u multirank
Universities compared. Your way.

The construction of rank groups:
Significant deviations from the average score

Results for a particular university depend on the judgments of those students who responded; compared to the complete sample of students at a university they are subject to uncertainty. How well they meet the "true" judgment of a department depends largely on the number of respondents and the range of their reviews.

How much "trust" you may have in such an average rating is expressed statistically by a so-called confidence interval. These confidence intervals can be utilized to incorporate the uncertainty of the judgment values in the ranking calculation. Instead of fixing limits for the average judgments and then determine the top and bottom groups, the length of the corresponding confidence interval is taken into account in the grouping procedure.

For university i, let $\overline{x_i}$ be the average judgment of its students, s_i the standard deviation of the judgments and n_i the number of cases, and $\bar{x}$ be the overall average of judgments. Then:

- if $\overline{x_i} - 1.96 \cdot \frac{s_i}{\sqrt{n_i}} > \bar{x} \Rightarrow$ assign university i "below average",
- if $\overline{x_i} + 1.96 \cdot \frac{s_i}{\sqrt{n_i}} < \bar{x} \Rightarrow$ assign university i "above average",

For the remaining institutions which are a now mixture of "really" intermediate judgments and of those that could not be assigned to one of the extreme groups because their confidence intervals were too long we use two additional limits besides the mean. These limits deviate 0.25 from the mean plus one standard error on the department level, a normalized variation between the departments in one field. They differ for each field. Let $\bar{x}$ be the overall average of judgments, s the standard deviation "between" the universities, i.e. of $\overline{x_i}, i = 1, \dots, n$, and n the number of universities ranked in this field:

- Lower limit: $\mathrm{L} = \bar{\mathrm{x}} - (0.25 + \left(1.96 \cdot \frac{\mathrm{s}}{\sqrt{\mathrm{n}}}\right))$,
- Upper limit: $\mathrm{U} = \bar{\mathrm{x}} + (0.25 + \left(1.96 \cdot \frac{\mathrm{s}}{\sqrt{\mathrm{n}}}\right))$.

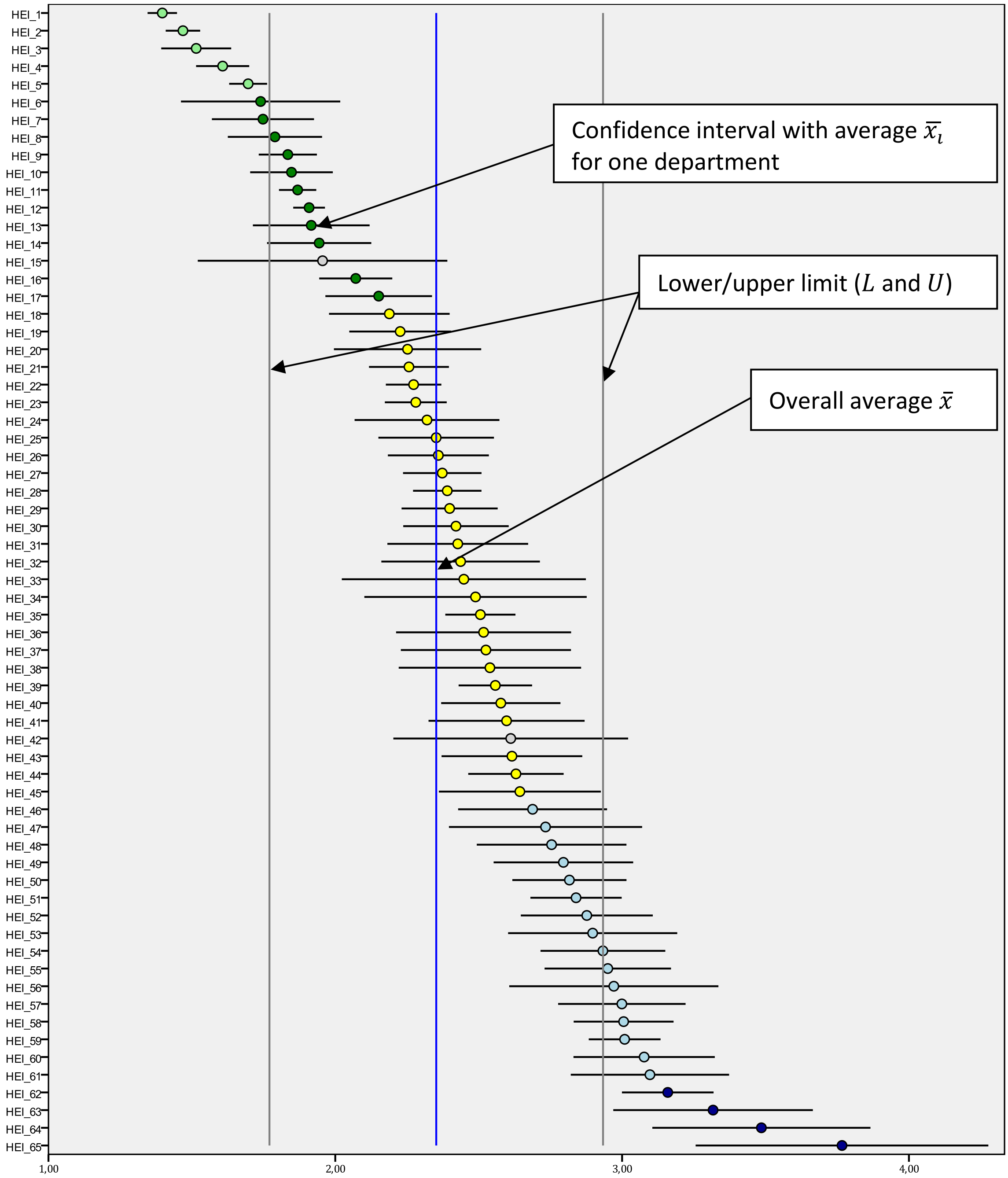
Departments with a confidence interval entirely between these two limits are, if they do not already belong to the top group, assigned to the middle group. However, if the confidence interval is so large that it extends across two of the three limits, its values are not included in the ranking because it is not clearly assignable to one group. In the example (see Figure 1 marked grey, HEI 15 and HEI 42. As no general statistical method for "proving" that something is intermediate exists, we had to develop something new. Testing this method with the data from several years of student surveys makes us confident that this is a method working well in practice.

As in U Multirank the number of HEIs compared is rather high two further groups are introduced, consisting of those HEIs beyond the lower and upper limit, so the rules for grouping are now:

- if $\overline{x_i} - 1.96 \cdot \frac{s_i}{\sqrt{n_i}} > U \Rightarrow$ assign university i to the "poor" group (dark blue in Figure 1),
- if $\overline{x_i} - 1.96 \cdot \frac{s_i}{\sqrt{n_i}} > \bar{x} \Rightarrow$ assign university i to the "below average" group (light blue in Figure 1),

- if $L < \bar{x}_i - 1.96 \cdot \frac{s_i}{\sqrt{n_i}}$ and $U > \bar{x}_i + 1.96 \cdot \frac{s_i}{\sqrt{n_i}}$ $\Rightarrow$ assign university i to the "average" group (yellow in Figure 1),
- if $\bar{x}_i + 1.96 \cdot \frac{s_i}{\sqrt{n_i}} < \bar{x} \Rightarrow$ assign university i to the "good" group (dark green in Figure 1)
- if $\bar{x}_i + 1.96 \cdot \frac{s_i}{\sqrt{n_i}} < L \Rightarrow$ assign university i to the "excellent" group (light green in Figure 1
- if $\bar{x}_i - 1.96 \cdot \frac{s_i}{\sqrt{n_i}} < L$ and $\bar{x}_i + 1.96 \cdot \frac{s_i}{\sqrt{n_i}} > \bar{x}$ or if $\bar{x}_i + 1.96 \cdot \frac{s_i}{\sqrt{n_i}} > U$ and $\bar{x}_i - 1.96 \cdot \frac{s_i}{\sqrt{n_i}} < \bar{x}$ $\Rightarrow$ university i is not ranked in this indicator group (grey in Figure 1).

Figure 1: Rank groups on the basis of confidence intervals



Advantages of this method are:

(a) Small departments still have the possibility to appear in the ranking.

(b) The validity of the group assignment even with small numbers of students is secured; doubtful cases are removed from the ranking in the subjective indicator.

As a consequence of this method departments with the same or similar averages, but different sized confidence intervals might be sorted into different ranking groups occasionally when they are placed on the border between an extreme group and the middle group. It is possible that a department ends up with a better mean in the middle group, while that with the worse mean is sorted into the top group.

Using the confidence intervals as basis for grouping ensures taking into account the homogeneity of the judgments, technically, however, the more trustworthy estimation of the "true" population mean in a statistical sense. In contrast to the grouping determined by quartile (where always a certain percentage of the universities is included in each group) or "top 10" lists in this process the size of the (top) group is not fixed, it results rather from the extent of internal variance within the different departments and the variance between them. If in a field the differences between individual universities are small and the responses in the departments are heterogeneous, very few institutions are to be found in the bottom or top group.

This method of comparing students' assessment of their own university which takes into account the relative assessment of a university compared to the average in that fields (on a particular indicator) and which refers to the degree of certainty of the means is more robust than just sorting universities simply by the mean scores. As this methods results in one group performing better than average, one lower than average and one in the middle it fits well to the he group approach of CHE ranking.