Math 146 3.3 - Measures of Relative Standing and Boxplots

Standardized Exam A: $50 \le x \le 250$ $\mu_1 = 187.4$ $\sigma_1 = 9.7$

Standardized Exam B: $0 \le x \le 80$ $\mu_2 = 63.3$ $\sigma_2 = 3.8$

Two students each take one of the standardized exams with summary stats above. Alice scores a 205 on Exam A, and Bob scores a 69 on Exam B. If only one scholarship is given to the highest score, which student is awarded the scholarship?

To convert values to a standardized scale, we calculate how many standard deviations from the mean a particular value is. This value is called the **z-score** (or standardized value).

$z = \frac{x-\bar{x}}{s}$ using sample statistics or $z = \frac{x-\mu}{\sigma}$ using population parameters

Compare the *z*-scores for both Alice and Bob above and determine the better score. (Round all *z*-scores to two decimal places.)

(**Note:** usual values will have *z*-scores $-2 \le z \le 2$, and unusual values have $z < -2$ or $z > 2$.)

BOXPLOTS AND THE FIVE NUMBER SUMMARY

The *Five Number Summary* divides an entire data set into four equal groups, where each group contains roughly 25% of the data values. These are called quartiles.

Five Number summary:

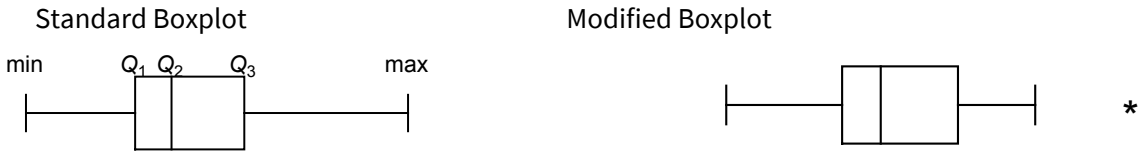
1. Minimum Value
2. First Quartile, Q_1 (median of the lower half)
3. Second Quartile, Q_2 (the median)
4. Third Quartile, Q_3 (median of the upper half)
5. Maximum Value

Use the five-number summary to compare home runs of Hank Aaron and Babe Ruth

Aaron: 11 12 13 20 24 26 27 29 30 32 34 34 38 39 39 40 40 44 44 44 44 45 47

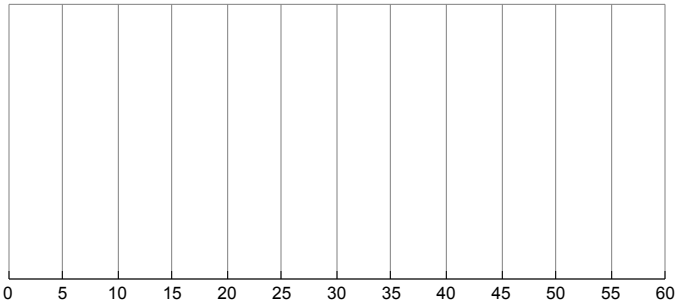
Ruth: 0 2 3 4 6 11 22 25 29 34 35 41 41 46 46 46 47 49 54 54 59 60

Boxplots are a graphical representation of the five-number-summary:



Define the **Interquartile Range** as **IQR** = $Q_3 - Q_1$. An outlier is any value higher than Q_3 by 1.5 × IQR, or below Q_1 by 1.5 × IQR. Outliers are indicated by a symbol.

Use your calculator to create side by side boxplots for the Aaron-Ruth data. Which is the better hitter; more consistent hitter?



PERCENTILES

Quartiles are a special case of quantiles or **percentiles**, e.g., $Q_1 = P_{25}$, $Q_2 = P_{50}$, $Q_3 = P_{75}$.

Example: The value that separates the bottom 30% from the top 70% is the 30 percentile or P_{30}.

1. Sort the data.
2. Calculate $L = n \cdot \frac{30}{100}$, where n is the number of values:
   a) if L is an integer, P_{30} is the mean of L^{th} value and $L + 1$ value.
   b) if L is not an integer, round up and use that L^{th} value as P_{30}.

Find the 65th percentile for Ruth.