characterizing data-

distribution: way values are spread-table or graph, describe center/shape

mean - $\bar{x} = \frac{\text{sum of values}}{\text{total}}$ mode - commonest value

median - middle value of sorted data set

outlier - value way higher/lower can of set mean

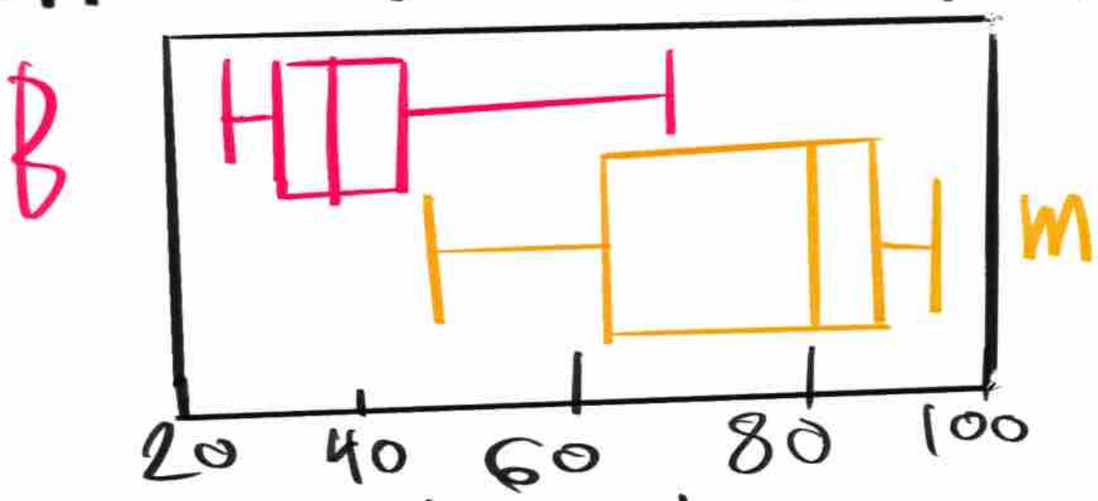
Shape: symmetry (line of symetry, mode = mean = median), skewness (how far off from symetry, tail/skew on side of tail), peaks (unimodal/bimodal)

Variation: how spread away from mean

Range: max/min, limited, min (lowest val), Q1 (median of lower half), Q2 (overall median), Q3 (upper half), max (highest value)

Box Plot - Q1-Q3, vert line through median lines min/max

| | mean | median | mode | min | Q1 | med | Q3 | max |
|---|---|---|---|---|---|---|---|---|
| B | 39.78 | 36 | 42 | 26 | 29 | 36 | 45 | 69 |
| M | 75 | 80 | 44 | 50 | 62 | 80 | 87.5 | 94 |




Standard deviation - Describes variation, dev from mean = data value - mean

$$\sqrt{\frac{\text{sum of deviations from mean}^2}{\text{total data values} - 1}}$$