

central tendency → tendency of a given set of obs to cluster around a single set of obs or middle value and the single value that represents given facilitates comparison

mean

$\bar{X} = \frac{\sum x_i}{N}$
 $\bar{X} = \frac{\sum f_i x_i}{N}$
 $\bar{X} = A + \frac{\sum f_i d_i}{N}$
 $d_i = x_i - A$

• $\bar{X} = \frac{\sum x_i}{N}$
 • $\bar{X} = \frac{\sum f_i x_i}{N}$
 • $\bar{X} = A + \frac{\sum f_i d_i}{N}$
 • $\bar{X} = \frac{\sum x_i}{N}$
 • $\bar{X} = \frac{\sum f_i x_i}{N}$
 • $\bar{X} = A + \frac{\sum f_i d_i}{N}$

• $\bar{X} = \frac{\sum x_i}{N}$
 • $\bar{X} = \frac{\sum f_i x_i}{N}$
 • $\bar{X} = A + \frac{\sum f_i d_i}{N}$

Median

→ positional average
 → middle value of obs after arranging in ascending or descending order
 → median for odd = $\frac{(n+1)}{2}$ th
 → even = $\frac{1}{2} \left(\left(\frac{n}{2} \right) \text{th} + \left(\frac{n}{2} + 1 \right) \text{th} \right)$
 → avg of middle two value see eg. module 148

mode

→ maximum number of value is repeated
 → most common value
 → distribution having more than one mode is called multimodal
 → all obs occurs equal time no mode affected by sampling
 → does not depend on all obs
 → for grouped data
 mode = $L + \frac{\left(\frac{f_b - f_{-1}}{2f_b - f_{-1} - f_{+1}} \right) \times h}{1}$

→ $\bar{X} \rightarrow$ middle most (odd)
 → $\bar{X} \rightarrow$ two middle most (even)
 Properties:
 → all obs const k
 $\bar{X} = k$
 → for ungrouped only for 2+ve obs
 data $\sum (x_i - \bar{X}) = 0$
 → for grouped $\sum f_i (x_i - \bar{X}) = 0$

$\bar{X} = A + \frac{\sum f_i d_i}{N}$

→ For grouped data
 $m = L + \left(\frac{\frac{f}{2} - C}{f - C} \right) \times h$
 steps:
 1) data cont.
 2) $\frac{f}{2}$
 3) median class
 Properties:
 → $y = a + bx$
 $y_{me} = a + b x_{me}$
 $\sum (x_i - \text{median}) = \text{minimum}$
 → median is not affected by sampling
 Empirical relationship
 mean - mode = 3 (mean - median)
 → $y = a + bx$
 $y_{me} = a + b x_{me}$
 perfectly skewed mean = median
 DGA moderately skewed
 -ve skew
 +ve skew

$$\rightarrow \bar{x} = \frac{n_1 x_1 + n_2 x_2}{n_1 + n_2}$$

Weighted
 $Am = \frac{\sum w_i x_i}{\sum w_i}$

Weighted Am and Hm for first n natural number with their sq. as weight

$$A_{weighted} = \frac{3n(n+1)}{2(2n+1)}$$

$$H_{weighted} = \frac{2n+1}{3}$$

Median also meets all criteria except not based on all obs
 → does not allow itself to mathematical treatment.
 → most appropriate for

Review
 → best measure → Am
 (all criteria meets)
 → mean → stable & highly reliable avg
 Drawback → affected by sampling, cannot be applied for open end classification

2nd Quartile = 5th decile = 50th percentile

Quartiles → 4 = parts
 $\rightarrow Q_1, Q_2, Q_3$
 1st quartile / lower 11 / 2nd / 3rd / the highest
 Deciles → 10 = parts
 $\rightarrow D_1, D_2, \dots, D_9$
 Percentile → 100 = parts

$$P = \frac{n(n+1)}{4} + n \text{ obs (Rank)}$$

when
 Quartile → P = 1, 2, 3
 Decile → P = 1, 2, ..., 9
 Percentile → P = 1, 2, ..., 99

→ Find Quartiles for grouped frequency distribution we can draw the ogive less than type.

When all obs are same
~~mean~~
 $Am = G = H = Hm$
 when obs distinct
 $Am > G > Hm$

Criteria for central tendency
 → properly defined
 → easy to comprehend
 → based on all obs
 → have mathematical properties
 → least affected by extreme presence