Appendix – Explaining percentile conversions

In order to compare different tests across different year levels and different years, the basic idea is to convert every score to a *percentile*. This value then gives the percentage of the other students in the same state in the same year who had a lower score.

All the percentiles are derived from the data provided at the state level, which is the 10th, 25th, 50th, 75th and 90th percentiles.

As an example, the state percentiles provided for Year 7 in 2021 for the Reading test are as follows.

Percentile	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Score	466	509	553	598	642

We can plot the points provided.

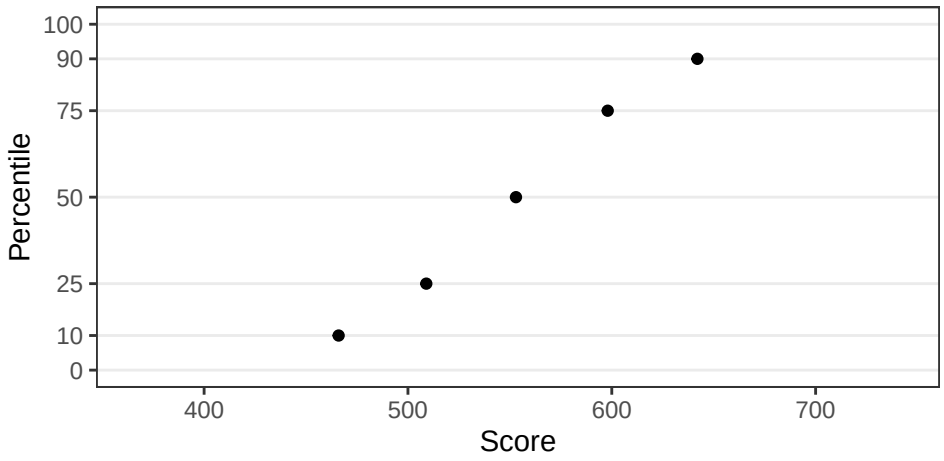

Figure 1: Known scores together with percentiles

The goal is to convert any possible score to a percentile, not just the five numbers provided. The final conversion graph extends the existing points and is shown here. An explanation of how this conversion graph is derived is shown below.

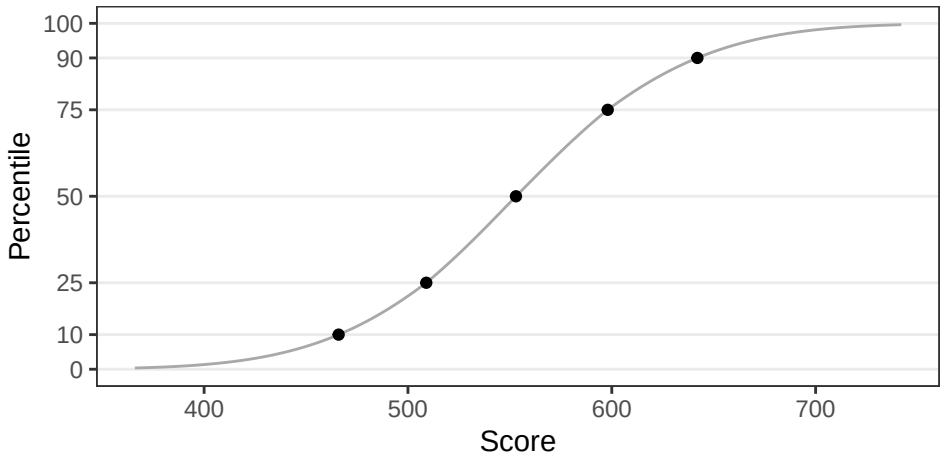

Figure 2: Final Score to Percentile Conversion

Explanation of Method

An initially appealing idea is simply to join the points with line segments. One significant issue with this is that it is unclear what to do with values lower than the 10th percentile or higher than the 90th percentile, since all percentiles must be between 0 and 100.

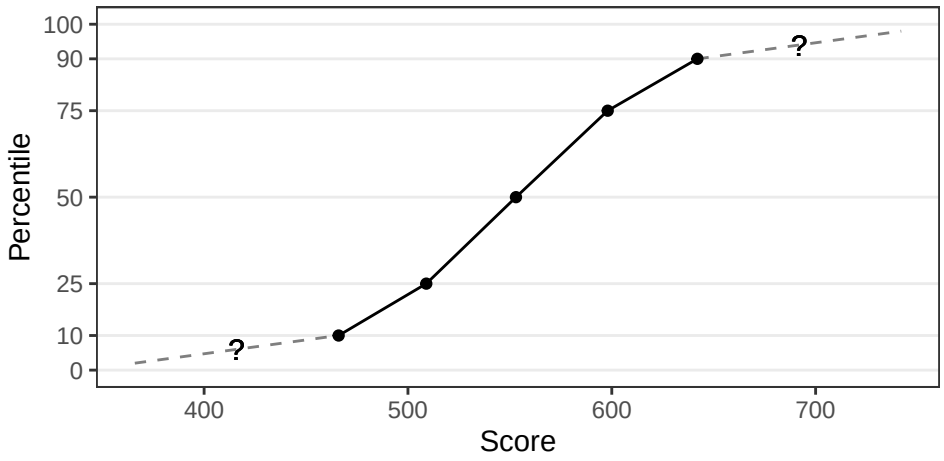

Figure 3: Attempted Conversion (line segments)

Instead of using line segments we can assume that scores across the state are likely to be approximately normally distributed. A normal distribution can be determined from just two 'percentile facts' and we have been provided five.

For scores below the 25th percentile, we can use two percentile facts: the 10th percentile is 466 and the 25th percentile is 509; this allows us to determine the mean and standard deviation of an underlying normal distribution.

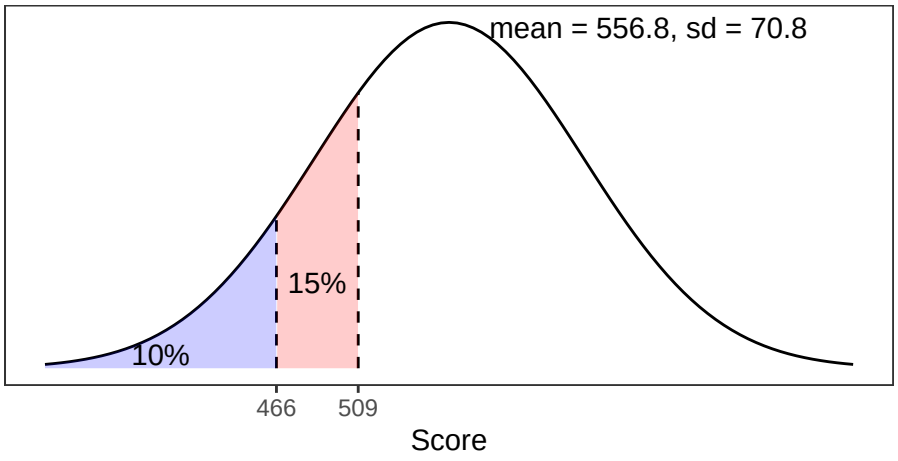

Figure 4: Normal distribution matching 10th and 25th percentiles

By considering areas, we can use this distribution to convert scores to percentiles for scores up to 509 (shown as the solid portion in the graph below).

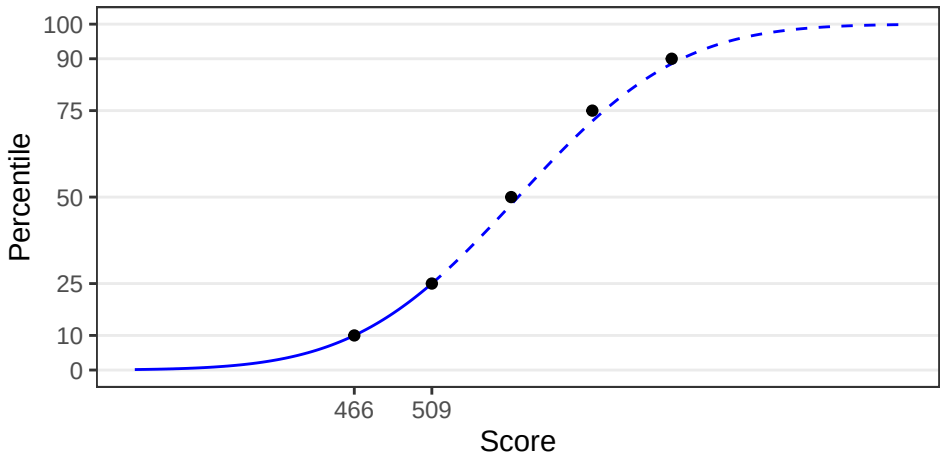

Figure 5: Mapping scores to percentiles (up to the 25th percentile)

The dashed portion of the curve does not pass cleanly through the other given points, because the underlying distribution is not perfectly normal across all the values.

Therefore, we use a similar strategy for values between the 25th and 50th percentiles (509 $\rightarrow$ 553), values between the 50th and 75th percentiles (553 $\rightarrow$ 598), and values greater than the 75th percentile (598) with the help of the 90th percentile (642), to generate the three curves shown below.

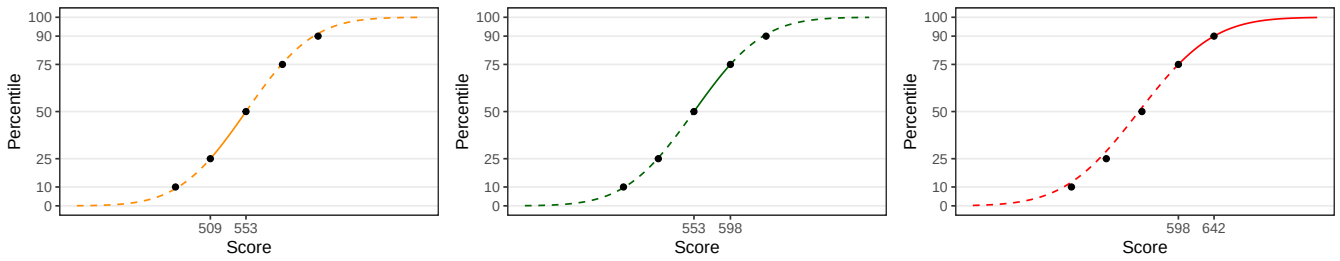

Figure 6: Separate mappings from scores to percentiles

These mappings are combined to create a single conversion of scores to percentile for Year 7 Reading scores in 2021. This process is repeated across all year levels, components and years.

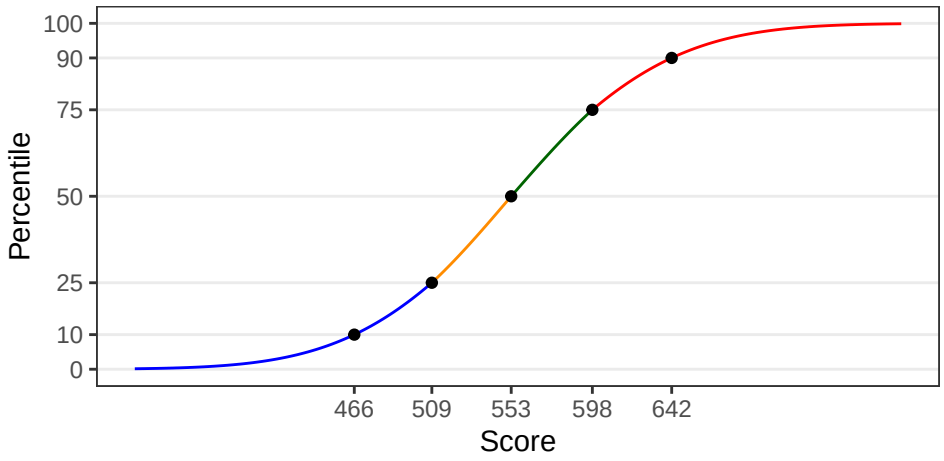

Figure 7: Combined mappings from scores to percentiles