Government of India
Department of Personnel and Training

Staff Selection Commission
CGO Complex, New Delhi

Normalization procedure for SSC exams from June 2025

1 Step-by-step procedure for converting raw scores into normalised scores

For each test/subject/area for which the examination is held in multiple shifts, the raw score for each candidate appearing for the test/subject is converted into ***normalised score*** in the following three steps.

Step 1 (Intermediate) Convert raw scores into percentile scores: the percentiles are calculated separately for each shift.

Step 2 (Intermediate) *Pull-back* the percentiles to the *marks scale*: this is done by first collating the data across all sessions into one table, then sorting the records in decreasing (increasing) order of percentiles, and finally filling in the gaps in the raw score table by interpolation.

Step 3 (Final) At the end of the previous step, each percentile value will have a *corresponding raw score value* for each session. This is combined to get the ***normalised score***.

This procedure will be done separately for each test/subject/area so that each candidate is assigned a score for each test/subject/area which has been opted for and at the end of this transformation procedure, each candidate will have a normalised score in each such test/subject/area.

The details of each of the three steps mentioned above are given below.

Step 1: Calculation of Percentile Scores: This first step is to be completed *separately for each shift.*

1. Record the number of candidates who have actually appeared in the shift. Denote this number by N.
2. Sort all the candidates in one shift in decreasing order of their marks.
3. Note the *raw marks* for each candidate. Suppose this is denoted by T. Count the number of candidates in that shift whose raw scores are *less than or equal to* T. Denote this number by m.
4. The percentile score for this candidate is then calculated as

$$P = \frac{m}{N}.$$

Note that the percentile so calculated will satisfy $0 \leq P \leq 1$.

5. The percentile P above can be rounded off to the requisite number of decimal places. It is recommended to do it till 8-th places of decimal.

1.1 Illustration

Suppose that the examination in a certain subject is held in two different shifts, $S1$ and $S2$, say. Let us consider six candidates A, B, C, D, E and F out of which A, B, C are from shift S1 and the other three are from shift $S2$.

Let the raw marks of the six candidates be x_A, x_B, x_C (shift $S1$ marks) and y_D, y_E, y_F (shift $S2$ marks).

For candidates A, B, C, the percentiles are calculated using the totality of marks obtained by candidates appearing in shift $S1$ (in the same subject) as explained above.

Similarly, for candidates D, E, F, the percentiles are calculated using the totality of marks obtained by candidates appearing in shift $S2$ (in the same subject).

Let the respective percentiles be denoted by $P_A, P_B, P_C, P_D, P_E, P_F$.

We would have a table which would look like the following. The terms in red colour indicate that these are the output of this step.

Shift S1			Shift S2		
Candidate	Raw score	Percentile	Candidate	Raw score	Percentile
A	x_A	P_A	D	y_D	P_D
B	x_B	P_B	E	y_E	P_E
C	x_C	P_C	F	y_F	P_F

Since the calculations for percentiles in any shift deppends only on the data from that shift alone, in essence, there is a separate table for each shift.

2 Step 2: Sorting the candidates using their percentiles after combining data from all sessions

1. The session-wise data is now to be combined (or concatanated) together.
2. During the concatenation, the separate percentile columns SHOULD be combined to create a single percentile column.

However, the columns for the shift-wise raw score should be kept separate.

The column identifying the candidate can also be combined.

3. All the records are to be sorted in decreasing order of the percentiles.

   In the *illustrative example* given in Section 1.1 above suppose that the percentiles of the six candidates satisfy
$$P_E > P_A > P_C = P_F > P_B > P_D.$$
   Then the table at the end of this sub-step would look as given below.

Candidate	Percentile	Raw score $S1$	Raw Score $S2$
E	P_E	–	y_E
A	P_A	x_A	–
C & F	$P_C = P_F$	x_C	y_F
B	P_B	x_B	–
D	P_D	–	y_D

4. Candidate C from shift $S1$ and candidate F from shift $S2$ have the same percentile. The relevant entries under "Raw Score $S1$" and "Raw Score $S2$" are the actual raw scores x_C and y_F respectively. [2]

5. Candidates A and B, appearing in shift $S1$, have a blank entry in column "Raw Score $S2$", as there is no corresponding candidate having exactly the same percentile from shift $S2$.

6. Similarly, Candidates D and E, appearing in shift $S2$, have a blank entry in column "Raw Score $S1$", as there is no corresponding candidate having exactly the same percentile from shift $S1$.

7. In the remaining part of this *Step 2*, the blank entries in the two "Raw Score" columns are to be filled up using linear interpolation. This is achieved as follows.

   - Consider a record (row) whose entry in the column "Raw Score $S1$" is blank. The blank will be replaced by a score X. We will find the value corresponding to X.
   - For this record or row, let the entry in "Percentile" column be P.
   - Let x_1 denote the first *non-blank* entry BELOW X. i.e. $x_1 < X$ and there is no other non-blank entry in the column between x_1 and X.
   - Let x_2 denote the first *non-blank* entry ABOVE X.[3] i.e. $x_2 > X$ and there is no other non-blank entry in the column between X and x_2.
   - Let p_1 be the entry in the "Percentile" column corresponding to x_1 from the column "Raw Score $S1$".

[2] This has the obvious interpretation that marks x_C of shift $S1$ are equivalent to marks y_F of shift $S2$, under this ***equipercentile method.***

[3] There may be several blank entries between x_1 and x_2.

- Let p_2 be the entry in the "Percentile" column corresponding to x_2 from the column "Raw Score $S1$". Note that P, p_1, p_2, x_1, x_2 are know values and the only unknown is X.
- The **interpolated score X** is then calculated as

$$X = x_1 + \frac{x_2 - x_1}{p_2 - p_1}(P - p_1).$$

8. All the blank entries in column "Raw Score $S1$" can now be replaced by the *interpolated values*.

9. The blank entries in column "Raw Score $S2$" are also replaced using the similar procedure.

2.1 Illustration (Continued)

At the end of this step, the table in the earlier illustrative example would look like the following, where the entries in red indicate the additions or output at the end of this step.

Candidate	Percentile	Raw score $S1$	Raw Score $S2$
E	P_E	X_E	y_E
A	P_A	x_A	Y_A
C & F	$P_C = P_F$	x_C	y_F
B	P_B	x_B	Y_B
D	P_D	X_D	y_D

3 Step 3: Calculation of the normalised score

At the end of the previous step, *there is a score assigned to each percentile value and each session.* The final step is to combine these to come up with a unique ***normalised score*** for each percentile value and hence *to each candidate.*

Corresponding to any candidate, if the percentile value is P, and corresponding to thie value P, if the raw scores for different sessions is $u_1, u_2, \ldots, u_t$ (t being the number of different sessions), then the normalised score Z is defined as

$$Z = \text{Average of } (u_1, u_2, \ldots, u_t) = \frac{u_1 + u_2 + \ldots u_t}{t}.$$

3.1 Illustration (Continued)

This is once again illustrated through our example. The **final** table would now be as follows, with the final column in red denoting the final ***normalised score.***

Candidate	Percentile	Raw Score $S1$	Raw Score $S2$	Normalised Score
E	P_E	X_E	y_E	$(X_E + y_E)/2$
A	P_A	x_A	Y_A	$(x_A + Y_A)/2$
C & F	$P_C = P_F$	x_C	y_F	$(x_C + y_F)/2$
B	P_B	x_B	Y_B	$(x_B + Y_B)/2$
D	P_D	X_D	y_D	$(X_D + y_D)/2$

****