

Measures of Central Tendency and Variability

*Tools in Descriptive Statistics for
Arrays*

Measures of Central Tendency



most *representative or typical* of all values in a group

“average”

MODE

- most frequent data point
- mode exists as a data point
- unaffected by extreme values
- useful for qualitative data
- may have more than 1 value

MEDIAN

- value that divides ranked data points into halves: 50% larger than it, 50% smaller
- may not exist as a data point in the set
- influenced by position of items, but not their values

MEAN

$$\bar{x} = \frac{\sum x}{N}$$

- most stable measure
- affected by extreme values
- may not exist as a data point in the set

Measures of Variability



spread or shape of the data points in a group
“homogeneous or heterogeneous”

RANGE

$$R_E = HV - LV$$

$$R_I = UL_{HV} - LL_{LV}$$

- easily calculated
- unstable measure
- does not describe middle values
- good for small data sets

MEAN/AVERAGE DEVIATION

$$MD = \frac{\sum(|x_i - \bar{x}|)}{N}$$

- average of the absolute deviation of data points from the mean
- takes all values into consideration
- use when there are few extreme values

VARIANCE

$$\sigma^2 = \frac{\sum(X_i - \bar{X})^2}{N}$$

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{N - 1}$$

- amplifies differences between each data point and the mean
- stable measure

STANDARD DEVIATION

$$\sigma = \sqrt{\frac{\sum(X_i - \bar{X})^2}{N}}$$

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N - 1}}$$

- affected by all values
- most stable measure
- useful in inferential statistics