(I) How to find the Mean, Mode, & median?

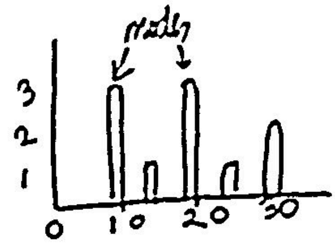
STEPS: #1 Put nos in ascending order, the number(s) that appear most = mode ("A" or "S") to represent series of numbers

#2 Add all the no.s (i.e. $\sum_{i=1}^{n} x_i$) — more than 1 possible

#3 Count the amount of numbers in the series (N)

#4 $\frac{\sum_{i=1}^{n} x_i}{N}$ = mean = average

#5 if N = odd
- YES → find number in middle of S = median
- NO → find the middle 2 no. of series & take their average = median



☆ Mean

what?
- average value
- measure of the center of dataset
- formula → ungrouped data: $\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots x_n}{n}$
  - grouped data: $\bar{x} = \frac{\sum f x}{n}$ — f = frequency in each class, x = midpoint of each class, n = total no. of classes

why?
- to summarize dataset
- most common statistic tool to measure central tendency / center of numerical dataset.
- why ∵ it incorporates score from every subject in the research study.
- useful ∵ without other tool, we can guess what next outcome can be (closely)

Appl'ns
- most common use in research, academics, sports.
  - eg in school: final grade is usually mean
  - Sports: eg baseball game, player's batting average represents total no. of hits divided by no. of times at bat = mean
- when unskewed or no outliers present
- filling missing data

Disadvantages
- Sensitive to extreme values/outliers (especially when sample size is small)
- not for nominal or nonnominal ordinal data.

☆ Median

what?
- it is a value separating higher half from lower half of data sample, population, or, probability distribution
- middle value of dataset

why?
- tells the middle of dataset
- when mean & median same → dataset is more or less evenly distributed from lowest to highest value.
- in presence of outliers, concepts of (mean, var, std-dev corrupted. ⇒ use median as it remains unaffected

Type of variable	Best measure of central tendency
Nominal	→ mode (nominal has mode)
ordinal	→ median (ordinal has median & mode)
Interval/Ratio (not skewed)	→ mean
Interval/ratio (skewed)	→ median

★ Standard Deviation

$$\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2\right)^2$$

Median Absolute Deviation

$$\text{median}\left[\underbrace{|x_i - \text{median}|}_{\text{dev}}\right]_{i=1}^{N}$$

abs-dev

→ not corrupted by outliers, else same as SD.