishers. ISBN: 0-03-061634-4. Pages ... 438-450 (Chapter 19).

Paul Barrett (2011) paul@pbarrett.net

6.2 percentiles

There are two common 'textbook' versions of interpreting a percentile:

Either:

Bartram, D. and Lindley, P.A. (1994) BPS Level A Open Learning Training Manual: Scaling Norms and Standardization, Module 2, part 1. London: BPS Publications (p.17)

The proportion of people scoring less than a particular score is called the percentile rank of the score. More commonly we refer to this as just the percentile.

Crocker, L., & Algina, J. (1986). Introduction to Classical and Modern Test Theory. New York: Holt, Rinehart and Winston. (p. 439)

Loosely speaking, the percentile rank corresponding to a particular raw score is interpreted as the percentage of examinees in the norm group who scored below the score of interest.

OR

Hinkle, D., Wiersma, W., & Jurs, S. (1994). Applied statistics for the behavioral sciences. (3rd ed.). Boston: Houghton Mifflin Company. (p. 49-50)

A percentile is the point in a distribution at or below which a given percentage of scores is found. For example, the 28th percentile of a distribution of scores is the point at or below which 28% of the scores fall.

Hays, W.L. (1994) Statistics, 5th Edition. Florida: Harcourt Brace. (p. 194)

In any frequency distribution of numerical scores, the percentile rank of any specific value is the percentage of the total cases that fall at or below in value.

Kiess, H.O. (1996) Statistical Concepts for the Behavioral Sciences. London: Allyn and Bacon (p. 46)

A percentile is the score at or below which a specified percentage of scores in a distribution falls.

As noted in one of my Technical Whitepapers entitled "**Percentiles and Percentile Ranks - confused or what?**" which can be downloaded from: <https://www.pbarrett.net/techpapers.html>, Technical whitepaper #3, the distinction seems to be all about whether you treat integer scores as discrete values, or as observed integer values from an underlying continuous real-valued score range.

IMPORTANT

But, I'm afraid this reasoning just doesn't work when dealing with integer test scores whose values change as a single, discrete unit (e.g. 0, 1, 2, 3, 4, 5, 6, 7, 8...).

The problem is, the "continuous" assumption calculation only uses half the frequencies at an observed score. So for the maximum possible score (or indeed any score value), only half the observed frequencies are used to express the final percentile. This is crazy. Really.

It is factually incorrect to state *the percentile for this score is the value below which n% of the norm-group score*, because half the people scoring the maximum observed score (or a particular score) are included in the cumulative proportion. Alternatively it is just as wrong to state *the percentile is the score at or below which n% of the group score*, because half the people scoring at the percentile score are missing from the calculation.

So, Stanscore 4 now computes the cumulative frequency distribution from the observed frequencies at each integer score, assuming an integer score is a discrete entity (no continuity assumption). This distribution is what you find if you computed a standard frequency distribution for a set of scores in say SPSS, STATISTICA, or another stats package. The percentiles of this distribution range between a potential minimum value of 0 (if there are no observed frequencies at 0), and **ALWAYS** 100.

The interpretation of these percentiles is very simple, accurate, and very clear. It is that **person X** with **percentile rank** of say 80 scores at or above n% of the group, or, n% of the group score at or below a percentile rank of 80.

A person receiving the maximum score would always have a percentile rank of 100, which would mean they score at and above 100% of the norm group.

The formula for "continuous score" percentiles is:

$$exact_proportion_i = \left(\left[\frac{(cf_{i-1} + 0.5 \cdot f_i)}{N} \right] \right)$$

The formula used for calculating percentiles in Stanscore 4 is:

$$PR_i = round \left(\left(\left[\frac{(cf_{i-1} + 0.5 \cdot f_i)}{N} \right] \cdot 100 \right) \right)$$

where

PR_i = integer percentile rank (the percentile) for score i

$round$ = round the calculation to the nearest integer

cf_{i-1} = the cumulative frequency for immediately preceding score $i - 1$

f_i = the frequency at score i

N = the sum of all frequencies across all scores.

taken from Crocker, L., and Algina, J. (1986) *Introduction to Classical and Modern Test Theory*. Harcourt Brace Jovanovich Publishers. ISBN: 0-03-061634-4. Pages ... 438-450 (Chapter 19, page 440).

I sincerely recommend you do not use this method anymore. Use the Discrete Score Frequency Distribution Percentiles.

Paul Barrett (2011) paul@pbarrett.net

6.3 stanines

standard nine scores. These are [normalized](#) scores which range between 1 and 9, with a mean of 5 and standard deviation of 2. The difference between stanine scores is approximately one half a standard deviation.

$$\text{z-score} = \frac{(\text{raw score} - \bar{x})}{s} \dots \text{where} \dots$$

\bar{x} = the norm sample mean

s = the norm sample standard deviation

then...

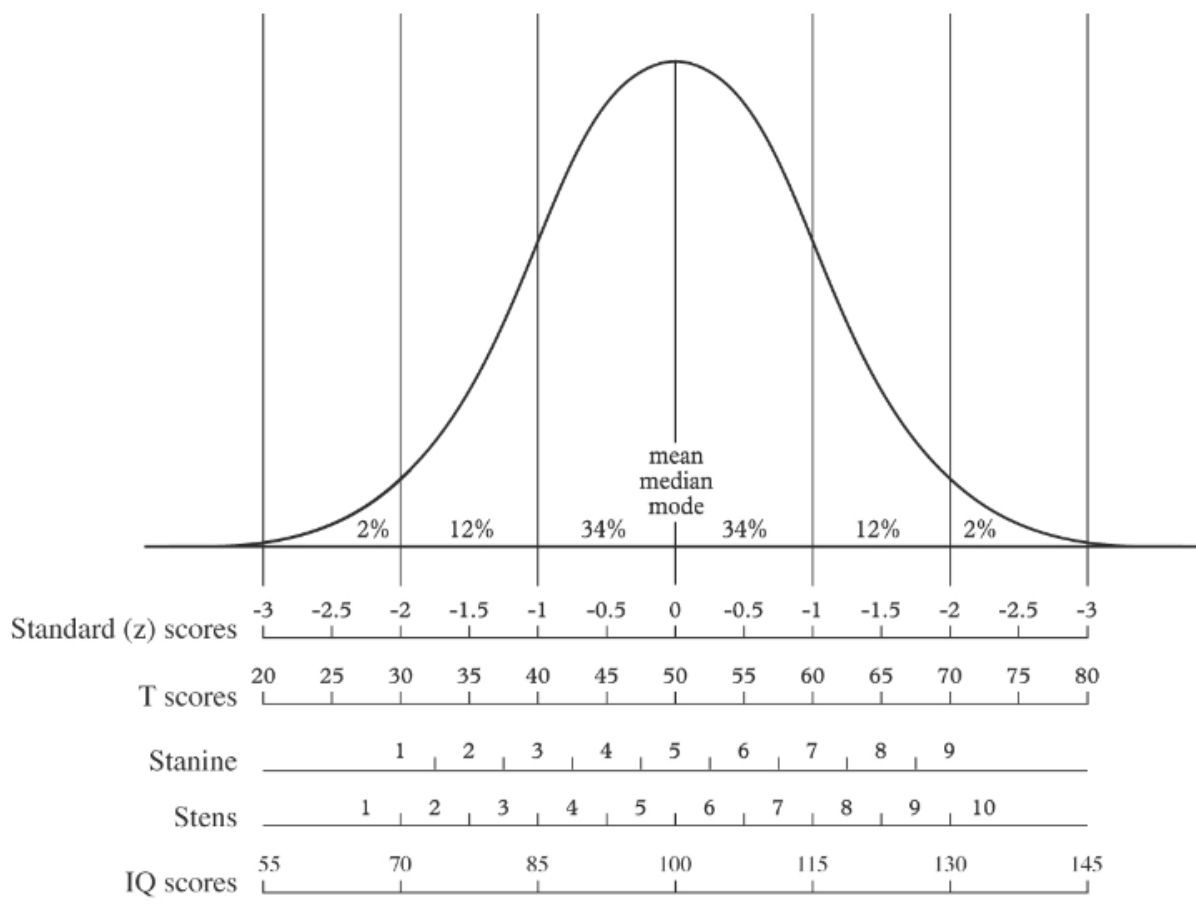
transformation equation = $(\text{z-score} \cdot \text{new_Std.dev}) + \text{new_mean}$

Stanine = $(\text{z-score} \cdot 2) + 5$ (range 1-9: integers)

Sten = $(\text{z-score} \cdot 2) + 5.5$ (range 1-10: integers)

T-score = $(\text{z-score} \cdot 10) + 50$ (nominal range 20-80: integers)

A nice way of visualizing the relationship between the different kinds of transformed *normalized* scores is:



Paul Barrett (2011) paul@pbarrett.net

6.4 standardised scores

Simply the raw score (or its deviation form) standardized against some other value. The other value is generally the standard deviation for a set of raw scores. Also, in psychometrics, the raw score is initially expressed as a "deviation score", prior to division by the standard deviation.

In formula terms ...

$$\text{z-score} = \frac{(\text{raw score} - \bar{x})}{s} \dots \text{where} \dots$$

\bar{x} = the norm sample mean

s = the norm sample standard deviation

then...

transformation equation = $(\text{z-score} \cdot \text{new_Std.dev}) + \text{new_mean}$

Stanine = $(\text{z-score} \cdot 2) + 5$ (range 1-9: integers)

Sten = $(\text{z-score} \cdot 2) + 5.5$ (range 1-10: integers)

T-score = $(\text{z-score} \cdot 10) + 50$ (nominal range 20-80: integers)

Unless the distribution of raw scores is actually "normal", then the percentages under the normal curve associated with each transformed score will not be symmetric around the distribution mean, hence the expected frequencies associated with each score might not be matched by the actual frequencies.

This is why [normalized](#) scores are used - where the proportions of scores in the observed frequency distribution are used to construct the transformed score, so as to best approximate/guarantee the expected frequencies associated with a transformed score.

Note that Nunnally and Bernstein (1994) page 309 give the incorrect formula for a stanine as that with a mean of 5.5 and an SD of 1.0. No other major text provides these definitional values of mean 5.5 and SD of just 1.0, including Cattell etc.

Nunnally, J.C., and Bernstein, I.H. (1994) *Psychometric Methods 3rd. Edition*. McGraw-Hill. ISBN: 0-0704-7849-X.

Paul Barrett (2011) paul@pbarrett.net

6.5 stens

"Standard 10" scores ... These are [normalized](#) scores which range between 1 and 10, with a mean of 5.5 and standard deviation of 2.

$$\text{z-score} = \frac{(\text{raw score} - \bar{x})}{s} \dots \text{where...}$$

\bar{x} = the norm sample mean

s = the norm sample standard deviation

then...

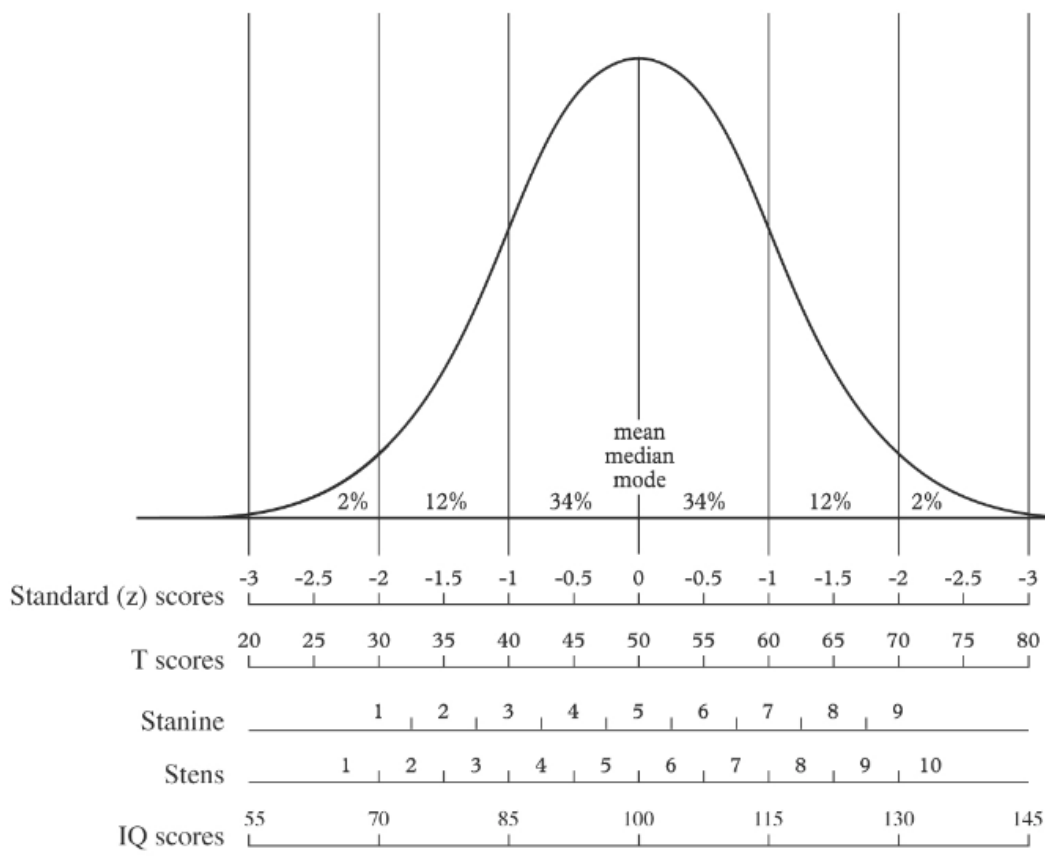
transformation equation = $(\text{z-score} \cdot \text{new_Std.dev}) + \text{new_mean}$

Stanine = $(\text{z-score} \cdot 2) + 5$ (range 1-9: integers)

Sten = $(\text{z-score} \cdot 2) + 5.5$ (range 1-10: integers)

T-score = $(\text{z-score} \cdot 10) + 50$ (nominal range 20-80: integers)

A nice way of visualizing the relationship between the different kinds of transformed normalized scores is:



Paul Barrett (2011) paul@pbarrett.net

6.6 T-scores

These are [normalized](#) scores which range between about 20 and 80, with a mean of 50 and standard deviation of 10.

$$\text{z-score} = \frac{(\text{raw score} - \bar{x})}{s} \dots \text{where...}$$

\bar{x} = the norm sample mean

s = the norm sample standard deviation

then...

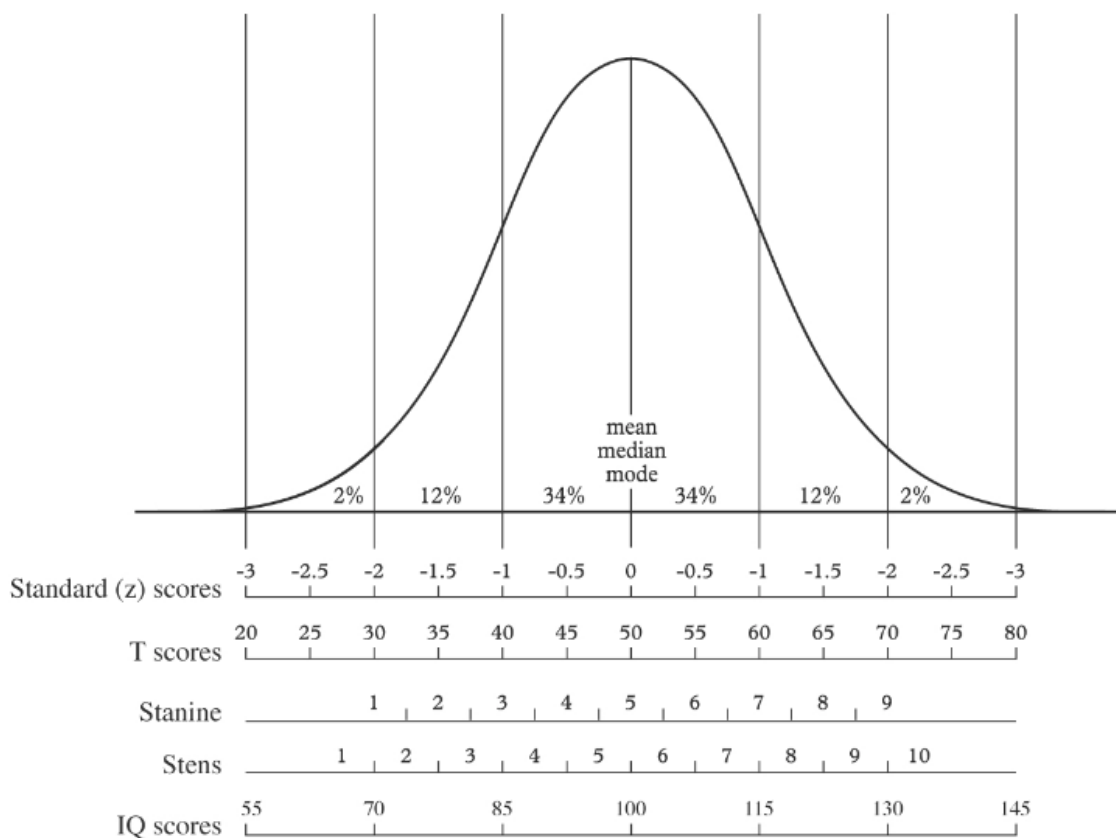
transformation equation = (z-score · new_Std.dev) + new_mean

Stanine = (z-score · 2) + 5 (range 1-9: integers)

Sten = (z-score · 2) + 5.5 (range 1-10: integers)

T-score = (z-score · 10) + 50 (nominal range 20-80: integers)

A nice way of visualizing the relationship between the different kinds of transformed normalized scores is:



6.7 z-scores

The z-score is simply the raw score expressed in standard units, where each raw score is subtracted from the mean of all scores, and divided by the standard deviation of the raw score distribution.

$$\text{z-score} = \frac{(\text{raw score} - \bar{x})}{s} \dots \text{where} \dots$$

\bar{x} = the norm sample mean

s = the norm sample standard deviation

then...

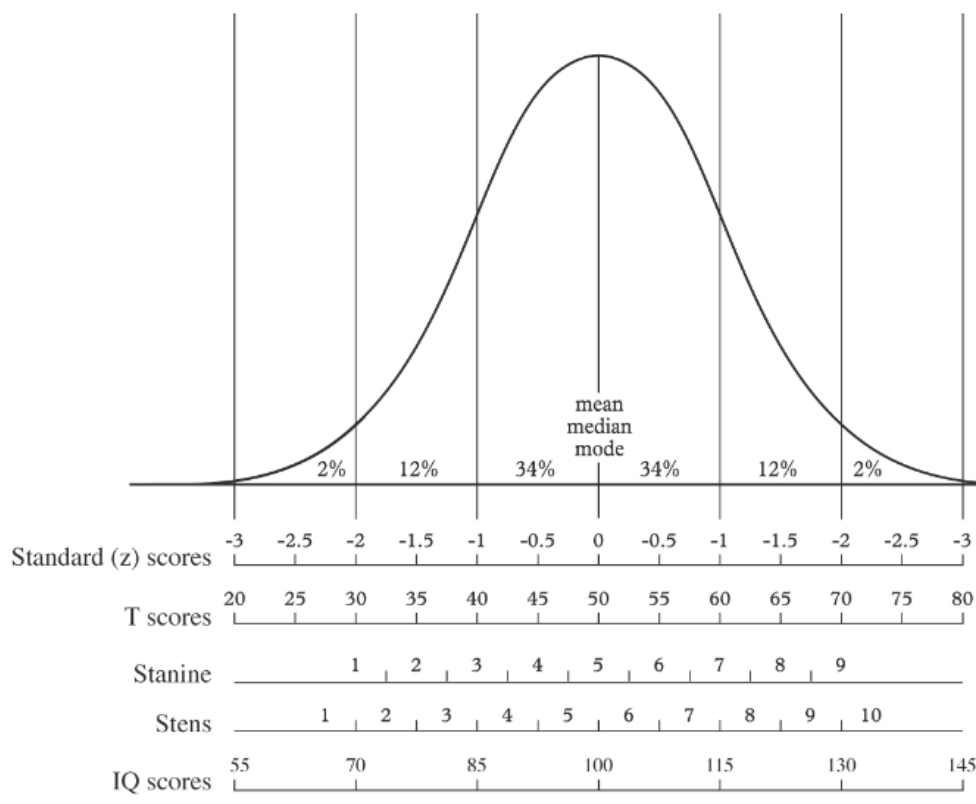
transformation equation = $(\text{z-score} \cdot \text{new_Std.dev}) + \text{new_mean}$

Stanine = $(\text{z-score} \cdot 2) + 5$ (range 1-9: integers)

Sten = $(\text{z-score} \cdot 2) + 5.5$ (range 1-10: integers)

T-score = $(\text{z-score} \cdot 10) + 50$ (nominal range 20-80: integers)

A nice way of visualizing the relationship between the different kinds of transformed normalized scores is:



Paul Barrett (2011) paul@pbarrett.net

Index

- A -

Access file input 18
ASCII file input 19

- E -

Excel file input 20
Excel xls vs xlsx format 27

- N -

normalised score 43

- P -

percentile 43
PsWin file input 21

- R -

results - frequency score distribution 32
results - input data table 30
results - sten/stanine lookup table 36

- S -

saving excel files 27
standard score 47
stanine 45
Stanscore 3 algorithm 40
Stanscore 4 - What's New 5
Stanscore 4 algorithm 41
Stanscore basics 7
Stanscore v1 and 2 algorithm 39
STATISTICA file input 21
sten 47

- T -

T-score 49

- U -

using the program 9

- Z -

z-score 50

